Signed Support of Multivariate Polynomials and Applications

PhD thesis by

Máté László Telek

Department of Mathematical Sciences
University of Copenhagen
Denmark

PhD School of Science - Faculty of Science - University of Copenhagen

Máté László Telek
Department of Mathematical Sciences, University of Copenhagen
Universitetsparken 5, 2100 København Ø, Denmark
mlt@math.ku.dk

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen, Denmark, on the 31st of March 2024.

Academic advisor: Elisenda Feliu
University of Copenhagen, Denmark

Assessment Committee: Henrik Granau Holm (chair)
University of Copenhagen, Denmark

Frédéric Bihan
Université Savoie Mont Blanc, France

Cordian Riener
UiT The Artic University of Norway, Norway

ISBN 978-87-7125-227-9

Acknowledgements

First and foremost, I want to thank my supervisor, Elisenda Feliu, for her valuable guidance, important advice, continuous support, and engaging conversations. I am truly thankful for everything you have done. I wish to thank all current and former members of the Feliu group: Oskar Henriksson, Nidhi Kaihnsa, Beatriz Pascual Escudero, AmirHosein Sadeghimanesh, and Angélica Torres. Having you around made my PhD years amazing, and I am thankful for each and every one of you.

I also want to thank Maurice Rojas for hosting me at Texas A&M during my research stay abroad. You made my time in Texas a pleasant and unforgettable experience. Over the past three years, I have had numerous opportunities to meet researchers from the applied algebraic geometry community. I wish to thank all these people for the inspiring mathematical and non-mathematical discussions.

Thanks should also go to my friends, who were always by my side, encouraging me and never letting me down. Last but certainly not least, I must express my deepest gratitude to my parents for providing me with unfailing support and endless love throughout my years of study.

Máté L. Telek
Copenhagen, March 2024

Summary

This thesis includes six papers that investigate three different areas: chemical reaction network theory, Descartes' rule of signs, and real tropicalization. A common thread among them is the significant role played by the signed support of multivariate polynomials.

Paper I and II focus on chemical reaction networks. In Paper I, we describe a general algorithm for verifying connectivity of the parameter region of multistationarity of a reaction network and apply it to several biologically relevant networks. In Paper II, our focus is on two families of phosphorylation networks, called n-site phosphorylation networks. We provide a proof showing that their parameter region of multistationarity is connected for every $n \in \mathbb{N}_{\geq 2}$.

In Paper III and IV, we present combinatorial conditions on the signed support that provide upper bounds on the number of connected components of the set in the positive real orthant where the polynomial takes negative values. We frame this problem as a generalization of Descartes' rule of signs to multivariate polynomials. The methods developed in Paper III and IV are crucial for the arguments used in Paper I and II.

In Paper V, we investigate the real tropicalization of semi-algebraic sets and show its relation to the signed support of the polynomials defining these sets. In Paper VI, we study the signed A-discriminant and show that it has a simple structure if the signed support satisfies some combinatorial conditions. In such cases, Viro's patchworking becomes applicable for determining all isotopy types of hypersurfaces in the positive real orthant with a prescribed signed support for their defining polynomials.

Dansk resumé

Denne afhandling omfatter seks artikler inden for tre forskellige forskningsfelter: kemisk reaktionsnetværksteori, Descartes' fortegnsregel og reel tropikalisering. En fællesnævner er den afgørende betydning, som den støtte med angivet fortegn af polynomier i flere variable spiller.

Artikel I og II omhandler kemiske reaktionsnetværker. I artikel I beskrives en algoritme til verificering af hvorvidt parameterregionen for multistationaritet af et reaktionsnetværk er sammenhængende, hvilken anvendes på flere relevante netværker inden for biologi. I artikel II er vores fokus rettet mod to familier af fosforyleringsnetværk kaldet n-steder fosforylerings netværk. Vi fremlægger et bevis for sammenhængen af parameterregionen for multistationaritet for hvert $n \in \mathbb{N}_{\geq 2}$.

I artiklerne III og IV præsenterer vi kombinatoriske betingelser for den støtte med angivet fortegn, som giver et øvre estimat for antallet af sammenhængskomponenter af delmængden af den positive reelle ortant, hvor polynomiet antager negative værdier. Vi fortolker dette som en generalisering af Descartes' fortegnsregel til multivariate polynomier. Metoderne fra artikel III og IV er afgørende for argumenterne i artikel I og II.

I artikel V undersøger vi den reelle tropikalisering af semi-algebraiske mængder og sætter den i relation til den støtte med angivet fortegn af de polynomier, som definerer mængden. I artikel VI studerer vi den fortegn A-diskriminant og beviser, at den har en simpel struktur under visse kombinatoriske betingelser til den støtte med angivet fortegn. I sådanne tilfælde kan Viro's patchworking anvendes til at bestemme alle isotopityper af hyperflader i den positive reelle ortant med foreskrevet støtte med angivet fortegn for de definerende polynomier.

Contents

Papers

1 Introduction

This thesis focuses on the fascinating field of real algebraic geometry, where the interplay between algebra, geometry, and topology uncovers insights into the structure of real solutions of polynomial equations and inequalities. In this work, we investigate how the set of exponent vectors of a polynomial together with the signs of the coefficients, called the *signed support*, influences properties of semi-algebraic sets defined by the polynomial. Specifically, we study generalizations of Descartes' rule of signs to multivariate polynomials and real tropicalization of semi-algebraic sets. The motivation to investigate these questions arose from analyzing mathematical models of chemical reaction networks.

1.1 From reaction networks to real algebraic geometry

The mathematical foundations of chemical reaction network theory go back to works of Feinberg, Horn, and Jackson in the 1970's [41, 43, 61]. To illustrate the concept of a reaction network, consider the phosphorylation process of a substrate S. This biological mechanism is represented by the reaction network

$$\mathrm{S}_0 + \mathrm{E} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{ES}_0 \xrightarrow{\kappa_3} \mathrm{S}_1 + \mathrm{E}. \tag{1.1}$$

Each arrow is a reaction and the letters $\mathrm{E}, \mathrm{S}_0, \mathrm{S}_1, \mathrm{ES}_0$ correspond to biological species. In the first reaction, labeled by κ_1, the unphosphorylated substrate S_0 binds with a kinase E to form an intermediate species ES_0. This intermediate species ES_0 can then dissociate in two different ways: either producing a phosphorylated substrate S_1 or reverting back to the unphosphorylated form S_0 without attaching a phosphate group. This mechanism is known as the Michaelis-Menten mechanism in the literature [32].

To model the evolution of the concentration of the species of a given reaction network over time, several mathematical models can be used, including both stochastic and deterministic ones [93]. Chemical reaction network theory aims to investigate these models and comprehend their local and global dynamical behaviors. In this thesis, we focus on deterministic models using an Ordinary Differential Equation (ODE) system.

A first step often taken in the examination of qualitative properties of the dynamical system involves exploring its *steady states*, that is, finding solutions of the ODE system such that the concentration of each species is constant. From a biological perspective, the existence of at least two stable steady states has been linked to cellular decision-making, switching, and the memory of cells [75, 87].

Under the assumption of mass-action kinetics, the functions in the ODE system become parametrized polynomial functions, offering opportunities for the application of methods from computer algebra and real algebraic geometry [38, 44]. In particular, the question of *multistationarity*, that is, the existence of multiple steady states, boils down to whether a parametrized polynomial equation system has at least two solutions for some choice of the parameters. Since these solutions represent concentrations of species, one is only interested in positive real solutions. The number of such solutions might depend on the values of the parameters and there might be open regions in the parameter space giving rise to equation systems with different number of positive real solutions. Finding these regions and determining the number of positive solutions within each region is a challenging problem in real algebraic geometry.

Several methods have been developed to decide whether a reaction network exhibits multistationarity. Classical results providing sufficient conditions to preclude multistationarity include the Deficiency Zero and Deficiency One Theorem [42]. Over the past two decades, researchers have devoted significant attention to the so-called Injectivity Criterion, an alternative condition used to rule out multistationarity [5, 6, 33, 34, 49, 50, 67, 83, 84, 88]. We discuss this important criterion in more detail in Section 2.2.

Beyond precluding multistationarity, there are methods to verify whether a reaction network is multistationary for some choice of the parameters. One common approach involves identifying a subnetwork with at least two steady states and then "lifting" these steady states to the original network [35, 51, 68]. Additionally, methods based on polyhedral geometric conditions can also be used to find parameters for which the reaction network is multistationary [12, 16, 56, 57]. For a comprehensive overview of methods for determining multistationarity, we refer to the survey article [69].

The above-mentioned methods offer a way to determine whether the *parameter region of multistationarity* is non-empty, but they do not provide additional information about the shape or size of this region. Theoretically, a semi-algebraic description of the multistationarity region can be found using existing algorithms such as Cylindrical Algebraic Decomposition and Real Quantifier Elimination (see [29] for a survey). However, due to the high complexity of these algorithms, their applicability is limited to handling reaction networks of moderate size [24]. Numerical methods [48, 60] might provide information about certain parts of the parameter space for somewhat larger networks, yet these methods also have their limitations.

Instead of providing a complete description of the parameter region of multista-

tionarity, it might be interesting to explore some of its properties, such as its shape. Probably the first article in this direction was [85], where the authors investigated the connectivity of such a parameter region for a specific phosphorylation network using homotopy continuation and topological data analysis techniques. In the same article, it was suggested that lack of connectivity may indicate that different biological mechanisms underlie multistationarity.

A projection of the parameter region of multistationarity onto a subset of the parameters was proven to be connected in [45] for the *2-site phosphorylation network*

$$\begin{aligned} &\mathrm{S}_0 + \mathrm{E} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{ES}_0 \xrightarrow{\kappa_3} \mathrm{S}_1 + \mathrm{E} \underset{\kappa_8}{\overset{\kappa_7}{\rightleftharpoons}} \mathrm{ES}_1 \xrightarrow{\kappa_9} \mathrm{S}_2 + \mathrm{E} \\ &\mathrm{S}_2 + \mathrm{F} \underset{\kappa_{11}}{\overset{\kappa_{10}}{\rightleftharpoons}} \mathrm{FS}_2 \xrightarrow{\kappa_{12}} \mathrm{S}_1 + \mathrm{F} \underset{\kappa_5}{\overset{\kappa_4}{\rightleftharpoons}} \mathrm{FS}_1 \xrightarrow{\kappa_6} \mathrm{S}_0 + \mathrm{F}. \end{aligned} \tag{1.2}$$

Each phosphorylation and dephosphorylation step in (1.2) follows the same mechanism as described in (1.1). By increasing the number of phosphorylation sites, one obtains a family of reaction networks known as *n-site phosphorylation networks*, indexed by $n \in \mathbb{N}$. This family of reaction networks has been extensively studied, with particular focus on questions about the number of positive steady states [47, 52, 57, 59, 109]. Extending the connectivity result for the 2-site phosphorylation network, in [46] it was proven that the projected parameter region of multistationarity is connected for the n-site phosphorylation network for every n. The question, whether the same is true for the whole parameter region of multistationarity, became the catalyst for my research journey.

Our first breakthrough came a few months later after we started investigating this question. Based on the results in [31], we gave sufficient conditions on the *critical polynomial* of a reaction network (see Section 2.3) that imply connectivity of the parameter region of multistationarity. To verify one of the conditions, we needed to show that the set of points in the positive real orthant, where the critical polynomial takes negative values, is connected. For the 2-site phosphorylation network (1.2), the critical polynomial was large with 15 variables and 400 monomials. Therefore, computing the number of connected components of the preimage of the negative real line using the existing semi-algebraic algorithms was out of reach. We found ourselves in search of a computable sufficient condition that could guarantee the connectivity of the preimage of the negative real line under a polynomial function —a puzzle that set the stage for this thesis.

1.2 Real algebraic geometry

The field of algebraic geometry is devoted to studying properties of solution sets of polynomial equalities, called algebraic varieties. In cases where one is interested in so-

lutions in an algebraically closed field, such as the field of complex numbers, there are several classical results providing information about these solution sets. However, in many applications, including chemical reaction networks, robotics or computer vision, only the real solutions, or even only the positive real solutions, are of interest. The structure of the real and complex algebraic varieties might differ significantly, for example, the polynomial $x^{2024} - 1$ has 2024 complex solutions by the Fundamental Theorem of Algebra, but only two of the solutions are real, $x = 1$ and $x = -1$. In real algebraic geometry, the focus lies on these real solutions.

In 1637, Descartes published his famous result, Descartes' rule of signs, that provides an upper bound on the number of positive real solutions of a univariate real polynomial in terms of the number of sign changes in the coefficient sequence of the polynomial. The above example, $x^{2024} - 1$, has only one sign change in its coefficient sequence, so the bound given by Descartes' rule of signs is one. It is known that Descartes' bound is sharp, that is, for any given sign sequence, there exists a compatible choice of coefficients such that the polynomial has as many positive real solutions as the number of sign changes [58]. Moreover, the number of positive real solutions has the same parity as the number of sign changes [54], and Descartes' rule of signs is valid for polynomials with real exponents [36, 110].

One possible way to generalize Descartes' rule of signs to multivariate polynomials is to bound the number of positive real solutions of an equation system involving n polynomial equations in n variables, also called a *square system*. In 1991, Khovanskii gave such a bound in terms of n and the number of monomials appearing in the polynomials [71]. Khovanskii's bound has been improved in [17, 19] and for specific equation systems in [4, 9, 13, 72, 76].

Based on Viro's patchworking for complete intersections, proven by Sturmfels in 1996 [101], Itenberg and Roy formulated their famous conjecture regarding the maximum number of (positive) real solutions of a system of n multivariate polynomial equations in n variables [65]. Their conjecture was based on the combinatorial properties of the set of exponent vectors and the signs of the coefficients of the polynomials. They showed that their bound is a lower bound of a possible upper bound, that is, for any prescribed set of exponent vectors and signs of the coefficients, there exist n polynomials for which the number of their common (positive) real solutions matches the combinatorial bound. The first non-trivial example supporting the conjecture was proposed by Sturmfels and proven by Lagarias and Richardson in 1997 [74]. Almost at the same time, Li and Wang gave a counterexample to the Itenberg-Roy conjecture [77].

There are a few special cases when a Descartes-type bound is known for the number of positive real solutions of a square system. The Injectivity Criterion [83], mentioned in Section 1.1, provides a condition when the number of positive real solutions is at most one. For systems involving n variables, n polynomials and a total number of $n + 2$ monomials, a sharp upper bound on the number of positive solutions was given

in [10, 11]. This bound is in terms of the number of sign changes of a certain sequence of numbers associated with the polynomial system.

An alternative approach to generalizing Descartes' rule of signs to the multivariate setting is to consider a polynomial in n variables and to bound topological invariants of the hypersurface defined by the positive real zero set of the polynomial. Upper bounds for the sum of the Betti-numbers were given in [18, 71], while bounds on the number of connected components of the hypersurface in [14, 15, 53, 76, 89]. Like Khovanskii's bound for square systems, these bounds rely on the number of variables and monomials in the polynomial.

For our applications, as discussed at the end of Section 1.1, we wished to have a similar bound. However, instead of bounding the number of connected components of a hypersurface, we aimed to bound the number of connected components of the complement of the hypersurface where the defining polynomial takes negative values. In particular, we were interested in the case where this upper bound is one. Our findings in that direction have been published in Paper III and Paper IV. We refer to Section 1.4 for a brief summary and Chapter 3 for a detailed discussion of the results.

It is worth mentioning that results regarding Descartes' rule of signs for the hypersurface case might have implications for square systems. For example, one of the techniques used in Paper III enables to derive novel bounds on the number of positive real solutions of polynomial equation systems with specific signed support. This, again, can be viewed as a generalization of Descartes' rule of signs. We provide these bounds in Section 3.4.

1.3 Tropical geometry

During the past two years, as I presented our results from Paper III at conferences and workshops, I often received comments that our results have a "tropical flavor". Inspired by these remarks, I started to explore the connection between our work and real tropical geometry.

One of the primary goals of tropical geometry is to establish a connection between algebraic and polyhedral geometry, enabling to transform an algebraic variety into a polyhedral object, called *tropical variety*, that mimics essential properties of its algebraic counterpart. This approach has been successful for varieties defined over algebraically closed fields with non-trivial valuation, such as the field of complex Puiseux series [64, 78].

Tropicalization of varieties over the field of real numbers traces back to Viro's patchworking [105]. For a fixed signed support, Viro described a method to construct a polyhedral complex, known as the *Viro diagram*, which has the same isotopy type as the

positive real zero set of *some* polynomial matching the signed support (see Section 4.3). In Paper VI, we showed that under conditions on the signed support similar to those in Paper III, Viro diagrams can be used to find *all* possible isotopy types, that is, for *all* polynomials matching the signed support, the positive real hypersurface is isotopic to one Viro diagram. It is important to note that this statement is not true in general, for a counterexample, we refer to Example VI.2.11.

The tropicalization of semi-algebraic sets goes back to the work by Alessandrini [1]. In his paper, Alessandrini investigated the logarithmic limit of semi-algebraic sets defined in the positive orthant $\mathbb{R}^n_{>0}$ and showed that it coincides with the real tropicalization. We will recall this fundamental result in Theorem 4.7. Only in certain special cases it is known how to compute the real tropicalization of semi-algebraic sets [2, 20, 25, 102, 104]. However, there exist polyhedral complexes that serve as approximations of these, either by containing the real tropicalization or being contained within it [20, 66, 104]. Such approximations are easier to compute and under certain additional assumptions they coincide with the real tropicalization. The proof techniques used in Paper III and Paper IV, allowed us to derive a novel approximation of real tropicalization and to identify cases when these two objects coincide. This is the content of Paper V.

1.4 Contribution to the state of the art

The contributions presented in this thesis can be divided into three main parts, focusing on the parameter region of multistationarity of a reaction network, on generalizing Descartes' rule of signs to multivariate polynomials, and on tropicalization of semi-algebraic sets. The signed support played a central role in each of these three topics.

Paper I contains an algorithm that checks a sufficient condition for the connectivity of the parameter region of multistationarity. The algorithm is based on Theorem 2.14 that relates the preimage of the negative real line under the critical polynomial to the parameter region of multistationarity. We applied the method to several biologically relevant networks, including the n-site phosphorylation network for $n = 2, 3$ (cf. (1.2)), and showed that their parameter region of multistationarity is connected. The case of $n > 3$ required a deeper investigation and new techniques. In Paper II, we showed that the critical polynomial for the n-site phosphorylation network can be written recursively. We used this recursive formula to show inductively that the parameter region of multistationarity is connected for the n-site phosphorylation network for every $n \in \mathbb{N}_{\geq 2}$.

Paper III and IV provide conditions on the signed support of a multivariate polynomial such that the set of points in the positive orthant $\mathbb{R}^n_{>0}$ where the polynomial takes negative values has at most one or two connected components. We phrased these results as generalization of Descartes' rule of signs to hypersurfaces. Furthermore, in

Paper IV it is shown that the problem of finding the number of such connected components can be reduced to the same problem for a polynomial in fewer monomials if all the exponent vectors with negative coefficients are contained in a face of the Newton polytope. A similar reduction is possible if the Newton polytope has two parallel faces containing all the exponent vectors. As an addition to the results in Paper III and IV, in Section 3.4 we describe conditions on the signed support of n polynomials in n variables such that these polynomials have at most 2 or infinitely many common positive real solutions. These statements can be seen as (partial) generalization of Descartes' rule of signs to square systems.

Paper V contains results about the real tropicalization of semi-algebraic subsets of $\mathbb{R}^n_{>0}$. We give a self-contained proof for the result that the negative normal cone (see Section 4.2 for a precise definition) approximates the real tropicalization. This proof is also valid for polynomials with real exponents. Furthermore, Paper V describes a cone that provides a better approximation of the real tropicalization, and that might be computable by means of existing algorithms. Moreover, in Paper V we discuss certain scenarios when these approximations coincide with the real tropicalization. Paper VI investigates the signed A-discriminant of a signed support. The connection to tropical geometry comes from the observation that if the signed A-discriminant has a "simple structure", then the isotopy types of the positive real hypersurfaces defined by polynomials matching the signed support can be described by a polyhedral object, the Viro diagram. In Paper VI, we provide conditions on the signed support that ensure that the signed A-discriminant has such a "simple structure".

1.5 Structure of the thesis

This thesis is based on six articles, which are collected after the four introductory chapters. The purpose of these chapters is to motivate and highlight the main contributions of the papers. Each paper has its own bibliography. The references used in Chapter 1-4 can be found in the bibliography after Chapter 4.

Chapter 2 starts by giving the necessary background on chemical reaction networks. We define the parameter region of multistationarity and discuss how the Jacobi determinant of a certain function can be used to gain information about this parameter region. The last part of Chapter 2 summarizes the results of Paper I and Paper II. In Chapter 3, we turn our investigation to Descartes' rule of signs. We recall this classical theorem for univariate polynomials and discuss possible generalizations to the multivariate setting. In Section 3.3, we discuss partial generalizations to the hypersurface case, based on the results in Paper III and Paper IV. In Section 3.4, we investigate a generalization of Descartes' rule of signs to square systems, these results have not been included in any publication. Section 4.1 and 4.2 contain a brief introduction to tropi-

cal geometry along with an overview of recent results on real tropicalization including findings discussed in Paper V. Section 4.3 is based on Paper VI, where we introduce the signed A-discriminant and discuss its relation to Viro's patchworking.

Last but not least we fix some notation we will use throughout the thesis. The cardinality of a finite set S will be denoted by $|S|$. For the index set $\{1, \dots, n\}$ we write $[n]$ for short, and for $I \subseteq [n]$ we write I^c for the complement of I in $[n]$. For two vectors $v, w \in \mathbb{R}^n$, $v \cdot w$ denotes their Euclidean scalar product and $v * w$ denotes their coordinate-wise product. The transpose of a matrix $M \in \mathbb{R}^{n \times m}$ is denoted by $M^\top$. For a differentiable function $f \colon \mathbb{R}^n \to \mathbb{R}^m$, we write $J_f(x)$ for the Jacobian matrix at $x \in \mathbb{R}^n$, which is the matrix of size $m \times n$ whose (i,j)-th entry equals the partial derivative $\frac{\partial f_i(x)}{\partial x_j}$. In the special case $m = 1$, we denote the Jacobian matrix $J_f(x)$ by $\nabla f(x)$.

2 Chemical reaction networks

The goal of this chapter is to provide more information and context for the statements discussed in Section 1.1. Specifically, we recall a formal definition of a reaction network and describe the ODE system that models the change in the concentration of the species in the network over time. We define the parameter region of multistationarity and discuss several methods for gaining information about it. Furthermore, we give an overview of the results of Paper I and Paper II.

2.1 Basic definitions

A (chemical) reaction network $(\mathcal{S}, \mathcal{R})$ is a collection of *reactions* $\mathcal{R} = \{R_1, \dots, R_r\}$ between *species* in a set $\mathcal{S} = \{\mathrm{X}_1, \dots, \mathrm{X}_n\}$. Each reaction is given by a linear combination of the species, that is, each reaction has the form

$$R_j\colon \quad \sum_{i=1}^{n} a_{i,j}\mathrm{X}_i \longrightarrow \sum_{i=1}^{n} b_{i,j}\mathrm{X}_i, \qquad j = 1, \dots, r,$$

where $a_{i,j}, b_{i,j}$ are non-negative integers, called *stoichiometric coefficients*. The species on the left-hand (resp. right-hand) side of a reaction with non-zero stoichiometric coefficient are called *reactants* (resp. *products*) of the reaction. The *reactant matrix* $A \in \mathbb{Z}^{n\times r}$ keeps track of the reactant species. Specifically, its (i,j)-th entry is given by the stoichiometric coefficients $a_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, r$. The *stoichiometric matrix* $N \in \mathbb{Z}^{n\times r}$ encodes the net production of the reactions. The (i,j)-th entry of N equals $b_{i,j} - a_{i,j}$, $i = 1, \dots, n$, $j = 1, \dots, r$.

Under the assumption of mass-action kinetics, introduced by Guldberg and Waage in 1864 [107, 108], the *reaction rate function* equals

$$v_\kappa(x) = \operatorname{diag}(\kappa)x^A = (\kappa_1 x_1^{a_{1,1}} \cdots x_n^{a_{n,1}}, \dots, \kappa_r x_1^{a_{1,r}} \cdots x_n^{a_{n,r}})^\top.$$

Here, x_i denotes the concentration of the species X_i and $\kappa = (\kappa_1, \dots, \kappa_r)^\top \in \mathbb{R}^r_{>0}$ is a vector of parameters called *reaction rate constants*. The ODE system modeling the

evolution of the concetration of the species over time has the form:

$$\dot{x} = f_\kappa(x), \qquad x \in \mathbb{R}^n_{\geq 0}, \tag{2.1}$$

where $f_\kappa(x) := N v_\kappa(x)$.

The positive and the non-negative orthants $\mathbb{R}^n_{>0}$, $\mathbb{R}^n_{\geq 0}$ are forward invariant with respect to the ODE system (2.1) [99, 106]. In other words, if the initial condition is contained in these orthants, the entire trajectory remains contained in them. This property is particularly important, since the trajectories encode concentration of species that should not be negative numbers.

Moreover, the trajectories of (2.1) are contained in certain affine subspaces. To see this, consider the *stoichiometrix subspace* S defined as the column space of the matrix N. Let s denote the rank of N and choose a full-rank matrix $W \in \mathbb{R}^{d\times n}$ whose rows form a basis of the left kernel of N. Note that by definition, we have $d = n - s$ and $\ker(W) = \operatorname{im}(N)$. Since $w \cdot \dot{x} = w \cdot (N v_\kappa(x)) = 0$ for every vector w in the left kernel of N, it follows immediately that the trajectories of (2.1) with initial condition $x(0) \in \mathbb{R}^n_{\geq 0}$ are contained in the affine subspace $x(0)+S$. A *stoichiometric compatibility class* is a set of the form

$$\mathcal{P}_c := \{x \in \mathbb{R}^n_{\geq 0} \mid Wx = c\} = \big(x(0) + S\big) \cap \mathbb{R}^n_{\geq 0},$$

for $c = Wx(0)$. This vector is called the *vector of total concentrations*, and it is usually treated as a parameter. The defining equations of $\mathcal{P}_c$ are called *conservation laws*, and the matrix W is called a *conservation matrix*.

Example 2.1. To illustrate the above definitions, we consider the following reaction network

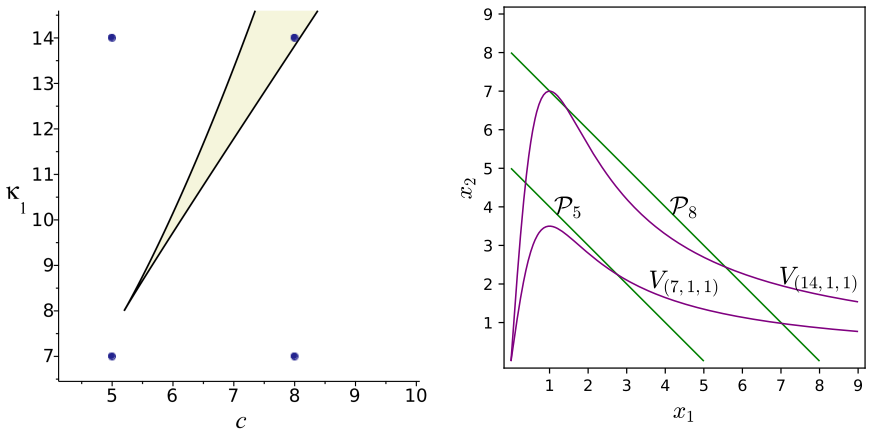
$$\mathrm{X}_1 \xrightarrow{\kappa_1} \mathrm{X}_2, \qquad \mathrm{X}_2 \xrightarrow{\kappa_2} \mathrm{X}_1, \qquad 2\mathrm{X}_1 + \mathrm{X}_2 \xrightarrow{\kappa_3} 3\mathrm{X}_1, \tag{2.2}$$

which served as the running example in Paper I. Its reactant, stoichiometric and conservation matrices have the form

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & 1 \end{pmatrix},$$

and the corresponding ODE system and the conservation laws are

$$\begin{aligned} \dot{x}_1 &= -\kappa_1 x_1 + \kappa_2 x_2 + \kappa_3 x_1^2 x_2, \qquad x_1 + x_2 = c, \\ \dot{x}_2 &= \kappa_1 x_1 - \kappa_2 x_2 - \kappa_3 x_1^2 x_2. \end{aligned}$$



(a) (b)

Figure 2.1: (a) Parameter region of multistationarity for the reaction network (2.2) sliced by the hyperplane $\kappa_2 = 1, \kappa_3 = 1$. (b) Steady state varieties (purple curves) and stoichiometric compatibility classes (green lines) for different choices of parameters corresponding to the points marked by blue dots in (a).

For fixed reaction rate constants $\kappa = (\kappa_1, \ldots, \kappa_r)^\top \in \mathbb{R}^r_{>0}$, the set of non-negative steady states of the ODE system (2.1)

$$V_\kappa := \{x \in \mathbb{R}^n_{\geq 0} \mid Nv_\kappa(x) = 0\}$$

is called the *steady state variety*. A pair of parameters (κ, c) *enables* multistationarity if the intersection $V_\kappa \cap \mathcal{P}_c \cap \mathbb{R}^n_{>0}$ contains at least two points.

Definition 2.2. Let $(\mathcal{S}, \mathcal{R})$ be a reaction network. The *parameter region of multistationarity* is defined as

$$\Omega := \{(\kappa, c) \in \mathbb{R}^r_{>0} \times \mathbb{R}^d \mid (\kappa, c) \text{ enables multistationarity}\}.$$

Example 2.3. The reaction network from (2.2) is small enough to apply Cylindrical Algebraic Decomposition [26] to decompose the parameter space into semi-algebraic sets, called *cells*, where the number of positive solutions of $f_\kappa(x) = 0$, $Wx = c$ is constant. We used the `Maple` [79] function `CellDecomposition` from the package `Parametric` [98] to compute such a cell decomposition. The code returned only one open cell with at least two positive solutions. This cell is given by the inequalities

$$\kappa_2 > 0, \quad \kappa_3 > 0, \quad \kappa_1 > 8\kappa_2, \quad \xi_3 < c < \xi_4, \tag{2.3}$$

where ξ_3, ξ_4 denote the 3rd and 4th root of the polynomial

$$4c^4\kappa_2\kappa_3^2 - c^2\kappa_1^2\kappa_3 - 20c^2\kappa_1\kappa_2\kappa_3 + 8c^2\kappa_2^2\kappa_3 + 4\kappa_1^3 + 12\kappa_1^2\kappa_2 + 12\kappa_1\kappa_2^2 + 4\kappa_2^3.$$

For fixed parameters $\kappa_2 = 1, \kappa_3 = 1$, this cell is depicted in Figure 2.1(a). The parameter pair $((14, 1, 1), 8)$ lies in the cell given by the inequalities in (2.3). Figure 2.1(b) shows the steady state variety $V_{(14,1,1)}$ and the stoichiometric compatibility class $\mathcal{P}_8$. Their intersection contains 3 points in $\mathbb{R}^2_{>0}$, thus $((14, 1, 1,), 8)$ enables multistationarity.

If $\kappa_2 = 1$, then from the inequalities in (2.3) follows that for $\kappa_1 < 8$ the parameter pair $((\kappa_1, 1, \kappa_3), c)$ does not enable multistationarity independently of the choices of κ_3 and c. Figure 2.1(b) illustrates that the steady state variety $V_{(7,1,1)}$ has at most one intersection point with $\mathcal{P}_c$ for all $c \in \mathbb{R}$.

The equation system $f_\kappa(x) = 0$, $Wx = c$, defining the points in $V_\kappa \cap \mathcal{P}_c$, is overdetermined, that is, it has more equations than variables. Following [31], we reduce it to a system in n equations and n variables as follows. Assume that the conservation matrix W is row reduced and let $i_1 < \cdots < i_d$ be the indices of the first non-zero coordinates of each row of W. We denote by $F_{\kappa,c}(x)$ the function obtained from $f_\kappa(x)$ by replacing the i_j-th entry $f_{\kappa,i_j}(x)$ by the affine linear function $w_j \cdot x - c_j$, $j = 1, \ldots, d$, where w_j denotes the j-th row of W. In [31], it was shown that

$$V_\kappa \cap \mathcal{P}_c = \big\{x \in \mathbb{R}^n_{\geq 0} \mid F_{\kappa,c}(x) = 0\big\}.$$

Since the rows of the Jacobian matrix $J_{f_\kappa}(x)$ are given by the gradients $\nabla f_{\kappa,i}(x)$ $i = 1, \ldots, n$, it follows that

$$J_{F_{\kappa,c}}(x) = M_\kappa(x), \tag{2.4}$$

where $M_\kappa(x)$ denotes the matrix that is obtained from $J_{f_\kappa}(x)$ by replacing the rows $i_1, \ldots, i_d$ by the rows of W. The matrix $M_\kappa(x)$ does not depend on the parameter c.

2.2 Jacobian critera for multistationarity

In this section, we recall several criteria that either preclude or verify multistationarity. These criteria are based on the Jacobian matrix of certain functions associated with the reaction network. First, consider the function

$$f_\kappa \colon \mathbb{R}^n_{>0} \to \mathbb{R}^n, \quad x \mapsto f_\kappa(x) = N \operatorname{diag}(\kappa) x^A$$

as introduced in (2.1). If f_κ is injective on the stoichiometric compatibility class $\mathcal{P}_c$, that is, $f_\kappa(x) \neq f_\kappa(x')$ for $x, x' \in \mathcal{P}_c$ with $x \neq x'$, then $V_\kappa \cap \mathcal{P}_c \cap \mathbb{R}^n_{>0}$ contains at most one point, and the pair (κ, c) does not enable multistationarity. This simple observation yields that the parameter region of multistationarity is empty if f_κ is injective on $\mathcal{P}_c$ for all $\kappa \in \mathbb{R}^r_{>0}$ and all $c \in \mathbb{R}^d$. In [49, Theorem 5.6], injectivity of f_κ has been related to injectivity of the Jacobian matrix $J_{f_\kappa}(x)$.

Theorem 2.4. *[49, Theorem 5.6] Let $(\mathcal{S}, \mathcal{R})$ be a reaction network, and let $S = \mathrm{im}(N)$ be its stoichiometric subspace. The following are equivalent:*

(inj) *f_κ is injective on $\mathcal{P}_c$ for all $\kappa \in \mathbb{R}^r_{>0}$ and all $c \in \mathbb{R}^d$.*

(jac) *$J_{f_\kappa}(x)$ is injective on S for all $\kappa \in \mathbb{R}^r_{>0}$ and all $x \in \mathbb{R}^n_{>0}$.*

Theorem 2.4 extends a previously known Injectivity Criterion for so-called *CFSTR* networks [33], for which $S = \mathbb{R}^n$. Later, Theorem 2.4 has been generalized to power-law kinetics [111], as well as to more general functions, similar to f_κ, but not necessarily related to reaction networks [83].

In the following, we elaborate on the condition (jac) in Theorem 2.4. First, we show that the set of such Jacobian matrices can be reparametrised. Let $A_1, \ldots, A_r$ denote the columns of the reactant matrix A. We will write $*$ for the coordinate-wise product of two vectors, and $\frac{1}{x} = (\frac{1}{x_1}, \ldots, \frac{1}{x_n})^\top$. Using this notation, for $v_\kappa(x) = \mathrm{diag}(\kappa)x^A$ we have

$$J_{v_\kappa}(x) = \mathrm{diag}(\kappa * x^A)A^\top \mathrm{diag}(\tfrac{1}{x}).$$

From the Chain Rule, it follows that

$$J_{f_\kappa}(x) = N J_{v_\kappa}(x) = N\,\mathrm{diag}(\kappa * x^A)A^\top \mathrm{diag}(\tfrac{1}{x}).$$

By setting $v = \kappa * x^A$ and $h = \frac{1}{x}$, we obtain the following lemma.

Lemma 2.5. *[83, Lemma 2.7] Let $(\mathcal{S}, \mathcal{R})$ be a reaction network. Let N be the stoichiometric matrix and A be the reactant matrix, as introduced in Section 2.1. Then it holds that*

$$\left\{J_{f_\kappa}(x) \mid \kappa \in \mathbb{R}^r_{>0},\, x \in \mathbb{R}^n_{>0}\right\} = \left\{N\,\mathrm{diag}(v)A^\top \mathrm{diag}(h) \mid v \in \mathbb{R}^r_{>0},\, h \in \mathbb{R}^n_{>0}\right\}.$$

From Lemma 2.5, it follows immediately that condition (jac) in Theorem 2.4 is equivalent to the condition

$$(lin) \qquad \ker(N\,\mathrm{diag}(v)A^\top \mathrm{diag}(h)) \cap S = \{0\} \quad \text{for all } (v,h) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}.$$

Recall that the stoichiometric subspace $S = \mathrm{im}(N)$ equals the kernel of the conservation matrix W. Therefore, $y \in \mathbb{R}^n$ is contained in $\ker(N\,\mathrm{diag}(v)A^\top \mathrm{diag}(h)) \cap S$ if and only if $N\,\mathrm{diag}(v)A^\top \mathrm{diag}(h)y = 0$ and $Wy = 0$. The rows of N might be linearly dependent. Similar to the reduction described at the end of Section 2.1, we systematically remove these linear dependences. Following [49], we assume that W has the form $W = \begin{pmatrix}\mathrm{Id}_d & \widetilde{W}\end{pmatrix}$, where Id_d denotes the $d \times d$ identity matrix and $\widetilde{W} \in \mathbb{R}^{d\times(n-d)}$. Such

a choice of W is always possible after reordering the species of the network. Let N' denote the matrix obtained from N by deleting the first d rows. We define the matrix

$$\Gamma_{v,h} := \begin{pmatrix} W \\ N' \operatorname{diag}(v) A^\top \operatorname{diag}(h) \end{pmatrix}.$$

Note that $\Gamma_{v,h}$ equals the Jacobian $M_\kappa(x)$ of the function $F_{\kappa,c}$ (cf. (2.4)) for $v = \kappa * x^A$ and $h = \frac{1}{x}$.

It holds that $N \operatorname{diag}(v) A^\top \operatorname{diag}(h) y = 0$, $Wy = 0$ if and only if $(\Gamma_{v,h})y = 0$. We refer to [49, Corollary 4.8] for a proof of this fact. Since $\Gamma_{v,h}$ is a square matrix, $\ker(\Gamma_{v,h}) = \{0\}$ if and only if $\det(\Gamma_{v,h}) \neq 0$. Therefore, the condition (*lin*) is equivalent to

$$(det') \qquad \det(\Gamma_{v,h}) \neq 0 \quad \text{for all } (v,h) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}.$$

For a matrix M of size $k \times m$ and two sets of indices $I \subseteq [k]$, $J \subseteq [m]$, we denote by $[M]_{I,J}$ the matrix obtained from M by selecting the entries in rows indexed by I, and columns indexed by J. Using the Laplace expansion on complementary minors (see e.g. [90, Theorem 2.4.1]), we rewrite $\det(\Gamma_{v,h})$ as

$$\sum_{\substack{I \subseteq [n] \\ |I| = s}} (-1)^{\sum_{j=d+1}^{n} j + \sum_{i \in I} i} \det\left([W]_{[d],I^c}\right) \det\left([N' \operatorname{diag}(v) A^\top]_{[s],I}\right) \prod_{i \in I} h_i. \tag{2.5}$$

Now, we apply the Cauchy-Binet formula (see e.g. [90, Theorem 2.3]) to get for each $I \subseteq [n]$, $|I| = s$:

$$\det\left([N' \operatorname{diag}(v) A^\top]_{[s],I}\right) = \sum_{\substack{J \subseteq [r] \\ |J| = s}} \det([N']_{[s],J}) \det([A]_{I,J}) \prod_{j \in J} v_j. \tag{2.6}$$

From (2.5) and (2.6), it follows that $\det(\Gamma_{v,h})$ is a homogeneous polynomial in v and h. Furthermore, the exponent vector of each of its monomials has only 0 or 1 entries. For such a polynomial, it is simple to decide whether it has a root in the positive orthant. Specifically, such a polynomial has a positive root if and only if it has both positive and negative coefficients [83, Lemma 2.12].

Building upon the above discussions, Theorem 2.4 can be extended with the following additional equivalent conditions.

Theorem 2.6. *[49, 83] Let $(\mathcal{S}, \mathcal{R})$ be a reaction network with stoichiometric subspace $S = \operatorname{im}(N)$. Assume that the conservation matrix W has the form $W = \begin{pmatrix} \mathrm{Id}_d & \widetilde{W} \end{pmatrix}$. The following are equivalent:*

(inj) f_κ is injective on $\mathcal{P}_c$ for all $(\kappa, c) \in \mathbb{R}^r_{>0} \times \mathbb{R}^d$.

(jac) $J_{f_\kappa}(x)$ *is injective on* S *for all* $(\kappa, x) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}$.

(lin) $\ker(N \operatorname{diag}(v) A^\top \operatorname{diag}(h)) \cap S = \{0\}$ *for all* $(v, h) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}$.

(det) *Viewed as a polynomial in* v *and* h, $\det(\Gamma_{v,h})$ *is non-zero and all of its non-zero coefficients have the same sign.*

(det′) $\det(\Gamma_{v,h}) \neq 0$ *for all* $(v, h) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}$.

(det″) $\det(M_\kappa(x)) \neq 0$ *for all* $(\kappa, x) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}$.

In [83], two additional conditions are described, which are equivalent to the conditions in Theorem 2.6. One of these conditions relies on minors of size $s \times s$ of the matrices N and A. The other condition is based on the signs of the vectors in $\ker(N)$ intersected with a specific subset of the row span of A. Since we do not use these conditions in this thesis and prefer to keep the explanation as simple as possible, we choose to omit them.

By Theorem 2.6, the determinant of $M_\kappa(x)$ can preclude multistationarity. In the following, we discuss a theorem from [31] that uses $\det(M_\kappa(x))$ both to verify and to preclude multistationarity in a reaction network. This theorem is applicable for reaction networks satisfying certain assumptions, which we will briefly recall.

A reaction network is *conservative* if each stoichiometric compatibility class is a compact set. This is equivalent to the existence of a vector in the left kernel of N with positive coordinates [8], which is easy to check using linear programming. A milder condition is that for every stoichiometric compatibility class $\mathcal{P}_c$, there is a compact set $K_c \subseteq \mathcal{P}_c$ such that all the trajectories of the ODE system (2.1) starting in $\mathcal{P}_c$ enter the compact set K_c in finite time and do not leave again. A network with this property is called *dissipative*. In general, it is hard to verify whether a reaction network is dissipative, see the discussion in [31, p. 11-12]. For simplicity, we consider only conservative networks in our applications.

A steady state $x \in V_\kappa \cap \mathcal{P}_c$ is called a *relevant boundary steady state* if some of the coordinates of x are zero and $\mathcal{P}_c \cap \mathbb{R}^n_{>0} \neq \emptyset$.

Theorem 2.7. *[31, Theorem 1] Assume that the reaction network is dissipative without relevant boundary steady states and let* $s = \operatorname{rk}(N)$. *Then for each parameter pair* $(\kappa, c) \in \mathbb{R}^r_{>0} \times \mathbb{R}^d$ *it holds:*

(A) *If* $(-1)^s \det(M_\kappa(x)) > 0$ *for all* $x \in V_\kappa \cap \mathcal{P}_c \cap \mathbb{R}^n_{>0}$, *then* (κ, c) *does not enable multistationarity.*

(B) *If* $(-1)^s \det(M_\kappa(x)) < 0$ *for some* $x \in V_\kappa \cap \mathcal{P}_c \cap \mathbb{R}^n_{>0}$, *then* (κ, c) *enables multistationarity.*

Unlike Theorem 2.6, in Theorem 2.7, the determinant of $M_\kappa(x)$ is evaluated only at points in the *incidence variety*

$$\mathcal{V} := \Big\{(x,\kappa) \in \mathbb{R}^n_{>0} \times \mathbb{R}^r_{>0} \mid x \in V_\kappa\Big\}. \tag{2.7}$$

In Section 2.3, we introduce a systematic way how to parametrize $\mathcal{V}$ using so-called *convex parameters*. Here, at the end of this section, we discuss another type of parametrization of $\mathcal{V}$. Such a parametrization is obtained by solving some of the equations $f_\kappa(x) = 0$ for a subset of the variables. To illustrate this, we consider the following example.

Example 2.8. We revisit the reaction network from (2.2). Its steady state variety is given by the equations:

$$-\kappa_1 x_1 + \kappa_2 x_2 + \kappa_3 x_1^2 x_2 = 0, \qquad \kappa_1 x_1 - \kappa_2 x_2 - \kappa_3 x_1^2 x_2 = 0.$$

By solving the second equation for x_2, we obtain the parametrization

$$\Psi\colon \mathbb{R}^4_{>0} \to \mathcal{V}, \quad (x_1,\kappa_1,\kappa_2,\kappa_3) \mapsto (x_1, \tfrac{\kappa_1 x_1}{\kappa_3 x_1^2+\kappa_2}, \kappa_1,\kappa_2,\kappa_3).$$

The network (2.2) is conservative, since $(1,1)$ is contained in the left kernel of N. It is easy to check that the only non-negative steady state with zero coordinates is $(0,0)$. The stoichiometric compatibility class containing $(0,0)$ does not intersect $\mathbb{R}^2_{>0}$. Thus, the network (2.2) does not have relevant boundary steady states and Theorem 2.7 applies.

Evaluating $\det(M_\kappa(x))$ at $(x,\kappa) = \Psi(x_1,\kappa_1,\kappa_2,\kappa_3)$, we get the rational function

$$\tfrac{\kappa_3^2 x_1^4 + (2\kappa_2\kappa_3 - \kappa_1\kappa_3)x_1^2 + \kappa_1\kappa_2 + \kappa_2^2}{\kappa_3 x_1^2 + \kappa_2}, \tag{2.8}$$

whose denominator is always positive. For the numerator of (2.8) to take negative values, we need $2\kappa_2\kappa_3 - \kappa_1\kappa_3 < 0$. Under this assumption, it is a so-called circuit polynomial and by [46, Theorem 2.2] (which is a direct specialization of [63, Theorem 3.8]) it takes negative values if and only if

$$-2\kappa_2\kappa_3 + \kappa_1\kappa_3 > \sqrt{2\kappa_3^2}\sqrt{2(\kappa_1\kappa_2 + \kappa_2^2)},$$

which is equivalent to

$$0 < (-2\kappa_2\kappa_3 + \kappa_1\kappa_3)^2 - 4\kappa_3^2(\kappa_1\kappa_2 + \kappa_2^2) = (\kappa_1 - 8\kappa_2)\kappa_1\kappa_3^2.$$

From this follows that (2.8) is negative for some $x_1 > 0$ if and only if $\kappa_1 > 8\kappa_2$. Now, we apply Theorem 2.7 to conclude that the projection of the parameter region of multistationarity onto the parameters $(\kappa_1,\kappa_2,\kappa_3)$ is given by the inequality $\kappa_1 > 8\kappa_2$. Specifically, for $\kappa \in \mathbb{R}^3_{>0}$ there exists $c \in \mathbb{R}$ such that (κ, c) enables multistationarity if and only if $\kappa_1 > 8\kappa_2$. Note that this is the same region as we computed in Example 2.3 using Cylindrical Algebraic Decomposition.

2.3 The critical polynomial

The aim of this section is to describe a convenient parametrization, called *convex parametrization*, of the incidence variety $\mathcal{V}$ (cf. (2.7)). Using this parametrization, we evaluate $\det(M_\kappa(x))$ (cf. Theorem 2.7) at points $(\kappa, x) \in \mathcal{V}$, and call the resulting function the *critical polynomial*.

Convex parameters in the context of chemical reaction networks were introduced by Clarke [28]. The idea behind such a parametrization is the observation that for $(\kappa, x) \in \mathbb{R}^r_{>0} \times \mathbb{R}^n_{>0}$ we have $x \in V_\kappa$ if and only if $\operatorname{diag}(\kappa)x^A \in \ker(N) \cap \mathbb{R}^r_{>0}$. The set $\ker(N) \cap \mathbb{R}^r_{\geq 0}$ is a closed convex pointed polyhedral cone, called the *flux cone*. Let $E^{(1)}, \dots, E^{(\ell)} \in \mathbb{R}^r$ denote a choice of the extreme vectors of the flux cone, and write them as columns of a matrix $E \in \mathbb{R}^{r \times \ell}$. Such a choice of extreme vectors is unique up to multiplication by a positive scalar. Every element in $\ker(N) \cap \mathbb{R}^r_{\geq 0}$ can be written as a non-negative linear combination of the extreme vectors. If E does not contain any zero row, by Proposition I.6.1 we have

$$\ker(N) \cap \mathbb{R}^r_{>0} = \Big\{ E\lambda = \sum_{i=1}^{\ell} \lambda_i E^{(i)} \mid \lambda \in \mathbb{R}^\ell_{>0} \Big\}.$$

The assumption that E does not have any zero row, is equivalent to $\ker(N) \cap \mathbb{R}^r_{>0} \neq \emptyset$. If a reaction network satisfies this assumption, we call it *consistent* [3]. For consistent networks, the map

$$\Psi \colon \mathbb{R}^n_{>0} \times \mathbb{R}^\ell_{>0} \mapsto \mathcal{V}, \quad (h, \lambda) \mapsto (\tfrac{1}{h}, (E\lambda) * h^A)$$

is surjective, i.e. a parametrization (Corollary I.6.2). From Lemma 2.5 with $v = E\lambda$ follows Lemma 2.9, which is a slight refinement of [30, Proposition 2].

Lemma 2.9. *For a consistent reaction network, it holds that*

$$\big\{ J_{f_\kappa}(x) \mid \kappa \in \mathbb{R}^r_{>0},\, x \in \mathbb{R}^n_{>0},\, x \in V_\kappa \big\} = \big\{ N \operatorname{diag}(E\lambda) A^\top \operatorname{diag}(h) \mid \lambda \in \mathbb{R}^\ell_{>0},\, h \in \mathbb{R}^n_{>0} \big\}.$$

Now, we do the same modification on $N \operatorname{diag}(E\lambda) A^\top \operatorname{diag}(h)$ as we did on $J_{f_\kappa}(x)$ at the end of Section 2.1. Assume that the conservation law matrix W is row reduced and let $i_1 < \cdots < i_d$ be the indices of the first non-zero coordinates of the rows of W. We denote by $M(h, \lambda)$ the matrix obtained from $N \operatorname{diag}(E\lambda) A^\top \operatorname{diag}(h)$ by replacing the rows $i_1, \dots, i_d$ by the rows of W. The polynomial

$$q \colon \mathbb{R}^n_{>0} \times \mathbb{R}^\ell_{>0} \to \mathbb{R}, \quad (h, \lambda) \mapsto q(h, \lambda) := (-1)^s \det M(h, \lambda), \tag{2.9}$$

is called the *critical polynomial*. In Section 2.4, we will use the critical polynomial to prove that the parameter region of multistationarity is path connected for certain reaction networks.

Example 2.10. We consider again the reaction network (2.2). A choice of the extreme vectors of the flux cone $\ker(N) \cap \mathbb{R}^3_{\geq 0}$ is given by the columns of the matrix

$$E = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

This can be computed for this small example by hand or by using the function `rays` in `SageMath` [103] or the function with the same name in `OSCAR` [37, 86]. With this choice of E, the critical polynomial is given by

$$q(h, \lambda) = h_1\lambda_2 - h_1\lambda_1 + h_2\lambda_1 + h_2\lambda_2. \tag{2.10}$$

To find the critical polynomial, one has to compute the determinant of the matrix $M(h, \lambda)$ with h and λ as symbolic variables. This task might be demanding if the matrix has many rows or there are many symbolic variables. Theorem 2.13 below might simplify such computations. In Paper II, we used this statement to provide a recursive expression of the critical polynomial for the n-site phosphorylation network, see Section 2.4 for more details. To state this result, first we recall the notion of a Gale dual matrix.

Definition 2.11. Let K be a field and $V \in K^{s\times n}$, $U \in K^{n\times d}$ be two matrices. The matrix U is called a *Gale dual* of V if $\mathrm{im}(U) = \ker(V)$ and $\ker(U) = \{0\}$.

For $I = \{i_1, \ldots, i_s\} \subseteq [n]$ with $i_1 < \cdots < i_s$ and $I^c = \{j_1, \ldots, j_d\}$, $j_1 < \cdots < j_d$, we denote by $\mathrm{sgn}(\tau_I) \in \{-1, 1\}$ the sign of the permutation that sends $(1, \ldots, n)$ to $(i_1, \ldots, i_s, j_1, \ldots, j_d)$.

Lemma 2.12. *(Corollary II.2.6, see also [83, Lemma 2.10] [70, Theorem 12.16]) Let $D(\lambda) \in \mathbb{R}(\lambda)^{n\times d}$ be a Gale dual of $N' \operatorname{diag}(E\lambda)A^\top \in \mathbb{R}(\lambda)^{s\times n}$. There exists $\delta(\lambda) \in \mathbb{R}(\lambda) \setminus \{0\}$ such that for all $I \subseteq [n]$ with $|I| = s$ it holds:*

$$\delta(\lambda) \det\left([D(\lambda)]_{I^c,[d]}\right) = \mathrm{sgn}(\tau_I) \det\left([N' \operatorname{diag}(E\lambda)A^\top]_{[s],I}\right). \tag{2.11}$$

Theorem 2.13. *(Theorem II.2.7) Let $D(\lambda) \in \mathbb{R}(\lambda)^{n\times d}$ be a Gale dual of $N' \operatorname{diag}(E\lambda)A^\top$. The critical polynomial (2.9) can be written as*

$$q(h, \lambda) = (-1)^{s(d+1)} \sum_{\substack{I\subseteq[n] \\ |I|=s}} \delta(\lambda) \det\left([W]_{[d],I^c}\right) \det\left([D(\lambda)]_{I^c,[d]}\right) \prod_{i\in I} h_i,$$

where $\delta(\lambda) \in \mathbb{R}(\lambda)$ as in (2.11).

The idea behind the proof of Theorem 2.13 is to use Laplace expansion on complementary minors on $M(h,\lambda)$ as in (2.5), and replace the minors $\det\left([N'\operatorname{diag}(E\lambda)A^\top]_{[s],I}\right)$, $I \subseteq [n], |I| = s$ by minors of a Gale dual matrix via Lemma 2.12.

The advantage of Theorem 2.13 is that the coefficients of $q(h,\lambda)$ (viewed as a polynomial in the h's) are given by minors of size $d \times d$. For the two families of reaction networks studied in Paper II, we have $d = 3$ for all networks in the family, which significantly simplifies the computation of the critical polynomial.

2.4 Connectivity of parameter region of multistationarity

The focus of Paper I and Paper II lies on showing that the parameter region of multistationarity is path connected for certain chemical reaction networks. Based on Theorem 2.7, we developed a sufficient criterion implying connectivity of the parameter region of multistationarity.

Theorem 2.14. *(Theorem I.2.4) Consider a conservative consistent reaction network without relevant boundary steady states. Assume that there exist species* $\mathrm{X}_1, \dots, \mathrm{X}_k$ *such that each* X_j *participates in exactly* 3 *reactions of the form*

$$\sum_{i=k+1}^{n} a_{i,j}X_i \underset{\kappa_{3j}}{\overset{\kappa_{3j-1}}{\rightleftharpoons}} \mathrm{X}_j \xrightarrow{\kappa_{3j-2}} \sum_{i=k+1}^{n} b_{i,j}X_i, \quad j = 1, \dots, k.$$

Let q *be the critical polynomial of the reduced network obtained by removing the reactions corresponding to* κ_{3j} *for* $j = 1, \dots, k$*. If*

(P1) $q^{-1}(\mathbb{R}_{<0})$ *is path connected, and*

(P2) the Euclidean closure of $q^{-1}(\mathbb{R}_{<0})$ *equals* $q^{-1}(\mathbb{R}_{\leq 0})$*,*

then the parameter region of multistationarity of both the reduced and the original network is path connected.

Instead of a convex parametrization, Theorem I.2.4 in Paper I is formulated in terms of any parametrization of the incidence variety (2.7). In this section, for the sake of simplicity, we focus only on the convex parametrization.

Since the critical polynomial q is a polynomial, the set

$$q^{-1}(\mathbb{R}_{<0}) = \left\{(h,\lambda) \in \mathbb{R}^n_{>0} \times \mathbb{R}^\ell_{>0} \mid q(h,\lambda) < 0\right\}$$

is a semi-algebraic set. Thus to verify condition (P1), one can apply semi-algebraic algorithms, which work well for small reaction networks. This is illustrated in the next example.

Example 2.15. Consider the reaction network (2.2). Its critical polynomial was computed in (2.10). The `Maple` [79] function `SamplePoints` from the package `RegularChains` [27] provides a sample point from each connected component of the set $q^{-1}(\mathbb{R}_{<0})$. For the polynomial (2.10), the code returned only one sample point, thus $q^{-1}(\mathbb{R}_{<0})$ is connected.

Our initial goal was to show connectivity of the parameter region of multistationarity of the n-site phosphorylation network

$$\mathrm{S}_i + \mathrm{E} \underset{\kappa_{6i+2}}{\overset{\kappa_{6i+1}}{\rightleftarrows}} \mathrm{ES}_i \xrightarrow{\kappa_{6i+3}} \mathrm{S}_{i+1} + \mathrm{E}, \quad \mathrm{S}_{i+1} + \mathrm{F} \underset{\kappa_{6i+5}}{\overset{\kappa_{6i+4}}{\rightleftarrows}} \mathrm{FS}_{i+1} \xrightarrow{\kappa_{6i+6}} \mathrm{S}_{i+6} + \mathrm{F}, \quad i = 0, \dots, n-1.$$

For $n = 2$ the network is displayed in (1.2), and its critical polynomial has 15 variables and 400 monomials. For polynomials of that size, the semi-algebraic algorithms checking (P1) become intractable. In Paper I, we used a sufficient condition from Paper III implying (P1) and (P2), which is based on separating hyperplanes of the signed support of the polynomial. Since this condition can be checked using linear programming, it is applicable to larger polynomials. For example, the critical polynomial of the ERK network (see Section I.3.2) has 21 variables and 18472 monomials, yet a separating hyperplane can be found in less than a minute using a computer. We will discuss this method in more detail in Chapter 3. Using the separating hyperplane condition and Theorem 2.14, we proved connectivity of the parameter region of multistationarity for the n-site phosphorylation network with $n = 2, 3$ and for several other biologically relevant reaction networks, see Table I.1.

In Paper II, we focused on two infinite families of phosphorylation networks, the n-site phosphorylation network discussed above and a slightly different family, called *n-site weakly irreversible phosphorylation network*

$$\begin{aligned} \mathrm{S}_i + \mathrm{K} &\underset{\kappa_{10i+2}}{\overset{\kappa_{10i+1}}{\rightleftarrows}} \mathrm{Y}_{2i+1} \xrightarrow{\kappa_{10i+3}} \mathrm{Y}_{2i+2} \underset{\kappa_{10i+5}}{\overset{\kappa_{10i+4}}{\rightleftarrows}} \mathrm{S}_{i+1} + \mathrm{K} \\ \mathrm{S}_{i+1} + \mathrm{F} &\underset{\kappa_{10i+7}}{\overset{\kappa_{10i+6}}{\rightleftarrows}} \mathrm{U}_{2i+1} \xrightarrow{\kappa_{10i+8}} \mathrm{U}_{2i+2} \underset{\kappa_{10i+10}}{\overset{\kappa_{10i+9}}{\rightleftarrows}} \mathrm{S}_i + \mathrm{F}, \end{aligned} \qquad i = 0, \dots, n-1.$$

For $n = 2$, the weakly irreversible phosphorylation network was investigated in [85], where based on numerical computations, it was suggested that the parameter region of multistationarity is connected. From the expression in Theorem 2.13, we derived a recursive formula for the critical polynomials. Based on the results from Paper III and Paper IV, we gave an inductive proof that these critical polynomials satisfy conditions (P1) and (P2) from Theorem 2.14. This provides a proof for connectivity of the parameter region for multistationarity of the two families of n-site phosphorylation networks for every $n \in \mathbb{N}_{\geq 2}$.

One should also note that the concepts discussed in Paper I are applicable not only for verifying the connectivity of the parameter region of multistationarity but also for

finding reaction networks whose parameter region of multistationarity has more than one connected component. In Section I.3.4, we showcased this idea by proving that for the *allosteric reciprocal enzyme regulation network*

$$\begin{gathered} \mathrm{S}+\mathrm{KL} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{SKL} \xrightarrow{\kappa_3} \mathrm{S_p}+\mathrm{KL}, \quad \mathrm{S_p}+\mathrm{P} \underset{\kappa_5}{\overset{\kappa_4}{\rightleftharpoons}} \mathrm{S_pP} \xrightarrow{\kappa_6} \mathrm{S}+\mathrm{P}, \\ \mathrm{K}+\mathrm{L} \underset{\kappa_8}{\overset{\kappa_7}{\rightleftharpoons}} \mathrm{KL}, \quad \mathrm{P}+\mathrm{L} \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} \mathrm{PL} \end{gathered} \tag{2.12}$$

from [100] the parameter region of multistationarity has exactly two connected components. The network (2.12) has 5 conservation laws, so the vector of total concentrations c is of length 5. Using Theorem 2.7 and a parametrization of the incidence variety (2.7), similar to the one used in Example 2.8, we showed that the pair $(\kappa, c) \in \mathbb{R}^{10}_{>0} \times \mathbb{R}^5$ does not enable multistationarity if $\kappa_3 = \kappa_6$. Furthermore, we proved that both sets

$$\left\{(\kappa, c) \in \mathbb{R}^{10}_{>0} \times \mathbb{R}^5 \mid \kappa_3 > \kappa_6\right\} \quad \text{and} \quad \left\{(\kappa, c) \in \mathbb{R}^{10}_{>0} \times \mathbb{R}^5 \mid \kappa_3 < \kappa_6\right\}$$

contain parameters that enable multistationarity. This observation implies that the parameter region of multistationarity has at least two connected components.

To show that the number of connected components is exactly two, we used Theorem I.2.4 that gives an upper bound on the number of connected components of the multistationarity region in terms of the number of connected components of $q^{-1}(\mathbb{R}_{<0})$. Here q denotes the critical polynomial associated with the reduced network obtained from (2.12) by removing the reactions κ_2, κ_5. We showed that to find such an upper bound it is enough to consider a restriction of q to a certain face of its Newton polytope (see Section 3.3 for further details). By applying Corollary III.3.13, we concluded that this restriction of q and, consequently, the parameter region of multistationarity, has at most two connected components. It is worth mentioning that the observation, which allows one to restrict q to a certain face of its Newton polytope in order to determine an upper bound on the number of connected components of $q^{-1}(\mathbb{R}_{<0})$, served as the motivation for Paper IV, where this concept was generalized.

Surprisingly, the structure of network (2.12) appears similar to the networks in Table I.1, which have connected parameter region of multistationarity. Exploring other biologically relevant networks with disconnected multistationarity region might help to characterize structural properties of reaction networks that imply (dis)connectivity of the parameter region of multistationarity. Understanding why one reaction network has a connected multistationarity region while another does not, poses an interesting question that warrants further investigation.

3

Descartes' rule of signs

The focus of this chapter is on the content of Paper III and Paper IV. In Section 3.1 and 3.2, we recall essential notions on signomials and convex geometry. In Section 3.3, we review the results from Paper III and IV. Specifically, we discuss how properties of the signed support of a polynomial can be used to provide bounds on the number of connected components of the set of points in the positive orthant, where the polynomial takes negative values. Additionally, in Section 3.4, we describe conditions on the signed support of an equation system of n polynomials in n variables ensuring that its number of positive real solutions is either infinite or at most two. The content of Section 3.4 is not part of Paper III and IV and has not been included in any publication, but follows directly from the argument used in Paper III.

3.1 Signomials

Consider a polynomial function with real exponents

$$f\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu, \tag{3.1}$$

where $\sigma(f) \subseteq \mathbb{R}^n$ is a finite set, called the *support* of f. Following [39, 97], in Paper III and IV we used the term *signomial* for such functions to emphasize that the domain of the function is restricted to the positive orthant and the exponent vectors $\mu \in \sigma(f)$ may have real entries.

We divide the support of f according to the signs of the coefficients into

$$\sigma_+(f) := \{\mu \in \sigma(f) \mid c_\mu > 0\} \quad \text{and} \quad \sigma_-(f) := \{\mu \in \sigma(f) \mid c_\mu < 0\}.$$

We call the elements of $\sigma_+(f)$ and $\sigma_-(f)$ *positive* and *negative exponent vectors of* f, respectively. For a set $S \subseteq \mathbb{R}^n$ we define the *restriction* of f to S as

$$f_{|S}\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto f_{|S}(x) := \sum_{\mu\in\sigma(f)\cap S} c_\mu x^\mu.$$

The set of common positive real roots of a collection of signomials $f_1, \dots, f_k \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ will be denoted by

$$V_{>0}(f_1, \dots, f_k) := \big\{x \in \mathbb{R}^n_{>0} \mid f_1(x) = \cdots = f_k(x) = 0\big\}.$$

Motivated by the problem discussed in Section 2.4, we were interested in showing that the preimage of the negative real line under a signomial f,

$$f^{-1}(\mathbb{R}_{<0}) = \big\{x \in \mathbb{R}^n_{>0} \mid f(x) < 0\big\}, \tag{3.2}$$

is a connected set. We call the connected components of (3.2) *negative connected components of* f, write $\mathcal{B}_0^-(f)$ for the set of negative connected components of f, and denote the cardinality of $\mathcal{B}_0^-(f)$ by $b_0(f^{-1}(\mathbb{R}_{<0}))$.

For a univariate signomial, using Descartes' rule of signs it is possible to derive a simple combinatorial condition, based on the signs of the coefficients, that ensures the existence of at most one negative connected component.

Theorem 3.1. *(Descartes' rule of signs, see e.g. [36, 92, 110]) Let* $g(t) = \sum_{i=1}^d c_i t^{\nu_i}$ *be a non-zero univariate signomial such that* $\nu_1 < \cdots < \nu_d$*. The number of positive real roots of* g *is bounded above by the number of sign changes in the coefficient sequence* $(c_1, \dots, c_d)$*.*

Corollary 3.2. *Let* $g(t) = \sum_{i=1}^d c_i t^{\nu_i}$ *be a non-zero univariate signomial whose coefficient sequence matches one of the following sign sequences*

$$(+\cdots+-\cdots-), \quad (-\cdots-+\cdots+), \quad (+\cdots+-\cdots-+\cdots+).$$

Then g *has at most one negative connected component.*

Thus, the univariate Descartes' rule of signs not only provides a bound on the number of positive real roots, but it also gives bounds on the number of negative connected components. This observation leads to the following generalization of Descartes' rule of signs to hypersurfaces.

Problem 3.3. *(Problem III.1.1) Let* $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ *be a signomial with* $f(x) = \sum_{\mu \in \sigma(f)} c_\mu x^\mu$*, and* $\sigma(f) \subseteq \mathbb{R}^n$ *a finite set. Find a (sharp) upper bound on the number of connected components of* $f^{-1}(\mathbb{R}_{<0})$ *based on the sign of the coefficients and the geometry of* $\sigma(f)$*.*

Problem 3.3 is invariant under affine transformations of the support $\sigma(f)$. To make this more precise, we recall Lemma III.2.3.

Lemma 3.4. *(Lemma III.2.3) Let* $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ *be a signomial. For an invertible matrix* $M \in \mathbb{R}^{n \times n}$ *and* $v \in \mathbb{R}^n$*, consider the map*

$$F_{M,v,f} \colon \mathbb{R}^n_{>0} \to \mathbb{R}, \qquad F_{M,v,f}(x) = x^v f(x^M).$$

There exists a homeomorphism between the connected components of $f^{-1}(\mathbb{R}_{<0})$ *and* $F^{-1}_{M,v,f}(\mathbb{R}_{<0})$*. Furthermore,*

$$\sigma_+\big(F_{M,v,f}\big) = M\sigma_+(f) + v \quad \text{and} \quad \sigma_-\big(F_{M,v,f}\big) = M\sigma_-(f) + v.$$

3.2 Background on convex and polyhedral geometry

The partial solutions for Problem 3.3, as presented in Paper III and IV, rely on polyhedral geometric conditions of the signomial's support. In Section 3.4, we will apply some results from convex geometry to find upper bounds on the number of positive real solutions of a polynomial equation system. Here, we provide a brief overview of the necessary background on these topics. In this section, we closely follow the standard books [70, 94, 112].

A set $C \subseteq \mathbb{R}^n$ is *convex* if for any $x, y \in C$ and $\lambda \in [0,1]$ we have $\lambda x + (1-\lambda)y \in C$. An *affine combination* of points $x_1, \ldots, x_m \in \mathbb{R}^n$ is a linear combination $\sum_{i=1}^m \lambda_i x_i$, $\lambda_1, \ldots, \lambda_m \in \mathbb{R}$ such that $\sum_{i=1}^m \lambda_i = 1$. The *affine hull* of a set C is

$$\mathrm{Aff}(C) = \Big\{ \sum_{i=1}^m \lambda_i x_i \mid x_1, \ldots, x_m \in C, \sum_{i=1}^m \lambda_i = 1 \Big\}.$$

This is the smallest affine linear subspace of $\mathbb{R}^n$ that contains C. The *dimension of a convex set* is the dimension of its affine hull. The Euclidean *interior* of $C \subseteq \mathbb{R}^n$ will be denoted by $\mathrm{int}(C)$ and the *boundary* of C by $\mathrm{bd}(C) = \overline{C} \setminus \mathrm{int}(C)$, where $\overline{C}$ denotes the closure of C in the Euclidean topology of $\mathbb{R}^n$. The *relative interior* of a convex set C, denoted as $\mathrm{relint}(C)$, is the interior of C interpreted as a subset of its affine hull.

An affine combination $\sum_{i=1}^m \lambda_i x_i$ is called a *convex combination* if additionally $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$. The *convex hull* of $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^n$ is defined as the set of all convex combinations of $x_1, \ldots, x_m$, that is

$$\mathrm{Conv}(\{x_1, \ldots, x_m\}) = \Big\{ \sum_{i=1}^m \lambda_i x_i \mid \lambda_1, \ldots, \lambda_m \in \mathbb{R}_{\geq 0}, \sum_{i=1}^m \lambda_i = 1 \Big\}.$$

Such a set is also called a *polytope*. The convex hull $\mathrm{Conv}(\{x_1, \ldots, x_m\})$ is a convex set. In fact, it is the smallest convex set (with respect to inclusion) that contains $x_1, \ldots, x_m$.

A *hyperplane* $\mathcal{H}_{v,a}$ is a set given by an affine linear equality, that is,

$$\mathcal{H}_{v,a} = \{x \in \mathbb{R}^n \mid v \cdot x = a\},$$

where $v \in \mathbb{R}^n \setminus \{0\}$ and $a \in \mathbb{R}$. Each hyperplane defines two *half-spaces*

$$\mathcal{H}^+_{v,a} = \{x \in \mathbb{R}^n \mid v \cdot x \geq a\}, \qquad \mathcal{H}^-_{v,a} = \{x \in \mathbb{R}^n \mid v \cdot x \leq a\}.$$

We denote by $\mathcal{H}^{+,\circ}_{v,a}$ and $\mathcal{H}^{-,\circ}_{v,a}$ the interior of these half-spaces respectively.

Theorem 3.5. *[94, Corollary 2.1.1, Theorem 11.5]*

(i) The intersection of a collection of closed half-spaces is convex.

(ii) A closed convex set is the intersection of the closed half-spaces that contain it.

A set $P \subseteq \mathbb{R}^n$ is a *polyhedron* if it is the intersection of finitely many half-spaces $\mathcal{H}^-_{v_1,a_1}, \ldots, \mathcal{H}^-_{v_k,a_k}$. An important theorem in polyhedral geometry states that a polytope is a bounded polyhedron.

Theorem 3.6. *[112, Theorem 1.1] A subset $P \subseteq \mathbb{R}^n$ is a polytope, i.e. the convex hull of a finite set $\{x_1, \ldots, x_m\} \subseteq \mathbb{R}^n$ if and only if P is a bounded polyhedron, i.e. P is bounded and the intersection of half-spaces $\mathcal{H}^-_{v_1,a_1}, \ldots, \mathcal{H}^-_{v_k,a_k} \subseteq \mathbb{R}^n$.*

For a polyhedron $P \subseteq \mathbb{R}^n$ and $v \in \mathbb{R}^n$, we define the *face* with normal vector v as

$$P_v := \left\{ x \in \mathbb{R}^n \mid v \cdot x = \max_{y \in P} v \cdot y \right\}. \tag{3.3}$$

The hyperplane $\mathcal{H}_{v,a} \subseteq \mathbb{R}^n$ with $a = \max_{y \in P} v \cdot y$ is called a *supporting hyperplane* of $P \subseteq \mathbb{R}^n$. It satisfies $P \subseteq \mathcal{H}^-_{v,a}$ and $P \cap \mathcal{H}_{v,a} = P_v$. A face of dimension $0, 1$ and $\dim(P) - 1$ is called a *vertex, edge* and *facet* respectively. We denote the set of vertices of a polytope P by $\mathrm{Vert}(P)$.

Proposition 3.7. *[112, Proposition 2.2]*

(i) Every polytope $P \subseteq \mathbb{R}^n$ is the convex hull of its vertices, i.e. $P = \mathrm{Conv}(\mathrm{Vert}(P))$.

(ii) If $P = \mathrm{Conv}(C)$ for a finite set $C \subseteq \mathbb{R}^n$, then $\mathrm{Vert}(P) \subseteq C$.

For a signomial f, its *Newton polytope* is defined as the convex hull of its support

$$\mathrm{N}(f) := \mathrm{Conv}(\sigma(f)).$$

From Proposition 3.7 it follows that vertices of $\mathrm{N}(f)$ correspond to monomials of f with non-zero coefficients.

Another ubiquitous class of convex sets that play an important role in this thesis is the class of *convex cones*. A subset $C \subseteq \mathbb{R}^n$ is a *cone* if it is closed under multiplication by positive scalars. A cone C is a *convex cone* if it is additionally a convex set.

Theorem 3.8. *[94, Corollary 2.5.1, Corollary 11.7.1]*

(i) The intersection of a collection of closed linear half-spaces is a convex cone.

(ii) A non-empty closed convex cone $C \subseteq \mathbb{R}^n$ is the intersection of the closed linear half-spaces $\mathcal{H}^-_{v_i,0}$ with $C \subseteq \mathcal{H}^-_{v_i,0}$.

A *convex polyhedral cone* is a subset of $\mathbb{R}^n$ that can be written as the intersection of a finite number of linear half-spaces $\mathcal{H}^-_{v_1,0}, \dots, \mathcal{H}^-_{v_m,0} \subseteq \mathbb{R}^n$. Similarly to polytopes, convex polyhedral cones have also an equivalent representation. The *conical hull* of $\{x_1, \dots, x_m\} \subseteq \mathbb{R}^n$ is defined as

$$\mathrm{Cone}(\{x_1, \dots, x_m\}) := \Big\{ \sum_{i=1}^{m} \lambda_i x_i \mid \lambda_1, \dots, \lambda_m \in \mathbb{R}_{\geq 0} \Big\}.$$

Theorem 3.9. *[112, Theorem 1.3] A subset $C \subseteq \mathbb{R}^n$ is a convex polyhedral cone, i.e. the intersection of linear half-spaces $\mathcal{H}^-_{v_1,0}, \dots, \mathcal{H}^-_{v_k,0} \subseteq \mathbb{R}^n$ if and only if C is the conical hull of a finite set $\{x_1, \dots, x_m\} \subseteq \mathbb{R}^n$.*

We finish this section by recalling some special classes of convex cones associated with convex sets. First, we discuss the *recession cone* of a convex set $C \subseteq \mathbb{R}^n$. This cone contains all vectors $y \in \mathbb{R}^n$ such that, for every $x \in C$, the half-line with direction vector y and endpoint x is contained in C. Thus, the recession cone is given by

$$0^+C := \{y \in \mathbb{R}^n \mid \forall x \in C \; \forall \lambda \geq 0 \colon x + \lambda y \in C\}.$$

The recession cone is a convex cone [94, Theorem 8.1]. This property of 0^+C will play a crucial role in the proof of Lemma 3.24. Furthermore, it is known that a non-empty closed convex set C is bounded if and only if $0^+C = \{0\}$ [94, Theorem 8.4].

Next, we pay our attention to *normal cones* of polytopes, which are defined as follows. Let $P \subseteq \mathbb{R}^n$ be a polytope and $F \subseteq P$ one of its faces. The set

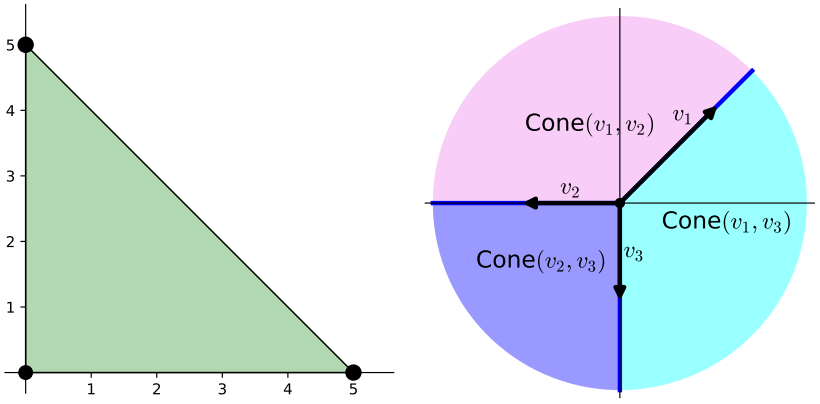
$$\mathcal{N}_P(F) := \big\{v \in \mathbb{R}^n \mid F \subseteq P_v\big\}. \tag{3.4}$$

is a convex cone, called the normal cone of F. The collection of the normal cones of all faces of a polytope is a *complete fan* $\mathcal{N}_P$ [112, Example 7.3], called the *normal fan* of P. The $(n-1)$-*skeleton* of $\mathcal{N}_P$ contains all cones $\mathcal{N}_P(F)$ with $\dim \mathcal{N}_P(F) \leq n-1$.

Example 3.10. The convex hull of the points $(0,0), (5,0), (0,5)$ is a triangle P, depicted in Figure 3.1(a), with 3 vertices and 3 edges. Since $\dim(P) = 2$, its edges and facets coincide. The normal cone of P consists of 3 one-dimensional cones spanned by the vectors $(1,1), (-1,0), (0,-1)$, and 3 two-dimensional cones

$$\mathrm{Cone}((1,1),(-1,0)), \quad \mathrm{Cone}((-1,0),(0,-1)), \quad \mathrm{Cone}((-1,0),(1,1)).$$

These cones are shown in Figure 3.1(b).



(a) (b)

Figure 3.1: Convex hull of $(0,0),(5,0),(0,5)$, and its normal cones.

3.3 Descartes' rule of signs for hypersurfaces

In this section, we elaborate further on Problem 3.3 and present the main results of Paper III and IV.

Definition 3.11. Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial. A hyperplane $\mathcal{H}_{v,a} \subseteq \mathbb{R}^n$ is a *separating hyperplane* of $\sigma(f)$ if

$$\sigma_-(f) \subseteq \mathcal{H}^+_{v,a} \quad \text{and} \quad \sigma_+(f) \subseteq \mathcal{H}^-_{v,a}.$$

If additionally $\sigma_-(f) \cap \mathcal{H}^{+,\circ}_{v,a} \neq \emptyset$, then $\mathcal{H}_{v,a}$ is called a *strict separating hyperplane*.

If $\mathcal{H}_{v,a}$ is a separating hyperplane of the support of a signomial f, then for each fixed $x \in \mathbb{R}^n_{>0}$ the univariate signomial

$$f_{v,x}\colon \mathbb{R}_{>0} \to \mathbb{R}, \quad t \mapsto f(t^v * x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu t^{v\cdot\mu} \tag{3.5}$$

has at most one sign change in its coefficient sequence. From the univariate Descartes' rule of signs (Theorem 3.1), it follows that

$$f_{v,x}(t) < 0, \text{ for all } t \in [1,\infty) \quad \text{if } x \in f^{-1}(\mathbb{R}_{<0}),$$

we refer to Lemma III.3.3 for more details on the proof of this fact. This observation provides an explicit method for constructing paths that are contained within $f^{-1}(\mathbb{R}_{<0})$.

For a face of the Newton polytope $\mathrm{N}(f)_v \subseteq \mathrm{N}(f)$, $v \in \mathbb{R}^n \setminus \{0\}$, as defined in (3.3), the hyperplane $\mathcal{H}_{v,a}$, $a = \max_{\mu\in\sigma(f)} v\cdot\mu$, satisfies $\sigma(f) \subseteq \mathcal{H}^-_{v,a}$. Thus, if $\sigma_-(f) \subseteq \mathrm{N}(f)_v$,

then $\mathcal{H}_{v,a}$ is a non-strict separating hyperplane of $\sigma(f)$. On the contrary, if $\sigma_-(f) \neq \emptyset$ and $\mathcal{H}_{v,a}$ is a non-strict separating hyperplane of $\sigma(f)$, then $F = \mathcal{H}_{v,a} \cap \mathrm{N}(f)$ is a face of the Newton polytope and $\sigma_-(f) \subseteq F$. We refer to Figure 3.2(a) for an illustration of non-strict and strict separating hyperplanes.

Theorem 3.12. *(Theorem III.3.6, Theorem IV.3.1) Let* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ *be a signomial.*

(i) If $\sigma(f)$ *has a strict separating hyperplane, then* $f^{-1}(\mathbb{R}_{<0})$ *is non-empty and contractible. In particular*

$$b_0(f^{-1}(\mathbb{R}_{<0})) = 1.$$

(ii) If $\sigma(f)$ *has a non-strict separating hyperplane* $\mathcal{H}_{v,a}$ *such that* $F = \mathcal{H}_{v,a} \cap \mathrm{N}(f)$ *is a face of* $\mathrm{N}(f)$*, then*

$$b_0(f^{-1}(\mathbb{R}_{<0})) = b_0(f^{-1}_{|F}(\mathbb{R}_{<0})).$$

The proof of Theorem 3.12 is based on arguments that use that the induced signomial (3.5) has at most one sign change in its coefficient sequence. If for some $v \in \mathbb{R}^n$ there are at most two sign changes in the coefficient sequence of (3.5), the induced signomial might still be used to derive bounds on the number of negative connected components.

Definition 3.13. Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial. A pair of parallel hyperplanes $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$, $a \geq b$ such that

$$\sigma_+(f) \subseteq \mathcal{H}^-_{v,a} \cap \mathcal{H}^+_{v,b} \qquad \text{and} \qquad \sigma_-(f) \subseteq \mathbb{R}^n \setminus \left(\mathcal{H}^{-,\circ}_{v,a} \cap \mathcal{H}^{+,\circ}_{v,b}\right),$$

is called a *pair of enclosing hyperplanes* of $\sigma_+(f)$. A pair of enclosing hyperplanes is *strict*, if

$$\sigma_-(f) \cap \mathcal{H}^{+,\circ}_{v,a} \neq \emptyset \quad \text{and} \quad \sigma_-(f) \cap \mathcal{H}^{-,\circ}_{v,b} \neq \emptyset.$$

Strict enclosing hyperplanes were used in Theorem III.3.8 to show that a signomial has at most two negative connected components. In Paper IV, the argument has been generalized to non-strict separating hyperplanes. Before we recall this statement, we introduce some notation. Given a pair of enclosing hyperplanes $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ of $\sigma_+(f)$, we define

$$A = (\mathcal{H}^+_{v,a} \cap \sigma_-(f)) \cup \sigma_+(f), \qquad B = (\mathcal{H}^-_{v,b} \cap \sigma_-(f)) \cup \sigma_+(f). \tag{3.6}$$

and the bipartite graph with vertex and edge sets defined as

$$\begin{aligned} \mathcal{B}_{A,B} &:= \mathcal{B}^-_0(f_{|A}) \sqcup \mathcal{B}^-_0(f_{|B}), \\ \mathcal{E}_{A,B} &:= \left\{(U,V) \mid U \in \mathcal{B}^-_0(f_{|A}), V \in \mathcal{B}^-_0(f_{|B})\colon U \cap V \neq \emptyset\right\}. \end{aligned} \tag{3.7}$$

Proposition 3.14. *(Proposition IV.2.6) Let* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ *be a signomial,* $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ *a pair of enclosing hyperplanes of* $\sigma_+(f)$*, and let* $A, B \subseteq \sigma(f)$ *as defined in* (3.6)*. Then*

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \leq C \leq b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) + b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big),$$

where C *denotes the number of connected components of the bipartite graph* $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$ *from* (3.7)*.*

Example 3.15. Consider the signomial $f = x_1^3x_2^3 - 5x_1^3x_2 + x_1x_2^3 + x_1^3 + x_1^2 - 5x_2$. A pair of enclosing hyperplanes $(\mathcal{H}_{v,3}, \mathcal{H}_{v,1})$ of $\sigma_+(f)$ with $v = (1,0)$ is depicted in Figure 3.2(a). The sets A, B as defined in (3.6) are given by

$$A = \{(3,1),(2,0),(3,0),(1,3),(3,3)\}, \quad B = \{(0,1),(2,0),(3,0),(1,3),(3,3)\}.$$

The hyperplane $\mathcal{H}_{-v,-1}$ is a strict separating hyperplane of $\sigma(f_{|B})$, which implies that $f_{|B}^{-1}(\mathbb{R}_{<0})$ is connected by Theorem 3.12(i), see Figure 3.2(b). Using Theorem 3.12(ii), we have that the number of connected components of $f_{|A}^{-1}(\mathbb{R}_{<0})$ is the same as the number of negative connected components of $x_1^3 - 5x_1^3x_2 + x_1^3x_2^3 = x_1^3(1 - 5x_2 + x_2^3)$. Since the latter polynomial is essentially univariate, it is easy to deduce that it has one negative connected component.

Thus, the bipartite graph $\mathcal{B}_{A,B}$ has two vertices corresponding to $f_{|A}^{-1}(\mathbb{R}_{<0})$ and $f_{|B}^{-1}(\mathbb{R}_{<0})$. An easy computation shows that $(1,1) \in f_{|A}^{-1}(\mathbb{R}_{<0}) \cap f_{|B}^{-1}(\mathbb{R}_{<0})$, which implies that the graph $\mathcal{B}_{A,B}$ has an edge between the two vertices. Using Proposition 3.14, we conclude that f has one negative connected component.

Note that if $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of strict enclosing hyperplanes of $\sigma(f)$, then $\mathcal{H}_{v,a}$ and $\mathcal{H}_{-v,-b}$ are strict separating hyperplanes of $\sigma(f_{|A})$ and $\sigma(f_{|B})$ respectively. Thus, both $f_{|A}$ and $f_{|B}$ have one negative connected component by Theorem 3.12(i).

Corollary 3.16. *(Theorem III.3.8) Let* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ *be a signomial such that* $\sigma_+(f)$ *has a pair of strict enclosing hyperplanes. Then*

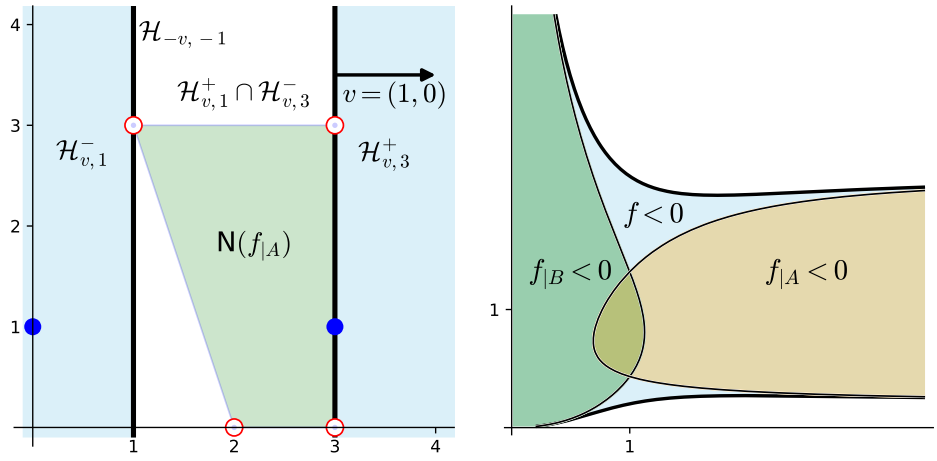
$$b_0(f^{-1}(\mathbb{R}_{<0})) \leq 2.$$

Theorem 3.12 has another important consequence, which allows us to reduce the problem of finding the number of negative connected components to the same problem but for polynomials in less monomials and variables.

Corollary 3.17. *(Theorem IV.3.6) Let* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ *be a signomial. Assume that there exists* $v \in \mathbb{R}^n$ *such that* $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$ *and*

$$b_0\big(f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0})\big) = b_0\big(f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0})\big) = 1.$$

If there exist negative exponent vectors $\beta_1 \in \mathrm{N}(f)_v$ *and* $\beta_2 \in \mathrm{N}(f)_{-v}$ *such that* $\mathrm{Conv}(\beta_1, \beta_2)$ *is an edge of* $\mathrm{N}(f)$*, then* $b_0\big(f^{-1}(\mathbb{R}_{<0})\big) = 1$*.*



(a) (b)

Figure 3.2: (a) Signed support of $f = x_1^3x_2^3 - 5x_1^3x_2 + x_1x_2^3 + x_1^3 + x_1^2 - 5x_2$, positive exponent vectors are depicted as red circles, negative exponent vectors as blue dots. The pair $(\mathcal{H}_{v,3}, \mathcal{H}_{v,1})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$. The green quadrilateral shows the Newton polytope of $f_{|A} = x_1^3x_2^3 - 5x_1^3x_2 + x_1x_2^3 + x_1^3 + x_1^2$. The hyperplane $\mathcal{H}_{-v,-1}$ is a strict separating hyperplane of the support of $f_{|B} = x_1^3x_2^3 + x_1x_2^3 + x_1^3 + x_1^2 - 5x_2$ (b) Preimage of the negative real line under f, $f_{|A}$ and $f_{|B}$.

Corollary 3.17 played a crucial role in Paper II, since it allowed us to show inductively that the critical polynomial (see Chapter 2) of the phosphorylation network considered in Paper II has one negative connected component.

Using the induced signomial (3.5) and adopting a similar approach as in the proofs of Theorem 3.12 and Proposition 3.14, one can show that the signomial has at most one negative connected component in the following cases.

Theorem 3.18. *(Theorem III.3.4, Corollary IV.2.13) Let* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ *be a signomial.*

(i) If $|\sigma_-(f)| \leq 1$, *then* $f^{-1}(\mathbb{R}_{<0})$ *is either empty or a logarithmically convex set. In particular*

$$b_0(f^{-1}(\mathbb{R}_{<0})) \leq 1.$$

(ii) If $|\sigma_+(f)| \leq 1$ *and* $\dim \mathrm{N}(f) \geq 2$, *then*

$$b_0(f^{-1}(\mathbb{R}_{<0})) = 1.$$

Remark 3.19. The condition $\dim \mathrm{N}(f) \geq 2$ in Theorem 3.18(ii) is necessary. The univariate signomial $-x^2 + 4x - 1$ has one positive coefficient but two negative connected components.

Besides separating and enclosing hyperplanes of the signed support, one can ensure that a signomial has one negative connected component if its signed support is separated by a simplex in a certain way. In the remaining of this section, we recall this result.

A *simplex* $P \subseteq \mathbb{R}^n$ is the convex hull of $n+1$ affinely independent points $\mu_0, \dots, \mu_n$, which are its vertices. The *negative vertex cone* at the vertex μ_k is defined as

$$P^{-,k} := \mu_k + \operatorname{Cone}(\mu_k - \mu_0, \dots, \mu_k - \mu_n).$$

We denote by P^- the union of the negative vertex cones $P^{-,0}, \dots, P^{-,n}$. We refer to Figure 3.3 for an illustration of a simplex and its negative vertex cones. If the vertices of the simplex are the standard basis vectors $e_1, \dots, e_n \in \mathbb{R}^n$ and the zero vector, then we call the simplex the standard n-simplex and write $\Delta_n := \operatorname{Conv}(\{0, e_1, \dots, e_n\})$.

Lemma 3.20. *(Lemma III.4.4) Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If $\sigma_-(f) \subseteq \Delta_n$ and $\sigma_+(f) \subseteq \Delta_n^-$, then f is a convex function.*

Sublevel sets of a convex function are convex sets [94, Theorem 4.6]. Thus, if a signomial f is a convex function, then f has one negative connected component, which is a convex set. Using an affine transformation, one can transform every n-dimensional simplex P into the standard n-simplex Δ_n (see e.g. Lemma III.4.5). By Lemma 3.4, such an affine transformation induces a homeomorphism between the corresponding negative connected components. This leads to the following result.

Theorem 3.21. *(Theorem III.4.6, Corollary IV.2.15) Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu \in \sigma(f)} c_\mu x^\mu$ be a signomial, and let $P \subseteq \mathbb{R}^n$ be an n-simplex.*

(i) If $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$, then $f^{-1}(\mathbb{R}_{<0})$ is either empty or contractible.

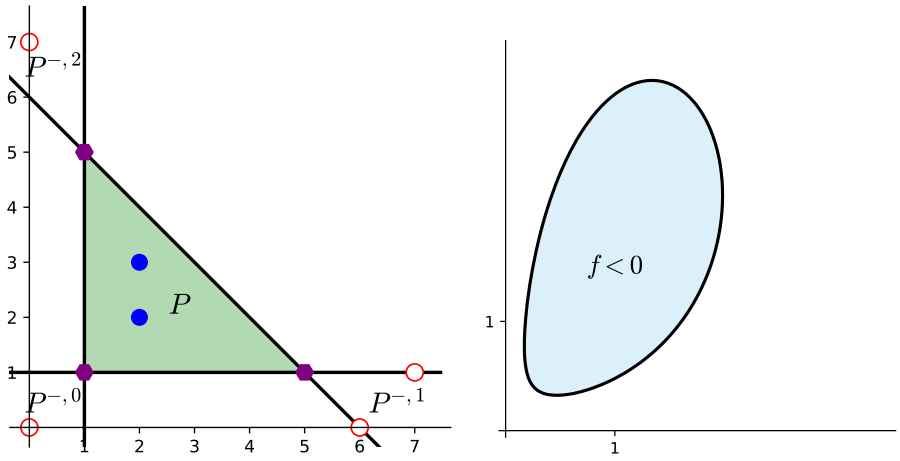
(ii) If $\sigma_+(f) \subseteq P$, $\sigma_-(f) \subseteq P^-$, $\sigma_-(f) \cap \operatorname{int}(P^-) \neq \emptyset$ and $n \geq 2$, then $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected.

Remark 3.22. In Theorem 3.21(ii) both conditions $\sigma_-(f) \cap \operatorname{int}(P^-) \neq \emptyset$ and $n \geq 2$ are necessary for $f^{-1}(\mathbb{R}_{<0})$ to be connected. If $n = 1$, then we consider the polynomial $f = -x^2 + 4x - 1$ as in Remark 3.19. The simplex $P = \operatorname{Conv}(0.5, 1.5)$ satisfies the conditions in Theorem 3.21(ii) , but $f^{-1}(\mathbb{R}_{<0})$ has two connected components. For the necessity of $\sigma_-(f) \cap \operatorname{int}(P^-) \neq \emptyset$, we refer to Example IV.2.16(b).

Example 3.23. Consider the signomial

$$f = x_1^7 x_2 + x_2^7 + x_1^6 - 3x_1^2 x_2^3 - 2x_1^2 x_2^2 + 1,$$

which has 4 positive and 2 negative exponent vectors, which are shown in Figure 3.3(a). The simplex $P := \operatorname{Conv}((1,1),(5,1),(1,5))$ separates $\sigma_+(f)$ and $\sigma_-(f)$ as required in Theorem 3.21(i). Therefore, $f^{-1}(\mathbb{R}_{<0})$ is a contractible set, see Figure 3.3(b).



(a) (b)

Figure 3.3: (a) Signed support of $f = x_1^7x_2 + x_2^7 + x_1^6 - 3x_1^2x_2^3 - 2x_1^2x_2^2 + 1$. Positive exponent vectors are marked by red circles, while the negative exponent vectors are marked by blue dots. The simplex $P = \mathrm{Conv}((0,0),(5,1),(1,5))$ and its negative vertex cones, $P^{-,0}, P^{-,1}, P^{-,2}$, separate $\sigma_+(f)$ and $\sigma_-(f)$ as required in Theorem 3.21(i). (b) Preimage of the negative real line under f.

3.4 Descartes' rule of signs for square systems

The goal of this section is to provide conditions on the signed support of signomials $f_1, \ldots, f_n$ in n variables such that their set of common positive roots $V_{>0}(f_1, \ldots, f_n)$ contains at most two or infinitely many points. Our strategy is to transform

$$V_{>0}(f_1, \ldots, f_n) = V_{>0}(f_1) \cap V_{>0}(f_2, \ldots, f_n)$$

such that $V_{>0}(f_1)$ becomes the boundary of a convex set and $V_{>0}(f_2, \ldots, f_n)$ is transformed to an affine linear subspace of $\mathbb{R}^n$. A similar strategy has been applied in [4, 13, 72, 76] to bound the number of points in $V_{>0}(f_1, \ldots, f_n)$. Assuming that $\sigma(f_2), \ldots, \sigma(f_n)$ are contained in the vertex set of an n-simplex (up to affine translation) and $\sigma(f_1)$ has m exponent vectors, the current best bound in the case $n = 2$ is

$$|V_{>0}(f_1, f_2)| \leq 2m - 2,$$

if $V_{>0}(f_1, f_2)$ is finite [4, Theorem 1]. For general n, the authors in [76] proved the bound $n + n^2 + \cdots + n^{m-1}$ on the number of isolated points in $V_{>0}(f_1, \ldots, f_n)$ under the assumption that $V_{>0}(f_2, \ldots, f_n)$ is smooth.

To prove the new bounds in Theorem 3.26 and Theorem 3.28 below, we need the following technical lemmata.

Lemma 3.24. *Let $C \subseteq \mathbb{R}^n$ be a closed convex set and $L \subseteq \mathbb{R}^n$ be an affine linear subspace. Then* $\mathrm{bd}(C) \cap L$ *is either infinite or contains at most two points.*

Proof. Our strategy is to show that the existence of three pairwise distinct points $x_1, x_2, x_3 \in \mathrm{bd}(C) \cap L$ implies that $\mathrm{bd}(C) \cap L$ contains infinitely many points. Since $C \cap L$ is convex, it contains $\mathrm{Conv}(x_1, x_2, x_3)$. If $\mathrm{relint}(\mathrm{Conv}(x_1, x_2, x_3)) \subseteq \mathrm{bd}(C) \cap L$, then we are done, since the relative interior of a convex set containing more than one point is infinite.

Thus, in the following, we assume $\mathrm{relint}(\mathrm{Conv}(x_1, x_2, x_3)) \cap \mathrm{int}(C) \cap L \neq \emptyset$. Let y be a point in this intersection such that $y \neq x_i$ for $i = 1, 2, 3$. Let $v_i := x_i - y$ denote the vector pointing from y into x_i for each $i = 1, 2, 3$. Note that $v_i \neq 0$, and

$$y + \lambda v_i = (1 - \lambda)y + \lambda x_i \in \mathrm{Aff}(x_1, x_2, x_3) \subseteq L \quad \text{for all } \lambda \in \mathbb{R}.$$

Furthermore, since $y \in \mathrm{Conv}(x_1, x_2, x_3)$, there exist $\mu_1, \mu_2, \mu_3 \geq 0$ with $\sum_{i=1}^3 \mu_i = 1$ and $\sum_{i=1}^3 \mu_i x_i = y$. A simple computation shows that $\sum_{i=1}^3 \mu_i v_i = \sum_{i=1}^3 \mu_i (x_i - y) = 0$. Thus, the linear span of v_1, v_2, v_3 is an at most two-dimensional linear subspace of $\mathbb{R}^n$.

We distinguish between two cases. First, we assume that there exists $i \in \{1, 2, 3\}$ and $\lambda_0 > 1$ such that $y + \lambda_0 v_i \in C$. Since C is convex and $y \in C$, we have

$$y + \lambda v_i = (1 - \tfrac{\lambda}{\lambda_0})y + \tfrac{\lambda}{\lambda_0}(y + \lambda_0 v_i) \in C, \quad \text{for all } 0 \leq \lambda \leq \lambda_0.$$

If for all $1 < \lambda \leq \lambda_0$ we have $y + \lambda v_i \in \mathrm{bd}(C)$, then $\mathrm{bd}(C) \cap L$ is infinite. If there exists $1 < \lambda_1 \leq \lambda_0$ such that $y + \lambda_1 v_i \in \mathrm{int}(C)$, then

$$x_i = (1 - \tfrac{1}{\lambda_1})y + \tfrac{1}{\lambda_1}(y + \lambda_1 v_i) \in \mathrm{Conv}(y, y + \lambda_1 v_i) \subseteq \mathrm{int}(C),$$

which contradicts $x_i \in \mathrm{bd}(C)$.

Now, we focus on the second case and assume that for all $i \in \{1, 2, 3\}$ and all $\lambda > 1$ $y + \lambda v_i \notin C$. In particular, v_1, v_2, v_3 are not contained in the recession cone 0^+C. If $v_i = \lambda v_j$ for $i \neq j$ and $\lambda > 1$, then $x_i = y + v_i = y + \lambda v_j \notin C$, which is a contradiction. If $v_i = \lambda v_j$ for $i \neq j$ and $0 < \lambda < 1$, then $\frac{1}{\lambda} v_i = v_j$, $\frac{1}{\lambda} > 1$ and by the same argument as above, we get a contradiction. From these, it follows that v_1, v_2, v_3 do not lie on a line. Thus, we can assume without loss of generality that $\mathrm{Conv}(v_1, v_3)$ and $\mathrm{Conv}(v_2, v_3)$ are two-dimensional cones.

In the next step of the proof, we show that at least one of the cones $\mathrm{Cone}(v_1, v_2)$, $\mathrm{Cone}(v_2, v_3)$, or $\mathrm{Cone}(v_1, v_3)$ does not intersect 0^+C and this cone has dimension two. If v_1 and v_2 are linearly independent, then $\mathrm{Cone}(v_1, v_2)$ is also two dimensional. Assume that there exist $w_1 \in \mathrm{Cone}(v_1, v_3), w_2 \in \mathrm{Cone}(v_2, v_3), w_3 \in \mathrm{Cone}(v_1, v_2)$ contained in 0^+C. Since $v_1, v_2, v_3, w_1, w_2, w_3$ are contained in the linear span of v_1, v_2, v_3, which is a linear subspace of dimension at most two, there exist $i, j, k \in \{1, 2, 3\}$ such that

$v_k \in \operatorname{Cone}(w_i, w_j)$. Since the recession cone is a convex cone [94, Theorem 8.1], we have $v_k \in \operatorname{Cone}(w_i, w_j) \subseteq 0^+C$, which is a contradiction.

If v_1 and v_2 are linearly dependent, that is $v_1 = \mu v_2$ for $\mu \in \mathbb{R} \setminus \{0\}$. From the above discussion it follows that $\mu < 0$. Assume that there exist $w_1 \in \operatorname{Cone}(v_1, v_3)$, $w_2 \in \operatorname{Cone}(v_2, v_3)$ contained in 0^+C. There exist $\lambda_1, \lambda_2, \lambda_3, \lambda_3' > 0$ such that $w_1 = \lambda_1 v_1 + \lambda_3 v_3$ and $w_2 = \lambda_2 v_2 + \lambda_3' v_3$. Since $\mu < 0$, there exists $t > 0$ with $\frac{\lambda_1}{\lambda_3}\mu + \frac{\lambda_2}{\lambda_3'}t = 0$. A direct computation shows

$$\tfrac{1}{\lambda_3(1+t)} w_1 + \tfrac{t}{\lambda_3'(1+t)} w_2 = \tfrac{1}{(1+t)}\big(\tfrac{\lambda_1}{\lambda_3}\mu v_2 + v_3 + \tfrac{\lambda_2 t}{\lambda_3'} v_2 + t v_3\big) = v_3$$

Thus, $v_3 \in \operatorname{Cone}(w_1, w_2) \subseteq 0^+C$, which is again a contradiction.

From the above arguments it follows that at least one of the cones $\operatorname{Cone}(v_1, v_2)$, $\operatorname{Cone}(v_2, v_3)$, $\operatorname{Cone}(v_1, v_3)$ does not intersect 0^+C and this cone has dimension two. We assume without loss of generality this holds for $\operatorname{Cone}(v_1, v_2)$. For all $v \in \operatorname{Cone}(v_1, v_2)$ and all $\lambda \gg 0$ we have $y + \lambda v \notin C$. Since C is closed and $y \in \operatorname{int}(C)$, there exists $\lambda_v > 0$ such that $y + \lambda_v v \in \operatorname{bd}(C)$. Thus, in that case we also have that $\operatorname{bd}(C) \cap L$ contains infinitely many points.

□

Lemma 3.25. *Let $C \subseteq \mathbb{R}^n$ be an open convex set of dimension n and let $f\colon C \to \mathbb{R}$ be a strictly convex function.*

(a) If $\{x \in C \mid f(x) < 0\} = \emptyset$, then $\{x \in C \mid f(x) = 0\}$ contains at most one point.

(b) If $\{x \in C \mid f(x) \leq 0\}$ is a closed subset of $\mathbb{R}^n$ with respect to the Euclidean topology, then

$$\operatorname{bd}\big(\{x \in C \mid f(x) \leq 0\}\big) = \{x \in C \mid f(x) = 0\}.$$

Proof. We prove part (a) by contradiction. Assume that there exists $x, y \in C$ with $x \neq y$ and $f(x) = f(y) = 0$. Since f is strictly convex it follows

$$f(\tfrac{1}{2}x + \tfrac{1}{2}y) < \tfrac{1}{2}f(x) + \tfrac{1}{2}f(y) = 0,$$

which contradicts $\{x \in C \mid f(x) < 0\} = \emptyset$.

To prove part (b), we distinguish two cases. If $\{x \in C \mid f(x) < 0\} = \emptyset$, then $\{x \in C \mid f(x) \leq 0\}$ is either empty or contains at most one point by (a), and the statement follows. If $\{x \in C \mid f(x) < 0\} \neq \emptyset$, then by [94, Theorem 7.6] the interior of $\{x \in C \mid f(x) \leq 0\}$ equals $\{x \in C \mid f(x) < 0\}$. Since $\{x \in C \mid f(x) \leq 0\}$ is closed, we have

$$\begin{aligned}\operatorname{bd}\big(\{x \in C \mid f(x) \leq 0\}\big) &= \{x \in C \mid f(x) \leq 0\} \setminus \{x \in C \mid f(x) < 0\} \\ &= \{x \in C \mid f(x) = 0\},\end{aligned}$$

which concludes the proof. □

A signomial f is a *binomial* if $\sigma(f)$ contains exactly 2 exponent vectors. Binomial equation systems are well studied. In the chemical reaction network literature for networks whose steady state variety is defined by binomial equations, there are easily checkable conditions to decide upon multistationarity [83, 91, 96]. It is well known that if $f_1, \dots, f_n$ are binomials, then $V_{>0}(f_1, \dots, f_n)$ contains at most one point or infinitely many (see e.g. [76, Theorem 4]). In Theorem 3.26, we replace one of the binomials by a signomial with arbitrarily many exponent vectors, assuming that only one exponent vector is negative. Under this assumption, we have the following bound.

Theorem 3.26. *Let $f_1, \dots, f_n \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be signomials such that $f_2, \dots, f_n$ are binomials, $\dim \mathrm{N}(f_1) = n$, and f_1 has exactly one negative exponent vector. Then $V_{>0}(f_1, \dots, f_n)$ is either infinite or contains at most two points.*

Proof. Let β denote the unique negative exponent vector of f_1 and define the signomial $g_1(x) := x^{-\beta} f_1(x)$. Note that $V_{>0}(g_1, f_2 \dots, f_n) = V_{>0}(f_1, f_2 \dots, f_n)$ and $\sigma_-(g_1) = \{0\}$. Denote the elements of $\sigma_+(g_1)$ by $\alpha_1, \dots, \alpha_m$ and let $\mathrm{Exp} \colon \mathbb{R}^n \to \mathbb{R}^n_{>0}$ and $\mathrm{Log} \colon \mathbb{R}^n_{>0} \to \mathbb{R}^n$ be the coordinate-wise exponential and natural logarithm functions. Consider the exponential sum

$$G(y) := g_1(\mathrm{Exp}(y)) = c_0 + \sum_{i=1}^{m} c_i e^{\alpha_i \cdot y},$$

where $c_0, \dots, c_m$ denote the coefficients of g_1.

In the proof of Theorem VI.4.11, we showed that the Hessian of G is positive definite for all $y \in \mathbb{R}^n$. We recall this computation for the sake of completeness. Consider the matrix

$$\tilde{A} = \begin{pmatrix} \alpha_1 & \dots & \alpha_m \end{pmatrix} \in \mathbb{R}^{n \times m}.$$

Since $c_i e^{\alpha_i \cdot y}$ is positive for $i = 1, \dots, m$, its square root is a real number. The Hessian of G at y is given by

$$\begin{aligned} \mathrm{Hess}_G(y) &= \sum_{i=1}^{m} (c_i e^{\alpha_i \cdot y}) \alpha_i \alpha_i^\top = \tilde{A} \operatorname{diag}\left((c_i e^{\alpha_i \cdot y})_{i=1,\dots,m} \right) \tilde{A}^\top \\ &= \left(\tilde{A} \operatorname{diag}\left((\sqrt{c_i e^{\alpha_i \cdot y}})_{i=1,\dots,m} \right) \right) \left(\tilde{A} \operatorname{diag}\left((\sqrt{c_i e^{\alpha_i \cdot y}})_{i=1,\dots,m} \right) \right)^\top. \end{aligned}$$

By assumption $\mathrm{N}(f_1)$ has dimension n. The same is true for $\mathrm{N}(g_1)$, since it is an affine translate of $\mathrm{N}(f_1)$. Since $\sigma_-(g_1) = \{0\}$, it follows that

$$n = \mathrm{rk}(\tilde{A}) = \mathrm{rk}(\tilde{A} \operatorname{diag}\left((\sqrt{c_i e^{\alpha_i \cdot y}})_{i=1,\dots,m} \right)),$$

in particular the rows of $\tilde{A} \operatorname{diag}\left((\sqrt{c_i e^{\alpha_i \cdot y}})_{i=1,\dots,m} \right)$ are linearly independent. From [62, Theorem 7.2.10] it follows that $\mathrm{Hess}_G(y)$ is positive definit. Using [23, Section 3.1.4], we conclude that G is a strictly convex function.

By [94, Theorem 4.6], the sublevel set

$$C := \{y \in \mathbb{R}^n \mid G(y) \leq 0\} = \operatorname{Log}\left(g_1^{-1}(\mathbb{R}_{\leq 0})\right)$$

is a convex set. Since G is a continuous function on $\mathbb{R}^n$, the set C is a closed subset of $\mathbb{R}^n$. From Lemma 3.25, it follows that

$$\operatorname{bd}(C) = \{y \in \mathbb{R}^n \mid G(y) = 0\} = \operatorname{Log}(V_{>0}(g_1)).$$

In the second part of the proof, we show that the image of $V_{>0}(f_2, \ldots, f_n)$ under Log is an affine linear subspace of $\mathbb{R}^n$. If there exists f_i, $i = 2, \ldots, n$, such that its coefficients have the same sign, then $V_{>0}(f_2, \ldots, f_n) = \emptyset$. In the following, we assume that this is not the case, and denote by γ_i (resp. β_i) the positive (resp. negative) exponent vector of f_i and $c_{\gamma_i}, c_{\beta_i}$ the corresponding coefficients of f_i for each $i = 2, \ldots, n$. The set

$$\begin{aligned}\operatorname{Log}(V_{>0}(f_2, \ldots, f_n)) &= \left\{ \operatorname{Log}(x) \mid x \in \mathbb{R}^n_{>0} \text{ and } c_{\gamma_i} x^{\gamma_i} = -c_{\beta_i} x^{\beta_i} \text{ for } i = 2, \ldots, n \right\} \\ &= \left\{ y \in \mathbb{R}^n \mid \log(c_{\gamma_i}) + \gamma_i \cdot y = \log(-c_{\beta_i}) + \beta_i \cdot y \text{ for } i = 2, \ldots, n \right\}\end{aligned}$$

is an affine linear subspace. Therefore, using Lemma 3.24 we conclude that $\operatorname{bd}(C) \cap \operatorname{Log}(V_{>0}(f_2, \ldots, f_n))$ contains at most 2 or infinitely many points. Since Log is a bijection, the theorem follows. □

In Theorem 3.28 below, we present another condition on the signed support that implies that a square system has either infinitely many or at most two positive solutions. Before proving this result, we state another technical lemma.

Lemma 3.27. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a strictly convex signomial. If $\sigma_-(f) \subseteq \operatorname{int}(\mathrm{N}(f))$, then $f^{-1}(\mathbb{R}_{\leq 0})$ is a closed convex subset of $\mathbb{R}^n$.*

Proof. Since f is strictly convex, $f^{-1}(\mathbb{R}_{<0}) = \emptyset$ implies that $f^{-1}(\mathbb{R}_{\leq 0})$ is a point, in particular a closed subset of $\mathbb{R}^n$, see Lemma 3.25. In the following, we assume $f^{-1}(\mathbb{R}_{<0}) \neq \emptyset$. We prove the statement in two steps. First, we show that

$$f^{-1}(\mathbb{R}_{\leq 0}) = \overline{f^{-1}(\mathbb{R}_{<0})}^{>0}, \tag{3.8}$$

where the closure is taken in the subspace topology of $\mathbb{R}^n_{>0} \subseteq \mathbb{R}^n$, see [73, Definition 1.2.82]. Since f is continuous, the set $f^{-1}(\mathbb{R}_{\leq 0})$ is closed in the subspace topology of $\mathbb{R}^n_{>0}$. Thus to prove (3.8), it is enough to show that for every $y \in f^{-1}(\mathbb{R}_{\leq 0})$ there exists a sequence in $f^{-1}(\mathbb{R}_{<0})$ converging to y. If $y \in f^{-1}(\mathbb{R}_{<0})$, then we take the constant sequence given by y. If $f(y) \in f^{-1}(\mathbb{R}_{\leq 0}) \setminus f^{-1}(\mathbb{R}_{<0})$, then $f(y) = 0$. Since f is a convex function, by [94, Theorem 4.1] we have for any fixed $x \in f^{-1}(\mathbb{R}_{<0})$ and for all $t \in (0, 1)$ that

$$f((1-t)y + tx) \leq (1-t)f(y) + tf(x) = tf(x) < 0.$$

Thus, $(1-t)y + tx \in f^{-1}(\mathbb{R}_{<0})$. As $t \to 0$, we have $(1-t)y + tx \to y$, which implies that y lies in $\overline{f^{-1}(\mathbb{R}_{<0})}^{>0}$.

In the second step, we show

$$\overline{f^{-1}(\mathbb{R}_{<0})}^{>0} = \overline{f^{-1}(\mathbb{R}_{<0})}, \tag{3.9}$$

where on the right-hand side the closure is taken in the Euclidean topology of $\mathbb{R}^n$. Since $\overline{f^{-1}(\mathbb{R}_{<0})}$ is closed, it contains all limit points of $f^{-1}(\mathbb{R}_{<0})$. In particular, it contains all limit points that are contained in $\mathbb{R}^n_{>0}$. Therefore, the left-hand side of (3.9) is contained in the right-hand side. To see the other inclusion, assume that there exists $y \in \overline{f^{-1}(\mathbb{R}_{<0})} \setminus \mathbb{R}^n_{>0}$. For such a y it holds that $y \in \mathbb{R}^n_{\geq 0} \setminus \mathbb{R}^n_{>0}$. Thus there exists $k \in [n]$ with $y_k = 0$. Since $y \in \overline{f^{-1}(\mathbb{R}_{<0})}$ there exists a sequence $\{x_n\}_n \subseteq f^{-1}(\mathbb{R}_{<0})$ with $x_n \to y$. Since the logarithm function is continuous, it follows that $\log(x_{n,k}) \to \log(y_k) = -\infty$. From the assumption $\sigma_-(f) \subseteq \operatorname{int}(\mathrm{N}(f))$ and Corollary V.3.6, it follows that $\operatorname{Log}(f^{-1}(\mathbb{R}_{<0}))$ is a bounded set, which contradicts $\log(x_{n,k}) \to -\infty$.

Combining (3.8) and (3.9), we conclude that $f^{-1}(\mathbb{R}_{\leq 0})$ is a closed subset of $\mathbb{R}^n$. From [94, Theorem 4.6], it follows that $f^{-1}(\mathbb{R}_{\leq 0})$ is a convex set. □

Theorem 3.28. *Let $f_1, \dots, f_n \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be signomials. Assume that there exists an n-simplex $P \subseteq \mathbb{R}^n$ and $v_2, \dots, v_n \in \mathbb{R}^n$ such that*

$$\begin{aligned} &\sigma_-(f_1) \subseteq P, \quad \sigma_+(f_1) \subseteq P^-, \quad \sigma(f_1) \cap (P \cup P^-) \neq \emptyset, \\ &\sigma_-(f_1) \subseteq \operatorname{int}(\mathrm{N}(f_1)), \qquad \sigma(f_i) \subseteq \operatorname{Vert}(v_i + P) \quad \text{for } i = 2, \dots n. \end{aligned}$$

Then $V_{>0}(f_1, f_2, \dots, f_n)$ is either infinite or contains at most two points.

Proof. First, we transform $f_1, \dots, f_n$ into more convenient signomials using affine transformations as in Lemma 3.4. Multiplying the equation f_i by x^{-v_i} does not change the number of positive solutions. Thus, we assume without loss of generality that $\sigma(f_i) \subseteq \operatorname{Vert}(P)$ for $i = 2, \dots, n$. From Lemma III.4.5 it follows that there exists an invertible matrix $M \in \mathbb{R}^{n \times n}$ and $v \in \mathbb{R}^n$ such that

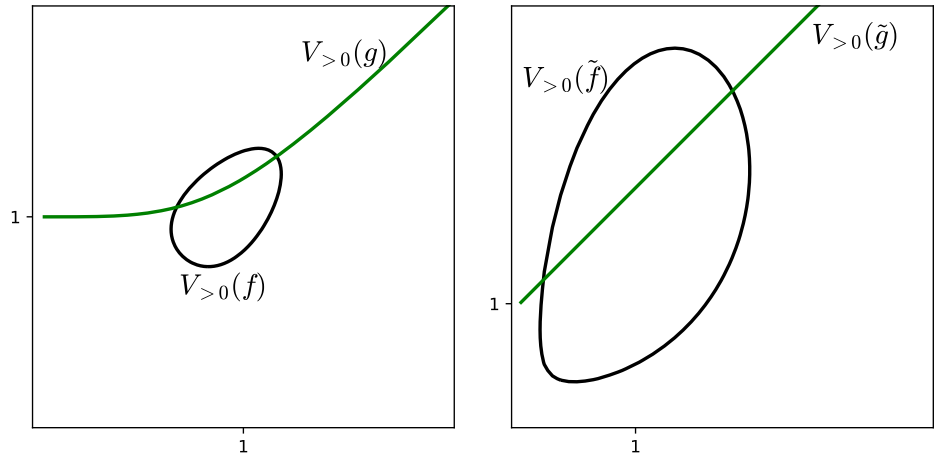
$$M\sigma_-(f_1) + v \subseteq \Delta_n, \quad M\sigma_+(f_1) + v \subseteq \Delta_n^-, \quad (M\sigma(f_1) + v) \cap (\Delta_n \cup \Delta_n^-) \neq \emptyset, \tag{3.10}$$

$$M\sigma(f_i) + v \subseteq \operatorname{Vert}(\Delta_n), \quad i = 2, \dots, n. \tag{3.11}$$

Consider the signomials

$$g_i \colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto x^v f_i(x^M), \quad i = 1, \dots, n.$$

Note that the number of points in $V_{>0}(f_1, \dots, f_n)$ is the same as in $V_{>0}(g_1, \dots, g_n)$. From $\sigma(g_i) \subseteq \operatorname{Vert}(\Delta_n)$, $i = 2, \dots, n$, it follows that the functions $g_2, \dots, g_n$ are affine linear and hence $V_{>0}(g_2, \dots, g_n)$ is the intersection of an affine linear subspace with $\mathbb{R}^n_{>0}$.



(a) (b)

Figure 3.4: Illustration of Example 3.29 (a) Positive zero sets of $f = x_1^7x_2 + x_2^7 + x_1^6 - 3x_1^2x_2^3 - 2x_1^2x_2^2 + 1$ and $g = x_1^5x_2 - x_1x_2^5 + x_1x_2$ (b) Positive zero sets of $\tilde{f} = x_1^{\frac{6}{4}} + x_1^{\frac{-1}{4}}x_2^{\frac{6}{4}} + x_1^{\frac{5}{4}}x_2^{\frac{-1}{4}} - 3x_1^{\frac{1}{4}}x_2^{\frac{2}{4}} - 2x_1^{\frac{1}{4}}x_2^{\frac{1}{4}} + x_1^{\frac{-1}{4}}x_2^{\frac{-1}{4}}$, $\tilde{g} = x_1 - x_2 + 1$.

Since $M\sigma_-(f_1) + v = \sigma_-(g_1)$ and $M\sigma_+(f_1) + v = \sigma_+(g_1)$ (cf. Lemma 3.4), from (3.10) follows that g_1 is a strictly convex function by [80, Theorem 7]. This implies that $g_1^{-1}(\mathbb{R}_{\leq 0})$ is a closed convex set by Lemma 3.27. From Lemma 3.25, it follows that $\mathrm{bd}(g_1^{-1}(\mathbb{R}_{\leq 0})) = V_{>0}(g_1)$. Now, the theorem follows from Lemma 3.24. □

Example 3.29. We revisit the polynomial

$$f = x_1^7x_2 + x_2^7 + x_1^6 - 3x_1^2x_2^3 - 2x_1^2x_2^2 + 1$$

from Example 3.23. Its positive and negative exponent vectors are separated by the simplex $P = \mathrm{Conv}((1,1),(5,1),1,5))$, cf. Figure 3.3(a). Consider the polynomial

$$g = x_1^5x_2 - x_1x_2^5 + x_1x_2.$$

Since $\sigma(g) \subseteq P$, from Theorem 3.28 follows that $V_{>0}(f) \cap V_{>0}(g)$ is infinite or contains at most two points. The sets $V_{>0}(f)$, $V_{>0}(g)$ are shown in Figure 3.4(a).

In the following, we illustrate the idea of the proof of Theorem 3.28. The affine map $\mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \frac{1}{4}x - (\frac{1}{4}, \frac{1}{4})$ transforms P into the standard 2-simplex. The signomials f, g are transformed into

$$\tilde{f} = x_1^{\frac{6}{4}} + x_1^{\frac{-1}{4}}x_2^{\frac{6}{4}} + x_1^{\frac{5}{4}}x_2^{\frac{-1}{4}} - 3x_1^{\frac{1}{4}}x_2^{\frac{2}{4}} - 2x_1^{\frac{1}{4}}x_2^{\frac{1}{4}} + x_1^{\frac{-1}{4}}x_2^{\frac{-1}{4}},$$
$$\tilde{g} = x_1 - x_2 + 1.$$

The function $\tilde{g}$ is affine linear and $\tilde{f}\colon \mathbb{R}^2_{>0} \to \mathbb{R}$ is a convex function by Lemma 3.20. Thus, $V_{>0}(\tilde{g})$ is an affine linear subspace and $V_{>0}(\tilde{f})$ is the boundary of a convex set, see Figure 3.4(b). From Lemma 3.24 follows that $V_{>0}(\tilde{f}) \cap V_{>0}(\tilde{g})$ is either infinite or it contains at most 2 points.

In Section 3.3, we observed that the existence of a separating hyperplane, a pair of enclosing hyperplanes, or a separating simplex of the signed support implies that the preimage of the negative real line has at most one or two connected components. Theorem 3.26 and Theorem 3.28 demonstrate that, under similar conditions on the signed support of signomials in a square system, the number of positive solutions is at most two. Could other polyhedral conditions on the signed support provide upper bounds on the number of positive solutions? Is it possible to use separating hyperplanes to establish such bounds? In Papers I and II, we observed that even very restrictive assumptions on the signed support can lead to interesting applications. My future plan is to pursue this path, and to investigate the connection between the signed support and upper bounds on the number of positive real solutions of square systems. Such results would be highly valuable in applications where bounding the number of positive real solutions is crucial, such as addressing questions of multistationarity in the study of chemical reaction networks.

4

Real tropicalization

This chapter offers an overview of the content of Paper V and Paper VI. To relate these two papers to previous results in tropical geometry, in Section 4.1, we discuss how to tropicalize algebraic varieties over the field of complex Puiseux series. In Section 4.2, our focus shifts to tropicalization of semi-algebraic sets over the field of real Puiseux series and discuss why the notion of real tropicalization is analogous to tropicalization of complex algebraic varieties. Afterward, we present the main results of Paper V. In Section 4.3, we elaborate on when it is possible to capture the isotopy type of a hypersurface with a polyhedral object, called the Viro diagram. For that, we study the signed reduced A-discriminant, which is the main object of Paper VI.

4.1 Tropical geometry

The main theorems of tropical geometry concern algebraic varieties over algebraically closed fields with non-trivial valuations. For simplicity, in this thesis, we only consider the field of Puiseux series with complex coefficients. The field of *complex Puiseux series* $\mathcal{C} = \mathbb{C}\{\{t\}\}$ contains formal power series of the form

$$c(t) = \sum_{k=k_0}^{\infty} c_k t^{\frac{k}{N}}, \quad \text{for } k_0 \in \mathbb{Z},\ N \in \mathbb{N},\ c_k \in \mathbb{C},\ c_{k_0} \neq 0. \tag{4.1}$$

The *valuation map* $\operatorname{val}\colon \mathcal{C}^\times \to \mathbb{R}$ is given by $\operatorname{val}(c(t)) = \frac{k_0}{N}$ [78, Example 2.1.3]. Here, we use the notation $\mathcal{C}^\times = \mathcal{C} \setminus \{0\}$. We denote the coordinate-wise valuation map by

$$\operatorname{val}\colon (\mathcal{C}^\times)^n \to \mathbb{R}^n, \quad x \mapsto \operatorname{val}(x) := (\operatorname{val}(x_1), \dots, \operatorname{val}(x_n)). \tag{4.2}$$

For a Laurent polynomial $f = \sum_{\mu \in \sigma(f)} c_\mu(t) x^\mu \in \mathcal{C}[x_1^\pm, \dots, x_n^\pm]$, we write

$$V_{\mathcal{C}}(f) := \left\{ x \in (\mathcal{C}^\times)^n \mid f(x) = 0 \right\} \tag{4.3}$$

for the set of its zeros in $(\mathcal{C}^\times)^n$, and define its *tropicalization*

$$\operatorname{Trop}(V_{\mathcal{C}}(f)) := \overline{\left\{ -\operatorname{val}(x) \mid x \in V_{\mathcal{C}}(f) \right\}}, \tag{4.4}$$

where the closure is with respect to the Euclidean topology in $\mathbb{R}^n$. The tropicalization $\mathrm{Trop}(V_{\mathcal{C}}(f))$ is a finite union of polyhedral sets. To see this, first we introduce the *tropical hypersurface* defined by f as the set

$$\mathrm{trop}(f) := \big\{ w \in \mathbb{R}^n \mid \max_{\mu \in \sigma(f)} (-\mathrm{val}(c_\mu(t)) + w \cdot \mu) \text{ is achieved at least twice} \big\}, \quad (4.5)$$

which is a finite union of polyhedral sets [78, Proposition 3.1.6]. In this thesis, we use the maximum to define tropical hypersurfaces as in [64]. This convention has the advantage that the logarithmic limit of a complex hypersurface equals its tropicalization (cf. Theorem 4.3). Some authors prefer using the minimum to define tropical hypersurfaces [78]. In such cases, the logarithmic limit is the reflection of the tropicalization through the origin.

In (4.4), we introduced the tropicalization of the hypersurface defined by a Laurent polynomial f, and in (4.5) the tropical hypersurface defined by f. The following theorem states that the two objects coincide.

Theorem 4.1. *(Kapranov's Theorem [40, Theorem 2.1.1] see also [78, Theorem 3.1.3]) Let $f \in \mathcal{C}[x_1^{\pm}, \dots, x_n^{\pm}]$ be a Laurent polynomial. Then*

$$\mathrm{Trop}(V_{\mathcal{C}}(f)) = \mathrm{trop}(f).$$

Theorem 4.1 can be generalized to varieties of higher codimension as follows. Let $I \subseteq \mathcal{C}[x_1^{\pm}, \dots, x_n^{\pm}]$ be an ideal and denote $V_{\mathcal{C}}(I)$ its vanishing locus in $(\mathcal{C}^\times)^n$

$$V_{\mathcal{C}}(I) := \big\{ x \in (\mathcal{C}^\times)^n \mid \forall f \in I \colon f(x) = 0 \big\}.$$

Similarly as in (4.4), the tropicalization of $V_{\mathcal{C}}(I)$ is defined as

$$\mathrm{Trop}(V_{\mathcal{C}}(I)) := \overline{\big\{ -\mathrm{val}(x) \mid x \in V_{\mathcal{C}}(I) \big\}}. \quad (4.6)$$

By the Fundamental Theorem of Tropical Algebraic Geometry, the tropicalization $\mathrm{Trop}(V_{\mathcal{C}}(I))$ is the intersection of the tropical hypersurfaces defined by the polynomials in the ideal I.

Theorem 4.2. *(Fundamental Theorem of Tropical Algebraic Geometry [78, Theorem 3.2.3]) For an ideal $I \subseteq \mathcal{C}[x_1^{\pm}, \dots, x_n^{\pm}]$, we have*

$$\mathrm{Trop}(V_{\mathcal{C}}(I)) = \bigcap_{f \in I} \mathrm{trop}(f).$$

By [78, Theorem 2.6.6], every ideal $I \subseteq \mathcal{C}[x_1^{\pm}, \dots, x_n^{\pm}]$ has a finite *tropical basis*, that is, there exist finitely many polynomials $g_1, \dots, g_m \in I$ such that

$$\mathrm{Trop}(V_{\mathcal{C}}(I)) = \bigcap_{i=1}^{m} \mathrm{trop}(g_i).$$

There exist algorithms based on Gröbner basis methods for computing the tropicalization of $V_{\mathcal{C}}(I)$, performing such computations in practice is highly challenging [21, 81].

In the rest of this section, we focus on hypersurfaces, where the coefficients of the defining polynomial are complex numbers. The field of complex Puiseux series $\mathcal{C}$ is a field extension of $\mathbb{C}$. Every complex number $z \in \mathbb{C}$ can be viewed as a constant Puiseux series zt^0. In particular, if $z \neq 0$, then $\operatorname{val}(z) = 0$. For a Laurent polynomial $f \in \mathbb{C}[x_1^{\pm}, \dots, x_n^{\pm}]$, we write

$$V_{\mathbb{C}}(f) := \big\{x \in (\mathbb{C}^{\times})^n \mid f(x) = 0\big\}.$$

For each $t > 0$, denote $\log_t \colon \mathbb{R}_{>0} \to \mathbb{R}$ the logarithm map with base t and let $|z|$ be the usual Archimedean absolute value of a complex number $z \in \mathbb{C}$. Consider the map

$$\operatorname{Log}_t \colon (\mathbb{C}^{\times})^n \to \mathbb{R}^n, \quad x \mapsto \big(\log_t(|x_1|), \dots, \log_t(|x_n|)\big). \tag{4.7}$$

The *logarithmic limit* of a set $S \subseteq (\mathbb{C}^{\times})^n$ is defined as

$$\mathcal{L}(S) := \lim_{t \to \infty} \operatorname{Log}_t(S), \tag{4.8}$$

see [1, Section 2] for a precise definition of the limit for a family of subsets of $\mathbb{R}^n$.

Theorem 4.3. *Let $f = \sum_{\mu \in \sigma(f)} c_\mu x^\mu \in \mathbb{C}[x_1^{\pm}, \dots, x_n^{\pm}]$ be a Laurent polynomial. The following sets coincide:*

(i) the logarithmic limit $\mathcal{L}(V_{\mathbb{C}}(f))$,

(ii) the tropicalization $\operatorname{Trop}(V_{\mathcal{C}}(f))$,

(iii) the tropical hypersurface $\operatorname{trop}(f)$,

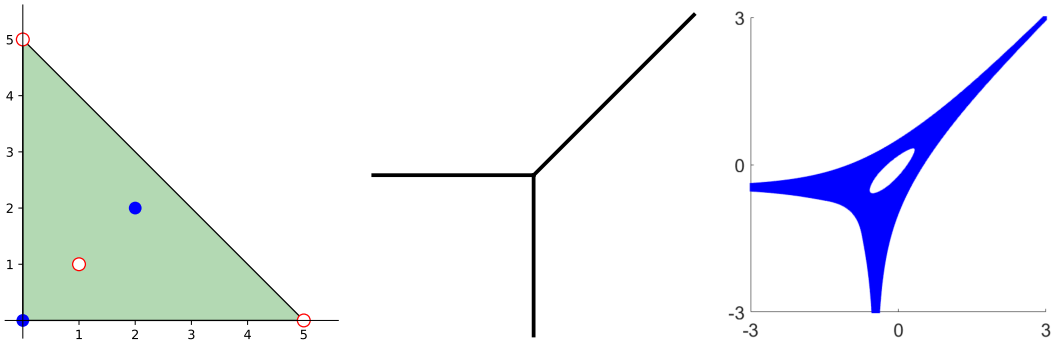
(iv) the $(n-1)$-skeleton of the outer normal fan of the Newton polytope of f.

Proof. In [82, Corollary 6.4], it has been shown that the logarithmic limit $\mathcal{L}(V_{\mathbb{C}}(f))$ and the tropicalization $\operatorname{Trop}(V_{\mathcal{C}}(f))$ coincide. The sets in (ii) and (iii) are equal by Kapranov's theorem (Theorem 4.1). Since $\operatorname{val}(c_\mu) = 0$ for all $\mu \in \sigma(f)$, it follows directly from the definition (cf. (4.5)) that $\operatorname{trop}(f)$ is the $(n-1)$-skeleton of the outer normal fan of $\mathrm{N}(f)$ (cf. Section 3.2). □

Example 4.4. To illustrate Theorem 4.3, we consider the polynomial

$$f = 10x_1^5 + 10x_2^5 - 33x_1^2x_2^2 + 10x_1x_2 - 1, \tag{4.9}$$

which will serve as a running example in the upcoming sections. The Newton polytope of f and the 1-skeleton of its outer normal fan are shown in Figure 4.1. By Theorem 4.3, this 1-skeleton coincide with the tropicalization $\operatorname{Trop}(V_{\mathcal{C}}(f))$ and with the logarithmic limit of $V_{\mathbb{C}}(f)$. For $t = e$, Euler's number, the set $\operatorname{Log}_e(V_{\mathbb{C}}(f))$ is depicted in Figure 4.1(c). The set $\operatorname{Log}_e(V_{\mathbb{C}}(f))$ is sometimes called the *amoeba* of f. To create Figure 4.1(c), we used [22].



(a) (b) (c)

Figure 4.1: (a) The Newton polytope of $f = 10x_1^5 + 10x_2^5 - 33x_1^2x_2^2 + 10x_1x_2 - 1$ from Example 4.4 (b) Logarithmic limit of $V_{\mathbb{C}}(f)$ (c) Image of $V_{\mathbb{C}}(f)$ under the map Log_t (4.7) for $t = e$, Euler's number.

4.2 Real tropical geometry

In this section, we study tropicalization of semi-algebraic sets and discuss the results of Paper V. The field of *real Puiseux series* $\mathcal{R} = \mathbb{R}\{\{t\}\}$ is a subfield of $\mathbb{C}\{\{t\}\}$ containing formal power series (cf. (4.1)) whose coefficients are real numbers. The valuation map $\mathrm{val}\colon (\mathcal{R}^\times)^n \to \mathbb{R}^n$ is the restriction of the valuation map of complex Puiseux series (4.2). It is known that $\mathcal{R}$ is a real closed field [7, Theorem 2.91]. An element $x(t) \in \mathcal{R}^\times$ is *positive* (resp. *negative*) if the coefficient of the smallest non-zero term $t^{\mathrm{val}(x(t))}$ is positive (resp. negative). We denote by $\mathcal{R}_{>0}$ the set of positive real Puiseux series.

Following the notation from Section 3.1, we denote by $\sigma_-(f)$ the set of exponent vectors of a polynomial $f \in \mathcal{R}[x_1, \ldots, x_n]$ whose corresponding coefficients are negative. For $f_1, \ldots, f_k \in \mathcal{R}[x_1, \ldots, x_n]$, we define the semi-algebriac set

$$S_{\mathcal{R}}(f_1, \ldots, f_k) := \big\{x \in \mathcal{R}^n_{>0} \mid f_1(x) < 0, \ldots, f_k(x) < 0\big\}, \tag{4.10}$$

and its *real tropicalization*

$$\mathrm{Trop}(S_{\mathcal{R}}(f_1, \ldots, f_k)) := \overline{\big\{-\mathrm{val}(x) \mid x \in S_{\mathcal{R}}(f_1, \ldots, f_k)\big\}}.$$

where the closure is taken in the Euclidean topology of $\mathbb{R}^n$. Note that this definition is analogous to the tropicalization of complex algebraic varieties (4.6). Similar to (4.5), we define

$$\mathrm{trop}^-(f) := \big\{w \in \mathbb{R}^n \mid \max_{\mu \in \sigma(f)} (-\mathrm{val}(c_\mu(t)) + w \cdot \mu) \text{ is achieved for some } \mu_0 \in \sigma_-(f)\big\}.$$

In [66], the authors proved a semi-algebraic analogue of the Fundamental Theorem of Tropical Algebraic Geometry (Theorem 4.2).

Theorem 4.5. *([66, Theorem 6.9], see also [20, Section 2.1]) Let $f_1, \dots, f_k \in \mathcal{R}[x_1, \dots, x_n]$ be polynomials and let $S = S_{\mathcal{R}}(f_1, \dots, f_k)$ be a semi-algebraic set as in (4.10). Then*

$$\operatorname{Trop}(S_{\mathcal{R}}(f_1, \dots, f_k)) = \bigcap_{f \leq 0 \text{ on } S} \operatorname{trop}^-(f), \tag{4.11}$$

where the intersection is over all polynomials f, which are non-positive on S.

It is known that the intersection on the right-hand side of (4.11) might be taken over finitely many polynomials. However, there exists no general algorithm in the literature that would determine such a finite list of polynomials. Therefore, it is valuable to have approximations of $\operatorname{Trop}(S_{\mathcal{R}}(f_1, \dots, f_k))$ which might be easier to compute. The following theorem provides such an approximation.

Theorem 4.6. *([66, Theorem 6.12]) Let $f_1, \dots, f_k \in \mathcal{R}[x_1, \dots, x_n]$ be polynomials and let $S_{\mathcal{R}}(f_1, \dots, f_k)$ be a semi-algebraic set as in (4.10). Then*

$$\operatorname{int}\Big(\bigcap_{i=1}^{k} \operatorname{trop}^-(f_i)\Big) \subseteq \operatorname{Trop}(S_{\mathcal{R}}(f_1, \dots, f_k)) \subseteq \bigcap_{i=1}^{k} \operatorname{trop}^-(f_i).$$

From now on, we focus on the case when the defining equations of the semi-algebraic set (4.10) have real coefficients. For $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_n]$, we set

$$S(f_1, \dots, f_k) := \big\{x \in \mathbb{R}^n_{>0} \mid f_1(x) < 0, \dots, f_k(x) < 0\big\}. \tag{4.12}$$

The logarithmic limit of $S(f_1, \dots, f_k)$ (cf. (4.8)) coincides with the real tropicalization of $S_{\mathcal{R}}(f_1, \dots, f_k)$.

Theorem 4.7. *[1, Corollary 4.6] Let $f_1, \dots, f_k \in \mathbb{R}[x_1, \dots, x_n]$ be polynomials, then*

$$\mathcal{L}(S(f_1, \dots, f_k)) = \operatorname{Trop}(S_{\mathcal{R}}(f_1, \dots, f_k)).$$

In Paper V, we studied logarithmic limits of sets of the form (4.12). We used the notation $\operatorname{Trop}(S(f_1, \dots, f_k))$ for the logarithmic limit. In the rest of the section, we follow this convention and write

$$\operatorname{Trop}(S(f_1, \dots, f_k)) := \mathcal{L}(S(f_1, \dots, f_k)).$$

This slight abuse of notation is justified by Theorem 4.7. All the results in Paper V are valid for polynomials $f_1, \dots, f_k$ with real exponents, that is, for signomials (cf. Section 3.1).

The *negative normal cone* of the Newton polytope of a signomial f contains all outer normal vectors of $\mathrm{N}(f)$ such that the corresponding face of $\mathrm{N}(f)$ contains a negative exponent vector of f,

$$\mathcal{N}_f^- := \{v \in \mathbb{R}^n \mid \mathrm{N}(f)_v \cap \sigma_-(f) \neq \emptyset\}.$$

Note that if $f \in \mathbb{R}[x_1, \dots, x_n]$, then $\mathcal{N}_f^- = \mathrm{trop}^-(f)$. The *actual negative normal cone* of a collection of signomials $f_1, \dots, f_k$ is given by

$$\Sigma(f_1, \dots, f_k) := \{v \in \mathbb{R}^n \mid \bigcap_{i=1}^{k} f_{i|\,\mathrm{N}(f_i)_v}^{-1}(\mathbb{R}_{<0}) \neq \emptyset\}.$$

It holds that $\Sigma(f) \subseteq \mathcal{N}_f^-$, and it is easy to construct examples where the inclusion is strict, see Figure 1 in Paper V or Example V.3.12.

The next theorem is a version of Theorem 4.6 for signomials, which additionally includes the actual negative normal cone.

Theorem 4.8. *(Theorem V.3.3) Let $f_1, \dots, f_k \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be signomials. Then, we have*

$$\mathrm{int}\,\big(\bigcap_{i=1}^{k} \mathcal{N}_{f_i}^-\big) \subseteq \Sigma(f_1, \dots, f_k) \subseteq \mathrm{Trop}(S(f_1, \dots, f_k)) \subseteq \bigcap_{i=1}^{k} \mathcal{N}_{f_i}^-.$$

Example 4.9. We revisit the polynomial

$$f = 10x_1^5 + 10x_2^5 - 33x_1^2x_2^2 + 10x_1x_2 - 1,$$

from Example 4.4. Its negative normal cone is spanned by the vectors $(0, -1)$ and $(-1, 0)$, see Figure 4.2 (cf. Figure 3.1 and Figure 4.1). Since the closure of $\mathrm{int}(\mathcal{N}_f^-)$ equals $\mathcal{N}_f^-$, it follows that all the inclusions in Theorem 4.8 are equalities. In particular, $\mathcal{N}_f^-$ coincide with the logarithmic limit of $S(f) = f^{-1}(\mathbb{R}_{<0})$.

It might happen that all the inclusions in Theorem 4.8 are strict, see Example V.3.12. However, if a signomial f satisfies some additional properties, one might use the negative normal cone or the actual negative normal cone to compute the real tropicalization of $S(f)$.

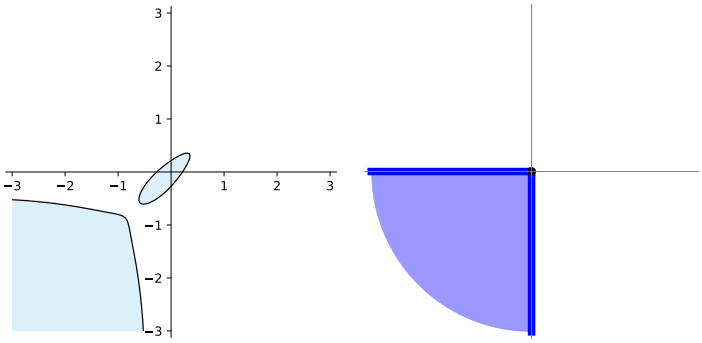
Proposition 4.10. *(Corollary V.3.6, Corollary V.3.10) Let $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial.*

(i) If $\sigma(f) = \mathrm{Vert}(\mathrm{N}(f))$, then

$$\mathrm{Trop}(S(f)) = \mathcal{N}_f^-.$$

(ii) If f has generic enough coefficients, then

$$\mathrm{Trop}(S(f)) = \Sigma(f).$$



(a) (b)

Figure 4.2: Illustration of Example 4.9, $f = 10x_1^5 + 10x_2^5 - 33x_1^2x_2^2 + 10x_1x_2 - 1$ (a) Image of $S(f) = f^{-1}(\mathbb{R}_{<0})$ under the map $\mathrm{Log}_t \colon \mathbb{R}^2_{>0} \to \mathbb{R}^2$ for $t = e$, Euler's number (b) Negative normal cone of the Newton polytope of f, which agrees with $\mathrm{Trop}(S(f))$ and with the logarithmic limit of $S(f)$.

4.3 Viro's patchworking and the signed A-discriminant

In Section 3.3, we used polyhedral geometric properties of the exponent vectors of a signomial f and the signs of its coefficients to provide bounds on the number of connected components of the set $f^{-1}(\mathbb{R}_{<0})$. In this section, we focus on the topology of the hypersurface

$$V_{>0}(f) = \big\{x \in \mathbb{R}^n_{>0} \mid f(x) = 0\big\}.$$

We fix a finite set of exponent vectors $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$, a *sign distribution* $\varepsilon \in \{1, -1\}^{n+k+1}$, and consider the family of signomials

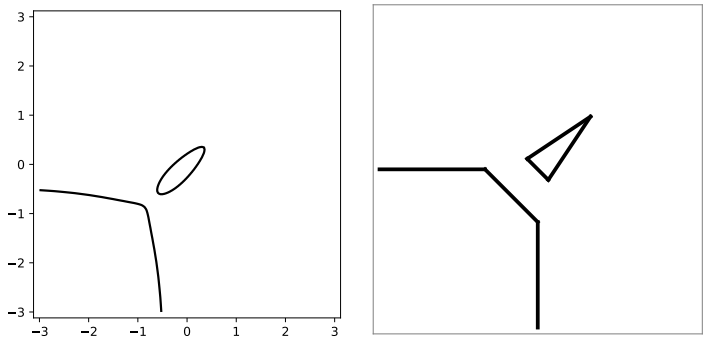
$$f_c \colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto \sum_{i=1}^{n+k+1} c_i x^{\alpha_i},$$

for $c \in \mathbb{R}^n$ such that $\mathrm{sign}(c) = \varepsilon$. We call $(\mathcal{A}, \varepsilon)$ a *signed support*. Similar to Chapter 3, we divide $\mathcal{A}$ into positive and negative exponent vectors

$$\mathcal{A}_+ := \{\alpha_i \in \mathcal{A} \mid \varepsilon_i = 1\}, \quad \mathcal{A}_- := \{\alpha_i \in \mathcal{A} \mid \varepsilon_i = -1\}.$$

For a face $F \subseteq \mathrm{Conv}(\mathcal{A})$, we define the restricted signed support $(\mathcal{A}_F, \varepsilon_F)$ as $\mathcal{A}_F := \mathcal{A} \cap F$ and ε_F containing the signs corresponding the elements in $\mathcal{A}_F$.

One of the roots of tropical geometry goes back to the 1980's when Viro showed it is possible to construct a polyhedral complex associated to a fixed signed support such that for some choice of the coefficients the positive hypersurface and the polyhedral complex have the same isotopy type. To be more precise, the polyhedral complex is



(a) (b)

Figure 4.3: (a) Image of the positive real zero set of $f = 10x_1^5+10x_2^5-33x_1^2x_2^2+10x_1x_2-1$ under the logarithm map $\mathrm{Log}_e\colon \mathbb{R}^2_{>0} \to \mathbb{R}^2$ (b) Viro diagram of the signed support of f with $h = (0,0,1,1,0)$

constructed as follows. Let $(\mathcal{A}, \varepsilon)$ be a signed support with $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$, and let $h \in \mathbb{R}^{n+k+1}$. The *Viro diagram* is defined as

$$\mathrm{Viro}_\varepsilon(\mathcal{A}, h) := \Big\{ v \in \mathbb{R}^n \mid \max_{i\in[n+k+1]} (v\cdot\alpha_i + h_i) \text{ is attained for some } \alpha \in \mathcal{A}_+ \text{ and } \alpha' \in \mathcal{A}_- \Big\}.$$

For $t \in \mathbb{R}_{>0}$, we consider the polynomial

$$f_{\varepsilon * t^h}\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto \sum_{i=1}^{n+k+1} \varepsilon_i t^{h_i} x^{\alpha_i}.$$

Theorem 4.11. *(Viro's patchworking) [105][55, Ch.11 Theorem 5.6][64, Theorem 2.19] Let $(\mathcal{A}, \varepsilon)$ be a signed support such that $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\} \subseteq \mathbb{Z}^n$, and let $h \in \mathbb{R}^{\mathcal{A}}$ be generic.*

Then $\mathrm{Viro}_\varepsilon(\mathcal{A}, h)$ is isotopic to $V_{>0}(f_{\varepsilon * t^h})$ for $t \gg 1$ sufficiently large.

The Viro diagram of the signed support of the running example f from (4.9) with $h = (0,0,1,1,0)$ has the same isotopy as $V_{>0}(f)$, see Figure 4.3. Conversely, in Example 4.9, the logarithmic image of a semi-algebraic set failed to detect bounded connected components. This observation might suggest that the "correct" notion of real tropicalization should involve Viro diagrams rather than the logarithmic limit construction. However, two main obstructions hinder this approach. First, it is unknown how to associate a Viro diagram to a fixed $c \in \mathbb{R}^n$, such that $V_{>0}(f_c)$ and $\mathrm{Viro}_\varepsilon(\mathcal{A}, h)$ are isotopic. Second, there exist examples where the isotopy type cannot be obtained be any Viro diagram, see Example VI.2.11.

In Paper VI, we addressed the question of the conditions on $(\mathcal{A}, \varepsilon)$ that ensure that all possible isotopy types are covered by Viro diagrams. To answer this question, we investigated the *signed reduced A-discriminant*. All the results in Paper VI are phrased in terms of exponential sums, but using the coordinate-wise exponential and logarithm maps $\operatorname{Exp}\colon \mathbb{R}^n \to \mathbb{R}^n_{>0}$, $\operatorname{Log}_e\colon \mathbb{R}^n_{>0} \to \mathbb{R}^n$, one can turn signomials to exponential sums and vice versa without changing the topology of the corresponding zero sets.

For a fixed signed support $(\mathcal{A}, \varepsilon)$, with $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\}$, consider the matrix

$$\hat{A} = \begin{pmatrix} 1 & \dots & 1 \\ \alpha_1 & \dots & \alpha_{n+k+1} \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+k+1)}. \tag{4.13}$$

If $\dim \operatorname{Conv}(\mathcal{A}) = n$, then $\hat{A}$ has full rank and its kernel has dimension k. Let $B \in \mathbb{R}^{(n+k+1)\times k}$ be a Gale dual of $\hat{A}$, see Definition 2.11.

The *signed A-discriminant* contains all the coefficients such that the hypersurface has a singularity in the positive orthant

$$\nabla_{\mathcal{A},\varepsilon} := \Big\{ c \in \mathbb{R}^{n+k+1} \mid \operatorname{sign}(c) = \varepsilon \text{ and } \exists x \in \mathbb{R}^n_{>0} \colon f_c(x) = \frac{\partial f_c(x)}{\partial x_1} = \dots = \frac{\partial f_c(x)}{\partial x_n} = 0 \Big\}.$$

For each face $F \subseteq \operatorname{Conv}(\mathcal{A})$, we define $\nabla_{\mathcal{A}_F,\varepsilon_F}$ analogously and define

$$\widetilde{\nabla}_{\mathcal{A}_F,\varepsilon_F} := \Big\{ (c_{\alpha_i})_{i=1,\dots,n+k+1} \in \mathbb{R}^{n+k+1} \mid \operatorname{sign}(c) = \varepsilon \text{ and } (c_{\alpha_i})_{\alpha_i \in \mathcal{A}_F} \in \nabla_{\mathcal{A}_F,\varepsilon_F} \Big\}.$$

This set contains all the coefficients $c \in \mathbb{R}^{n+k+1}$ with $\operatorname{sign}(c) = \varepsilon$ such that the hypersurface $V_{>0}(f_{c|F})$ has a singularity.

It is known [95, Theorem 3.8], see also [14, Proposition 2.9], that the isotopy type of the hypersurfaces is constant in the connected components of

$$\mathbb{R}^k \setminus \bigcup_{F \subseteq \operatorname{Conv}(\mathcal{A}) \text{ a face}} B^\top \operatorname{Log}(|\widetilde{\nabla}_{\mathcal{A}_F,\varepsilon_F}|).$$

The *signed reduced A-discriminant* is defined as

$$\Gamma_\varepsilon(A, B) := B^\top \operatorname{Log}(|\nabla_{\mathcal{A},\varepsilon}|).$$

Under similar conditions on the signed support as in Section 3.3, the complement of the signed reduced A-discriminant has at most two connected components. Recall that an affine hyperplane $\mathcal{H}_{v,a} \subseteq \mathbb{R}^n$ separates $(\mathcal{A}, \varepsilon)$ if $\mathcal{A}_+$ and $\mathcal{A}_-$ are contained in different half-spaces given by $\mathcal{H}_{v,a}$. We call a separating hyperplane *non-trivial* if $\mathcal{H}_{v,a} \cap \mathcal{A} \neq \mathcal{A}$.

Theorem 4.12. *(Theorem VI.3.3) Let $(\mathcal{A}, \varepsilon)$ be a signed support. Then $\Gamma_\varepsilon(A, B) = \emptyset$ if and only if $(\mathcal{A}, \varepsilon)$ has a non-trivial separating hyperplane.*

If for all faces $F \subseteq \mathrm{Conv}(\mathcal{A})$ the restricted signed support $(\mathcal{A}_F, \varepsilon_F)$ has a non-trivial separating hyperplane, then all hypersurfaces $V_{>0}(f_c)$ with $c \in \mathbb{R}^{n+k+1}$, $\mathrm{sign}(c) = \varepsilon$ have the same isotopy type (Theorem VI.3.5). In that case, any Viro diagram $\mathrm{Viro}_\varepsilon(\mathcal{A}, h)$ for generic $h \in \mathbb{R}^{n+k+1}$ gives that isotopy type.

In the special case when $\mathcal{A}$ contains exactly $n+3$ exponent vectors we characterized conditions on $(\mathcal{A}, \varepsilon)$ implying that $\mathbb{R}^k \backslash \Gamma_\varepsilon(A, B)$ has at most two connected components. Recall that for an n-simplex $P \subseteq \mathbb{R}^n$, we denote by P^- the union of its negative vertex cones (cf. Section 3.3).

Theorem 4.13. *(Theorem VI.4.11, Theorem VI.4.12) Let $(\mathcal{A}, \varepsilon)$ be a signed support such that $\dim \mathrm{Conv}(\mathcal{A}) = n$ and $|\mathcal{A}| = n + 3$. The complement of the signed reduced A-discriminant has at most two connected components if*

(i) $|\mathcal{A}_-| = 1$, *or*

(ii) $\mathcal{A}_+ \subseteq P$, $\mathcal{A}_- \subseteq P^-$, *and* $\mathcal{A} \cap \mathrm{int}\left(P \cup P^-\right) \neq \emptyset$.

Under the conditions in Theorem 4.13, and assuming that $(\mathcal{A}_F, \varepsilon_F)$ has a non-trivial separating hyperplane for all proper faces $F \subsetneq \mathrm{Conv}(\mathcal{A})$, the isotopy type of $V_{>0}(f_c)$ is given by a Viro diagram (Corollary VI.4.13). One should note that in Paper VI we did not address how to compute such a Viro diagram for a given $V_{>0}(f_c)$. Addressing this question and characterizing other properties of $(\mathcal{A}, \varepsilon)$ that imply $\mathbb{R}^k \setminus \Gamma_\varepsilon(A, B)$ has two connected components are interesting questions that require further research.

Bibliography

[1] D. Alessandrini. Logarithmic limit sets of real semi-algebraic sets. *Adv. Geom.*, 13(1):155–190, 2013.

[2] X. Allamigeon, S. Gaubert, and M. Skomra. Tropical spectrahedra. *Discrete. Comput. Geom.*, 63(3):507–548, 2020.

[3] D. Angeli, P. De Leenheer, and E. D. Sontag. A petri net approach to the study of persistence in chemical reaction networks. *Math. Biosci.*, 210(2):598–618, 2008.

[4] M. Avendaño. The number of roots of a lacunary bivariate polynomial on a line. *J. Symb. Comput.*, 44(9):1280–1284, 2009. Effective Methods in Algebraic Geometry.

[5] M. Banaji and G. Craciun. Graph-theoretic criteria for injectivity and unique equilibria in general chemical reaction systems. *Adv. Appl. Math.*, 44(2):168–184, 2010.

[6] M. Banaji, P. Donnell, and S. Baigent. P matrix properties, injectivity, and stability in chemical reaction systems. *SIAM J. Appl. Math.*, 67(6):1523–1547, 2007.

[7] S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry*. Algorithms and Computation in Mathematics. Springer Berlin Heidelberg, 2007.

[8] A. Ben-Israel. Notes on linear inequalities, I: The intersection of the nonnegative orthant with complementary orthogonal subspaces. *J. Math. Anal. Appl.*, 9(2):303–314, 1964.

[9] F. Bihan. Polynomial systems supported on circuits and dessins d'enfants. *J. Lond. Math. Soc. (2)*, 75(1):116–132, 2007.

[10] F. Bihan and A. Dickenstein. Descartes' rule of signs for polynomial systems supported on circuits. *Int. Math. Res. Notices.*, 39(22):6867–6893, 2017.

[11] F. Bihan, A. Dickenstein, and J. Forsgård. Optimal Descartes' rule of signs for systems supported on circuits. *Math. Ann.*, 381:1283–1307, 2021.

[12] F. Bihan, A. Dickenstein, and M. Giaroli. Lower bounds for positive roots and regions of multistationarity in chemical reaction networks. *J. Algebra*, 542:367–411, 2019.

[13] F. Bihan and B. El Hilany. A sharp bound on the number of real intersection points of a sparse plane curve with a line. *J. Symbolic Comput.*, 81:88–96, 2017.

[14] F. Bihan, T. Humbert, and S. Tavenas. New bounds for the number of connected components of fewnomial hypersurfaces. *arXiv*, 2208.04590, 2022.

[15] F. Bihan, J. M. Rojas, and F. Sottile. On the sharpness of fewnomial bounds and the number of components of fewnomial hypersurfaces. In A. Dickenstein, F.-O. Schreyer, and A. J. Sommese, editors, *Algorithms in Algebraic Geometry*, pages 15–20. Springer New York, 2008.

[16] F. Bihan, F. Santos, and P.-J. Spaenlehauer. A polyhedral method for sparse systems with many positive solutions. *SIAM J. Appl. Algebra Geom.*, 2(4):620–645, 2018.

[17] F. Bihan and F. Sottile. New fewnomial upper bounds from Gale dual polynomial systems. *Mosc. Math. J.*, 7(3):387–407, 2007.

[18] F. Bihan and F. Sottile. Betti number bounds for fewnomial hypersurfaces via stratified Morse theory. *Proc. Am. Math. Soc.*, 137(9):2825–2833, 2009.

[19] F. Bihan and F. Sottile. Fewnomial bounds for completely mixed polynomial systems. *Adv. Geom.*, 11(3):541–556, 2011.

[20] G. Blekherman, F. Rincón, R. Sinn, C. Vinzant, and J. Yu. Moments, sums of squares, and tropicalization. *arXiv*, 2203.06291, 2022.

[21] T. Bogart, A. N. Jensen, D. Speyer, B. Sturmfels, and R. R. Thomas. Computing tropical varieties. *J. Symb. Comput.*, 42(1):54–73, 2007. Effective Methods in Algebraic Geometry.

[22] D. V. Bogdanov, A. A. Kytmanov, and T. M. Sadykov. Algorithmic computation of polynomial amoebas. In V. P. Gerdt, W. Koepf, W. M. Seiler, and E. V. Vorozhtsov, editors, *Computer Algebra in Scientific Computing*, pages 87–100. Springer International Publishing, 2016.

[23] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[24] R. Bradford, J. H. Davenport, M. England, H. Errami, V. Gerdt, D. Grigoriev, C. Hoyt, K. Košta, O. Radulescu, T. Sturm, and A. Weber. Identifying the parametric occurrence of multiple steady states for some biological networks. *J. Symb. Comput.*, 98:84–119, 2020. Special Issue on Symbolic and Algebraic Computation: ISSAC 2017.

[25] M. Brandenburg, G. Loho, and R. Sinn. Tropical positivity and determinantal varieties. *Algebr. Comb.*, 6(4):999–1040, 2023.

[26] B. F. Caviness and J. R. Johnson. *Quantifier Elimination and Cylindrical Algebraic Decomposition*. Springer Vienna, 1998.

[27] C. Chen, J. H. Davenport, F. Lemaire, M. M. Maza, N. Phisanbut, B. Xia, R. Xiao, and Y. Xie. Solving semi-algebraic systems with the RegularChains library in Maple. In S. Raschau, editor, *Proceedings of the Fourth Internationa Conference on Mathematical Aspects of Computer Science and Information Sciences (MACIS 2011)*, pages 38–51, 2011.

[28] B. L. Clarke. *Stability of Complex Reaction Networks*, pages 1–215. John Wiley and Sons, Ltd, 1980.

[29] G. E. Collins. Quantifier elimination by cylindrical algebraic decomposition — twenty years of progress. In B. F. Caviness and J. R. Johnson, editors, *Quantifier Elimination and Cylindrical Algebraic Decomposition*, pages 8–23. Springer Vienna, 1998.

[30] C. Conradi, E. Feliu, and M. Mincheva. On the existence of Hopf bifurcations in the sequential and distributive double phosphorylation cycle. *Math. Biosci. Eng.*, 17(1):494–513, 2020.

[31] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *PLoS Comput. Biol.*, 13(10):e1005751, 2017.

[32] A. Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Elsevier Science, 2014.

[33] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks: I. the injectivity property. *SIAM J. Appl. Math.*, 65(5):1526–1546, 2005.

[34] G. Craciun, L. D. García-Puente, and F. Sottile. Some geometrical aspects of control points for toric patches. In M. Dæhlen, M. Floater, T. Lyche, J.-L. Merrien, K. Mørken, and L. L. Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, pages 111–135. Springer Berlin Heidelberg, 2010.

[35] G. Craciun and Feinberg M. Multiple equilibria in complex chemical reaction networks: extensions to entrapped species models. *IEE P. Syst. Biol.*, 153:179–186(7), 2006.

[36] D. R. Curtiss. Recent extentions of Descartes' rule of signs. *Ann. Math.*, 19(4):251–278, 1918.

[37] W. Decker, C. Eder, C. Fieker, M. Horn, and M. Joswig, editors. *The OSCAR book*. 2024.

[38] A. Dickenstein. Biochemical reaction networks: an invitation for algebraic geometers. In J. A. de la Peña, J. A. López-Mimbela, M. Nakamura, and J. Petean, editors, *MCA 2013*, volume 656, pages 65–83. Contemporary Mathematics, 2016.

[39] R. J. Duffin and E. L. Peterson. Geometric programming with signomials. *J. Optimiz. Theory. App.*, 11(1):3–35, 1973.

[40] M. Einsiedler, M. Kapranov, and D. Lind. Non-archimedean amoebas and tropical varieties. *J. für die Reine und Angew. Math.*, 2006(601):139–157, 2006.

[41] M. Feinberg. Complex balancing in general kinetic systems. *Arch. Ration. Mech. Anal.*, 49(3):187–194, 1972.

[42] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.*, 42(10):2229–68, 1987.

[43] M. Feinberg. *Foundations of Chemical Reaction Network Theory*, volume 202. Springer, 2019.

[44] E. Feliu. On the role of algebra in models in molecular biology. *J. Math. Biol.*, 80:1159–1161, 2020.

[45] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürück. The kinetic space of multistationarity in dual phosphorylation. *J. Dyn. Differ. Equ.*, 34:825–852, 2022.

[46] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. Parameter region for multistationarity in $\boldsymbol{n}-$site phosphorylation networks. *SIAM J. Appl. Dyn. Syst.*, 22(3):2024–2053, 2023.

[47] E. Feliu, A. D. Rendall, and C. Wiuf. A proof of unlimited multistability for phosphorylation cycles. *Nonlinearity*, 33(11):5629–5658, 2020.

[48] E. Feliu and A. Sadeghimanesh. Kac-Rice formulas and the number of solutions of parametrized systems of polynomial equations. *Math. Comput.*, 91:2739–2769, 2022.

[49] E. Feliu and C. Wiuf. Preclusion of switch behavior in networks with mass-action kinetics. *Appl. Math. Comput.*, 219(4):1449–1467, 2012.

[50] E. Feliu and C. Wiuf. A computational method to preclude multistationarity in networks of interacting species. *Bioinformatics*, 29(18):2327–2334, 2013.

[51] E. Feliu and C. Wiuf. Simplifying biochemical models with intermediate species. *J. R. Soc. Interface*, 10(87):20130484, 2013.

[52] D. Flockerzi, K. Holstein, and C. Conradi. N-site Phosphorylation Systems with 2N-1 Steady States. *Bull. Math. Biol.*, 76(8):1892–1916, 2014.

[53] J. Forsgård, M. Nisse, and J. M Rojas. New subexponential fewnomial hypersurface bounds. *arXiv*, 1710.00481, 2017.

[54] C. F. Gauß. Beweis eines algebraischen Lehrsatzes. *J. Reine. Angew. Math.*, 3:1–4, 1828.

[55] I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants*. Mathematics (Boston, Mass.). Birkhäuser, 1994.

[56] M. Giaroli, F. Bihan, and A. Dickenstein. Regions of multistationarity in cascades of Goldbeter-Koshland loops. *J. Math. Biol.*, 78(4):1115–1145, 2019.

[57] M. Giaroli, R. Rischter, M. Pérez Millán, and A. Dickenstein. Parameter regions that give rise to 2[n/2] +1 positive steady states in the n-site phosphorylation system. *Math. Biosci. Eng.*, 16(6):7589–7615, 2019.

[58] D. J. Grabiner. Descartes' rule of signs: Another construction. *Am. Math. Mon.*, 106(9):854–856, 1999.

[59] J. Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophys. J.*, 93:3828–3834, 2007.

[60] H. A. Harrington, D. Mehta, H. M. Byrne, and J. D. Hauenstein. Decomposing the parameter space of biological networks via a numerical discriminant approach. In J. Gerhard and I. Kotsireas, editors, *Maple in Mathematics Education and Research*, pages 114–131. Springer International Publishing, 2020.

[61] F. Horn and R. Jackson. General mass action kinetics. *Arch. Rational Mech. Anal.*, 47:81–116, 1972.

[62] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1990.

[63] S. Iliman and T. de Wolff. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Res. Math. Sci.*, 3(9), 2016.

[64] I. Itenberg, G. Mikhalkin, and E. I. Shustin. *Tropical Algebraic Geometry*. Springer Science, 2 edition, 2009.

[65] I. Itenberg and M. F. Roy. Multivariate Descartes' rule. *Beitr. Algebra. Geom.*, 37(2):337–346, 1996.

[66] P. Jell, C. Scheiderer, and J. Yu. Real Tropicalization and Analytification of Semialgebraic Sets. *Int. Math. Res. Notices*, 2022(2):928–958, 2020.

[67] B. Joshi and A. Shiu. Simplifying the Jacobian Criterion for precluding multistationarity in chemical reaction networks. *SIAM J. Appl. Math.*, 72(3), 2011.

[68] B. Joshi and A. Shiu. Atoms of multistationarity in chemical reaction networks. *J. Math. Chem.*, 51(1):153–178, 2013.

[69] B. Joshi and A. Shiu. A survey of methods for deciding whether a reaction network is multistationary. *Math. Model. Nat. Phenom.*, 10(5):47–67, 2015.

[70] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Springer, 2013.

[71] A. G. Khovanskii. *Fewnomials*, volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991.

[72] P. Koiran, N. Portier, and S. Tavenas. A wronskian approach to the real τ-conjecture. *J. Symb. Comput.*, 68:195–214, 2015. Effective Methods in Algebraic Geometry.

[73] S. Kumaresan. *Topology of Metric Spaces*. Alpha Science International, 2005.

[74] J. C. Lagarias and T. J. Richardson. Multivariate Descartes rule of signs and Sturmfels's challenge problem. *Math. Intell.*, 19:9–15, 1997.

[75] M. Laurent and N. Kellershohn. Multistability: a major means of differentiation and evolution in biological systems. *Trends. Biochem. Sci.*, 24(11):418–22, 1999.

[76] T. Y. Li, J. M. Rojas, and X. Wang. Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces. *Discrete Comput. Geom.*, 30(3):379–414, 2003.

[77] T. Y. Li and X. Wang. On multivariate Descartes' rule - a counterexample. *Beitr. Algebra. Geom.*, 39(1):1–5, 1998.

[78] D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2015.

[79] Maplesoft, a division of Waterloo Maple Inc. `Maple`. https://www.maplesoft.com, 2023.

[80] C. D. Maranas and C. A. Floudas. All solutions of nonlinear constrained systems of equations. *J. Global. Optim.*, 7:143–182, 1995.

[81] T. Markwig and Y. Ren. Computing tropical varieties over fields with valuation. *Foun. Comut. Math.*, 20:783–800, 2020.

[82] G. Mikhalkin. Decomposition into pairs-of-pants for complex algebraic hypersurfaces. *Topology*, 43(5):1035–1065, 2004.

[83] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16:69–97, 2016.

[84] S. Müller and G. Regensburger. Generalized mass action systems: Complex balancing equilibria and sign vectors of the stoichiometric and kinetic-order subspaces. *SIAM J. Appl. Math.*, 72(6):1926–1947, 2012.

[85] K. M. Nam, B. M. Gyori, S. V. Amethyst, D. J. Bates, and J. Gunawardena. Robustness and parameter geography in post-translational modification systems. *Plos. Comput. Biol.*, 16(5):1–50, 2020.

[86] OSCAR – Open Source Computer Algebra Research system, version 0.12.2-dev, 2023.

[87] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of escherichia coli. *Nature.*, 427:737–740, 2004.

[88] C. Pantea, H. Koeppl, and G. Craciun. Global injectivity and multiple equilibria in uni- and bi-molecular reaction networks. *Discrete Continuous Dyn. Syst. Ser. B.*, 17(6):2153–2170, 2012.

[89] D. Perrucci. Some bounds for the number of components of real zero sets of sparse polynomials. *Discrete Comput Geom.*, 34:475–495, 2003.

[90] V. V. Prasolov and S. Ivanov. *Problems and Theorems in Linear Algebra*. History of Mathematics. American Mathematical Society, 1994.

[91] M. Pérez Millán, A. Dickenstein, A. Shiu, and C. Conradi. Chemical reaction systems with toric steady states. *Bull. Math. Biol.*, 74:1027–1065, 2012.

[92] Q. I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials*. London Mathematical Society monographs. Clarendon Press, 2002.

[93] P. Érdi and J. Tóth. *Mathematical Models of Chemical Reactions Theory and Applications of Deterministic and Stochastic Models*. Manchester University Press, 1989.

[94] R. T. Rockafellar. *Convex analysis*. Princeton University Press, 1972.

[95] J. M. Rojas and K. Rusek. A-discriminants for complex exponents and counting real isotopy types. *arXiv*, 1612.03458, 2017.

[96] A. Sadeghimanesh and E. Feliu. The multistationarity structure of networks with intermediates and a binomial core network. *Bull. Math. Biol.*, 81:2428–2462, 2019.

[97] I. Sahidul and A. M. Wasim. *Fuzzy Geometric Programming Techniques and Applications*. Springer, 2019.

[98] L. Songxin, J. Gerhard, D. J. Jeffrey, and G. Moroz. A package for solving parametric polynomial systems. *ACM Commun. Comput. Algebra*, 43(3/4):61–72, 2009.

[99] E. D. Sontag. Structure and stability of certain chemical networks and applications to the kinetic proofreading model of t-cell receptor signal transduction. *IEEE Trans. Automat. Contr.*, 46(7):1028–1047, 2001.

[100] R. Straube and C. Conradi. Reciprocal enzyme regulation as a source of bistability in covalent modification cycles. *J. Theor. Biol.*, 330:56–74, 2013.

[101] B. Sturmfels. Viro's theorem for complete intersections. *Ann. Scuola. Norm-Sci.*, 21:377–386, 1994.

[102] L. F. Tabera. On real tropical bases and real tropical discriminants. *Collect. Math.*, 66:77–92, 2015.

[103] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.2)*, 2021. `https://www.sagemath.org`.

[104] C. Vinzant. Real radical initial ideals. *J. Algebra*, 352(1):392–407, 2009.

[105] O. Y. Viro. *Constructing real algebraic varieties with prescribed topology*. Dissertation, LOMI, Leningrad, an english translation by the author is available at https://arxiv.org/abs/math/0611382 edition, 1983.

[106] A. I. Vol'pert. Differential equations on graphs. *Math. USSR Sb.*, 17(4):571–582, 1972.

[107] P. Waage and C. M. Guldberg. Studier over affiniteten. *Forhandlinger I Videnskabs-selskabet I Christiania*, 35:35–45, 1864.

[108] P. Waage and C. M. Guldberg. Studies concerning affinity. *J. Chem. Educ.*, 63(12):1044, 1986.

[109] L. Wang and E. D. Sontag. On the number of steady states in a multiple futile cycle. *J. Math. Biol.*, 57:29–52, 2008.

[110] X. Wang. A simple proof of Descartes's rule of signs. *Am. Math. Mon.*, 111:525–526, 2004.

[111] C. Wiuf and E. Feliu. Power-law kinetics and determinant criteria for the preclusion of multistationarity in networks of interacting species. *SIAM J. Appl. Dyn. Syst.*, 12(4):1685–1721, 2013.

[112] G. M. Ziegler. *Lectures on Polytopes*. Springer, 2007.

Papers

I

Topological descriptors of the parameter region of multistationarity: deciding upon connectivity

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Elisenda Feliu
Department of Mathematical Sciences
University of Copenhagen

Publication details

Published in PLOS Computational Biology 19(3):1–38, 2023
DOI: https://doi.org/10.1371/journal.pcbi.1010970

TOPOLOGICAL DESCRIPTORS OF THE PARAMETER REGION OF MULTISTATIONARITY: DECIDING UPON CONNECTIVITY

MÁTÉ L. TELEK AND ELISENDA FELIU

Abstract. Switch-like responses arising from bistability have been linked to cell signaling processes and memory. Revealing the shape and properties of the set of parameters that lead to bistability is necessary to understand the underlying biological mechanisms, but is a complex mathematical problem. We present an efficient approach to address a basic topological property of the parameter region of multistationary, namely whether it is connected. The connectivity of this region can be interpreted in terms of the biological mechanisms underlying bistability and the switch-like patterns that the system can create.

We provide an algorithm to assert that the parameter region of multistationarity is connected, targeting reaction networks with mass-action kinetics. We show that this is the case for numerous relevant cell signaling motifs, previously described to exhibit bistability. The method relies on linear programming and bypasses the expensive computational cost of direct and generic approaches to study parametric polynomial systems. This characteristic makes it suitable for mass-screening of reaction networks.

Although the algorithm can only be used to certify connectivity, we illustrate that the ideas behind the algorithm can be adapted on a case-by-case basis to also decide that the region is not connected. In particular, we show that for a motif displaying a phosphorylation cycle with allosteric enzyme regulation, the region of multistationarity has two distinct connected components, corresponding to two different, but symmetric, biological mechanisms.

Keywords: reaction network, mass-action kinetics, steady states, phosphorylation cycle

1. Introduction

Bistable switches are frequently observed and studied in living systems, and have been linked to cellular decision making and memory processes [1, 2]. These switches arise in different forms; one common form in parametric systems is that of *hysteresis* [3], that is, the system is monostable for small or large values of a parameter, and has two or more stable steady states for intermediate values. When the parameter changes slowly enough to allow the system to remain approximately at steady state, the resulting steady states depend on whether the parameter is increased or decreased. This is illustrated in Fig. 1(a), which displays a hypothetical system with three steady states. When the parameter increases from low to high, the steady state goes through a bifurcation at a critical parameter value $\tau_{\max}$, after which the system discontinuously settles to another region of the output space. If the parameter value is decreased again, the system remains at the high steady state value, that is, it does not return to the low steady state value immediately. This will first happen if the parameter is decreased beyond another critical value $\tau_{\min} < \tau_{\max}$. This behavior confers the switch with *robustness*: after a change of level of steady state value takes place, small fluctuations in the parameter will not reverse the change. The larger the interval $[\tau_{\min}, \tau_{\max}]$ where the system has several steady states, the higher the degree of robustness of the change. More complicate switches can arise if, for example, the system has more than one steady state (is *multistationary*) in two intervals of the parameter, as illustrated in Fig. 1(b,c). Irreversible switches are obtained if the relevant interval is of the form $(0, \tau_{\max}]$. Fig. 1(a) and Fig. 1(b,c) are qualitatively different in one important aspect: the parameter region where the system has multistationarity is connected in Fig. 1(a) and disconnected (has two disjoint pieces) in Fig. 1(b,c).

This phenomenon appears also in higher dimensions, where instead of a curve of steady states, there is a steady state manifold with "bends", and instead of having one parameter being varied, a vector of parameters is changed along a curve. The shape of the *parameter region of*

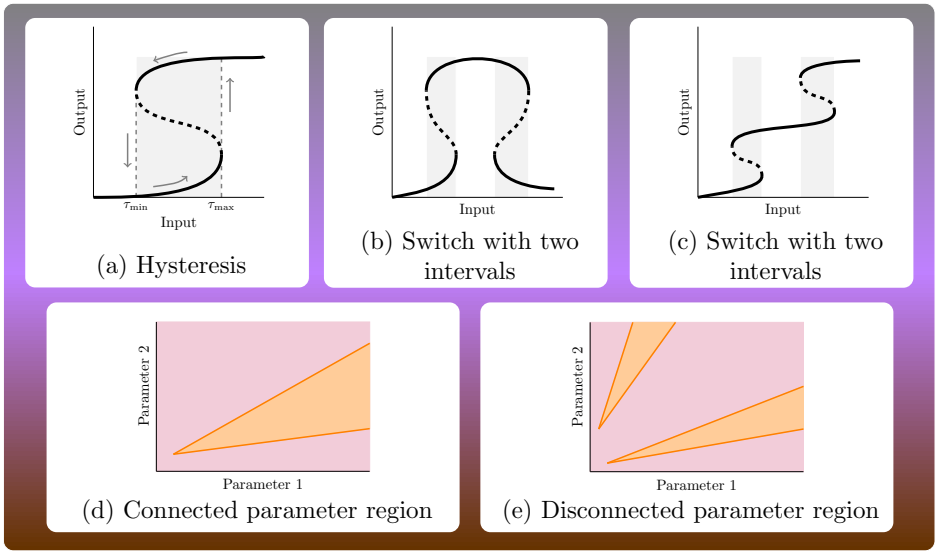


FIGURE 1. (a-c) Input-output curves for hypothetical systems. Input is thought to be a parameter of the system that is varied, and the output is the concentration of a species at steady state. Dashed lines correspond to unstable steady states, and solid lines to stable steady states. (a) displays a simple hysteresis switch; (b-c) show input-output curves for systems where bistability arises in two disjoint intervals of input. (d-e) For a system with two parameters, the system has more than one positive steady state in the orange regions, and one in the purple regions. In panel (d) the multistationarity region is connected, while in (e) it has two connected components.

multistationarity (or *multistationarity region* for short), and specifically the number of path-connected components, modulates the type of switches that can arise. If the multistationarity region is path connected, as in Fig. 1(d), then any two parameter values in the region can be joined by a continuous path completely included in the region, typically giving rise to simple hysteresis switches (if the system has three steady states for parameters in the region). If, on the contrary, the region has two path-connected components, as in Fig. 1(e), then any path joining two parameter points in different regions, necessarily goes through parameter points where the system does not have several steady states, allowing for complex switches to arise.

Mathematically, understanding the connectivity of a region is a basic topological property of a set, and the number of connected components is called the 0th Betti number of the set. Higher order Betti numbers describe the shape of the set in more detail, for instance, the first Betti number is the number of "holes" of the set. Tools from topological data analysis can infer the Betti numbers of a set from sample data points. By generating points in the multistationarity region, properties of the shape of the region have been explored for a specific dual phosphorylation system (Fig. 2(h)) in [4], where it has also been suggested that lack of connectivity may indicate that different biological mechanisms underlie multistationarity.

In this work, we address connectivity of the multistationarity region for polynomial systems describing the steady states of biochemical reaction networks. We achieve this by using exact symbolic tools and theoretical results relating the multistationarity region with the region where a polynomial attains negative values. Specifically, we work in the framework of chemical reaction network theory [5, 6], where extensive work has been done to decide whether a network exhibits multistationarity, i.e. whether the multistationarity region is non-empty [7]. More recent work focuses on understanding and finding the multistationarity region, but here progress is scarce and often restricted to special systems, e.g. [8–14].

Finding and studying the region where a polynomial system has more than one positive solution is a mathematical problem that belongs to the realm of semi-algebraic geometry and

quantifier elimination [15, 16]. Although there are generic methods to address this question, these have high complexity, and fail for realistic networks, even of moderate size. To overcome these difficulties, methods targeting the specificities of reaction networks systems have been developed, and some partial results, for example describing the projection of the multistationarity region onto a subset of the parameters, have been developed [9–11].

Here we present a new algorithm to assert that the multistationarity region is connected without explicitly finding it, which bypasses the use of computationally expensive algorithms from semi-algebraic geometry and quantifier elimination. In fact, we reduce the problem to computing the determinant of a symbolic matrix, and finding a point in the feasible region of a system of linear inequalities. This makes the algorithm successful for networks of moderate size.

We apply our algorithm to numerous motifs in cell signaling, known to exhibit multistationarity, and conclude that these have connected multistationarity regions. These systems are shown in Fig. 2, where for all subfigures but (c) and (h), we confirm that the region is connected. For the system in Fig. 2(h), previously suggested to have a connected multistationarity region [4], our method is inconclusive. For system (c), modeling the enzymatic phosphorylation of a substrate S with allosteric regulation of the enzymes [17], the algorithm is also inconclusive. In this case, a detailed inspection of the system employing ideas similar to those behind our algorithm, allows us to assert that the region has exactly two path-connected components. These components are contained in the subset of parameters where $\kappa_3 > \kappa_6$ or $\kappa_6 > \kappa_3$ respectively. These two parameters are the catalytic constants of the phosphorylation and dephosphorylation processes, respectively. Therefore, if for example κ_6 increases from a small value to a value larger than κ_3, then the multistationarity region will be crossed twice, and a non-simple hysteresis switch arises.

The fact that the remaining networks in Figure 2 have a connected multistationarity region, does not forbid complex switches, as a parameter path could still enter and exit the multistationarity region several times. However, in this scenario we can assert that for any pair of parameter values of multistationarity, there is a path connecting them and completely included in the multistationarity region. This implies that the conditions yielding to multistationarity vary continuously, and we cannot separate the multistationarity region into two disjoint sets, each corresponding to a distinct biological mechanism.

Our algorithm builds on two previous results. First, in [9], the authors associated with each reaction network a function, whose signs are closely related to the multistationarity region. Second, in [18], we developed a new criterion to determine the connectivity of a set described as the preimage of the negative real half-line by a polynomial map. We connect these two key ingredients in Theorem 2.2 to give a criterion for connectivity of the multistationarity region. Our results establish a stronger property, namely path-connectivity of the region, which in turn imply connectivity.

The criterion relies on computing the determinant of a matrix with symbolic entries, and this task might become unfeasible for large matrices with many variables. To bypass this, we show that the network can be reduced by removing some reactions and connectivity of the multistationarity region for the reduced network can be translated into the original network, see Theorem 2.4.

This paper is organized as follows. We first establish the framework and background material, and present our algorithm for connectivity of the multistationarity region. We proceed to apply our algorithm to the networks in Fig. 2, and while doing so, we illustrate the strengths and limitations of the algorithm. To keep the exposition concise, we compile the proofs of the main theorems in Section 6 at the end.

2. Results

2.1. **Theory.** The results of this work are framed in the context of chemical reaction network theory, a formalism to study reaction networks that goes back to the 70s with the works of Horn,

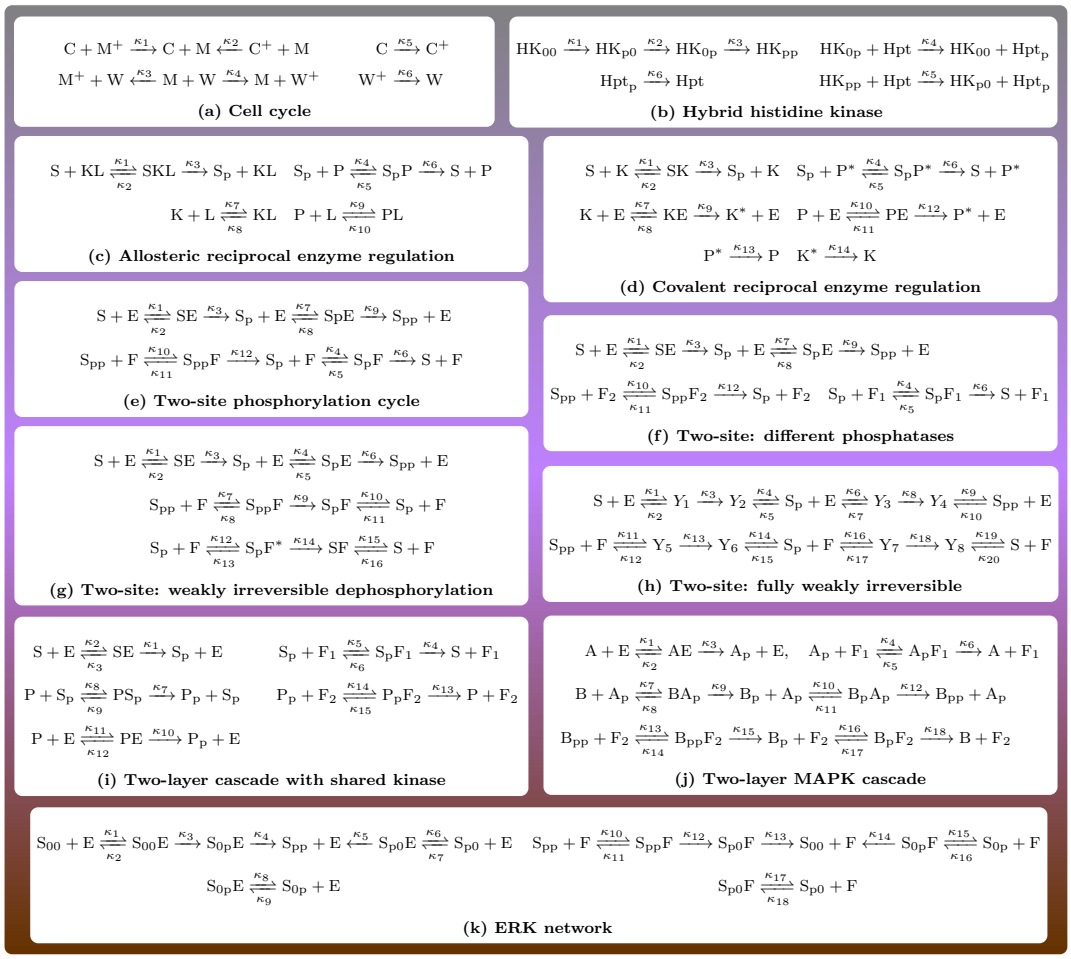


FIGURE 2. Reaction networks arising in cell signaling. The subindex 'p' indicates a phosphorylated site. When writing 'p' and 'pp' it is assumed the substrate has two phosphorylation sites, and phosphorylation/dephosphorylation is ordered. When writing '0p' for example, it means the substrate also has two sites numbered 1 and 2, and the second one is phosphorylated. All networks are known to be multistationary. For all networks but (c) and (h), the multistationarity region is path connected. For network (c), the multistationarity region has two path-connected components, while for network (h) our approach is inconclusive.

Jackson and Feinberg [5, 6]. To help the unfamiliar reader, and to fix the notation, we start with a brief introduction. This part ends with the main theoretical result to decide whether the set of parameters where multistationarity arises is path connected (and hence connected). To keep the exposition simple for non-experts, the proofs of the statements are given in Section 6 at the end.

Reaction networks. A **reaction network** $(\mathcal{S}, \mathcal{R})$ is a collection of *reactions* $\mathcal{R} = \{R_1, \dots, R_r\}$ between *species* in a set $\mathcal{S} = \{X_1, \dots, X_n\}$. In our applications, species will be proteins, such as kinases and substrates. The reactions will encode events such as complex formation or posttranslational modifications.

Formally, each reaction connects to linear combinations of species, that is, it has the form:

$$R_j: \quad a_{1j}X_1 + \cdots + a_{nj}X_n \longrightarrow b_{1j}X_1 + \cdots + b_{nj}X_n, \qquad j = 1, \dots, r,$$

where the coefficients a_{ij}, b_{ij} are non-negative integer numbers. The net production of each species when the reactions takes place is encoded in the **stoichiometric matrix**

$$N = \big[b_{ij} - a_{ij}\big]_{\substack{i=1,\dots,n\\ j=1,\dots,r}} \in \mathbb{R}^{n\times r}.$$

We do not consider reactions where the reactant and product are equal, hence N has no zero columns.

For illustration purposes, we consider a reaction network that is small enough to get a good feeling about the formal concepts but large enough to display the relevant features. To construct such a reaction network, the article [19] was particularly helpful. Realistic networks will be considered in Section 3 of this work. We refer to the following reaction network as the *running example*:

$$\mathrm{X}_1 \to \mathrm{X}_2, \qquad \mathrm{X}_2 \to \mathrm{X}_1, \qquad 2\mathrm{X}_1 + \mathrm{X}_2 \to 3\mathrm{X}_1. \tag{1}$$

Here, two species are related by three reactions, so the corresponding stoichiometric matrix has two rows and three columns, see Fig. 3.

Mathematical modeling offers us tools to get insights into the dynamics of the network and understand the temporal changes in the concentrations of the species of the network. By encoding the concentrations of $X_1, \dots, X_n$ into the vector $x = (x_1, \dots, x_n) \in \mathbb{R}^n_{\geq 0}$, and under the assumption of *mass-action kinetics*, the evolution of the concentrations of the species over time is modeled by the ODE system:

$$\dot{x} = f_\kappa(x), \qquad x \in \mathbb{R}^n_{\geq 0}, \tag{2}$$

where $f_\kappa(x) := N v_\kappa(x)$ with the *rate function* $v_\kappa(x)$ given for $x \in \mathbb{R}^n_{\geq 0}$ by

$$v_\kappa(x) = (\kappa_1\, x_1^{a_{11}} \cdots x_n^{a_{n1}}, \dots, \kappa_n\, x_1^{a_{1n}} \cdots x_n^{a_{nn}})^\top \qquad \in \mathbb{R}^r_{\geq 0}. \tag{3}$$

Here $\kappa = (\kappa_1, \dots, \kappa_r) \in \mathbb{R}^r_{>0}$ is the vector of *reaction rate constants*, which are parameters of the system. Observe that each component of $v_\kappa(x)$ corresponds to one reaction, and it is obtained by considering the coefficients of the reactant of the reaction as exponents. The function $v_\kappa(x)$ and the associated ODE system of the running example are shown in Fig. 3.

The ODE system (2) is forward invariant on **stoichiometric compatibility classes** [20], that is, for any initial condition x_0 the dynamics takes place in the stoichiometric compatibility class of x_0, which is the set $(x_0 + S) \cap \mathbb{R}^n_{\geq 0}$, where S denotes the vector space spanned by the columns of N. To work with stoichiometric compatibility classes it is more convenient to have equations for them. These are obtained by considering a full rank matrix $W \in \mathbb{R}^{(n-s)\times n}$ such that $W N = 0$, where s is the rank of the stoichiometric matrix N. Then for each parameter vector $c \in \mathbb{R}^{n-s}$, the associated stoichiometric compatibility class is the set

$$\mathcal{P}_c := \{x \in \mathbb{R}^n_{\geq 0} \mid Wx = c\}. \tag{4}$$

This class is the set $(x_0 + S) \cap \mathbb{R}^n_{\geq 0}$ for any initial condition x_0 satisfying $Wx_0 = c$. Such a matrix W is called a **matrix of conservation relations**, any equation defining a class is a conservation relation, and the parameter vector c is called a *vector of total concentrations*.

For the running example, we have $n = 2, s = 1$, and hence W has one row, as given in Fig. 3. The figure depicts also the stoichiometric compatibility classes for $c = 2, 3, 4$, which are compact. In general, a reaction network is called **conservative** if each stoichiometric compatibility class is a compact subset of $\mathbb{R}^n_{\geq 0}$, and this is the same as asking that there is a vector with all entries positive in the left-kernel of N [21]. Under this assumption, all trajectories of the ODE system (2) are completely contained in a compact subset, and hence no component can go to infinity. For the purposes of this work, it will be enough to require a milder condition, namely that there is a compact set that all trajectories enter in finite time and do not leave again. In this case, the reaction network is said to be **dissipative**. See [9] for ways to verify that a network is dissipative. For simplicity, our applications are conservative networks.

We denote the set of non-negative steady states of the ODE system (2) by

$$V_\kappa := \{x \in \mathbb{R}^n_{\geq 0} \mid Nv_\kappa(x) = 0\}, \tag{5}$$

and call it the **steady state variety**. The **positive steady state variety** consists then of the points in V_κ with all entries positive and is denoted by $V_{\kappa,>0}$. A steady state $x \in V_\kappa$ is a **boundary steady state**, if one of its coordinates equals zero. For our results, we will need that stoichiometric compatibility classes intersecting $\mathbb{R}^n_{>0}$ do not contain boundary steady states. We call therefore a boundary steady state **relevant** if it belongs to a class $\mathcal{P}_c$ with $\mathcal{P}_c \cap \mathbb{R}^n_{>0} \neq \emptyset$.

We say that a pair of parameters (κ, c) **enables multistationarity** if the intersection of the positive steady state variety $V_{\kappa,>0}$ with the stoichiometric compatibility class $\mathcal{P}_c$ contains more than one point. The set of parameters that enable multistationarity form the **multistationarity region**. We show in the right panel of Fig. 3 some stoichiometric compatibility classes, steady state varieties and their intersection, showing that multistationarity arises.

The running example:

$\mathrm{X}_1 \xrightarrow{\kappa_1} \mathrm{X}_2, \qquad \mathrm{X}_2 \xrightarrow{\kappa_2} \mathrm{X}_1, \qquad 2\mathrm{X}_1 + \mathrm{X}_2 \xrightarrow{\kappa_3} 3\mathrm{X}_1.$

Stoichiometric matrix N, reaction rate function $v_\kappa(x)$ and a full rank matrix W such that $WN = 0$:

$$N = \begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \qquad W = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

$v_\kappa(x) = (\kappa_1 x_1, \kappa_2 x_2, \kappa_3 x_1^2 x_2)$

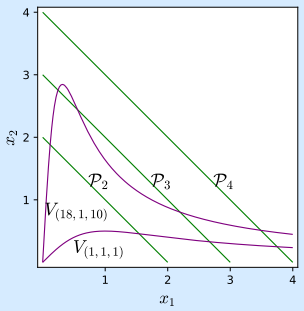
ODE system:

$$\dot{x}_1 = \kappa_3 x_1^2 x_2 - \kappa_1 x_1 + \kappa_2 x_2,$$
$$\dot{x}_2 = -\kappa_3 x_1^2 x_2 + \kappa_1 x_1 - \kappa_2 x_2.$$

Equation of the stoichiometric compatibility class $\mathcal{P}_c$:

$$x_1 + x_2 = c.$$

Steady state varieties V_κ and stoichiometric compatibility classes $\mathcal{P}_c$ for different choices of parameters (κ, c):



$(\kappa, c) = ((18, 1, 10), 3)$ enables multistationarity,
$(\kappa, c) = ((18, 1, 10), 4)$ does not enable multistationarity.
For $\kappa = (1, 1, 1)$ there is no total concentration c such that $((1, 1, 1), c)$ enables multistationarity.

FIGURE 3. Illustration of the relevant objects for the running example given in (1).

We now recall the main theorem from [9] that is key to describe the multistationarity region. The theorem requires choosing a matrix of conservation relations W that is row reduced. Then, if $i_1 < \cdots < i_{n-s}$ are the indices of the first non-zero coordinates of each row of W, we construct the matrix $M_\kappa(x)$ from the Jacobian of $f_\kappa(x)$ by replacing the i_jth row by the jth row of W (for $j = 1, \ldots, n-s$).

Theorem 2.1 (Multistationarity)**.** *[9, Theorem 1] Consider a reaction network that is dissipative and does not have relevant boundary steady states. For each vector of reaction rate constants $\kappa \in \mathbb{R}^r_{>0}$ and vector of total concentrations $c \in \mathbb{R}^{n-s}$, it holds:*

(A) *If $(-1)^s \det(M_\kappa(x)) > 0$ for all $x \in V_{\kappa,>0} \cap \mathcal{P}_c$, then the parameter pair (κ, c) does not enable multistationarity.*
(B) *If $(-1)^s \det(M_\kappa(x)) < 0$ for some $x \in V_{\kappa,>0} \cap \mathcal{P}_c$, then the parameter pair (κ, c) enables multistationarity.*

Our running example is dissipative, as it is conservative, and it has no relevant boundary steady states. We also find that

$$(-1)^s \det(M_\kappa(x)) = \kappa_3 x_1^2 - 2\kappa_3 x_1 x_2 + \kappa_1 + \kappa_2.$$

This expression is negative for $\kappa = (18, 1, 10)$ and $x^* \approx (0.2448, 2.7552) \in V_\kappa$. Since $Wx^* = 3$, we can conclude using Theorem 2.1, that the intersection of $V_{(18,1,10)}$ and $\mathcal{P}_3$ contains more than one point. This is exactly what Fig. 3 indicates.

Parametrizations. In Theorem 2.1, it is crucial to evaluate the determinant of $M_\kappa(x)$ at points in the incidence set:

$$\mathcal{V} := \{(x,\kappa) \in \mathbb{R}^n_{>0} \times \mathbb{R}^r_{>0} \mid x \in V_\kappa\}. \tag{6}$$

Therefore, we need to be able to describe the points in $\mathcal{V}$ in a useful way. This is done by considering *parametrizations* of $\mathcal{V}$. Loosely speaking, a parametrization is a function whose image set consists precisely of the points in $\mathbb{R}^n_{>0} \times \mathbb{R}^r_{>0}$ that belong to $\mathcal{V}$. Formally, we define a **parametrization of** $\mathcal{V}$ as a surjective analytic map

$$\Phi\colon \mathcal{D} \to \mathcal{V}.$$

In practice, $\mathcal{D}$ is the positive orthant of some $\mathbb{R}^k$ and Φ is described by polynomials or quotients of polynomials such that their denominators do not vanish on $\mathcal{D}$. Below, we discuss how to choose a parametrization and show that there is always at least one.

Using Theorem 2.1, one can show (see Lemma 6.4) that the multistationarity region is closely related to the preimage of the negative real half-line under the polynomial map

$$g\colon \mathcal{V} \to \mathbb{R}, \quad (x,\kappa) \mapsto (-1)^s \det(M_\kappa(x)), \tag{7}$$

which can be described using a parametrization

$$g \circ \Phi\colon \ \mathcal{D} \to \mathcal{V} \to \mathbb{R}, \qquad \xi \mapsto g(\Phi(\xi)). \tag{8}$$

Following [22], we call the function $g \circ \Phi$ a **critical function**, and observe that it depends on the choice of the parametrization. We are ready to present the main theoretical result of this work, namely a criterion for connectivity of the multistationarity region. The statement tells us that we can look at the number of path-connected components of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$, that is, of the set of values ξ where $g \circ \Phi$ is negative. The proof of the following theorem can be found in Section 6.

Theorem 2.2 (Deciding connectivity)**.** *Consider a reaction network that is dissipative and does not have relevant boundary steady states. Let $g \circ \Phi$ be a critical function as in* (8) *such that the closure of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$ equals $(g \circ \Phi)^{-1}(\mathbb{R}_{\leq 0})$.*

Then the number of path-connected components of the multistationarity region is at most the number of path-connected components of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$.

In particular, if $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$ is path connected, then the multistationarity region is path connected.

2.2. **Algorithm for checking path connectivity.** Theorem 2.2 gives a theoretical criterion to decide upon connectivity, from which one can establish an algorithm for connectivity with the following steps:

(Step 1) Check that the reaction network is dissipative and does not have relevant boundary steady states.

(Step 2) Find a parametrization Φ of $\mathcal{V}$ and compute the critical function $g \circ \Phi$.

(Step 3) Check that $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$ is path connected and its closure equals $(g \circ \Phi)^{-1}(\mathbb{R}_{\leq 0})$.

The important point is that each of these steps can be addressed computationally, and hence the algorithm can be carried through without manual intervention, at least for networks of moderate size. We proceed to describe each of these steps in detail.

(Step 1) has already been described in detail in [9], as it consists of verifying that the conditions to apply Theorem 2.1 hold. Computable criteria that are sufficient to ensure that the properties hold are presented in [9]. These are, however, not necessary and hence it might not always be possible to decide upon this step.

To verify dissipativity, the first attempt is to show that the reaction network is conservative by finding a row vector $w \in \mathbb{R}^n_{>0}$ such that $w\,N = 0$. This can be checked by solving the system of linear equalities:

$$w\,N_i = 0, \quad \text{for all } i = 1,\dots,r \quad \text{and} \quad w_j > 0 \quad \text{for all } j = 1,\dots,n, \tag{9}$$

where N_i denotes the ith column of N. We already noticed that the running example is conservative, by choosing

$$w = (1,1) \ \in \mathbb{R}^2_{>0}. \tag{10}$$

A sufficient criterion to preclude the existence of relevant boundary steady states arises by using siphons, that is, subsets of species such that for all species in the set and all reactions producing them, there is a species in the reactant also in the set, see [23, Theorem 2], [25, Proposition 2], [26]. In a nutshell, the criterion requires that for each minimal *siphon* it is possible to choose $w \in \mathbb{R}^n_{\geq 0}$ with $wN = 0$ and such that the positive entries of w correspond exactly to the species in the siphon. For more details, we refer to [9, 23, 25]. Note that this criterion also relies on solving linear inequalities. Our running example has only one siphon, namely $\{X_1, X_2\}$. As the two entries of w in (10) are positive, the criterion holds and the network does not have relevant boundary steady states.

(Step 2) asks for the choice of a parametrization of $\mathcal{V}$ and the computation of the critical function. To find a parametrization systematically, we consider so-called *convex parameters* introduced by Clarke in [27]. Since then, they have been applied to study reaction networks, for example to detect Hopf bifurcations and study bistability [28–31].

The idea behind convex parameters is the simple observation that the rate function $v_\kappa(x)$ has to be in the *flux cone*:

$$\mathcal{F} := \{v \in \mathbb{R}^r_{\geq 0} \mid Nv = 0\} = \ker(N) \cap \mathbb{R}^r_{\geq 0}$$

for each $(x, \kappa) \in \mathcal{V}$. It is easy to see that $\mathcal{F}$ is a convex polyhedral cone containing no lines. Using software with packages for polyhedral sets (see Methods), one can compute a minimal collection of generators $E_1, \ldots, E_\ell \in \mathbb{R}^r$ of $\mathcal{F}$.

These generators are often called *extreme vectors* of the cone. Their choice is unique up to multiplication by a positive number. Since the flux cone $\mathcal{F}$ does not contain lines, each of its elements can be written as a non-negative linear combination of *extreme vectors* [32, Corollary 18.5.2], that is for each $v \in \mathcal{F}$ there exists some $\lambda = (\lambda_1, \ldots, \lambda_\ell) \in \mathbb{R}^\ell_{\geq 0}$ such that

$$v = \sum_{i=1}^{\ell} \lambda_i E_i = E\lambda$$

where $E \in \mathbb{R}^{r \times \ell}$ denotes the matrix with columns $E_1, \ldots, E_\ell$. We call E a **matrix of extreme vectors**. This gives rise to the following *convex parametrization*:

$$\Psi \colon \mathbb{R}^n_{>0} \times \mathbb{R}^\ell_{>0} \to \mathcal{V}, \quad (h, \lambda) \mapsto \ (\tfrac{1}{h}, \mathrm{diag}((h^{A_1}, \ldots, h^{A_r}))E\lambda), \tag{11}$$

where $A_1, \ldots, A_r$ denote the columns of the matrix $A := \left[a_{ij}\right] \in \mathbb{R}^{n \times r}$ of the coefficients of the reactants of the reactions, h^{A_j} is short notation for $h_1^{a_{1j}} \cdots h_n^{a_{nj}}$, $\mathrm{diag}(v)$ is the diagonal matrix with diagonal entries given by v, and $1/h$ is taken component-wise.

In Corollary 6.2(a) in Section 6, we show that Ψ is surjective if E does not have a row where all the entries are equal to zero, and hence Ψ is a parametrization of $\mathcal{V}$. This restriction is not relevant for our purposes: a zero row of E is equivalent to $\ker(N)$ not having any positive vector, and hence there is no positive steady state of the ODE system (2), see Corollary 6.2(b). In particular, the reaction network cannot be multistationary. In the rest of the work, we assume that E does not have a zero row and when this holds, we say that the network is **consistent**[23] (consistent networks are called *dynamically nontrivial* in other works, e.g. [24]).

For the convex parametrization, the critical function $g \circ \Psi$ can be represented in a direct way using the following observation. The Jacobian of $f_\kappa(x)$ evaluated at $\Psi(h, \lambda)$ equals

$$\tilde{J}(h, \lambda) := N \,\mathrm{diag}(E\lambda) A^\top \,\mathrm{diag}(h), \tag{12}$$

for each $(h, \lambda) \in \mathbb{R}^n_{>0} \times \mathbb{R}^\ell_{>0}$, see [28]. We construct the matrix $\tilde{M}(h, \lambda)$ from $\tilde{J}(h, \lambda)$ as above: if W is row reduced and $i_1 < \cdots < i_{n-s}$ are the indices of the first non-zero coordinates of each

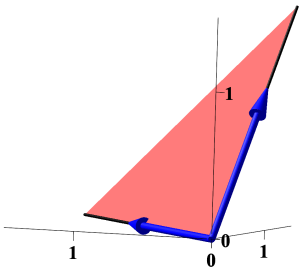


FIGURE 4. The flux cone of the running example in $\mathbb{R}^3$. The extreme vectors $E_1 = (1,0,1)$, $E_2 = (1,1,0)$ are shown in blue.

row, replace the $i_j$throw of $\tilde{J}(h,\lambda)$ by the jth row of W. Then, it holds

$$(g \circ \Psi)(h,\lambda) = (-1)^s \det \tilde{M}(h,\lambda). \tag{13}$$

From this equality, one computes $g \circ \Psi$ directly using symbolic software. Since the entries of $\tilde{M}(h,\lambda)$ are polynomials in (h,λ), so is $g \circ \Psi$. In the following, we call this polynomial the **critical polynomial**.

Let us find the critical polynomial $g \circ \Psi$ for the running example. The flux cone $\mathcal{F}$ and its extreme vectors are displayed in Fig. 4. Now, all we have to do is to compute the matrix product in (12)

$$\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1+\lambda_2 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} = \begin{bmatrix} (\lambda_1-\lambda_2)h_1 & (\lambda_1+\lambda_2)h_2 \\ -(\lambda_1-\lambda_2)h_1 & -(\lambda_1+\lambda_2)h_2 \end{bmatrix},$$

and replace the first row by $(1,1)$. After taking the determinant and multiplying by $(-1)^s = -1$, we obtain the critical polynomial:

$$(g \circ \Psi)(h_1,h_2,\lambda_1,\lambda_2) = h_1\lambda_2 - h_1\lambda_1 + h_2\lambda_1 + h_2\lambda_2. \tag{14}$$

The above discussion shows that we can always find a suitable parametrization and compute the critical polynomial. In some cases, other types of parametrizations arise by parametrizing each $V_{\kappa,>0}$ separately. This is done by first trying to express some variables among $x_1,\dots,x_n$, $\kappa_1,\dots,\kappa_r$ in terms of the others using the equations in (5). For finding these expressions, a computer algebra system such as `SageMath` [33] or `Maple` [34] can be useful. Once such an expression is found, one should check whether it gives a well-defined surjective analytic map, that is a parametrization. If all the components of the parametrization Φ are quotients of polynomials with positive denominators, then $g \circ \Phi$ is a quotient of polynomials too, and its denominator is positive.

Let us see how this works in practice for the running example. We see from the ODE system in Fig. 3 that positive steady states are characterized by

$$x_2 = \tfrac{\kappa_1 x_1}{\kappa_3 x_1^2 + \kappa_2}.$$

This expression gives the parametrization

$$\Phi\colon\ \mathbb{R}^4_{>0} \to \mathcal{V} \subseteq \mathbb{R}^2_{>0} \times \mathbb{R}^3_{>0}, \qquad (x_1,\kappa_1,\kappa_2,\kappa_3) \mapsto (x_1, \tfrac{\kappa_1 x_1}{\kappa_3 x_1^2+\kappa_2}, \kappa_1,\kappa_2,\kappa_3).$$

Combining Φ with g from (7), we get the critical function:

$$(g \circ \Phi)(x_1,\kappa_1,\kappa_2,\kappa_3) = \tfrac{\kappa_3^2 x_1^4 - \kappa_1\kappa_3 x_1^2 + 2\kappa_2\kappa_3 x_1^2 + \kappa_1\kappa_2 + \kappa_2^2}{\kappa_3 x_1^2 + \kappa_2}. \tag{15}$$

In general, there is no guarantee that such a parametrizations for each κ can be found. However, there are broad classes of reaction networks allowing such a parametrization, for example networks with toric steady states [35] and post-translational modification systems [36] to name a few. As we always can find a critical function using the convex parametrization, one

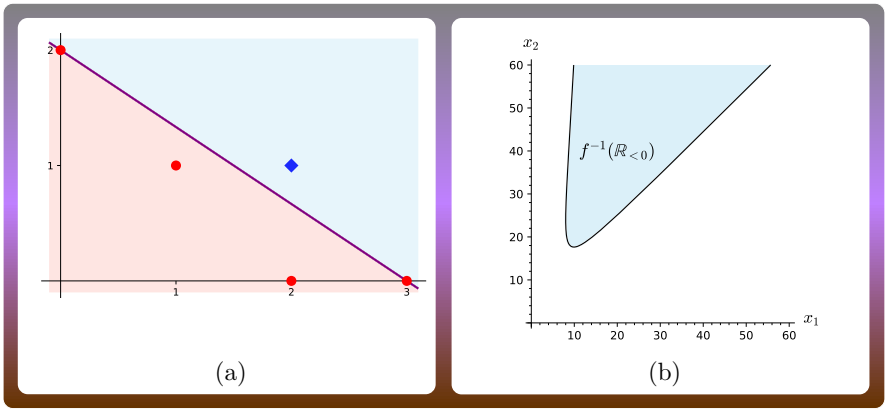


FIGURE 5. (a) A strict separating hyperplane (in purple) of the support of $f(x_1, x_2) = x_1^2 - x_1^2 x_2 + 2x_1 x_2 + x_1^3 + x_2^2$. Red dots correspond to positive exponents, the blue square corresponds to the only negative exponent. (b) The preimage of the negative real half-line under f.

might wonder what the value of these other type of parametrizations is. The point is that with this type, the reaction rate constants are still present in the parametrization, and this is useful to get information about what parameter values yield to multistationarity. This was the theme of [9], and we will explore this advantage later in the application of our algorithm to the network with allosteric reciprocal enzyme regulation in Fig. 2(c).

Finally, we discuss how to address **(Step 3)**, now that we know how to compute the critical polynomial/function. To check whether the preimage of the negative real half-line under a critical function is path connected is in general hard and depends strongly on the parametrization. As we discussed in (Step 2), critical functions are in practice polynomials or rational functions with positive denominator. In the latter case, we can restrict to the numerator of the rational function. To verify the conditions in Theorem 2.2, it is then enough to study the preimage of the negative real half-line under a polynomial function restricted to the positive orthant.

Recall that a polynomial function can be written as

$$f\colon\ \mathbb{R}^k_{>0} \to \mathbb{R}, \qquad f(x) = \sum_{\mu \in \sigma(f)} c_\mu x_1^{\mu_1} \dots x_k^{\mu_k}, \quad \text{with } c_\mu \neq 0,$$

and $\sigma(f) \subseteq \mathbb{N}^k$ is a finite set, called the *support* of f. To determine whether the *preimage of the negative real half-line*

$$f^{-1}(\mathbb{R}_{<0}) = \{x \in \mathbb{R}^k_{>0} \mid f(x) < 0\}$$

is path connected, one can use methods from real algebraic geometry [15, Remark 11.19], [37, Section 3]. These methods work well for polynomials in few variables, but they scale poorly. If the polynomial has many variables, the computation is unfeasible.

In [18], the authors of the present work gave a sufficient criterion for deciding that $f^{-1}(\mathbb{R}_{<0})$ is path connected, based on the geometry of the support and the sign of the coefficients. We call an exponent $\mu \in \sigma(f)$ positive (resp. negative) if the corresponding coefficient c_μ is positive (resp. negative). We write $\sigma_+(f)$ (resp. $\sigma_-(f)$) for the set of positive (resp. negative) exponents of a polynomial f. For example, the polynomial $f(x_1, x_2) = x_1^2 - x_1^2 x_2 + 2x_1 x_2 + x_1^3 + x_2^2$ has four positive exponents $(2,0), (1,1), (3,0), (0,2)$ and one negative exponent $(2,1)$. These exponents are depicted in Fig. 5.

A hyperplane in $\mathbb{R}^k$ is the set of solutions $\mu \in \mathbb{R}^k$ of a linear equation

$$v \cdot \mu = a,$$

where $v \in \mathbb{R}^k \setminus \{0\}$, $a \in \mathbb{R}$ and $v \cdot \mu$ denotes the Euclidean scalar product of two vectors. Each hyperplane has two sides, which are described by the linear inequalities

$$v \cdot \mu \leq a, \qquad \text{and} \qquad v \cdot \mu \geq a.$$

A hyperplane is called *strictly separating* if the positive and negative exponents of f are on different sides of the hyperplane and not all the negative exponents are on this hyperplane. For a geometric interpretation, we refer to Fig. 5.

Strict separating hyperplanes can be used to decide upon path connectivity of the multistationarity region using Theorem 2.2 via the following theorem.

Theorem 2.3 (Preimage of negative real half-line)**.** *[18, Theorem 3.9] Let $f\colon \mathbb{R}^k_{>0} \to \mathbb{R}$ be a polynomial function. If there exists a strict separating hyperplane of the support of f, then $f^{-1}(\mathbb{R}_{<0})$ is path connected and its closure equals $f^{-1}(\mathbb{R}_{\leq 0})$.*

For the running example, the supports of the two critical functions (14) and (15) form a quadrilateral. In both cases, there is only one negative exponent, which is at a corner (vertex) of the quadrilateral. Hence, one can easily find strict separating hyperplanes for the supports of each polynomial. To get a geometric intuition, we investigate the numerator of (15). Its support lives in a 2 dimensional subspace of $\mathbb{R}^4$, so we can project it onto the plane $\mathbb{R}^2$ and find a strict separating hyperplane there. The projected support is precisely that depicted in Fig. 5(a).

Therefore, Theorem 2.3 holds for the running example and any of the two critical functions, and then, by Theorem 2.2 we conclude that the multistationarity region of the running example is path connected.

Note that a strict separating hyperplane of the support of f exists if the following system of linear inequalities has a solution $(v, a) \in \mathbb{R}^{k+1}$:

$$v \cdot \alpha \leq a, \quad \text{for all } \alpha \in \sigma_+(f) \tag{16}$$

$$v \cdot \beta \geq a, \quad \text{for all } \beta \in \sigma_-(f) \tag{17}$$

$$\sum\nolimits_{\beta \in \sigma_-(f)} (v \cdot \beta - a) > 0. \tag{18}$$

For the polynomial in Fig. 5, $v = (2, 3)$, and $a = 6$ form a solution to the system.

In practice, we determine whether the system of linear inequalities (16)-(18) has a solution as follows. First, we construct the polyhedral cone $C \subseteq \mathbb{R}^{k+1}$ defined by the inequalities (16)-(17). Second, we pick a point (v, a) in the relative interior of C. If there exists $\beta \in \sigma_-(f)$ such that $v \cdot \beta > a$, then (v, a) satisfies also inequality (18) and a strict separating hyperplane exists. If such β does not exist, then a simple argument gives that $\sigma_-(f)$ is contained in the hyperplane defined by any $(w, b) \in C$, i.e. $w \cdot \beta = b$ for all $\beta \in \sigma_-(f)$ and all $(w, b) \in C$. Therefore a strictly separating hyperplane does not exist.

Theorem 2.3 gives a way to assert that the multistationarity region is path connected, but it is not informative if that is not the case. In [18] additional results are given to include two path-connected components. One of these results will be used to show that the multistationarity region of network in Fig. 2(c) has two path-connected components.

Model reduction for the simplification of the computations. Finding the critical function or the critical polynomial requires the computation of the determinant of a symbolic matrix, which can have a high computational cost if the matrix is large or the entries are long expressions in the symbolic variables. The next theorem shows that it is possible to remove certain reverse reactions from the network, and use a critical function for the reduced network to study the multistationarity region of the original network, thereby reducing (often dramatically) the computational cost.

Theorem 2.4 (Reduction and connectivity)**.** *Consider a conservative reaction network $(\mathcal{S}, \mathcal{R})$ without relevant boundary steady states. Assume that there exist species $\mathrm{X}_1, \ldots, \mathrm{X}_k \in \mathcal{S}$ such*

that each X_j *participates in exactly* 3 *reactions of the form*

$$a_{k+1,j}X_{k+1}+\cdots+a_{n,j}X_n \underset{\kappa_{3j}}{\overset{\kappa_{3j-1}}{\rightleftharpoons}} \mathrm{X}_j \xrightarrow{\kappa_{3j-2}} b_{k+1,j}X_{k+1}+\cdots+b_{n,j}X_n, \quad j=1,\dots,k.$$

Let $\tilde{g}\circ\tilde{\Phi}$ *be a critical function of the reduced network obtained by removing the reactions corresponding to* κ_{3j} *for* $j=1,\dots,k$*. Assume that the closure of* $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ *equals* $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{\leq 0})$*.*

Then the number of path-connected components of the multistationarity region for both the reduced and the original reaction network is at most the number of path-connected components of $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})$*.*

In particular, if $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ *is path connected, then the multistationarity region of the original network* $(\mathcal{S},\mathcal{R})$ *is path connected.*

The theorem might look a bit technical, but it is simply saying that it is enough to apply the algorithm to a smaller network obtained by removing the reverse reactions κ_{3j}, and the conclusions can be translated to the original network. Removal of reverse reactions contribute to the reduction of the computational cost as each of them gives an extreme vector to the flux cone (see Lemma 6.3(a) in Section 6). Making reversible reactions irreversible removes this extreme vector and thereby the matrix $\tilde{M}(h,\lambda)$ depends on one less variable.

To illustrate Theorem 2.4, we consider the reaction network representing a signaling **cascade with shared kinase** in Fig. 2(i). This reaction network describes the phosphorylation of two substrates S and P with one phosphorylation site. The phosphorylation of S is catalyzed by a kinase E, while the phosphorylation of P is catalyzed both by E and by the phosphorylated form of S. The dephosphorylation processes are governed by two different phosphatases F_1 and F_2 [38].

One checks using the above criteria that this network is conservative and has no relevant boundary steady states. Then, Theorem 2.2 for connectivity of the multistationarity region can be applied. A matrix of extreme vectors, formed by a minimal collection of extreme vectors generating the flux cone is

$$E=\begin{bmatrix}
0&0&0&0&0&0&0&1\\
0&0&0&0&0&0&1&1\\
0&0&0&0&0&0&1&0\\
0&0&0&0&0&0&0&1\\
0&0&0&0&0&1&0&1\\
0&0&0&0&0&1&0&0\\
1&0&0&0&0&0&0&0\\
1&0&0&1&0&0&0&0\\
0&0&0&1&0&0&0&0\\
0&0&0&0&1&0&0&0\\
0&0&1&0&1&0&0&0\\
0&0&1&0&0&0&0&0\\
1&0&0&0&1&0&0&0\\
1&1&0&0&1&0&0&0\\
0&1&0&0&0&0&0&0
\end{bmatrix}.$$

The highlighted column extreme vectors correspond to the 5 reversible reactions of the type in Theorem 2.4 (the κ_{3j} in the theorem). Computing the determinant of (12) takes approximately 1.5 minutes. The critical polynomial has 20 variables and 5312 terms.

Following Theorem 2.4, we remove the reactions corresponding to $\kappa_3,\kappa_6,\kappa_9,\kappa_{12},\kappa_{15}$. The reduced network has the form

$$\begin{aligned}
&\mathrm{S}+\mathrm{E}\xrightarrow{\kappa_2}\mathrm{SE}\xrightarrow{\kappa_1}\mathrm{S_p}+\mathrm{E} \qquad && \mathrm{S_p}+\mathrm{F_1}\xrightarrow{\kappa_5}\mathrm{S_pF_1}\xrightarrow{\kappa_4}\mathrm{S}+\mathrm{F_1}\\
&\mathrm{P}+\mathrm{S_p}\xrightarrow{\kappa_8}\mathrm{PS_p}\xrightarrow{\kappa_7}\mathrm{P_p}+\mathrm{S_p} && \mathrm{P}+\mathrm{E}\xrightarrow{\kappa_{11}}\mathrm{PE}\xrightarrow{\kappa_{10}}\mathrm{P_p}+\mathrm{E}\\
&\mathrm{P_p}+\mathrm{F_2}\xrightarrow{\kappa_{14}}\mathrm{P_pF_2}\xrightarrow{\kappa_{13}}\mathrm{P}+\mathrm{F_2}.
\end{aligned}$$

A matrix of extreme vectors is now

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

The extreme vectors are the non-highlighted vectors in the matrix E above, with the entries corresponding to the reverse reactions removed (every third row of E). Computing the critical polynomial for the reduced network takes only 4 seconds. This critical polynomial is much simpler than the one for the full network. It has 15 variables and 204 terms.

This example illustrates that the reduction in Theorem 2.4 might reduce the computational cost substantially. On one hand, the computation of the critical polynomial is faster and, on the other, the critical polynomial itself has less variables and terms, and therefore checking **(Step 3)** becomes faster as well. In the next section, we investigate networks where the benefit of applying network reduction and Theorem 2.4 is more dramatical. For example, for two of the networks, computing the critical polynomial for the full network turned out to be infeasible, but the computation became possible for the reduced network. By means of Theorem 2.4, we could assert connectivity of the multistationarity region for the full network (see Table 1 for more detail). This illustrates that Theorem 2.4 allows us to apply our approach to networks that were originally too large.

An important observation is that the existence of a strict separating hyperplane for a network or for a reduced version of it like in Theorem 2.4 are independent. That is, if we cannot find a strict separating hyperplane for the reduced network, it could still be that it exists for the original network. Also, the existence of this hyperplane depends on the choice of critical function, that is, of the parametrization.

Algorithm for path connectivity. We conclude this section by giving a procedure that checks a sufficient criterion for connectivity of the multistationarity region with no user intervention. Since most of the steps rely on solving linear inequalities, we implemented the algorithm using the computer algebra system `SageMath` [33]. The code is given in the **Supporting Information**. We would like to emphasize that the multistationarity region could still be path connected, even if our algorithm terminates inconclusively.

Algorithm 2.5. ***Input:*** *a reaction network*

(Step 1) *Check that the reaction network is conservative and that it does not have relevant boundary steady states using siphons.*

(Step 2) *Compute the convex parametrization map Ψ if the network is consistent, and the critical polynomial $g \circ \Psi$ from* (13).

(Step 3) *Decide whether a strict separating hyperplane of the support of $g \circ \Psi$ exists.*

(Step 4) *Eventually repeat Steps 1-3 with a reduced network as in Theorem 2.4.*

Output: *'The parameter region of multistationarity is path connected' or 'The algorithm is inconclusive'.*

3. Investigating connectivity in relevant biochemical networks

We now demonstrate that Algorithm 2.5 is useful for realistic networks and that the number of connected components of the multistationarity region can be understood for several relevant networks in cell signaling of moderate size.

We start by going through the algorithm with two small networks: first, with the module regulating the cell cycle shown in Fig. 2(a), and then, with the simplified hybrid histidine kinase

network in Fig. 2(b). The corresponding matrices and the critical polynomials become more complicated than for the running example, but still, are small enough to be displayed here.

Afterwards we analyze the rest of the networks in Fig. 2. Additionally, we consider the extensions of Fig. 2(e-f), with two phosphorylation sites, to several phosphorylation sites, and explore the strengths and weaknesses of the algorithm. When increasing the network size, the computation of the critical polynomial becomes unfeasible and we apply the reduction from Theorem 2.4.

We summarize the main properties of all the applications of the algorithm discussed in this work in Table 1. Table 1 shows the number of species, reactions, and extreme vectors of the reaction network. If the critical polynomial can be computed, it shows the number of positive and negative exponents of the critical polynomial and whether a strict separating hyperplane of the support exists. The same computations are repeated with the reduced network of Theorem 2.4, and we report the same data except the number of species and reactions.

3.1. **Small networks.**

Cell cycle regulating module. We consider the model proposed in [39] for the second module that regulates the cell's transition from G2 phase to M-phase [40], which is shown in Fig. 2(a). This model has been analyzed for bistability in [30]. With the order of species $\mathrm{C}, \mathrm{C}^+, \mathrm{M}, \mathrm{M}^+, \mathrm{W}, \mathrm{W}^+$, the stoichiometric matrix, a matrix of conservation relations and a matrix of extreme vectors are:

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}, \quad W = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

The sum of the three rows of W gives a positive vector and hence the network is conservative. We further verified that the network has no relevant boundary steady states. Furthermore, E has no zero row, so the network is consistent and the critical polynomial can be found using (13). The matrix $\tilde{M}(h,\lambda)$ is found by replacing the first, third and fifth rows of $N \operatorname{diag}(E\lambda) A^\top \operatorname{diag}(h)$ by the rows of W:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ \lambda_2 h_1 & -\lambda_2 h_2 & -\lambda_2 h_3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ -\lambda_3 h_1 & 0 & \lambda_3 h_3 & -\lambda_3 h_4 & \lambda_3 h_5 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & \lambda_1 h_3 & 0 & \lambda_1 h_5 & -\lambda_1 h_6 \end{bmatrix}.$$

Since $s = 3$, the negative of the determinant of $\tilde{M}(h,\lambda)$ gives the critical polynomial:

$$(g \circ \Psi)(h,\lambda) = (-h_1h_3h_5 + h_1h_4h_5 + h_2h_4h_5 + h_2h_3h_6 + h_1h_4h_6 + h_2h_4h_6)\lambda_1\lambda_2\lambda_3.$$

A strict separating hyperplane exists, for example $v \cdot (h_1, \ldots, h_6, \lambda_1, \lambda_2, \lambda_3) = 2$ with

$$v = (1,0,1,0,1,0,0,0,0).$$

Indeed, $(g \circ \Psi)(h,\lambda)$ has 6 monomials, all with exponent $(1,1,1)$ for λ, and for h they have exponents $(1,0,1,0,1,0)$, $(1,0,0,1,1,0)$, $(0,1,0,1,1,0)$, $(0,1,1,0,0,1)$, $(1,0,0,1,0,1)$, $(0,1,0,1,0,1)$. The first exponent is negative, and the scalar product with v returns the value 3, which is strictly larger than 2. The other exponents correspond to positive coefficients, and their scalar product with v give the values $2,1,1,1,0$ respectively. All of them are smaller or equal to 2. Therefore, the condition for being a strict separating hyperplane holds, and we conclude that the multistationarity region of this network is path connected.

Hybrid histidine kinase. The hybrid histidine kinase network in Fig. 2(b) comprises a hybrid histidine kinase HK with the domain REC embedded, and separate histidine phospho-transfer domain Hpt. This reaction network has been studied in [41], where it was shown that the network displays multistationarity and a (complicate) description of the set of parameters with 3 steady states is given. It is further known that there is a choice of total concentrations such that the network is multistationary if and only if $\kappa_3 > \kappa_1$, see [9] (this set is the projection of the multistationarity region onto the space of reaction rate constants). It was not known whether the full multistationarity region in κ and c is path connected.

With the order of species $\mathrm{HK}_{00}, \mathrm{HK}_{\mathrm{p}0}, \mathrm{HK}_{0\mathrm{p}}, \mathrm{HK}_{\mathrm{pp}}, \mathrm{Hpt}, \mathrm{Hpt}_{\mathrm{p}}$, the stoichiometric matrix N, a matrix of conservation relations W, and a matrix E whose columns are a minimal set of extreme vectors are

$$N = \begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}, \quad W = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

The network is conservative, consistent, and has no relevant boundary steady states. By replacing the first and fifth rows of $N\,\mathrm{diag}(E\lambda)A^\top\,\mathrm{diag}(h)$ by the rows of W we find the matrix $\tilde{M}(h,\lambda)$ and its determinant gives the critical polynomial:

$$\begin{aligned}(g\circ\Psi)(h,\lambda) = {} & h_1h_2h_3h_5\lambda_1^3\lambda_2 + 3\,h_1h_2h_3h_5\lambda_1^2\lambda_2^2 + 2\,h_1h_2h_3h_5\lambda_1\lambda_2^3 + h_1h_2h_4h_5\lambda_1^2\lambda_2^2 \\ & + h_1h_2h_4h_5\lambda_1\lambda_2^3 - h_2h_3h_4h_5\lambda_1^3\lambda_2 - h_2h_3h_4h_5\lambda_1^2\lambda_2^2 + h_1h_2h_3h_6\lambda_1^3\lambda_2 \\ & + 2\,h_1h_2h_3h_6\lambda_1^2\lambda_2^2 + h_1h_2h_3h_6\lambda_1\lambda_2^3 + h_1h_2h_4h_6\lambda_1^3\lambda_2 + 2\,h_1h_2h_4h_6\lambda_1^2\lambda_2^2 \\ & + h_1h_2h_4h_6\lambda_1\lambda_2^3 + h_1h_3h_4h_6\lambda_1^3\lambda_2 + 2\,h_1h_3h_4h_6\lambda_1^2\lambda_2^2 + h_1h_3h_4h_6\lambda_1\lambda_2^3 \\ & + h_2h_3h_4h_6\lambda_1^3\lambda_2 + 2\,h_2h_3h_4h_6\lambda_1^2\lambda_2^2 + h_2h_3h_4h_6\lambda_1\lambda_2^3.\end{aligned}$$

The polynomial has 8 variables and 17 positive and 2 negative coefficients. A strict separating hyperplane of its support is given by the equation

$$(-5,-5,-1,0,5,0,0,0)\cdot\mu = -3.$$

Using Theorem 2.2, it follows that the multistationarity region for the hybrid histidine kinase network is path connected.

3.2. **Phosphorylation cycles.** We investigate models for phosphorylation and dephosphorylation of a substrate S with m binding sites with processes catalyzed by a kinase E and one or more phosphatases F.

Sequential and distributive phosphorylation cycles. We first assume that phosphorylation and dephosphorylation occurs sequentially and distributively [42]: the kinase E catalyzes the phosphorylation one site at a time in a given order, while the phosphatase F dephosphorylates in the reverse order, also one site at a time. Under these assumptions, the network for $m = 2$ sites is shown in Fig. 2(e).

The dynamics of phosphorylation cycles have been intensively studied, e.g. [14, 28, 42–52]. In particular, it is known that they are multistationary for $m \geq 2$, and there are choices of parameter values where they have $m + 1$ steady states for m even, and m for m odd [44], with half of them plus one being asymptotically stable [47]. It has been conjectured that these networks can have up to $2m - 1$ steady states, but this has only been established for small m [48]. These networks are in the class of post-translational modification networks, which are conservative and consistent, and by the results in [25], since the underlying substrate network is strongly path connected, they do not have relevant boundary steady states.

The phosphorylation cycle with $m = 2$ binding sites has $n = 9$ species and the flux cone has $\ell = 6$ extreme vectors. Then, the corresponding critical polynomial $g \circ \Psi$ has 15 variables. It is a big polynomial with 288 positive and 112 negative exponents. Despite the large number of

exponents, a strict separating hyperplane of its support can be found in less than a second. Our algorithm (and Theorem 2.2) can be applied to conclude that the multistationarity region for the sequential and distributive phosphorylation cycle with two binding sites is path connected.

By increasing the number m of binding sites, the reaction network size increases systematically. For example, for $m = 3$, the reaction network becomes

$$\mathrm{S} + \mathrm{E} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} \mathrm{ES} \xrightarrow{\kappa_3} \mathrm{S_p} + \mathrm{E} \underset{\kappa_8}{\overset{\kappa_7}{\rightleftharpoons}} \mathrm{ES_p} \xrightarrow{\kappa_9} \mathrm{S_2} + \mathrm{E} \underset{\kappa_{14}}{\overset{\kappa_{13}}{\rightleftharpoons}} \mathrm{ES_{pp}} \xrightarrow{\kappa_{15}} \mathrm{S_{ppp}} + \mathrm{E}$$

$$\mathrm{S_{ppp}} + \mathrm{F} \underset{\kappa_{17}}{\overset{\kappa_{16}}{\rightleftharpoons}} \mathrm{FS_{ppp}} \xrightarrow{\kappa_{18}} \mathrm{S_{pp}} + \mathrm{F} \underset{\kappa_{11}}{\overset{\kappa_{10}}{\rightleftharpoons}} \mathrm{FS_{pp}} \xrightarrow{\kappa_{12}} \mathrm{S_p} + \mathrm{F} \underset{\kappa_5}{\overset{\kappa_4}{\rightleftharpoons}} \mathrm{FS_p} \xrightarrow{\kappa_6} \mathrm{S} + \mathrm{F}.$$

The critical polynomial $g \circ \Psi$ has 2560 positive and 1536 negative exponents. The algorithm confirms that the multistationarity region is path connected in 96 seconds.

For $m = 4$, we could not compute the critical polynomial for the original network due to computer memory constraints. This shows that the computation of the critical polynomial is the bottleneck of the algorithm. After removing all reverse reactions as in Theorem 2.4, we could compute the critical polynomial, but it does not have a strict separating hyperplane. Therefore, for this family of networks we know that the multistationarity region is path connected for $m = 2, 3$, but it is unknown for $m \geq 4$. Previous work has shown that the projection of the multistationarity region onto the reaction rate constants κ is path connected for all $m \geq 2$, see [53], so we conjecture that the full multistationarity region is path connected for all $m \geq 2$.

Phosphorylation cycles with different phosphatases. Different mechanisms for multisite phosphorylation have been observed, and in particular, phosphorylation and dephosphorylation of the different sites of a phosphate might not be catalyzed by the same kinase or phosphatase e.g. [54, 55]. If all steps are carried out by different enzymes, then multistationarity does not arise (see [38] for $m = 2$). Therefore, we consider the scenario where the phosphorylation occurs sequentially and the kinase acts in a distributive way, but we assume that each dephosphorylation step is governed by **different phosphatases** $F_1, \ldots, F_m$ (see Fig. 2(f) for $m = 2$). These networks are also conservative and do not have relevant boundary steady states for any m.

For $m = 2$ and $m = 3$, the algorithm finds a strict separating hyperplane, and hence the multistationarity region is path connected. For $m = 4$, the computation of the critical polynomial via the symbolic determinant in (Step 2) of the algorithm was too demanding, and the computer used for the tests ran out of memory. We employed the reduction approach given in Theorem 2.4 and removed eight reactions. The critical polynomial of the reduced network is significantly simpler: it has 22 variables and 178 monomials. Its support has a strict separating hyperplane. Thus, the multistationarity region of the original network is path connected by Theorem 2.4 and Theorem 2.3.

Weakly irreversible phosphorylation cycles. The two-site sequential and distributive phosphorylation network given in Fig. 2(e) assumes that each phosphorylation step proceeds via a Michaelis-Menten mechanism. This is referred to as *strong irreversibility* in [43]. More plausible mechanisms have been argued to include the complex formation of the product with the enzyme, as for example a mechanism of this form would allow:

$$\mathrm{S} + \mathrm{E} \rightleftharpoons \mathrm{SE} \to \mathrm{S_pE} \rightleftharpoons \mathrm{S_p} + \mathrm{E}.$$

A model incorporating this **weak irreversibility** at the dephosphorylation stage was proposed and analyzed for bistability in [56, Scheme 2]. In the model, dephosphorylation of ERK by the phosphatase MKP3 proceeds as shown in Fig. 2(g) [57]. For this network, our algorithm concludes that the multistationarity region is path connected.

A model with **full weak irreversibility**, that is, for both the phosphorylation and dephosphorylation processes, is shown in Fig. 2(h). The shape of the multistationarity region for this type of models was analyzed in [43], where it was concluded by means of a numerical approach that the multistationarity region in some aggregated steady state parameters is connected. Neither for the original network nor for the reduced network, a strict separating hyperplane exists,

TABLE 1. **Summary of the algorithm on selected systems**

Reaction network	n	r	ℓ	$\#\sigma_+(g\circ\Psi)$	$\#\sigma_-(g\circ\Psi)$	sep. hyp.	$\tilde{\ell}$	$\#\sigma_+(\tilde{g}\circ\tilde{\Psi})$	$\#\sigma_-(\tilde{g}\circ\tilde{\Psi})$	sep. hyp.
(a) Cell cycle	6	6	3	5	1	YES	3	5	1	YES
(b) Hybrid histidine kinase	6	6	2	17	2	YES	2	17	2	YES
(c) Allosteric regulation	9	10	5	168	8	NO	3	42	2	NO
(d) Covalent regulation	12	14	7	1856	32	YES	3	116	2	YES
(e) 2-site phosph. cycle	9	12	6	288	112	YES	2	18	7	YES
3-site phosph. cycle	12	18	9	2560	1536	YES	3	40	24	YES
4-site phosph. cycle	15	24	12	??	??	??	4	75	54	NO
(f) 2-site: different phosphat.	10	12	6	304	48	YES	2	19	3	YES
3-site: diff. phosphat.	14	18	9	3264	960	YES	3	51	15	YES
4-site: diff. phosphat.	18	24	12	??	??	??	4	127	51	YES
(g) 2-site: weak. irrev. dephos.	11	16	8	2176	640	YES	4	136	40	YES
(h) 2-site: fully weak. irrev.	13	20	10	16320	3648	NO	6	1020	228	NO
(i) Two-layer, shared kinase	12	15	8	5088	224	YES	3	195	9	YES
(j) Two layer MAPK cascade	14	18	9	5120	1408	YES	3	80	22	YES
(k) ERK network	12	18	9	15040	3432	YES	5	1374	340	YES

The columns of the table indicate: network; n = number of species; r = number of reactions; ℓ = number of extreme vectors of the flux cone; $\#\sigma_\pm(g\circ\Psi)$ = number of positive/negative exponents of the critical polynomial; existence of a strict separating hyperplane; $\tilde{\ell}$ = number of extreme vectors of the flux cone of the reduced network of Theorem 2.4; $\#\sigma_\pm(\tilde{g}\circ\tilde{\Psi})$ = number of positive/negative exponents of the critical polynomial of the reduced network; existence of strict separating hyperplane for the reduced network. The number of variables of the critical polynomial is $n+\ell$ for the original network and $n+\tilde{\ell}$ for the reduced network. ?? means that the computation could not be performed due to computer memory loss. The labels (a)-(k) refer to the networks in Fig. 2.

and hence our algorithm is inconclusive. It remains thus open to be confirmed whether the multistationarity region is path connected.

Extracellular signal-regulated kinase (ERK) network. Dual-site phosphorylation and dephosphorylation of extracellular signal-regulated kinase has an important role in the regulation of many cellular activities [58], and a better knowledge of the dynamical properties of the ERK network might facilitate the prediction of this network's response to environmental changes or drug treatments [59]. This network, analyzed in [60–62] and shown in Fig. 2(k), comprises as well phosphorylation of a substrate in two sites, but not in a distributive and sequential way.

Using Algorithm 2.5, we conclude that the multistationarity region for the ERK network is path connected.

3.3. **Signaling cascades.** We next investigate two types of signaling cascades comprising phosphorylation cycles, which are known to be multistationary.

Shared kinase. We consider first a two-layer signaling cascade with two single phosphorylation at each stage. The phosphorylated substrate of the first layer acts as the kinase of the second layer. We consider additionally that the kinase of the first layer also can act as kinase for the second layer, so the kinase is *shared* for the two layers, as shown in Fig. 2(i). Without this shared kinase, the cascade would not display multistationarity.

Algorithm 2.5 finds a strict separating hyperplane and we conclude that the multistationarity region for the cascade is path connected.

Two-layer MAPK-cascade. Huang and Ferrell proposed a model for the MAPK cascade consisting of three layers, the first one being a single phosphorylation cycle, while the last two are dual phosphorylation cycles with phosphorylation and dephosphorylation proceeding sequentially and distributive [63]. This network has bistability and also oscillations, and in fact for both properties only the first two layers of the cascade are required.

The network with two layers is shown in Fig. 2(j), and Algorithm 2.5 can be employed to conclude that the multistationarity region is path connected.

The full network with the three layers is large, with 22 species, 30 reactions, the rank of the stoichiometric matrix is 15, and the matrix E has 15 extreme vectors. Due to the computational cost, the computation of the critical polynomial was not possible for the original network, but was possible for the reduced network using the `Julia` package `SymbolicCRN.jl` [64]. However, a strict separating hyperplane does not exist, so our algorithm is inconclusive.

3.4. **Reciprocal enzyme regulation.** Finally, we consider two multistationary networks comprising single phosphorylation cycles, where the kinase and the phosphatase are subject to reciprocal regulation, both proposed and studied in [17].

Covalent regulation. We first consider the case where reciprocal regulation is via covalent modification catalyzed by the same enzyme, see Fig. 2(d). By means of Algorithm 2.5 we find that the multistationarity region for the the cascade is path connected.

Allosteric regulation: An example with two path-connected components. The other mechanism of reciprocal regulation considered in [17] is via allosteric regulation: it is assumed that there is an allosteric effector L that binds both the phosphatase and the kinase, see Fig. 2(c). After performing a quasi-steady-state approximation, the authors of [17] show that a necessary condition for multistationarity is that $\kappa_3 > \kappa_6$. This network is conservative, consistent, and has no relevant boundary steady states.

For all the applications we have seen so far, we could either conclude that the multistationarity region is path connected, or our approach was inconclusive. For this network our approach was inconclusive as well, however, by employing other theoretical results from [18] in conjunction with Theorem 2.4 and the approaches in [9], we were able to conclude that the multistationarity region has exactly two path-connected components, revealing two mechanisms underlying multistationarity.

The approach is as follows. On one hand, using the method to find parameter regions in κ from [9] relying on Theorem 2.1, we show in Section 6 that the two sets of parameters $\{(\kappa, c) \in \mathbb{R}^{10}_{>0} \times \mathbb{R}^5_{\geq 0} \,|\, \kappa_3 > \kappa_6\}$ and $\{(\kappa, c) \in \mathbb{R}^{10}_{>0} \times \mathbb{R}^5_{\geq 0} \,|\, \kappa_3 < \kappa_6\}$ both contain parameters that yield multistationarity, that is, these two regions both intersect the multistationarity region. To be precise, the condition $\kappa_3 > \kappa_6$ yields multistationarity if additionally the Michaelis-Menten constant for phosphorylation, $K_1 = \frac{\kappa_2+\kappa_3}{\kappa_1}$, is large enough relative to that for the dephosphorylation, $K_2 = \frac{\kappa_5+\kappa_6}{\kappa_4}$. Symmetrically, the condition $\kappa_6 > \kappa_3$ requires $K_2 \gg K_1$ for multistationarity to arise.

We also show that if $\kappa_3 = \kappa_6$, then the network cannot display multistationarity, no matter what the other parameters are. Therefore, any two points in each of the two sets above cannot be joined by a continuous path inside the multistationarity region, as any such path should cross at least one point where $\kappa_3 = \kappa_6$. Hence, the full multistationarity region cannot be path connected: it has *at least* two path-connected components.

On the other hand, we show also in Section 6 that the multistationarity region of the reduced network has *at most* two path-connected components using [18]. Hence, by Theorem 2.4, the original network in Fig. 2(c) has at most two path-connected components as well.

Putting the two pieces together, we conclude that the multistationarity region has precisely two connected components: one in which the catalytic rate of the phosphorylation step is larger than the catalytic rate of dephosphorylation, that is $\kappa_3 > \kappa_6$, and the other where the inequality is reversed $\kappa_6 > \kappa_3$. The second region was missed in [17] as it is outside the regime where the quasi-steady-state approximation employed there is valid.

To illustrate the type of switches that may arise from this system, we have considered the reduced model (for which there are also two connected components), and selected two parameter values α_1, α_2 in different path-connected components of the multistationarity region. The two parameter values are identical except for the three parameters governing the dephosphorylation event: κ_4, κ_6 and the total amount of P, see Fig. 6. We choose the path through the two points,

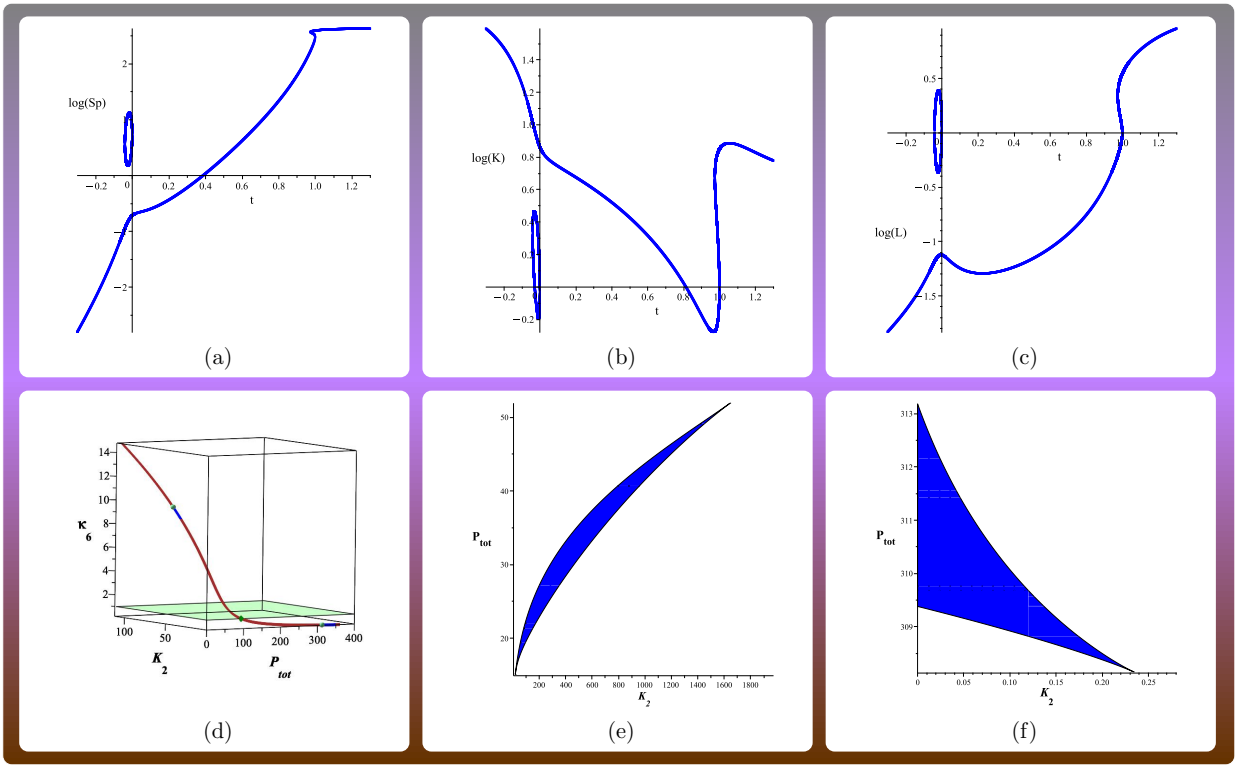


FIGURE 6. Input-output curves (bifurcation diagrams) for the reduced reciprocal allosteric regulation system in Fig. 2(c) (that is, with $\kappa_2 = \kappa_5 = 0$ for simplicity). The following parameters are fixed: $\kappa_1 = 1, \kappa_3 = 1, \kappa_7 = 1, \kappa_8 = 1, \kappa_9 = 1, \kappa_{10} = 0.1$, $\mathrm{L_{tot}} = 72$, $\mathrm{K_{tot}} = 62$, $\mathrm{S_{tot}} = 426$. In (a)-(d), the bifurcation parameter t, which can be negative, describes a path for $(\kappa_4, \kappa_6, \mathrm{P_{tot}})$: $\beta(t) = (\left(\frac{1}{5}\right)^t \left(\frac{510}{41}\right)^{1-t}, 10^t \left(\frac{51}{256}\right)^{1-t}, 17^t \left(\frac{5307}{17}\right)^{1-t})$. Subfigures (a)-(c) show the bifurcation diagrams for the concentration of $\mathrm{S_p}$, K and L at steady state. In the intervals with three steady states, the one in the middle is unstable. Subfigure (d) shows the path in the three-dimensional space $(K_2, \mathrm{P_{tot}}, \kappa_6)$. The blue regions indicate the region of the path that belongs to the multistationarity region, and the displayed plane $\kappa_6 = 1$ separates the two regions. Subfigures (e)-(f) show the multistationarity regions when we fix $\kappa_6 = 9.5$ and $\kappa_6 = 0.195$ respectively. These are two slices of the two path-connected components, obtained by keeping only two free parameters.

which component-wise is given as $\beta(t)_i = \alpha_{1,i}^t \alpha_{2,i}^{1-t}$. At $t = 0$ the path is α_1, while at $t = 1$ it is α_2. Fig. 6(a-c) shows bifurcation diagrams, where t is the perturbed parameter, which perturbs simultaneously K_2, κ_6 and the total amount of P, and we display the logarithm of the concentration of the phosphorylated substrate $\mathrm{S_p}$, kinase K and ligand L at steady state respectively. Note that by the choice of path we have made, t can take any real value, also negative. For the three concentrations, we obtain saddle-node bifurcations: a usual switch for larger values of t, while for smaller values of t, no hysteresis effect arises and the response curve has two components. In panel (d) of Fig. 6 a path in the three dimensional parameter space joining the two points is given. We see that the path enters and exists the multistationarity region twice, corresponding to the two path-connected components. The shape of these two components is displayed in Fig. 6(e-f) after slicing the three-dimensional space by fixing κ_6 for illustration purposes.

4. Discussion

Determining topological properties of semi-algebraic sets, that is, sets described by polynomial equalities and inequalities, is a highly complex problem that requires, for general sets,

computationally expensive algorithms that scale poorly with the number of variables [15]. For reaction networks with mass-action kinetics, the multistationarity region is a semi-algebraic set, and hence its description might not be straightforward.

Here we presented an approach to determine a basic topological property of a set, namely its connectivity. Non-connectivity of the multistationarity region may indicate different biological mechanisms underlying the existence of multiple steady states, and additionally, may give the cell the possibility to operate on complex switches as shown in Fig. 1. Our algorithm is to our knowledge the first to address the problem of connectivity in an effective and conclusive way. This is done by relying on linear programming and polyhedral geometry algorithms, rather than on semi-algebraic approaches, which reduces dramatically the computational cost. Additionally, our approach provides a symbolic proof of connectivity, and does not require numerical approaches, which unavoidably cannot explore the whole parameter space.

Although our algorithm might terminate inconclusively, even if the multistationarity region is path connected, we have shown that it is often applicable: for many motifs, the multistationarity region is connected because a strict separating hyperplane of the support of the critical polynomial exists. This came as a surprise to us, and might indicate a hidden feature present in realistic systems and brings up the question: What are the characteristics of the reaction networks from cell signaling that ensure that the support of the corresponding critical polynomial has a strict separating hyperplane?

It would certainly be relevant to understand the answer to this question, to bypass finding the critical polynomial, a step that is prohibitive for larger networks. This was illustrated with the networks of phosphorylation cycles with several phosphorylation sites, which revealed the computational boundaries of the algorithm. In several cases, the computation of the critical polynomial was not possible on a common computer. However, it was possible to compute the critical polynomial of the reduced network and still we were able to conclude that the multistationarity region is path connected.

Despite covering many networks, the algorithm remains inconclusive for some relevant networks, where we cannot decide whether the multistationarity region is connected, meaning that further investigations are required. For the m-site sequential and distributive systems, the projection of the multistationarity region onto the set of reaction rate constants is known to be path connected for all m [53], and this makes us believe that the same holds for the full region. In fact, based on the evidence gathered from the tested networks, we conjecture that if the projection onto the set of reaction rate constants κ is path connected, so is the multistationarity region. This would provide an additional strategy to study connectivity, which in particular might give a way to show that the fully weakly irreversible phosphorylation cycle studied in [4], see Fig. 2(h), is indeed path connected.

For the network in Fig. 2(c), where the algorithm was inconclusive, the network had two path-connected components. The strategy to show this was to combine knowledge about the projection of the multistationarity region onto the set of reaction rate constants, and a bound on the number of connected components of the reduced network found using ideas similar to those in the proof of [18, Theorem 3.9]. This example opens up for new directions for counting path-connected components and understanding underlying features of reaction networks where the multistationarity region is disconnected. On one hand, it would be interesting to devise algorithms that can assert that the multistationarity region is disconnected, and ideally, count or give bounds on the number of path-connected components. On the other hand, one might wonder what network characteristics might give rise to disconnected multistationarity regions. For the reduced network of Fig. 2(c), the multistationarity region is no longer disconnected after deleting a reaction or a species, so this network can be viewed as a minimal motif with this property. A proper investigation of minimal networks with disconnected multistationarity region would require a better understanding on how the connectivity of the multistationarity region changes upon modifications on the network, in the spirit of Theorem 2.4.

We would like to point out that the proof of the key theorem, namely Theorem 2.2, is based on relating the multistationarity region to the preimage of the negative real half-line by the critical

polynomial. When a strict separating hyperplane of the support of the critical polynomial exists, it is not only known that this preimage is path connected, but also that it is contractible [18, Theorem 3.9]. This implies that all Betti numbers of the preimage set are zero. We conjecture that when this is the case, the multistationarity region is contractible, and hence topologically very simple (for example, it has no holes). However, this cannot be directly deduced by our arguments in the proof of Theorem 2.2.

To conclude, we propose the application of our work in the design of synthetic circuits displaying predefined switches. Indeed, by combining algorithms to determine multistationarity with the study of the connectivity of the multistationarity region, one can systematically study small networks and search for a desired input-output curve shape. This approach would adhere to related work already done in this direction, e.g. [65, 66].

5. Methods

We implemented Algorithm 2.5 in `SageMath` 9.2 [33]. A Jupyter notebook containing the code can be found in the Supporting Information. The computations for the networks in Table 1 were run on a Windows 10 computer with Intel Core i5-10310U CPU @ 1.70GHz 2.21 GHz processor and 8GB RAM.

In our implementation, each species is represented by a *symbolic variable*, and each reaction is represented by a list containing two *symbolic expressions* in the variables. For example, to run the code for the running example (1), one has to type:

```
X1,X2 = var('X1,X2')
species = [X1,X2]
reactions = [[X1,X2],[X2,X1],[2*X1+X2,3*X1]]
F = CheckConnectivity(species,reactions)
```

The output is given in the following format:

```
n = 2
r = 3
The reaction network is conservative.
There are no relevant boundary steady states.
l = 2
Number of positive coefficients:  3
Number of negative coefficients:  1
The support set has a strict separating hyperplane.
All the conditions are satisfied.
We conclude that the parameter region of multistationarity
is path connected.
```

Although we used `SageMath` to compute the extreme vectors, the same computation can also be done with `Polymake` [67] (as a standalone [68] or as part of the `Oscar` project in `Julia` [69]). These programs compute extreme vectors of arbitrary pointed cones. For open cones of the form $\ker(N) \cap \mathbb{R}^n_{>0}$, there exist specific algorithms designed in the context of stoichiometric network analysis and metabolic network analysis. See for example [70].

As discussed above, the bottleneck of Algorithm 2.5 is to compute the determinant of the symbolic matrix $\tilde{M}(h, \lambda)$ in (13). In the tested examples, the `Julia` package `SymbolicCRN.jl` [64] is more efficient in computing the symbolic determinant than our implementation using `SageMath`. Using `SymbolicCRN.jl`, we were able to compute the critical polynomial for the reduced MAPK cascade with three layers, which was not possible with `SageMath`. For the 4-site phosphorylation networks in Table 1, we could not compute the critical polynomial using neither `Julia` nor `SageMath`.

The parameter regions of multistationarity displayed in Fig. 6(e-f) are found using cylindrical algebraic decomposition in `Maple`, using the command `CellDecomposition` from the packages `RootFinding[Parametric]` and `RegularChains`.

6. Proof of the results

6.1. **Convex parameters.** In this subsection we show basic results on the flux cone, and specially that positive combinations of extreme vectors of $\ker(N)\cap\mathbb{R}^r_{\geq 0}$ parametrize the positive part of the cone.

Proposition 6.1. *Let $E\in\mathbb{R}^{r\times\ell}$ be a matrix of extreme vectors for the flux cone $\ker(N)\cap\mathbb{R}^r_{\geq 0}$.*

(a) *The relative interior of $\ker(N)\cap\mathbb{R}^r_{\geq 0}$ equals $\{E\lambda \mid \lambda\in\mathbb{R}^\ell_{>0}\}$ and contains $\ker(N)\cap\mathbb{R}^r_{>0}$.*
(b) *$\ker(N)\cap\mathbb{R}^r_{>0}\neq\emptyset$ if and only if E does not have any zero row.*
(c) *If E does not have any zero row, then $\{E\lambda \mid \lambda\in\mathbb{R}^\ell_{>0}\}=\ker(N)\cap\mathbb{R}^r_{>0}$.*

Proof. (a) The first equality is proven in [71, Section XVIII, Theorem 1]. To show that $\ker(N)\cap\mathbb{R}^r_{>0}$ is included in $\{E\lambda \mid \lambda\in\mathbb{R}^\ell_{>0}\}$, let $v\in\ker(N)\cap\mathbb{R}^r_{>0}$ and $\mathbf{1}\in\mathbb{R}^\ell$ the vector with coordinates equal to 1. By [32, Corollary 18.5.2], there exists $\lambda\in\mathbb{R}^\ell_{\geq 0}$ such that $v=E\lambda$. Since all coordinates of v are positive, it is possible to choose $\epsilon>0$ such that $E\lambda-\epsilon E\mathbf{1}$ has positive coordinates. Thus, $E\big(\lambda-\epsilon\mathbf{1}\big)$ belongs to the flux cone. Using [32, Corollary 18.5.2] again, there exist $\mu\in\mathbb{R}^\ell_{\geq 0}$ such that $E\big(\lambda-\epsilon\mathbf{1}\big)=E\mu$. By reordering, we have

$$v=E\lambda=E(\mu+\epsilon\mathbf{1})$$

where $\mu+\epsilon\mathbf{1}\in\mathbb{R}^\ell_{>0}$. This shows (a).

(b) Follows easily from the equality $\ker(N)\cap\mathbb{R}^r_{\geq 0}=\{E\lambda \mid \lambda\in\mathbb{R}^\ell_{\geq 0}\}$.

(c) If E does not have a zero row, then $\{E\lambda \mid \lambda\in\mathbb{R}^\ell_{>0}\}\subseteq\ker(N)\cap\mathbb{R}^r_{>0}$. Together with (a), we obtain the equality of sets. □

Corollary 6.2. *Let $E\in\mathbb{R}^{r\times\ell}$ be a matrix of extreme vectors for the flux cone $\ker(N)\cap\mathbb{R}^r_{\geq 0}$. Recall the set $\mathcal{V}$ of steady states given in (6).*

(a) *If E does not have any zero row, then the convex parametrization map $\Psi\colon\mathbb{R}^n_{>0}\times\mathbb{R}^\ell_{>0}\to\mathcal{V}$ from (11) is surjective.*
(b) *If E has a zero row, then $\mathcal{V}=\emptyset$.*

Proof. First, we observe that for $(x,\kappa)\in\mathcal{V}$ the vector $v_\kappa(x)$ lies in $\ker(N)\cap\mathbb{R}^r_{>0}$ by (3). Now, part (b) follows directly from Proposition 6.1(b). To show that Ψ is surjective, let $(x,\kappa)\in\mathcal{V}$. By Proposition 6.1(a), there exist $\lambda\in\mathbb{R}^\ell_{>0}$ such that $v_\kappa(x)=E\lambda$. By letting $h=1/x$, using the definition of $v_\kappa(x)$ in (3) and the definition of Ψ, one easily sees that $\Psi(h,\lambda)=(x,\kappa)$, and hence, (x,κ) is in the image of Ψ, concluding the proof. □

Lemma 6.3. *Let $(\mathcal{S},\mathcal{R})$ be a reaction network such that the last two reactions R_{r-1} and R_r are reverse to each other. Let $(\tilde{\mathcal{S}},\tilde{\mathcal{R}})$ be the reduced network obtained by removing the reaction R_r.*

(a) *The vector $(0,\dots,0,1,1)\in\mathbb{R}^r_{\geq 0}$ is an extreme vector of the flux cone of $(\mathcal{S},\mathcal{R})$.*
(b) *If v is an extreme vector of the flux cone of $(\tilde{\mathcal{S}},\tilde{\mathcal{R}})$, then $(v,0)$ is an extreme vector of the flux cone of $(\mathcal{S},\mathcal{R})$.*

In particular, the flux cone of $(\mathcal{S},\mathcal{R})$ has more extreme vectors than the flux cone of $(\tilde{\mathcal{S}},\tilde{\mathcal{R}})$.

Proof. Before starting the proof, we recall some definitions and statements. The support of a vector $u\in\mathbb{R}^r$ is the set of indices where the vector is non-zero: $\operatorname{supp}(u):=\{i\in\{1,\dots,r\} \mid u_i\neq 0\}$. A vector $u\in\ker(N)\cap\mathbb{R}^r_{\geq 0}$ is said to have minimal support if for all $w\in\ker(N)\cap\mathbb{R}^r_{\geq 0}$ with $\operatorname{supp}(w)\subseteq\operatorname{supp}(u)$ we have $\operatorname{supp}(w)=\operatorname{supp}(u)$. A vector $u\in\ker(N)\cap\mathbb{R}^r_{\geq 0}$ is an extreme vector if and only if it has minimal support [72, Proposition 5], see also [73, Definition 1] and [74, Proposition 5].

(a) Since the reactions R_{r-1} and R_r are reverse to each other, for the last two columns of N it holds that $N_{r-1}=-N_r$, which implies

$$N(0,\dots,0,1,1)^\top=N_{r-1}+N_r=0.$$

Thus, $(0,\dots,0,1,1)$ belongs to the flux cone. If $(0,\dots,0,1,1)$ did not have minimal support, then $(0,\dots,0,1)$ or $(0,\dots,0,1,0)$ would be contained in the flux cone, but that would imply that $N_r = 0$ and N has a zero column, which cannot be the case.

(b) Let $\tilde{N} \in \mathbb{R}^{n\times(r-1)}$ denote the stoichiometric matrix of the reduced network and hence $N = \begin{pmatrix}\tilde{N} & N_r\end{pmatrix} \in \mathbb{R}^{n\times r}$. Since

$$N\begin{pmatrix}v\\0\end{pmatrix} = \tilde{N}v = 0,$$

we have that $(v,0)$ is contained in the flux cone of $(\mathcal{S},\mathcal{R})$. To show that $(v,0)$ is an extreme vector, we show that it has minimal support. Let $w \in \ker(N)\cap\mathbb{R}^r_{\geq 0}$ such that $\operatorname{supp}(w) \subseteq \operatorname{supp}((v,0))$. Then $w_n = 0$, and hence $(w_1\dots,w_{r-1}) \in \ker(\tilde{N})\cap\mathbb{R}^{r-1}_{\geq 0}$. Since v is an extreme vector of $\ker(\tilde{N})\cap\mathbb{R}^{r-1}_{\geq 0}$ and $\operatorname{supp}((w_1\dots,w_{r-1})) \subseteq \operatorname{supp}(v)$, it follows that $\operatorname{supp}(w) = \operatorname{supp}((w_1\dots,w_{r-1})) = \operatorname{supp}(v) = \operatorname{supp}((v,0))$. Hence $(v,0)$ has minimal support and is therefore an extreme vector. □

6.2. **Proof of Theorem 2.2.** Theorem 2.2 is a direct consequence of two technical lemmas. To state the lemmas, we use the notation of the main text, and additionally we write $\Omega \subseteq \mathbb{R}^r_{>0}\times\mathbb{R}^d$ for the parameter region of multistationarity:

$$\Omega := \big\{(\kappa,c)\in\mathbb{R}^r_{>0}\times\mathbb{R}^d \mid \#\big(V_\kappa\cap\mathcal{P}_c\cap\mathbb{R}^n_{>0}\big)\geq 2\big\}.$$

We will write Ω as the image of a subset of $\mathcal{V}$ under the map

$$\pi\colon \mathcal{V}\to\mathbb{R}^r_{>0}\times\mathbb{R}^d, \qquad (x,\kappa)\mapsto(\kappa,Wx), \tag{19}$$

and then compare path-connected components. Recall that W is a matrix of conservation relations.

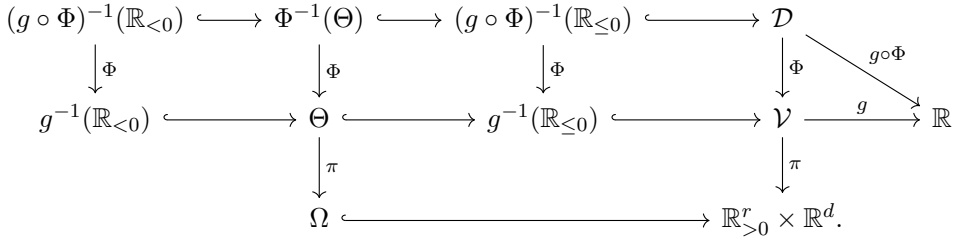
By Theorem 2.1, Ω is closely related to the preimage of the negative real half-line under the critical function given in (8) for a parametrization $\Phi\colon\mathcal{D}\to\mathcal{V}$:

$$g\circ\Phi\colon\mathcal{D}\to\mathbb{R}.$$

So we introduce the set:

$$\Theta := g^{-1}(\mathbb{R}_{\leq 0})\cap\pi^{-1}(\Omega) \quad \subseteq\mathcal{V}. \tag{20}$$

We summarize all the relevant functions that play a role in the proof of Theorem 2.2 in the following diagram:



We denote by $b_0(X)$ the number of path-connected components of a set $X\subseteq\mathbb{R}^k$ [75, Definition 3.3.7]. Note that X is path connected if and only if $b_0(X)=1$.

Lemma 6.4. *For a dissipative reaction network without relevant boundary steady states, it holds that*

$$\pi(\Theta) = \Omega.$$

In particular, $b_0(\Omega)\leq b_0(\Theta)$.

Proof. By definition of Θ, we have $\pi(\Theta)\subseteq\Omega$. To show the reverse inclusion, consider $(\kappa,c)\in\Omega$. All we need is to find a point (x^*,κ) such that $\pi(x^*,\kappa)=(\kappa,c)$ and $g(x^*,\kappa)\leq 0$. From the definition of V_κ in (5) and of $\mathcal{P}_c$ in (4), it follows that

$$\pi^{-1}(\kappa,c) = \{(x^*,\kappa)\in\mathcal{V} \mid x^*\in V_\kappa\cap\mathcal{P}_c\cap\mathbb{R}^n_{>0}\}.$$

Since (κ, c) enables multistationarity as it belongs to Ω, we have $\pi^{-1}(\kappa, c)$ has at least two elements and hence Theorem 2.1(A) cannot hold. It follows that $\pi^{-1}(\kappa, c)$ contains at least one point where (x^*, κ) with $g(x^*, \kappa) \leq 0$. As by construction $\pi(x^*, \kappa) = (\kappa, c)$, we have the inclusion.

The second part follows from the fact that continuous images of path connected sets are path connected [75, Theorem 3.3.5]. □

Lemma 6.5. *Consider a dissipative reaction network without relevant boundary steady states. If the closure of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$ equals $(g \circ \Phi)^{-1}(\mathbb{R}_{\leq 0})$, then*

$$b_0(\Theta) \leq b_0((g \circ \Phi)^{-1}(\mathbb{R}_{<0})).$$

Proof. For all $(x, \kappa) \in \mathcal{V}$ with $g(x, \kappa) < 0$, Theorem 2.1(B) gives that $\pi(x, \kappa) \in \Omega$. Thus, by definition of Θ, it holds that

$$g^{-1}(\mathbb{R}_{<0}) \subseteq \Theta \subseteq g^{-1}(\mathbb{R}_{\leq 0}).$$

By taking preimages under Φ, it follows that

$$(g \circ \Phi)^{-1}(\mathbb{R}_{<0}) \subseteq \Phi^{-1}(\Theta) \subseteq (g \circ \Phi)^{-1}(\mathbb{R}_{\leq 0}).$$

Since $(g \circ \Phi)^{-1}(\mathbb{R}_{\leq 0})$ is the closure of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$, every point $\eta \in \Phi^{-1}(\Theta)$ is contained in the closure of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$. Using the Curve Selecting Lemma [76], there exists a continuous path

$$\gamma \colon [0, 1] \to \mathbb{R}^n$$

such that $\gamma(0) = \eta$ and $\gamma\big((0, 1)\big) \subseteq (g \circ \Phi)^{-1}(\mathbb{R}_{<0})$. Thus, there exists a continuous path between η and one of the path-connected components of $(g \circ \Phi)^{-1}(\mathbb{R}_{<0})$. Therefore, $b_0(\Phi^{-1}(\Theta)) \leq b_0((g \circ \Phi)^{-1}(\mathbb{R}_{<0}))$.

Since Φ is surjective, it follows that $\Phi(\Phi^{-1}(\Theta)) = \Theta$. Since a continuous image of a path connected set is path connected [75, Theorem 3.3.5], we conclude that $b_0(\Theta) \leq b_0(\Phi^{-1}(\Theta))$. □

6.3. **Proof of Theorem 2.4.** To prove Theorem 2.4, we need to show that under the hypotheses of the theorem, we have

$$b_0(\Omega) \leq b_0((\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})).$$

The proof is based on inductive application of the following statement.

Proposition 6.6. *Let $(\mathcal{S}, \mathcal{R})$ be a conservative reaction network without relevant boundary steady states. Assume that the species X_n participates in exactly 3 reactions of the form*

$$a_1 X_1 + \cdots + a_{n-1} X_{n-1} \underset{\kappa_r}{\overset{\kappa_{r-1}}{\rightleftharpoons}} \mathrm{X}_n \xrightarrow{\kappa_{r-2}} b_1 X_1 + \cdots + b_{n-1} X_{n-1}.$$

Let $(\tilde{\mathcal{S}}, \tilde{\mathcal{R}})$ denote the reduced network obtained by removing the reaction corresponding to κ_r and $\tilde{\Theta}$ denote the set in (20) for $(\tilde{\mathcal{S}}, \tilde{\mathcal{R}})$. It holds that

$$b_0(\Theta) \leq b_0(\tilde{\Theta}).$$

Proof. For every object corresponding to a network, we write $\tilde{}$ to indicate that it corresponds to the reduced network, e.g. the reaction rate constants in the reduced network are denoted by $\tilde{\kappa}_1, \ldots, \tilde{\kappa}_{r-1}$.

As in the proof of Lemma 6.3, the stoichiometric matrices N and $\tilde{N}$ satisfy $N_{r-1} = -N_r$, $\tilde{N}_j = N_j$ for $j = 1, \ldots, r-1$, and $\mathrm{rk}(N) = \mathrm{rk}(\tilde{N})$. Recall that $W \in \mathbb{R}^{d \times n}$ is a row reduced full rank matrix such that $WN = 0$. It follows that $W\tilde{N} = 0$, so W is a matrix of conservation relations for $(\tilde{\mathcal{S}}, \tilde{\mathcal{R}})$ and we use this matrix in the definition of $\tilde{\pi}$, analogously to (19). Using (9), we also have that $(\tilde{\mathcal{S}}, \tilde{\mathcal{R}})$ is conservative if and only if $(\mathcal{S}, \mathcal{R})$ is conservative.

Following the proof of [77, Prop. 1] we introduce the maps:

$$\eta \colon \mathbb{R}^r_{>0} \to \mathbb{R}^{r-1}_{>0}, \qquad \kappa \mapsto (\kappa_1, \ldots, \kappa_{r-2}, \tfrac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}}),$$
$$\tilde{\eta} \colon \mathbb{R}^{r-1}_{>0} \to \mathbb{R}^{r-1}_{>0}, \qquad \tilde{\kappa} \mapsto (\kappa_1, \ldots, \kappa_{r-2}, \tfrac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}}).$$

For $x \in \mathbb{R}^n$, let $x^a = x_1^{a_1} \cdots x_{n-1}^{a_{n-1}}$ where $a = (a_1, \ldots, a_{n-1}, 0)$.

We start by relating the steady states of the two networks under the hypothesis that $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$. Recall that $(\kappa, c) \in \Omega$ if and only if the equation system

$$f_\kappa(x) = N v_\kappa(x) = 0, \qquad Wx = c \tag{21}$$

has at least two positive solutions. Redundant linear relations among the equations $f_\kappa(x) = 0$ arise from linear dependencies of the rows of N. To remove these redundancies, we consider $I = \{i_1, \ldots, i_d\}$ to be the set of indices of the first non-zero coordinates of each row of W, $i_1 < \cdots < i_d$. For all $j = 1, \ldots, d$, we replace the i_jth row of $f_\kappa(x)$ by the jth row of $Wx - c$. If we denote the resulting function by $h_{\kappa,c}(x)$, then $x^* \in \mathbb{R}^n_{>0}$ is a solution of (21) if and only if $h_{\kappa,c}(x^*) = 0$ holds. Thus, the parameter pair (κ, c) enables multistationarity if and only if $h_{\kappa,c}(x) = 0$ has at least two positive solutions.

Since x_n appears linearly in $h_{\kappa,c}(x)$, there exist vectors $z(\kappa), v(x,\kappa)$ such that

$$h_{\kappa,c}(x) = \begin{pmatrix} z(\kappa) \\ -(\kappa_r + \kappa_{r-2}) \end{pmatrix} x_n + \begin{pmatrix} v(x,\kappa) \\ \kappa_{r-1} x^a \end{pmatrix}. \tag{22}$$

Specifically:

$$\begin{aligned} &\text{if } i = i_j \in I: \quad z_i(\kappa) = W_{j,n}, \qquad v_i(x,\kappa) = -c_j + W_{j,1}x_1 + \cdots + W_{j,n-1}x_{n-1}, \\ &\text{if } i \in \{1, \ldots, n-1\} \setminus I: \quad z_i(\kappa) = \kappa_r a_i + \kappa_{r-2} b_i, \qquad v_i(x,\kappa) = -\kappa_{r-1} a_i x^a + u_i(x,\kappa), \end{aligned}$$

where $u(x,\kappa)$ is chosen such that (22) holds.

We define $\tilde{h}_{\tilde{\kappa},c}(x)$ analogously for the reduced network and write

$$\tilde{h}_{\tilde{\kappa},c}(x) = \begin{pmatrix} \tilde{z}(\tilde{\kappa}) \\ -\tilde{\kappa}_{r-2} \end{pmatrix} x_n + \begin{pmatrix} \tilde{v}(x,\tilde{\kappa}) \\ -\tilde{\kappa}_{r-1} x^a \end{pmatrix},$$

which under the assumption that $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$, we have

$$\begin{aligned} \text{if } i = i_j \in I: &\qquad \tilde{z}_i(\tilde{\kappa}) = z_i(\kappa), &&\qquad \tilde{v}_i(x,\tilde{\kappa}) = v_i(x,\kappa), \\ \text{if } i \in \{1, \ldots, n-1\} \setminus I: &\qquad \tilde{z}_i(\tilde{\kappa}) = \tilde{\kappa}_{r-2} b_i, &&\qquad \tilde{v}_i(x,\tilde{\kappa}) = -\tilde{\kappa}_{r-1} a_i x^a + u_i(x,\tilde{\kappa}). \end{aligned}$$

It then holds that

$$\tfrac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}} z(\kappa) x^a + v(x,\kappa) = \tfrac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}} \tilde{z}(\tilde{\kappa}) x^a + \tilde{v}(x,\tilde{\kappa}). \tag{23}$$

For $i \in I$, it is straightforward to see that this equality holds. If $i \notin I$, then the equation reduces to the equality

$$\begin{aligned} \tfrac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}}(\kappa_r a_i + \kappa_{r-2} b_i) - \kappa_{r-1} a_i &= \tfrac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}} \kappa_{r-2}(b_i - a_i) = \tfrac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}} \tilde{\kappa}_{r-2}(b_i - a_i) \\ &= \tilde{\kappa}_{r-1} b_i - \tilde{\kappa}_{r-1} a_i. \end{aligned}$$

With this in place, consider the matrices

$$B(\kappa) = \begin{pmatrix} \mathrm{Id}_{n-1} & \frac{z(\kappa)}{\kappa_r + \kappa_{r-2}} \\ 0 & -\frac{1}{\kappa_r + \kappa_{r-2}} \end{pmatrix}, \quad \text{and} \quad \tilde{B}(\tilde{\kappa}) = \begin{pmatrix} \mathrm{Id}_{n-1} & \frac{\tilde{z}(\tilde{\kappa})}{\tilde{\kappa}_{r-2}} \\ 0 & -\frac{1}{\tilde{\kappa}_{r-2}} \end{pmatrix},$$

where Id_{n-1} is the identity matrix of size $n-1$. Using (23), we have

$$\begin{aligned} B(\kappa) h_{\kappa,c}(x) = \begin{pmatrix} \mathrm{Id}_{n-1} & \frac{z(\kappa)}{\kappa_r + \kappa_{r-2}} \\ 0 & -\frac{1}{\kappa_r + \kappa_{r-2}} \end{pmatrix} \begin{pmatrix} z(\kappa) x_n + v(x,\kappa) \\ -(\kappa_r + \kappa_{r-2}) x_n + \kappa_{r-1} x^a \end{pmatrix} = \\ \begin{pmatrix} \frac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}} z(\kappa) x^a + v(x,\kappa) \\ x_n - \frac{\kappa_{r-1}}{\kappa_r + \kappa_{r-2}} x^a \end{pmatrix} = \begin{pmatrix} \frac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}} \tilde{z}(\tilde{\kappa}) x^a + \tilde{v}(x,\tilde{\kappa}) \\ x_n - \frac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}} x^a \end{pmatrix} = \tilde{B}(\tilde{\kappa}) \tilde{h}_{\tilde{\kappa},c}(x). \end{aligned} \tag{24}$$

Since the matrices $B(\kappa)$ and $\tilde{B}(\tilde{\kappa})$ are invertible, it follows from (24) that every positive solution of $h_{\kappa,c}(x) = 0$ is a solution of $h_{\tilde{\kappa},c}(x) = 0$. In particular, the reduced network does not have relevant boundary steady states.

Additionally, since $s = \tilde{s}$, and the Jacobian of $h_{\kappa,c}(x)$ is precisely the matrix $M_\kappa(x)$ (and analogous for the reduced network), taking the Jacobian and determinant of both sides of (24) yields:

$$-\tfrac{1}{\kappa_r+\kappa_{r-2}} g(x,\kappa) = -\tfrac{1}{\tilde{\kappa}_{r-2}} \tilde{g}(x,\tilde{\kappa}). \tag{25}$$

Since $\kappa_r, \kappa_{r-2}, \tilde{\kappa}_{r-2} > 0$, it follows that $g(x,\kappa)$ and $\tilde{g}(x,\tilde{\kappa})$ have the same sign if $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$.

Finally, we can easily see that for $(x^*,\kappa) \in \mathcal{V}$ and $(x^*,\tilde{\kappa}) \in \tilde{\mathcal{V}}$ with $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$ the following holds:

$$\pi(x^*,\kappa) \in \Omega \quad \text{if and only if} \quad \tilde{\pi}(x^*,\tilde{\kappa}) \in \tilde{\Omega}. \tag{26}$$

For the forward implication, if $\pi(x^*,\kappa) \in \Omega$, then by definition, the parameter pair (κ, c) with $c = Wx^*$ enables multistationarity, that is $h_{\kappa,c}(x) = 0$ has at least two positive solutions. Hence so does $h_{\tilde{\kappa},c}(x) = 0$, and $(\tilde{\kappa}, c) = \tilde{\pi}(x^*,\tilde{\kappa}) \in \tilde{\Omega}$. The reverse implication of (26) follows analogously.

With these preliminaries in place, consider the maps

$$\begin{aligned} F\colon \mathcal{V} &\to \mathbb{R}^{n-1}_{>0} \times \mathbb{R}^{r-1}_{>0}, & (x,\kappa) &\mapsto ((x_1,\dots,x_{n-1}),\eta(\kappa)),\\ \tilde{F}\colon \tilde{\mathcal{V}} &\to \mathbb{R}^{n-1}_{>0} \times \mathbb{R}^{r-1}_{>0}, & (\tilde{x},\tilde{\kappa}) &\mapsto ((\tilde{x}_1,\dots,\tilde{x}_{n-1}),\tilde{\eta}(\tilde{\kappa})). \end{aligned}$$

By considering the steady state equation of X_n, for $(x,\kappa) \in \mathcal{V}$ and $(\tilde{x},\tilde{\kappa}) \in \tilde{\mathcal{V}}$ the nth coordinate is uniquely determined by the rest of the coordinates, that is:

$$x_n = \tfrac{\kappa_{r-1}}{\kappa_r+\kappa_{r-2}}\, x^a, \qquad \tilde{x}_n = \tfrac{\tilde{\kappa}_{r-1}}{\tilde{\kappa}_{r-2}}\, \tilde{x}^a. \tag{27}$$

This implies that

$$F(x,\kappa) = \tilde{F}(\tilde{x},\tilde{\kappa}) \Leftrightarrow \begin{cases} x = \tilde{x} \\ \eta(\kappa) = \tilde{\eta}(\tilde{\kappa}). \end{cases} \tag{28}$$

Additionally, the restriction of F to a fixed κ or of $\tilde{F}$ to a fixed $\tilde{\kappa}$ are injective functions.

Now, the first step towards the proof of the theorem is to show that Θ satisfies

$$\Theta = F^{-1}\big(\tilde{F}(\tilde{\Theta})\big). \tag{29}$$

To show the inclusion $\subseteq$, let $(x,\kappa) \in \Theta \subseteq \mathcal{V}$. By the definition of Θ in (20), it holds $\pi(x,\kappa) \in \Omega$ and $g(x,\kappa) \le 0$. Since $\tilde{\eta}$ is surjective, there exists $\tilde{\kappa} \in \mathbb{R}^{r-1}_{>0}$ such that $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$. As $(x,\kappa) \in \mathcal{V}$, we have $(x,\tilde{\kappa}) \in \tilde{\mathcal{V}}$ by (24), and hence $\tilde{\pi}(x,\tilde{\kappa}) \in \tilde{\Omega}$ by (26). Now (25) implies also that $\tilde{g}(x,\tilde{\kappa}) \le 0$, thus $(x,\tilde{\kappa}) \in \tilde{\Theta}$ by definition. As $F(x,\kappa) = \tilde{F}(x,\tilde{\kappa})$, the inclusion $\subseteq$ follows.

For the reverse inclusion $\supseteq$, let $(x,\kappa) \in F^{-1}\big(\tilde{F}(\tilde{\Theta})\big)$. Then there exists $(\tilde{x},\tilde{\kappa}) \in \tilde{\Theta}$ such that $F(x,\kappa) = \tilde{F}(\tilde{x},\tilde{\kappa})$. By (28) it follows that $x = \tilde{x}$ and $\eta(\kappa) = \tilde{\eta}(\tilde{\kappa})$. We now use (26) and (25) again, to show that $(x,\kappa) \in \Theta$.

From (29) follows that $F(x,\kappa) \in \tilde{F}(\tilde{\Theta})$ for all $(x,\kappa) \in \Theta$. As a next step of the proof, we show that there exists a continuous path between any two points $(x,\kappa),(x',\kappa') \in \Theta$, if $F(x,\kappa)$ and $F(x',\kappa')$ lie in the same path-connected component of $\tilde{F}(\tilde{\Theta})$. So let

$$\gamma\colon [0,1] \to \mathbb{R}^{n-1}_{>0} \times \mathbb{R}^{r-1}_{>0}, \qquad t \mapsto (y(t),\alpha(t))$$

be a continuous path such that $\gamma(0) = F(x,\kappa)$, $\gamma(1) = F(x',\kappa')$ and $\operatorname{im}\gamma \subset \tilde{F}(\tilde{\Theta})$. We now extend γ to a path in $\mathbb{R}^n_{>0} \times \mathbb{R}^r_{>0}$ by defining

$$\begin{aligned} \Gamma\colon [0,1] &\to \mathbb{R}^n_{>0} \times \mathbb{R}^r_{>0},\\ t &\mapsto \Big(\big(y(t),\alpha_{r-1}(t)y(t)^a\big),\big(\alpha_1(t),\dots,\alpha_{r-2}(t),\alpha_{r-1}(t)(\kappa_r+\alpha_{r-2}(t)),\kappa_r\big)\Big). \end{aligned}$$

Here κ_r is fixed, and is the last component of the original parameter vector κ. Using $\gamma(0) = F(x,\kappa)$, that is, $y(0) = (x_1,\dots,x_{n-1})$ and $\alpha(0) = \eta(\kappa)$, and the recovery property of the nth coordinate (27), we have

$$\Gamma(0) = \Big((x_1,\dots,x_{n-1},\tfrac{\kappa_{r-1}}{\kappa_r+\kappa_{r-2}}x^a),(\kappa_1,\dots,\kappa_{r-2},\tfrac{\kappa_{r-1}}{\kappa_r+\kappa_{r-2}}(\kappa_r+\kappa_{r-2}),\kappa_r)\Big) = (x,\kappa).$$

Furthermore, $\operatorname{im}\Gamma \subseteq F^{-1}(\tilde{F}(\tilde{\Theta})) = \Theta$, since for all $t \in [0,1]$:

$$F(\Gamma(t)) = \Big(y(t),(\alpha_1(t),\dots,\alpha_{r-2}(t),\tfrac{\alpha_{r-1}(t)(\kappa_r+\alpha_{r-2}(t))}{\kappa_r+\alpha_{r-2}(t)})\Big) = \gamma(t) \ \in \tilde{F}(\tilde{\Theta}). \tag{30}$$

We have now connected $\gamma(0)$ to the point $\Gamma(1)$ via a continuous path in Θ. We now construct a continuous path between $\Gamma(1)$ and (x',κ') in Θ. First note that by (30),

$$F(\Gamma(1)) = \gamma(1) = F(x',\kappa').$$

Thus, both $\Gamma(1)$ and (x',κ') are contained in the set $F^{-1}(\gamma(1))$, which is

$$\Big\{\Big(x',(\kappa_1',\dots,\kappa_{r-2}',\alpha_{r-1}(1)(\beta+\kappa_{r-2}'),\beta)\Big) \mid \beta \in \mathbb{R}_{>0}\Big\}.$$

Since this set is path connected, there exists a continuous path

$$\Lambda\colon [0,1] \to F^{-1}(\gamma(1))$$

such that $\Lambda(0) = \Gamma(1)$ and $\Lambda(1) = (x',\kappa')$. This path is contained in $\Theta = F^{-1}(\tilde{F}(\tilde{\Theta}))$, since $F(\Lambda(t)) = \gamma(1) \in \tilde{F}(\tilde{\Theta})$ for all $t \in [0,1]$.

To finish the proof of the proposition, let $U_1,\dots,U_k$ be the path-connected components of Θ and let $\tilde{U}_1,\dots,\tilde{U}_{\tilde{k}}$ be the path-connected components of $\tilde{\Theta}$. Since a continuous image of a path connected set is path connected, $\tilde{F}(\tilde{U}_1),\dots,\tilde{F}(\tilde{U}_{\tilde{k}})$ are path connected. Moreover, from $\tilde{\Theta} = \cup_j \tilde{U}_j$ follows that $\tilde{F}(\tilde{\Theta}) = \cup_j \tilde{F}(\tilde{U}_j)$.

If $k > \tilde{k}$, then there must exist $j \in \{1,\dots,\tilde{k}\}$ and $i_1 \neq i_2 \in \{1,\dots,k\}$ and $(x,\kappa) \in U_{i_1}, (x',\kappa') \in U_{i_2}$ such that $F(x,\kappa), F(x',\kappa') \in \tilde{F}(\tilde{U}_j)$. By the above argument, there exist a continuous path between (x,κ) and (x',κ'). This contradicts that $i_1 \neq i_2$. Therefore, $b_0(\Theta) = k \leq \tilde{k} = b_0(\tilde{\Theta})$ as desired. □

Proof of Theorem 2.4. We conclude the proof of Theorem 2.4 using Proposition 6.6. For $i = 1,\dots k$, let $(\mathcal{S}_i,\mathcal{C}_i)$ denote the reaction network obtained by removing the reverse reactions corresponding to $j = 1,\dots,i$. Furthermore, let $\tilde{\Theta}_i$ be the set corresponding to the network $(\mathcal{S}_i,\mathcal{C}_i)$ as defined in (20).

Since the closure of $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ equals $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{\leq 0})$,

$$b_0(\tilde{\Theta}_k) \leq b_0((\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0}))$$

by Lemma 6.5. Applying Proposition 6.6 inductively, one has that

$$b_0(\Theta) \leq b_0(\tilde{\Theta}_1) \leq \cdots \leq b_0(\tilde{\Theta}_k).$$

Now, we use Lemma 6.4 to conclude that the multistationarity region Ω of the network $(\mathcal{S},\mathcal{R})$ satisfies:

$$b_0(\Omega) \leq b_0((\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})).$$

□

6.4. **Number of path-connected components of the multistationarity region for Fig. 2(c).** We aim at showing that the multistationarity region Ω of the network in Fig. 2(c) has exactly two path-connected components. We do this in two steps. First, we show that $b_0(\Omega) \geq 2$, and then that $b_0(\Omega) \leq 2$, giving the equality.

Showing that $b_0(\Omega) \geq 2$**.** We choose the order of species $\mathrm{S}, \mathrm{S_p}, \mathrm{K},\ \mathrm{P},\ \mathrm{KL},\ \mathrm{PL}, \mathrm{SKL}, \mathrm{S_pP}, \mathrm{L}$, giving the following stoichiometric matrix, and choice of matrix of conservation relations:

$$N = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 & 0 & -1 & 1 \\ -1 & 1 & 1 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 & 1 \end{bmatrix},$$

$$W = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

The network is conservative, consistent, and has no relevant boundary steady states (determined using the siphon criterion). Solving the steady states equations for x_2, x_6, x_7, x_8, x_9 gives that any positive steady state satisfies

$$x_2 = \frac{(\kappa_5+\kappa_6)\kappa_1\kappa_3 x_1 x_5}{\kappa_6(\kappa_2+\kappa_3)\kappa_4 x_4}, \quad x_6 = \frac{\kappa_8\kappa_9 x_4 x_5}{\kappa_7\kappa_{10}x_3}, \quad x_7 = \frac{\kappa_1 x_1 x_5}{\kappa_2+\kappa_3}, \quad x_8 = \frac{\kappa_1\kappa_3 x_1 x_5}{\kappa_6(\kappa_2+\kappa_3)}, \quad x_9 = \frac{\kappa_8 x_5}{\kappa_7 x_3}.$$

It is convenient to introduce the Michaelis-Menten constants of the two enzymatic processes of the network:

$$K_1 = \frac{\kappa_2+\kappa_3}{\kappa_1}, \qquad K_2 = \frac{\kappa_5+\kappa_6}{\kappa_4}.$$

With this notation, and by letting ξ be the vector of free variables, that is $\xi_1 = x_1, \xi_2 = x_3, \xi_3 = x_4, \xi_4 = x_5$, we obtain the parametrization $\Phi\colon \mathbb{R}^4_{>0} \times \mathbb{R}^{10}_{>0} \to \mathcal{V}$ given by

$$\Phi(\xi,\kappa) = \left(\xi_1, \frac{\kappa_3 K_2 \xi_1\xi_4}{\kappa_6 K_1 \xi_3}, \xi_2, \xi_3, \xi_4, \frac{\kappa_8\kappa_9\xi_3\xi_4}{\kappa_7\kappa_{10}\xi_2}, \frac{\xi_1\xi_4}{K_1}, \frac{\kappa_3\xi_1\xi_4}{\kappa_6 K_1}, \frac{\kappa_8\xi_4}{\kappa_7\xi_2}\right).$$

The critical function $g\circ\Phi$ is a quotient of polynomials with denominator $K_1\kappa_6\kappa_7\xi_2^2\xi_3$, and numerator a multiple of $\kappa_1\kappa_4$. Since these expressions are positive for all positive ξ and κ, the sign of $(g\circ\Phi)(\xi,\kappa)$ depends only on the sign of the numerator divided by $\kappa_1\kappa_4$, which we denote by $p_\kappa(\xi)$. If we view $p_\kappa(\xi)$ as a polynomial in ξ_1,ξ_2,ξ_3,ξ_4, then there are two monomials $\xi_1\xi_2^2\xi_3^2\xi_4$ and $\xi_1\xi_2^2\xi_3\xi_4^2$ with coefficients

$$\kappa_6\kappa_7\kappa_8\kappa_9 K_1(\kappa_6-\kappa_3), \qquad \text{and} \qquad \kappa_3\kappa_7\kappa_8\kappa_9 K_2(\kappa_3-\kappa_6).$$

The coefficient of the other monomials of $p_\kappa(\xi)$ are sums of products of the parameters, and are positive. Thus, the value of the critical function $(g\circ\Phi)(\xi,\kappa)$ is positive for all $\xi\in\mathbb{R}^4_{>0}$ if $\kappa_3 = \kappa_6$. Now, one can apply part A of Theorem 2.1 to conclude that (κ, c) does not enable multistationarity if $\kappa_3 = \kappa_6$.

As indicated in the main text, it is enough to show now that in the cases $\kappa_3 < \kappa_6$ and $\kappa_3 > \kappa_6$, the network can be multistationary. We show it is possible to choose $K_1, K_2, \kappa_1, \kappa_4, \kappa_7, \kappa_8, \kappa_9, \kappa_{10}$ such that the polynomial $p_\kappa(\xi)$ takes negative values for some $\xi\in\mathbb{R}^4_{>0}$. To this end, we need to employ some standard techniques relating the signs a polynomial attains and its Newton polytope, and refer the reader for example to [49, Section 2.2]. Since the negative monomials of $p_\kappa(\xi)$ are contained in a face of the Newton polytope of $p_\kappa(\xi)$, it is enough to show that $p_\kappa(\xi)$ restricted to that face takes negative values. The restricted polynomial is given by $\kappa_1\kappa_7\xi_1\xi_2$ times

$$\begin{aligned} q_\kappa(\xi) :=& \kappa_6 K_1 \xi_3^2\big(\kappa_6\kappa_7\kappa_{10}\xi_2^2 + \kappa_8\kappa_9(\kappa_6-\kappa_3)\xi_2\xi_4 + \kappa_6\kappa_8\kappa_9\xi_3\xi_4\big) \\ &+ \kappa_3 K_2\xi_4^2\big(\kappa_3\kappa_7\kappa_{10}\xi_2^2 + \kappa_8\kappa_9(\kappa_3-\kappa_6)\xi_2\xi_3 + \kappa_3\kappa_8\kappa_9\xi_3\xi_4\big). \end{aligned}$$

This is a polynomial with exactly one negative coefficient, for any choice of $\kappa_3 \neq \kappa_6$.

If $\kappa_6 - \kappa_3 < 0$, by choosing K_2 very small (or K_1 large), the polynomial multiplying K_1 determines the sign of $q_\kappa(\xi)$. By letting $\xi_2 = 1, \xi_3 = \frac{1}{\xi_4}$, the polynomial multiplying $\kappa_6 K_1 \xi_3^2$ becomes

$$\kappa_6\kappa_7\kappa_{10} + \kappa_8\kappa_9(\kappa_6 - \kappa_3)\xi_4 + \kappa_6\kappa_8\kappa_9.$$

Hence, for any $\xi_4 > 0$ large enough, this polynomial is negative, and so is $q_\kappa(\xi)$. Therefore, $p_\kappa(\xi)$ also attains negative values.

If $\kappa_3 - \kappa_6 < 0$, all we need to do is to let K_1 be small enough, or K_2 large enough and repeat the argument. We conclude that $b_0(\Omega) \geq 2$.

Showing that $b_0(\Omega) \leq 2$. We now apply Theorem 2.4 to the reduced network:

$$\mathrm{S} + \mathrm{KL} \xrightarrow{\kappa_1} \mathrm{SKL} \xrightarrow{\kappa_3} \mathrm{S_p} + \mathrm{KL}, \qquad \mathrm{K} + \mathrm{L} \underset{\kappa_8}{\overset{\kappa_7}{\rightleftharpoons}} \mathrm{KL}$$

$$\mathrm{S_p} + \mathrm{P} \xrightarrow{\kappa_2} \mathrm{S_pP} \xrightarrow{\kappa_6} \mathrm{S} + \mathrm{P}, \qquad \mathrm{P} + \mathrm{L} \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} \mathrm{PL}.$$

A matrix of extreme vectors is

$$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

The critical polynomial $\tilde{g} \circ \tilde{\Phi}$ in the variables $(\lambda_1, \lambda_2, \lambda_3, h_1, \ldots, h_9)$ has 2 negative monomials corresponding to the exponents

$$\beta_1 := (1,1,3,0,1,0,1,1,0,1,0,1), \quad \beta_2 := (1,1,3,1,0,0,1,1,0,0,1,1),$$

and it has 42 monomials with positive coefficients.

In the following, we write

$$(\tilde{g} \circ \tilde{\Phi})(y) = -c_{\beta_1} y^{\beta_1} - c_{\beta_2} y^{\beta_2} + \sum_{\alpha \in \sigma_+(\tilde{g} \circ \tilde{\Phi})} c_\alpha y^\alpha,$$

where $\sigma_+(\tilde{g} \circ \tilde{\Phi})$ denotes the set of positive monomials and c_α are all positive. For the vector $v = (-6,0,0,0,0,0,2,2,0,0,0,2)$ it holds that

$$\begin{aligned} v \cdot \beta_1 &= 0, & v \cdot \beta_2 &= 0, \\ v \cdot \alpha &\leq 0, & &\text{for all} \quad \alpha \in \sigma_+(\tilde{g} \circ \tilde{\Phi}), \end{aligned} \tag{31}$$

where for exactly two monomials $\alpha_1, \alpha_2 \in \sigma_+(\tilde{g} \circ \tilde{\Phi})$ there is equality in (31). Define the polynomial

$$h(y) := c_{\alpha_1} y^{\alpha_1} + c_{\alpha_2} y^{\alpha_2} - c_{\beta_1} y^{\beta_1} - c_{\beta_2} y^{\beta_2}.$$

Using [18, Corollary 3.13], it follows that $b_0(h^{-1}(\mathbb{R}_{<0})) \leq 2$.

The next step is to show that $b_0((\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})) \leq b_0(h^{-1}(\mathbb{R}_{<0}))$. First, we observe that $(\tilde{g} \circ \tilde{\Phi})(y) \geq h(y)$ for all $y \in \mathbb{R}^{12}_{>0}$ and hence

$$(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0}) \subseteq h^{-1}(\mathbb{R}_{<0}). \tag{32}$$

If $y, y' \in (\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ belong to the same path-connected component of $h^{-1}(\mathbb{R}_{>0})$, then we build a path between y and y' contained in $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ as follows. Since y and y' lie in the same path-connected component of $h^{-1}(\mathbb{R}_{<0})$, by (32) there exists a continuous path

$$\gamma \colon [0,1] \to h^{-1}(\mathbb{R}_{<0})$$

such that $\gamma(0) = y$, $\gamma(1) = y'$. Note that the image of γ is not necessarily contained in $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$.

For each $s \in [0,1]$, we define the function $f_s \colon \mathbb{R}_{>0} \to h^{-1}(\mathbb{R}_{<0})$ as

$$\begin{aligned} f_s(t) = (\tilde{g} \circ \tilde{\Phi})(\gamma(s) \circ t^v) &= -c_{\beta_1}\gamma(s)^{\beta_1} t^{v\cdot\beta_1} - c_{\beta_2}\gamma(s)^{\beta_2} t^{v\cdot\beta_2} + \sum_{\alpha \in \sigma_+(\tilde{g}\circ\tilde{\Phi})} c_\alpha \gamma(s)^\alpha t^{v\cdot\alpha}, \\ &= h(\gamma(s)) + p_s(t), \end{aligned}$$

where $\gamma(s) \circ t^v = (\gamma_1(s)t^{v_1}, \ldots, \gamma_{12}(s)t^{v_{12}})$, and from (31) $p_s(t)$ is a sum of generalized monomials in t with all exponents negative and all coefficients positive. Hence the leading coefficient of $f_s(t)$ is $h(\gamma(s)) < 0$, all the other coefficients of $f_s(t)$ are positive, and $f_s(t)$ has a unique positive real root $T_s > 0$. Since the roots of a polynomial depend continuously on the coefficients, T_s depends continuously on s. Therefore, $T := \max_{s\in[0,1]} T_s$ exists.

For each $s \in [0,1]$ and $t_0 > \max\{1, T\}$ it holds that $(\tilde{g} \circ \tilde{\Phi})(\gamma(s) \circ t_0^v) = f_s(t_0) < 0$. Using this observation, we define the path:

$$\Gamma \colon [0,1] \to (\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0}), \qquad s \mapsto \gamma(s) \circ t_0^v.$$

Now, we connect the points $y, \Gamma(0)$, and $y', \Gamma(1)$ using respectively the paths:

$$\begin{aligned} \gamma_1 \colon [1, t_0] &\to (\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0}) & \qquad \gamma_2 \colon [1, t_0] &\to (\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0}) \\ t &\mapsto y \circ t^v & \qquad t &\mapsto y' \circ t^v. \end{aligned}$$

Indeed, $\gamma_1(1) = y \circ 1^v = y$, $\gamma_1(t_0) = y \circ t_0^v = \gamma(0) \circ t_0^v = \Gamma(0)$, and the image of γ_1 is contained in $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$, as $(\tilde{g} \circ \tilde{\Phi})(y \circ t^v)$ is negative at $t = 1$, has negative leading coefficient, and has at most one positive real root. The analogous argument shows that γ_2 connects y' and $\Gamma(1)$ in $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$.

The concatenated path $\gamma_2^{-1}\,\Gamma\,\gamma_1$ gives a continuous path from y, y' contained in $(\tilde{g}\circ\tilde{\Phi})^{-1}(\mathbb{R}_{<0})$. We conclude that the number of path-connected components of $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ is less or equal than the number of path-connected components of $h^{-1}(\mathbb{R}_{<0})$. Hence the inequality $b_0((\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})) \leq b_0(h^{-1}(\mathbb{R}_{>0}))$ holds.

This worked for the reduced network. In order to lift the result to the original network, we need to verify the extra condition of Theorem 2.4, namely that the closure of $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})$ is $(\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{\leq 0})$. To see this, it is enough to show that the gradient $\nabla(\tilde{g} \circ \tilde{\Phi})(y)$ is different from zero for all $y \in \mathbb{R}^{12}_{>0}$. Since the sixth entry of both β_1 and β_2 are zero and the exponent $(1,1,3,1,1,1,0,0,1,0,0)$ corresponds to a positive monomial of $\tilde{g} \circ \tilde{\Phi}$, the last entry of $\nabla(\tilde{g} \circ \tilde{\Phi})(y)$ cannot be zero. Using Theorem 2.4 we conclude that:

$$b_0(\Omega) \leq b_0((\tilde{g} \circ \tilde{\Phi})^{-1}(\mathbb{R}_{<0})) \leq b_0(h^{-1}(\mathbb{R}_{<0})) \leq 2$$

as desired.

7. Supporting information

An accompanying Jupyter notebook contains the code of the algorithm, written in `SageMath` 9.2 [33].

Acknowledgments

We thank Sebastian Manecke and Oskar Henriksson for useful discussions about extreme vectors, and Carsten Wiuf for comments on the manuscript. The authors acknowledge funding from the Independent Research Fund of Denmark.

References

[1] Ozbudak EM, Thattai M, Lim HN, Shraiman BI, van Oudenaarden A. Multistability in the lactose utilization network of Escherichia coli. Nature. 2004;427:737–740.

[2] Laurent M, Kellershohn N. Multistability: a major means of differentiation and evolution in biological systems. Trends Biochem Sci. 1999;24(11):418–22.

[3] Tyson JJ, Chen KC, Novak B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol. 2003;15:221–231.

[4] Nam KM, Gyori BM, Amethyst SV, Bates DJ, Gunawardena J. Robustness and parameter geography in post-translational modification systems. Plos Comput Biol. 2020;16(5):1–50. doi:10.1371/journal.pcbi.1007573.

[5] Feinberg M. Complex Balancing in General Kinetic Systems. Arch Rational Mech Anal. 1972;49:187–194.

[6] Horn FJM, Jackson R. General mass action kinetics. Arch Rational Mech Anal. 1972;47:81–116.

[7] Joshi B, Shiu A. A survey of methods for deciding whether a reaction network is multistationary. Math Model Nat Phenom. 2015;10(5):47–67. doi:10.1051/mmnp/201510504.

[8] Feliu E, Sadeghimanesh A. Kac-Rice formulas and the number of solutions of parametrized systems of polynomial equations. Mathematics of Computation. 2022;91:2739–2769.

[9] Conradi C, Feliu E, Mincheva M, Wiuf C. Identifying parameter regions for multistationarity. Plos Comput Biol. 2017;13(10):1–25.

[10] Bihan F, Dickenstein A, Giaroli M. Sign conditions for the existence of at least one positive solution of a sparse polynomial system. Adv Math. 2020;375:107412. doi:10.1016/j.aim.2020.107412.

[11] Conradi C, Iosif A, Kahle T. Multistationarity in the space of total concentrations for systems that admit a monomial parametrization. Bull Math Biol. 2019;81(10):4174–4209. doi:10.1007/s11538-019-00639-4.

[12] Otero-Muras I, Banga JR, Alonso AA. Characterizing Multistationarity Regimes in Biochemical Reaction Networks. PLoS ONE. 2012;7(7):e39194. doi:10.1371/journal.pone.0039194.t003.

[13] Bradford R, Davenport JH, England M, Errami H, Gerdt V, Grigoriev D, et al. Identifying the parametric occurrence of multiple steady states for some biological networks. J Symb Comput. 2020;98:84–119. doi:https://doi.org/10.1016/j.jsc.2019.07.008.

[14] Conradi C, Mincheva M. Catalytic constants enable the emergence of bistability in dual phosphorylation. J R S Interface. 2014;11(95). doi:10.1098/rsif.2014.0158.

[15] Basu S, Pollack R, Roy MF. Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics). Springer-Verlag; 2006.

[16] Benedetti R, Loeser F, Risler JJ. Bounding the number of connected components of a real algebraic set. Discrete Comput Geom. 1991;6(3):191–209. doi:10.1007/BF02574685.

[17] Straube R, Conradi C. Reciprocal enzyme regulation as a source of bistability in covalent modification cycles. J Theor Biol. 2013;330:56–74.

[18] Feliu E, Telek ML. On generalizing Descartes' rule of signs to hypersurfaces. Adv Math. 2022;408(A).

[19] Tang X, Xu H. Multistability of Small Reaction Networks. SIAM J Appl Dyn Syst. 2021;20(2):608–635. doi:10.1137/20M1358761.

[20] Vol'pert AI. Differential equations on graphs. Math USSR Sb. 1972;17(4):571–582.

[21] Ben-Israel A. Notes on linear inequalities, I: The intersection of the nonnegative orthant with complementary orthogonal subspaces. J Math Anal Appl. 1964;9(2):303–314.

[22] Dickenstein A, Millán MP, Shiu A, Tang X. Multistationarity in Structured Reaction Networks. Bull Math Biol. 2019;81:1527–1581.

[23] Angeli D, Sontag E. A Petri Net approach to the study of persistence in chemical reaction networks. Math Biosci. 2008;210:598–618. doi:10.1016/j.mbs.2007.07.003.

[24] Banaji M. Counting chemical reaction networks with NAUTY. arXiv 2017;1705.10820. doi:10.48550/arXiv.1705.10820.
[25] Marcondes de Freitas M, Feliu E, Wiuf C. Intermediates, catalysts, persistence, and boundary steady states. J Math Biol. 2017;74:887–932.
[26] Shiu A, Sturmfels B. Siphons in chemical reaction networks. Bull Math Biol. 2010;72(6):1448–1463.
[27] Clarke BL. In: Stability of Complex Reaction Networks. John Wiley and Sons, Ltd; 1980. p. 1–215.
[28] Conradi C, Feliu E, Mincheva M. On the existence of Hopf bifurcations in the sequential and distributive double phosphorylation cycle. Math Biosci Eng. 2020;17(1):494–513.
[29] Errami H, Eiswirth M, Grigoriev D, Seiler W, Sturm T, Weber A. Detection of Hopf Bifurcations in Chemical Reaction Networks Using Convex Coordinates. J Comput Phys. 2015;291:279–302.
[30] Domijan M, Kirkilionis M. Bistability and oscillations in chemical reaction networks. J Math Biol. 2009;59:467–501. doi:10.1007/s00285-008-0234-7.
[31] Errami H, Eiswirth M, Grigoriev D, Seiler WM, Sturm T, Weber A. Detection of Hopf bifurcations in chemical reaction networks using convex coordinates. J Comput Phys. 2015;291:279 – 302. doi:10.1016/j.jcp.2015.02.050.
[32] Rockafellar RT. Convex analysis. Princeton University Press; 1972.
[33] The Sage Developers. `SageMath`, the Sage Mathematics Software System (Version 9.2); 2021.
[34] Maplesoft, a division of Waterloo Maple Inc. `Maple`; 2021. https://www.maplesoft.com.
[35] Pérez Millán M, Dickenstein A, Shiu A, Conradi C. Chemical Reaction Systems with Toric Steady States. Bull Math Biol. 2012;74:1027–1065.
[36] Thomson M, Gunawardena J. The rational parameterisation theorem for multisite post-translational modification systems. J Theor Biol. 2009;261(4):626–636.
[37] Basu S. Algorithms in Real Algebraic Geometry: A Survey. Panor Synthèses. 2017;51:107–153.
[38] Feliu E, Wiuf C. Enzyme-sharing as a cause of multi-stationarity in signalling systems. J Roy Soc Interface. 2012;9(71):1224–1232.
[39] Slepchenko BM, Terasaki M. Cyclin aggregation and robustness of bio-switching. Mol Biol Cell. 2003;14(11):4695–4706.
[40] Tyson JJ, Chen K, Novak B. Network dynamics and cell physiology. Nat Rev Mol Cell Biol. 2001;2(12):908–916.
[41] Kothamachu VB, Feliu E, Cardelli L, Soyer OS. Unlimited multistability and Boolean logic in microbial signalling. J Roy Soc Interface. 2015;12(108):215–234.
[42] Salazar C, Hofer T. Multisite protein phosphorylation–from molecular mechanisms to kinetic models. FEBS J. 2009;276(12):3177–3198.
[43] Thomson M, Gunawardena J. Unlimited multistability in multisite phosphorylation systems. Nature. 2009;406:274–277.
[44] Wang L, Sontag ED. On the number of steady states in a multiple futile cycle. J Math Biol. 2008;57:29–52.
[45] Giaroli M, Rischter R, Pérez Millán M, Dickenstein A. Parameter regions that give rise to 2[n/2] +1 positive steady states in the n-site phosphorylation system. Math Biosci Eng. 2019;16(6):7589–7615.
[46] Conradi C, Flockerzi D, Raisch J. Multistationarity in the activation of a MAPK: parametrizing the relevant region in parameter space. Math Biosc. 2008;211(1):105–131.
[47] Feliu E, Rendall AD, Wiuf C. A proof of unlimited multistability for phosphorylation cycles. Nonlinearity. 2020;33(11):5629–5658. doi:10.1088/1361-6544/ab9a1e.
[48] Flockerzi D, Holstein K, Conradi C. N-site Phosphorylation Systems with 2N-1 Steady States. Bull Math Biol. 2014;76(8):1892–1916. doi:10.1007/s11538-014-9984-0.
[49] Feliu E, Kaihnsa N, de Wolff T, Yürück O. The Kinetic Space of Multistationarity in Dual Phosphorylation. J Dyn Differ Equ. 2022;34:825–852.

[50] Gunawardena J. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. Biophys J. 2007;93:3828–3834.
[51] Gunawardena J. Multisite protein phosphorylation makes a good threshold but can be a poor switch. Proc Natl Acad Sci USA. 2005;102:14617–14622.
[52] Bihan F, Dickenstein A, Giaroli M. Sign conditions for the existence of at least one positive solution of a sparse polynomial system. Adv Math. 2020;375:107412. doi:10.1016/j.aim.2020.107412.
[53] Feliu E, Kaihnsa N, de Wolff T, Yürük O. Parameter region for multistationarity in $n-$site phosphorylation networks. SIAM J Appl Dyn Syst. 2023;To appear.
[54] Cohen P. The regulation of protein function by multisite phosphorylation–a 25 year update. Trends Biochem Sci. 2000;25(12):596–601.
[55] Xiao L, Gong LL, Yuan D, Deng M, Zeng XM, Chen LL, et al. Protein phosphatase-1 regulates Akt1 signal transduction pathway to control gene expression, cell survival and differentiation. Cell Death Differ. 2010;17(9):1448–1462.
[56] Markevich NI, Hoek JB, Kholodenko BN. Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades. J Cell Biol. 2004;164:353–359.
[57] Zhao Y, Zhang ZY. The mechanism of dephosphorylation of extracellular signal-regulated kinase 2 by mitogen-activated protein kinase phosphatase 3. J Biol Chem. 2001;276(34):32382–32391.
[58] Shaul YD, Seger R. The MEK/ERK cascade: From signaling specificity to diverse functions. BBA-Mol Cell Res. 2007;1773(8):1213–1226. doi:10.1016/j.bbamcr.2006.10.005.
[59] Futran AS, Link AJ, Seger R, Shvartsman SY. ERK as a model for systems biology of enzyme kinetics in cells. Curr Biol. 2013;23(21):R972–R979.
[60] Rubinstein BY, Mattingly HH, Berezhkovskii AM, Shvartsman SY. Long-term dynamics of multisite phosphorylation. Mol Biol Cell. 2016;27(14):2331–2340.
[61] Obatake N, Shiu A, Tang X, A T. Oscillations and bistability in a model of ERK regulation. J Math Biol. 2019;79:1515–1549.
[62] Conradi C, Obatake N, Shiu A, Tang X. Dynamics of ERK regulation in the processive limit. J Math Biol. 2021;82(32). doi:10.1007/s00285-021-01574-6.
[63] Huang CY, Ferrell JE. Ultrasensitivity in the mitogen-activated protein kinase cascade. Proc Natl Acad Sci USA. 1996;93:10078–10083.
[64] Brustenga L. Package "SymbolicCRN.jl" for reaction networks in `Julia`; 2021. Available from: `https://github.com/LauraBMo/SymbolicCRN.jl`.
[65] Leon M, Woods ML, Fedorec AJ, Barnes CP. A computational method for the investigation of multistable systems and its application to genetic switches. BMC Syst Biol. 2016;10(1):130.
[66] Otero-Muras I, Banga JR. Multicriteria global optimization for biocircuit design. BMC Syst Biol. 2014;8(1):113. doi:10.1186/s12918-014-0113-3.
[67] Gawrilow E, Joswig M. `Polymake`: a framework for analyzing convex polytopes. In: Polytopes—combinatorics and computation (Oberwolfach, 1997). vol. 29 of DMV Sem. Birkhäuser, Basel; 2000. p. 43–73.
[68] Polymake; Available from: `https://polymake.org`.
[69] The Oscar Computer Algebra System for `Julia`; Available from: `https://oscar.computeralgebra.de/`.
[70] Klapper I, Szyld DB, Zhao K. Metabolic Networks, Elementary Flux Modes, and Polyhedral Cones. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2021. Available from: `https://epubs.siam.org/doi/abs/10.1137/1.9781611976533`.
[71] Gerstenhaber M. Theory of convex polyhedral cones. In: Koopmans TC, editor. Activity analysis of production and allocation. Wiley; 1951. p. 298–317.
[72] Müller S, Regensburger G. Elementary Vectors and Conformal Sums in Polyhedral Geometry and their Relevance for Metabolic Pathway Analysis. Front Genet. 2016;7:90.
[73] Gagneur J, Klamt S. Computation of elementary modes: A unifying framework and the new binary approach. BMC bioinform. 2004;5:175. doi:10.1186/1471-2105-5-175.

[74] Fukuda K, Prodon A. Double description method revisited. In: Deza M, Euler R, Manoussakis I, editors. Combinatorics and Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg; 1996. p. 91–111.
[75] Singh TB. Introduction to Topology. Springer Singapore; 2019.
[76] Łojasiewicz S. On semi-analytic and subanalytic geometry. Banach Cent Publ. 1995;34(1):89–104.
[77] Feliu E, Wiuf C. Simplifying biochemical models with intermediate species. J R Soc Interface. 2013;10(87):20130484. doi:10.1098/rsif.2013.0484.

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: mlt@math.ku.dk

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: efeliu@math.ku.dk

II

Connectivity of Parameter Regions of Multistationarity for Multisite Phosphorylation Networks

Nidhi Kaihnsa
Department of Mathematical Sciences
University of Copenhagen

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Publication details

Submitted (2024)
Available on arXiv: https://doi.org/10.48550/arXiv.2403.16556

CONNECTIVITY OF PARAMETER REGIONS OF MULTISTATIONARITY FOR MULTISITE PHOSPHORYLATION NETWORKS

NIDHI KAIHNSA AND MÁTÉ L. TELEK

Abstract. The parameter region of multistationarity of a reaction network contains all the parameters for which the associated dynamical system exhibits multiple steady states. Describing this region is challenging and remains an active area of research. In this paper, we concentrate on two biologically relevant families of reaction networks that model multisite phosphorylation and dephosphorylation of a substrate at n sites. For small values of n, it had previously been shown that the parameter region of multistationarity is connected. Here, we extend these results and provide a proof that applies to all values of n. Our techniques are based on the study of the critical polynomial associated with these reaction networks together with polyhedral geometric conditions of the signed support of this polynomial.

Keywords: phosphorylation networks, connectivity, Newton polytope, Gale duality, signed support
2020 MSC: 92xx, 52Bxx

1. Introduction

Within the framework of reaction network theory [7], the change in concentration of the species over time is modeled by a parametrized ordinary differential equation (ODE) system. In this paper, we focus on a fundamental property of ODE systems: the existence of multiple steady states, also known as *multistationarity*. Having multistationarity is a precursor to *multistability*, which has been linked to cellular decision-making, switching, and the memory of cells [17, 22].

Under the assumption of mass-action kinetics, the functions in the ODE system become polynomials, which are parametrized by two types of parameters: reaction rate constants and total concentrations. Identifying the *parameter region of multistationarity* is equivalent to describing the set of parameters for which the polynomial equation system has at least two positive real solutions. While existing symbolic methods, such as Cylindrical Algebraic Decomposition and Quantifier Elimination, may offer a complete description of this region, their high algorithmic complexity limits their applicability to reaction networks of moderate size [2]. On the other hand, numerical methods are able to handle larger reaction networks and offer insights into specific parts of the parameter space, but they do not provide information about the entire parameter space [10, 14, 21].

Even though an exact description of the multistationarity region is usually out of reach, having some partial information about its properties, such as its shape, could have significant biological implications. For instance, connectivity of the parameter region of multistationarity has been associated with robustness, and the lack of connectivity has been suggested to indicate that multistationarity arises for different biological reasons [21]. Further research into the connectivity of the multistationarity region has been carried out in [5, 8, 9, 26]. In this paper, we continue this line of research and investigate the parameter region of multistationarity for two infinite families of reaction networks modeling phosphorylation mechanisms.

Phosphorylation networks play a crucial role in cell signaling processes [13, 15]. These mechanisms typically involve a substrate denoted by S, along with a kinase K and a phosphatase F, responsible for catalyzing the phosphorylation and dephosphorylation of S, respectively. Assuming that both phosphorylation and dephosphorylation occur in a sequential and distributed manner, and the substrate S has $n \in \mathbb{N}$ phosphorylation sites, the corresponding reaction network takes the

following form:

$$(\mathcal{N}_{1,n}) \qquad \begin{aligned} S_i + K &\underset{\kappa_{6i+2}}{\overset{\kappa_{6i+1}}{\rightleftarrows}} Y_{i+1} \xrightarrow{\kappa_{6i+3}} S_{i+1} + K \\ S_{i+1} + F &\underset{\kappa_{6i+5}}{\overset{\kappa_{6i+4}}{\rightleftarrows}} U_{i+1} \xrightarrow{\kappa_{6i+6}} S_i + F, \end{aligned} \qquad i = 0, \dots, n-1.$$

We refer to network $(\mathcal{N}_{1,n})$ as the *strongly irreversible n-site phosphorylation network*.

In [8], it was shown that the projection of the parameter region of multistationarity onto the reaction rate constants is connected for $n = 2$. This result was later generalized for all $n \geq 2$ in [9]. Subsequently, it was shown with the aid of a computer that the entire multistationarity region is connected for $n = 2, 3$ in [26], and for $n = 4, 5, 6, 7$ in [25]. One of the main contribution of the current article is a proof showing that for $(\mathcal{N}_{1,n})$, the parameter region of multistationarity is connected for every $n \geq 2$.

The second family of phosphorylation networks considered in this paper is the *weakly irreversible n-site phosphorylation network*, which is

$$(\mathcal{N}_{2,n}) \qquad \begin{aligned} S_i + K &\underset{\kappa_{10i+2}}{\overset{\kappa_{10i+1}}{\rightleftarrows}} Y_{2i+1} \xrightarrow{\kappa_{10i+3}} Y_{2i+2} \underset{\kappa_{10i+5}}{\overset{\kappa_{10i+4}}{\rightleftarrows}} S_{i+1} + K \\ S_{i+1} + F &\underset{\kappa_{10i+7}}{\overset{\kappa_{10i+6}}{\rightleftarrows}} U_{2i+1} \xrightarrow{\kappa_{10i+8}} U_{2i+2} \underset{\kappa_{10i+10}}{\overset{\kappa_{10i+9}}{\rightleftarrows}} S_i + F, \end{aligned} \qquad i = 0, \dots, n-1.$$

The name of these networks originates from the fact that once a phosphate group is attached, the product $S_{i+1} + K$ might rebind to form the intermediate species Y_{2i+2}. Analogously, for dephosphorylation, the product $S_i + F$ might rebind to U_{2i+2}. In [21], it has been argued that allowing this product rebinding for phosphorylation systems is more realistic compared to the mechanism represented in $(\mathcal{N}_{1,n})$. For $(\mathcal{N}_{2,n})$ with $n = 2$, the authors in [21] investigated the shape of the parameter region of multistationarity numerically, and their methods strongly indicated that this region is connected. In [25], a symbolic proof was provided showing that the multistatinarity region of $(\mathcal{N}_{2,2})$ is indeed connected.

In this paper we prove that for both families $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$ the parameter region of multistationarity is connected for every $n \geq 2$. Our approach builds upon [4], where the authors associated a polynomial, called the *critical polynomial*, with reaction networks satisfying some mild assumptions. Subsequently, in [26, Theorem 4], the authors gave conditions on the critical polynomial that imply the connectivity of the parameter region of multistationarity.

One of the main challenges faced is that a direct computation of the critical polynomial for the networks $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$ becomes infeasible even for relatively small n. In [26], it was shown that it is enough to compute the critical polynomial of a reduced version of the reaction network. In [25], this reduced critical polynomial was successfully computed for $(\mathcal{N}_{1,n})$ when $n = 2, 3, 4, 5, 6, 7$ and for $(\mathcal{N}_{2,n})$ when $n = 2$. Even for the reduced critical polynomial, the computation becomes intractable for larger values of n. To overcome this difficulty, in Section 2.2, we derive a formula for the critical polynomial (Theorem 2.7) using Gale dual matrices that allows us to work with the critical polynomial for all n. We believe this result is interesting in its own right, as it might simplify the computation of the critical polynomial for general networks.

In Section 2.3, we recall several results on the structure of the Newton polytope and the signed support of a polynomial that we later employ to study the critical polynomial. We explain our approach to prove connectivity of the parameter region of multistationarity in Section 2.4. In Sections 3 and 4 we prove connectivity for $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$ respectively for every $n \geq 2$.

2. **Preliminaries**

In this section we introduce basic definitions and results regarding reaction networks, the critical polynomial, and the signed support of polynomials. These will be used later on.

2.1. **Reaction networks and multistationarity.** A *reaction network* over species $X_1, \dots, X_m$ is a collection of *reactions* of the form $\sum_{i=1}^m a_{ij} X_i \to \sum_{i=1}^m b_{ij} X_i$ for $j = 1, \dots, r$, where a_{ij}, b_{ij} are non-negative integers. Each reaction is weighted by a positive parameter $\kappa_j \in \mathbb{R}_{>0}$, called *reaction rate constant*. By $\kappa := (\kappa_1, \dots, \kappa_r) \in \mathbb{R}^r_{>0}$, we denote the reaction rate vector.

The net production of the species along each reaction is encoded in the *stoichiometric matrix* $N \in \mathbb{Z}^{m \times r}$, which is defined as

$$N := \big(b_{ij} - a_{ij}\big) \in \mathbb{Z}^{m \times r}.$$

We consider also the *reactant matrix*, given by

$$A := \big(a_{ij}\big) \in \mathbb{Z}^{m \times r}.$$

Using these matrices and under the assumption of *mass-action kinetics*, the ODE system that models the evolution of the concentration of the species is

$$\dot{x} = N \operatorname{diag}(\kappa) x^A, \tag{1}$$

where $x = (x_1, \dots, x_m) \in \mathbb{R}^m_{\geq 0}$ is a vector representing the concentration of the species $X_1, \dots, X_m$, and $x^A := \big(\prod_{i=1}^m x_i^{a_{i1}}, \dots, \prod_{i=1}^m x_i^{a_{ir}}\big)^\top \in \mathbb{R}^r_{\geq 0}$.

For the ODE system in (1), the trajectories are contained in affine linear subspaces of $\mathbb{R}^m$ called *stoichiometric compatibility classes*. These are affine translates of the image of the stoichiometric matrix N. The dimension of this subspace will be denoted by $s = \operatorname{rk}(N)$. We represent $\operatorname{im}(N)$ by linear equations determined by a full-rank matrix $W \in \mathbb{R}^{d \times m}$ such that $WN = 0$ and $d = m - s$. Such a matrix is called a *conservation law matrix*. Using W, we define the stoichiometric compatibility classes as

$$\mathcal{P}_T := \{x \in \mathbb{R}^m_{\geq 0} \mid Wx = T\},$$

where $T \in \mathbb{R}^d$ is called the *total concentration vector*. A reaction network is *conservative*, if every stoichiometric compatibility class is a compact set. Equivalently, there exists a vector with only positive coordinates in the left kernel of N (cf. [1]).

Given a reaction network, the set of *steady states* is obtained by the common zeros of the polynomials on the right-hand side of (1). As we are only interested in the non-negative steady states, for fixed reaction rate constants κ, we consider the *steady state variety*

$$V_\kappa := \{x \in \mathbb{R}^m_{\geq 0} \mid N \operatorname{diag}(\kappa) x^A = 0\}. \tag{2}$$

A steady state $x \in V_\kappa$ is a *relevant boundary steady state* if one of the coordinates of x equals zero and the stoichiometric compatibility class containing x intersects the positive orthant $\mathbb{R}^m_{>0}$.

A parameter pair $(\kappa, T) \in \mathbb{R}^r_{>0} \times \mathbb{R}^d$ *enables multistationarity* if $V_\kappa \cap \mathcal{P}_T \cap \mathbb{R}^m_{>0}$ contains at least two points. The *parameter region of multistationarity* is thus defined as

$$\Omega := \{(\kappa, T) \in \mathbb{R}^r_{>0} \times \mathbb{R}^d \mid (\kappa, T) \text{ enables multistationarity}\}. \tag{3}$$

In this article, our main goal is to show that the set Ω is path connected for the two families of networks $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$.

2.2. **Critical polynomial.** In [26], the authors described a sufficient condition based on a polynomial that implies connectivity of the parameter region of multistationarity. In this section, we recall this statement (Theorem 2.2) and elaborate on how to compute this polynomial.

Given a reaction network with stoichiometric matrix N, the set $\mathcal{C} := \ker(N) \cap \mathbb{R}^r_{\geq 0}$ is a convex polyhedral cone, called the *flux cone* [3]. A minimal collection of generators $\{E^{(1)}, \dots, E^{(\ell)}\} \subseteq \mathbb{R}^r$ of $\mathcal{C}$ is called a choice of *extreme vectors*. The extreme vectors of $\mathcal{C}$ are unique up to multiplication by positive scalars [24]. A matrix $E \in \mathbb{R}^{r \times \ell}$, whose columns are given by a choice of extreme vectors, is called an *extreme matrix*. If E does not have a zero row, the reaction network is called *consistent*. This property is equivalent to $\ker(N) \cap \mathbb{R}^r_{>0} \neq \emptyset$.

In Proposition 2.1, we establish a condition that ensures that a basis of $\ker(N)$ gives a choice of extreme vectors of $\mathcal{C}$. We denote by $[r]$ the set $\{1,\dots,r\}$, and for a vector $u \in \mathbb{R}^r$, we write

$$\operatorname{supp}(u) := \big\{ i \in [r] \mid u_i \neq 0 \big\}.$$

A vector $v \in \mathcal{C}\setminus\{0\}$ is an extreme vector if and only if v is *support-minimal* [19], i.e. for all non-zero $w \in \mathcal{C}$ it holds

$$\operatorname{supp}(w) \subseteq \operatorname{supp}(v) \quad \text{implies} \quad \operatorname{supp}(w) = \operatorname{supp}(v). \tag{4}$$

Using this we now prove the following result.

Proposition 2.1. *Consider a reaction network with stoichiometric matrix $N \in \mathbb{Z}^{m\times r}$. Assume that $E^{(1)},\dots,E^{(\ell)} \in \mathbb{R}^r_{\geq 0}$ is a basis of $\ker(N)$, and for every $k \in [\ell]$ there exists $j_k \in [r]$ such that*

$$j_k \in \operatorname{supp}(E^{(k)}) \setminus \Big(\bigcup_{\substack{i=1\\ i\neq k}}^{\ell} \operatorname{supp}(E^{(i)}) \Big). \tag{5}$$

Then $\{E^{(1)},\dots,E^{(\ell)}\}$ is a choice of extreme vectors of the flux cone $\mathcal{C}$.

Proof. By (4), it is enough to show that $E^{(1)},\dots,E^{(\ell)}$ are the only support-minimal vectors in $\mathcal{C}$ up to multiplication by a scalar. Consider $w \in \mathcal{C} \subset \ker(N)$ with $w \neq 0$. By assumption, there exist $a_1,\dots,a_\ell \in \mathbb{R}$ such that $w = \sum_{k=1}^{\ell} a_k E^{(k)}$. As j_k satisfies (5), $w_{j_k} = a_k E^{(k)}_{j_k}$ for $k \in [\ell]$. Since $E^{(k)}_{j_k} > 0$, we conclude that $a_1,\dots,a_\ell \geq 0$ and hence,

$$\operatorname{supp}(E^{(k)}) \subseteq \operatorname{supp}(w) \quad \text{if } a_k \neq 0. \tag{6}$$

If $\operatorname{supp}(w) \subseteq \operatorname{supp}(E^{(k)})$ for some k, then $\operatorname{supp}(w) = \operatorname{supp}(E^{(k)})$. Consequently, $E^{(k)}$ is support-minimal, and hence $E^{(k)}$ is an extreme vector for all $k \in [\ell]$.

If w is an extreme vector of $\mathcal{C}$, then there exists $k \in [\ell]$ with $a_k \neq 0$. From (6) and the support-minimality of w, it follows that $\operatorname{supp}(E^{(k)}) = \operatorname{supp}(w)$. If there exist two distinct $k_1, k_2 \in [\ell]$ such that a_{k_1}, a_{k_2} are non-zero, then $\operatorname{supp}(E^{(k_1)}) = \operatorname{supp}(w) = \operatorname{supp}(E^{(k_2)})$, which contradicts (5). Thus, there exists exactly one non-zero a_k and $w = a_k E^{(k)}$. □

Consider a reaction network with stoichiometric matrix $N \in \mathbb{Z}^{m\times r}$, reactant matrix $A \in \mathbb{Z}^{m\times r}$, a conservation law matrix $W \in \mathbb{R}^{d\times m}$ and an extreme matrix $E \in \mathbb{R}^{r\times \ell}$. We assume that W is row reduced and that $1,\dots,d$ are the indices of the first non-zero entries of the rows of W. In Sections 3 and 4, we will choose the conservation law matrix for various networks such that this extra assumption is satisfied (cf. (18), (29)). For $h = (h_1,\dots,h_m) \in \mathbb{R}^m_{>0}$, $\lambda = (\lambda_1,\dots,\lambda_\ell) \in \mathbb{R}^\ell_{>0}$, we define

$$M(h,\lambda) := \begin{pmatrix} W \\ N' \operatorname{diag}(E\lambda) A^\top \operatorname{diag}(h) \end{pmatrix} \in \mathbb{R}^{m\times m}, \tag{7}$$

where $N' \in \mathbb{R}^{s\times r}$ is the matrix obtained from N by deleting the first d rows. Following [6, 26], we consider the following function

$$q\colon \mathbb{R}^m_{>0} \times \mathbb{R}^\ell_{>0} \to \mathbb{R}, \quad (h,\lambda) \mapsto q(h,\lambda) := (-1)^s \det M(h,\lambda), \tag{8}$$

and call $q(h,\lambda)$ the *critical polynomial*. Theorem 2.2 exploits this polynomial to establish path connectivity of the parameter region of multistationarity.

Theorem 2.2. *[26, Theorem 4] Consider a conservative consistent reaction network without relevant boundary steady states. Assume that there exist species $\mathrm{X}_1,\dots,\mathrm{X}_k$ such that each X_j participates in exactly 3 reactions of the form*

$$\sum_{i=k+1}^{m} a_{i,j} X_i \underset{\kappa_{3j}}{\overset{\kappa_{3j-1}}{\rightleftharpoons}} \mathrm{X}_j \xrightarrow{\kappa_{3j-2}} \sum_{i=k+1}^{m} b_{i,j} X_i, \quad j = 1,\dots,k.$$

Let q be the critical polynomial of the reduced network obtained by removing the reactions corresponding to κ_{3j} for $j = 1, \dots, k$. If

(P1) $q^{-1}(\mathbb{R}_{<0})$ is path connected, and
(P2) the Euclidean closure of $q^{-1}(\mathbb{R}_{<0})$ equals $q^{-1}(\mathbb{R}_{\leq 0})$,

then the parameter regions of multistationarity of both the reduced and the original network are path connected.

To simplify the computation of the critical polynomial, we derive a formula using *Gale dual matrices* (Theorem 2.7).

Definition 2.3 (Gale dual matrix)**.** Let K be a field and $V \in K^{s\times m}$, $U \in K^{m\times d}$ be two matrices. The matrix U is *Gale dual* of V if $\operatorname{im}(U) = \ker(V)$ and $\ker(U) = \{0\}$.

To compute a Gale dual matrix of V, one simply needs to find a basis of $\ker(V)$ and write these vectors as the columns of U. Once a Gale dual matrix is obtained, the maximal minors of V can be computed by determining the maximal minors of U using Lemma 2.4 below.

For a set $I \subseteq [k]$, we denote its complement by $I^c := [k] \setminus I$ and its cardinality by $|I|$. If $I = \{i_1, \dots, i_p\} \subseteq [k]$ and $I^c = \{j_1, \dots, j_{k-p}\}$ with $i_1 < \cdots < i_p$ and $j_1 < \cdots < j_{k-p}$, we define $\operatorname{sgn}(\tau_I) \in \{\pm 1\}$ to be the sign of the permutation τ_I that sends $(1, \dots, m)$ to $(i_1, \dots, i_p, j_1, \dots, j_{k-p})$. Additionally, consider any general matrix $Y \in K^{\ell_1\times\ell_2}$ and some set of indices $L_1 \subseteq [\ell_1]$ and $L_2 \subseteq [\ell_2]$. By $[Y]_{L_1,L_2}$ we denote the sub-matrix of Y given by the rows and the columns indexed by the elements of L_1 and L_2 respectively.

Lemma 2.4. *[20, Lemma 2.10] [16, Theorem 12.16] Let K be a field, $V \in K^{s\times m}$ a matrix of rank $s < m$ and let $U \in K^{m\times d}$ be a Gale dual matrix of V. There exists $\delta \in K \setminus \{0\}$ such that for all $I \subseteq [m]$ with $|I| = s$ it holds:*

$$\delta \det\left([U]_{I^c,[d]}\right) = \operatorname{sgn}(\tau_I) \det\left([V]_{[s],I}\right). \tag{9}$$

In particular, $\delta \in K \setminus \{0\}$ is independent of $I \subseteq [m]$.

Using the Laplace expansion on complementary minors (see e.g. [23, Theorem 2.4.1]) for the matrix $M(h, \lambda)$ in (7), we can rewrite $q(h, \lambda)$ as:

$$q(h,\lambda) = (-1)^s \sum_{\substack{I\subseteq[m]\\|I|=s}} (-1)^{\sum_{j=d+1}^{m} j + \sum_{i\in I} i} \det\left([W]_{[d],I^c}\right) \det\left([N' \operatorname{diag}(E\lambda) A^\top]_{[s],I}\right) \prod_{i\in I} h_i. \tag{10}$$

Remark 2.5. Viewed as a polynomial in h , the coefficients of $q(h, \lambda)$ are given by maximal minors of $W \in \mathbb{R}^{d\times m}$ and of $N' \operatorname{diag}(E\lambda)A^\top \in \mathbb{R}(\lambda)^{s\times m}$. Therefore, if $\operatorname{rk}(N' \operatorname{diag}(E\lambda)A^\top) < s$, then $q(h, \lambda)$ is the zero polynomial.

In the following, we treat $\lambda_1, \dots, \lambda_\ell$ as symbolic variables and view $N' \operatorname{diag}(E\lambda)A^\top$ as a matrix with entries in the field of rational functions $\mathbb{R}(\lambda)$. Furthermore, we assume that $N' \operatorname{diag}(E\lambda)A^\top$ has full rank s.

Corollary 2.6. *Let $D(\lambda) \in \mathbb{R}(\lambda)^{m\times d}$ be a Gale dual matrix of $N' \operatorname{diag}(E\lambda)A^\top \in \mathbb{R}(\lambda)^{s\times m}$. There exists $\delta(\lambda) \in \mathbb{R}(\lambda) \setminus \{0\}$ such that for all $I \subseteq [m]$ with $|I| = s$ it holds:*

$$\delta(\lambda) \det\left([D^\top(\lambda)]_{[d],I^c}\right) = \operatorname{sgn}(\tau_I) \det\left([N' \operatorname{diag}(E\lambda)A^\top]_{[s],I}\right). \tag{11}$$

The above corollary now establishes a way to simplify the computation of the critical polynomial by computing the determinant of minors of size d using the Gale dual matrices.

Theorem 2.7. *Let $D(\lambda) \in \mathbb{R}(\lambda)^{m\times d}$ be a Gale dual matrix of $N' \operatorname{diag}(E\lambda)A^\top$. The critical polynomial* (8) *can be written as*

$$q(h,\lambda) = (-1)^{s(d+1)} \sum_{\substack{I\subseteq[m]\\|I|=s}} \delta(\lambda) \det\left([W]_{[d],I^c}\right) \det\left([D^\top(\lambda)]_{[d],I^c}\right) \prod_{i\in I} h_i,$$

where $\delta(\lambda) \in \mathbb{R}(\lambda)\setminus\{0\}$ satisfies (11).

Proof. Let $I = \{i_1,\dots,i_s\} \subseteq [m]$, $I^c = \{j_1,\dots,j_d\} \subseteq [m]$ with $i_1 < \cdots < i_s$ and $j_1 < \cdots < j_d$. In the first part of the proof, we compute the sign of the permutation τ_I. The number of inversions in τ is given by

$$v := \sum_{k=1}^{d} \big(m - j_k - (d-k)\big) = ds - \sum_{k=1}^{d} j_k + \sum_{k=1}^{d} k, \tag{12}$$

and therefore $\operatorname{sgn}(\tau_I) = (-1)^v$.

We substitute the expression of $\det\left([N' \operatorname{diag}(E\lambda)A^\top]_{[s],I}\right)$ from (11) in (10). The total power of (-1) can, therefore, be computed as:

$$s + \sum_{k=d+1}^{m} k + \sum_{k=1}^{s} i_k + v = s + 2\sum_{k=1}^{m} k - 2\sum_{k=1}^{d} j_k + ds \equiv s(d+1) \mod 2.$$

This concludes the proof. □

2.3. **Signed supports of polynomials.** To establish connectivity in the parameter region of multistationarity via Theorem 2.2, it is enough to show that the properties (P1) and (P2) hold for the critical polynomial. To verify the two properties, we will exploit the structure of the Newton polytope of the critical polynomial.

Consider a multivariate polynomial function

$$f\colon \mathbb{R}^m_{>0} \to \mathbb{R}, \quad x \mapsto f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu,$$

where $c_\mu \in \mathbb{R}\setminus\{0\}$ and $\sigma(f) \subseteq \mathbb{Z}^m$ is a finite set, called the *support of* f. An exponent vector $\mu \in \sigma(f)$ is called *positive* (resp. *negative*) if the corresponding coefficient c_μ is positive (resp. negative). We write:

$$\sigma_+(f) := \{\mu \in \sigma(f) \mid c_\mu > 0\} \quad \text{and} \quad \sigma_-(f) := \{\mu \in \sigma(f) \mid c_\mu < 0\}.$$

The Newton polytope of f, denoted by $\mathrm{NP}(f)$ is given by the convex hull of $\sigma(f)$. For $S \subseteq \mathbb{R}^m$, we denote the *restriction* of f to S by

$$f_{|S}(x) = \sum_{\mu\in\sigma(f)\cap S} c_\mu x^\mu.$$

Following [25], we say that f has *one negative connected component* if

$$f^{-1}(\mathbb{R}_{<0}) = \left\{x \in \mathbb{R}^n_{>0} \mid f(x) < 0\right\}$$

is a connected set. If the Euclidean closure of $f^{-1}(\mathbb{R}_{<0})$ equals $f^{-1}(\mathbb{R}_{\le 0})$, then f *satisfies the closure property*. These are the terminologies for conditions (P1) and (P2) in Theorem 2.2.

For $v \in \mathbb{R}^m\setminus\{0\}$ and $a \in \mathbb{R}$, we define the *hyperplane* $\mathcal{H}_{v,a} := \{\mu \in \mathbb{R}^m \mid v\cdot\mu = a\}$, and the following two *half-spaces*:

$$\mathcal{H}^+_{v,a} = \{\mu \in \mathbb{R}^m \mid v\cdot\mu \ge a\}, \qquad \mathcal{H}^-_{v,a} = \{\mu \in \mathbb{R}^m \mid v\cdot\mu \le a\}.$$

We will denote by $\mathcal{H}^{+,\circ}_{v,a}$ and $\mathcal{H}^{-,\circ}_{v,a}$ the two *open half-spaces*:

$$\mathcal{H}^{+,\circ}_{v,a} = \{\mu \in \mathbb{R}^m \mid v \cdot \mu > a\}, \qquad \mathcal{H}^{-,\circ}_{v,a} = \{\mu \in \mathbb{R}^m \mid v \cdot \mu < a\}.$$

A hyperplane $\mathcal{H}_{v,a}$ is called a *separating hyperplane* of $\sigma(f)$ if $\sigma_-(f) \subseteq \mathcal{H}^+_{v,a}$ and $\sigma_+(f) \subseteq \mathcal{H}^-_{v,a}$. Additionally, we call $\mathcal{H}_{v,a}$ a *strict separating hyperplane* if there exists some $\mu \in \sigma_-(f)$ such that $v \cdot \mu > a$. The following proposition now recalls a result in [11], that establishes the properties (P1) and (P2) of f based on the existence of a strict separating hyperplane.

Proposition 2.8. *[11, Theorem 3.6] For a polynomial function $f\colon \mathbb{R}^m_{>0} \to \mathbb{R}$, $f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$, if $\sigma(f)$ has a strict separating hyperplane, then f has one negative connected component and satisfies the closure property.*

For a polytope $P \subseteq \mathbb{R}^m$, the *face* with normal vector $v \in \mathbb{R}^m$ is given by

$$P_v := \big\{\mu \in P \mid v \cdot \mu = \max_{\nu\in P} v \cdot \nu\big\}.$$

We call two faces $P_v, P_w \subseteq P$ *parallel* if $v = -w$. An *edge* (resp. *vertex*) of P is a face of dimension 1 (resp. 0). We write $\mathrm{Vert}(P)$ for the set of vertices of P. For a Newton polytope, we denote the face with normal vector v by $\mathrm{NP}_v(f)$. If $F \subseteq \mathrm{NP}(f)$ is an edge and $F \cap \sigma(f) \subseteq \sigma_-(f)$, we call F a *negative edge of* $\mathrm{NP}(f)$.

Theorem 2.9. *[25, Theorem 3.1] Let $f\colon \mathbb{R}^m_{>0} \to \mathbb{R}$, $f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a polynomial function. If there exists a proper face $\mathrm{NP}_v(f) \subsetneq \mathrm{NP}(f)$ such that $\sigma_-(f) \subseteq \mathrm{NP}_v(f)$, then f satisfies the closure property. If additionally $f_{|\,\mathrm{NP}_v(f)}$ has one negative connected component, then f also has one negative connected component.*

The following result from [25] provides a condition for splitting the polynomial into two parts and establishes that if both smaller polynomials have one negative connected component, then so does the original polynomial.

Theorem 2.10. *[25, Theorem 3.6] Let $f\colon \mathbb{R}^m_{>0} \to \mathbb{R}$, $f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a polynomial function. Assume that there exist parallel faces $\mathrm{NP}_v(f), \mathrm{NP}_{-v}(f) \subseteq \mathrm{NP}(f)$ such that $\sigma(f) \subseteq \mathrm{NP}_v(f) \cup \mathrm{NP}_{-v}(f)$ and both $f_{|\,\mathrm{NP}_v(f)}$ and $f_{|\,\mathrm{NP}_{-v}(f)}$ have one negative connected component. If there exist $\mu_0 \in \mathrm{NP}_v(f) \cap \sigma_-(f)$ and $\mu_1 \in \mathrm{NP}_{-v}(f) \cap \sigma_-(f)$ such that $\mathrm{Conv}(\mu_0, \mu_1)$ is an edge of $\mathrm{NP}(f)$, then f has one negative connected component.*

Example 2.11. Let $c_1, \dots, c_7 \in \mathbb{R}_{>0}$ and consider the polynomial

$$f = c_1x_1 + c_2x_1x_2 - c_3x_2 - c_4 - c_5x_1x_3 - c_6x_1x_2x_3 + c_7x_2x_3,$$

which has 3 positive and 4 negative exponent vectors:

$$\sigma_+(f) = \{(1,0,0), (1,1,0), (0,1,1)\}, \quad \sigma_-(f) = \{(0,0,0), (0,1,0), (1,0,1), (1,1,1)\}.$$

These exponent vectors are shown in red and blue colour respectively in Figure 1. For $e_3 = (0,0,1)$, the faces

$$F = \mathrm{NP}_{-e_3}(f) = \mathrm{Conv}((0,0,0), (1,0,0), (0,1,0), (1,1,0)),$$
$$G = \mathrm{NP}_{e_3}(f) = \mathrm{Conv}((1,0,1), (0,1,1), (1,1,1))$$

are parallel and their union contains $\sigma(f)$. The restricted polynomials are given by

$$f_{|F} = c_1x_1 + c_2x_1x_2 - c_3x_2 - c_4, \qquad f_{|G} = -c_5x_1x_3 - c_6x_1x_2x_3 + c_7x_2x_3.$$

Since the hyperplanes $\mathcal{H}_{-e_1,-0.5}$ and $\mathcal{H}_{e_1,0.5}$ with $e_1 = (1,0,0)$ are strict separating hyperplanes of $\sigma(f_{|F})$ and $\sigma(f_{|G})$ respectively, Proposition 2.8 implies that both $f_{|F}$ and $f_{|G}$ have one negative connected component.

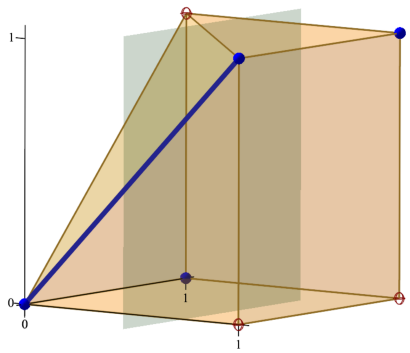


FIGURE 1. An illustration of Example 2.11. The positive and negative exponent vectors $\sigma_+(f)$, $\sigma_-(f)$ are marked with red circles and blue dots respectively. The grey hyperplane with normal vector $e_1 = (1,0,0)$ is a strict separating hyperplane of $\sigma(f_{|F})$ and $\sigma(f_{|G})$, for $F = \mathrm{NP}_{-e_3}(f)$, $G = \mathrm{NP}_{e_3}(f)$. The blue thick edge $\mathrm{Conv}((0,0,0),(1,0,1))$ connects negative exponent vectors of $f_{|F}$ and $f_{|G}$

For the negative exponent vectors $\mu_0 = (0,0,0)$, $\mu_1 = (1,0,1)$, $\mathrm{Conv}(\mu_0, \mu_1)$ is an edge of $\mathrm{NP}(f)$. Using Theorem 2.10, we conclude that f has one negative connected component. For an illustration, we refer to Figure 1.

We conclude the section with a couple of results that will be used in Section 3.2 and 4.2. In Section 3.2, we deal with polynomial functions with the special property that $\sigma(f) \subseteq \{0,1\}^m$ and each $\mu \in \sigma(f)$ has exactly d zero entries, for some fixed $d \in \mathbb{N}$. For $I \subseteq [m]$ with $|I| = d$, we write $z_I \subset \mathbb{R}^m$ for the vector with $(z_I)_i = 0$ for $i \in I$ and $(z_I)_i = 1$ for $i \notin I$. Moreover, we denote by $e_1, \ldots, e_m \in \mathbb{R}^m$ the standard basis vectors in $\mathbb{R}^m$.

Proposition 2.12. *For $d, m \in \mathbb{N}$ with $d < m$, let $P \subseteq \mathbb{R}^m$ be a polytope such that*

$$\mathrm{Vert}(P) \subseteq \big\{ z_I \in \{0,1\}^m \mid I \subseteq [m],\ |I| = d \big\},$$

and let $J_1, J_2 \subseteq [m]$ such that $|J_1| = |J_2| = d$. If $|J_1 \cap J_2| = d-1$ and $z_{J_1}, z_{J_2} \in P$, then $\mathrm{Conv}(z_{J_1}, z_{J_2})$ is an edge of P.

Proof. Since $|J_1 \cap J_2| = d-1$, we have that $J_1 = (J_1 \cap J_2) \cup \{j_1\}$, $J_2 = (J_1 \cap J_2) \cup \{j_2\}$ for $j_1 \neq j_2$. Let $v := e_{j_1} + e_{j_2} + 2\sum_{i \in (J_1 \cup J_2)^c} e_i$. For every $z_I \in \mathbb{R}^m$, it holds:

$$\begin{aligned} v \cdot z_I &= 2|(J_1 \cup J_2)^c| + 1 = 2(m-d-1)+1, \quad \text{if } I = J_1 \text{ or } I = J_2 \\ v \cdot z_I &< 2|(J_1 \cup J_2)^c| + 1 = 2(m-d-1)+1, \quad \text{if } I \neq J_1 \text{ or } I \neq J_2. \end{aligned} \tag{13}$$

From (13), it follows that $v \cdot \mu \leq 2(m-d-1)+1$ for all $\mu \in P$ with equality if and only if $\mu \in \mathrm{Conv}(z_{J_1}, z_{J_2})$. Thus, $\mathrm{Conv}(z_{J_1}, z_{J_2})$ is an edge of P. □

In Section 4.2, we will work with polynomial functions f such that $\sigma(f) \subseteq \{0,1\}^m \times \mathbb{R}^\ell$. To find edges of the Newton polytope for such polynomials the following proposition will be particularly helpful.

Proposition 2.13. *Let $\mathrm{pr}_1 \colon \mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R}^m$ and $\mathrm{pr}_2 \colon \mathbb{R}^m \times \mathbb{R}^\ell \to \mathbb{R}^\ell$ be projections onto the first m and onto the last ℓ coordinates respectively. Let $P \subseteq \mathbb{R}^m \times \mathbb{R}^\ell$ be a polytope, $x_1, x_2 \in \mathrm{Vert}(\mathrm{pr}_1(P))$ and $y \in \mathrm{Vert}(\mathrm{pr}_2(P))$. If $\mathrm{Conv}(x_1, x_2)$ is an edge of $\mathrm{pr}_1(P)$ such that*

$$\mathrm{pr}_1(\mathrm{Vert}(P)) \cap \mathrm{Conv}(x_1, x_2) = \{x_1, x_2\},$$

then $\mathrm{Conv}\big((x_1, y), (x_2, y)\big)$ is an edge of P.

Proof. Let $v \in \mathbb{R}^m$ be a normal vector of the face $\mathrm{Conv}(x_1, x_2) \subseteq \mathrm{pr}_1(\mathrm{Vert}(P))$. Since $\mathrm{pr}_1(\mathrm{Vert}(P)) \cap \mathrm{Conv}(x_1, x_2) = \{x_1, x_2\}$, for $z \in \mathrm{Vert}(P)$ we have

$$v \cdot \mathrm{pr}_1(z) \leq v \cdot x_1 = v \cdot x_2,$$

with equality if and only if $\mathrm{pr}_1(z) \in \{x_1, x_2\}$.

Let $w \in \mathbb{R}^\ell$ be a normal vector of $\mathrm{Vert}(\mathrm{pr}_2(P)) = \{y\}$. For $z \in \mathrm{Vert}(P)$, it holds

$$w \cdot \mathrm{pr}_2(z) \leq w \cdot y$$

with equality if and only if $\mathrm{pr}_2(z) = y$.

From the above inequalities, it follows that for $z \in \mathrm{Vert}(P)$

$$(v, w) \cdot z \leq (v, w) \cdot (x_1, y) = (v, w) \cdot (x_2, y)$$

with equality if and only if $z = (x_1, y)$ or $z = (x_2, y)$. So, $\mathrm{Conv}((x_1, y), (x_2, y))$ is an edge of P. □

2.4. **Overview of the approach.** In this section, we briefly summarize the approach we will use in Sections 3 and 4 to establish connectivity of the parameter region of multistationarity in the networks $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$, for every $n \geq 2$. We aim for this method to be applicable to other families of reaction networks. Therefore, we outline the key steps and how to address them.

Our arguments rely on Theorem 2.2, which applies to conservative consistent reaction networks without relevant boundary steady states. Both network families $(\mathcal{N}_{1,n})$ and $(\mathcal{N}_{2,n})$ are post-translational modification networks, which are conservative and consistent. From [18, Corollary 2] both networks do not have any relevant boundary steady states. So we can apply Theorem 2.2 and only consider the networks obtained from $(\mathcal{N}_{1,n})$ (resp. $(\mathcal{N}_{2,n})$) by removing the reversible reactions corresponding to $\kappa_{6i+2}, \kappa_{6i+5}$ (resp. $\kappa_{10i+2}, \kappa_{10i+7}$), $i = 0, \ldots, n-1$. After this modification, the reduced strongly irreversible phosphorylation network is given by

$$(\mathcal{F}_{1,n}) \qquad \begin{aligned} S_i + K &\xrightarrow{\kappa_{4i+1}} Y_{i+1} \xrightarrow{\kappa_{4i+2}} S_{i+1} + K \\ S_{i+1} + F &\xrightarrow{\kappa_{4i+3}} U_{i+1} \xrightarrow{\kappa_{4i+4}} S_i + F, \end{aligned} \qquad i = 0, \ldots, n-1,$$

and the reduced weakly irreversible phosphorylation network has the form

$$(\mathcal{F}_{2,n}) \qquad \begin{aligned} S_i + K &\xrightarrow{\kappa_{8i+1}} Y_{2i+1} \xrightarrow{\kappa_{8i+2}} Y_{2i+2} \underset{\kappa_{8i+4}}{\overset{\kappa_{8i+3}}{\rightleftharpoons}} S_{i+1} + K \\ S_{i+1} + F &\xrightarrow{\kappa_{8i+5}} U_{2i+1} \xrightarrow{\kappa_{8i+6}} U_{2i+2} \underset{\kappa_{8i+8}}{\overset{\kappa_{8i+7}}{\rightleftharpoons}} S_i + F, \end{aligned} \qquad i = 0, \ldots, n-1.$$

By Theorem 2.2, it is enough to show that the critical polynomial q_n of the reduced network satisfies properties (P1) and (P2), that is, it has one negative connected component and satisfies the closure property. To that end, first in Sections 3.1 and 4.1, we provide recursive formulas for the stoichiometric matrix $N_n \in \mathbb{R}^{m \times r}$, reactant matrix $A_n \in \mathbb{R}^{m \times r}$, a conservation law matrix $W_n \in \mathbb{R}^{d \times m}$ and an extreme matrix $E_n \in \mathbb{R}^{r \times \ell}$ of the networks $(\mathcal{F}_{1,n})$ and $(\mathcal{F}_{2,n})$. Moreover, we compute a Gale dual matrix $D_n(\lambda) \in \mathbb{R}(\lambda)^{m \times d}$ of $N_n' \operatorname{diag}(E_n \lambda) A_n^\top \in \mathbb{R}(\lambda)^{s \times m}$. We use these to write q_n in a recursive form.

For both families $(\mathcal{F}_{1,n})$ and $(\mathcal{F}_{2,n})$, we have $d = 3$ for every $n \in \mathbb{N}$. Thus, using Theorem 2.7, we can compute the coefficients of the critical polynomial q_n by computing minors of size 3 of the matrices W_n and $D_n(\lambda)$. In Sections 3.2 and 4.2, we conduct these computations, focusing on the signs of the coefficients of q_n. For $n = 1$, the critical polynomial q_1 has only positive coefficients for both $(\mathcal{F}_{1,n})$ and $(\mathcal{F}_{2,n})$. Thus, the parameter region of multistationarity is empty by [4, Theorem 1]. For $n = 2$, the polynomial q_2 satisfies (P1) and (P2) according to [26] for $(\mathcal{F}_{1,n})$, and as shown in [25] for $(\mathcal{F}_{2,n})$.

To prove that $q_n, n \geq 3$ satisfies the closure property, we show that $\sigma_-(q_n)$ is contained in a proper face $F \subsetneq \mathrm{NP}(q_n)$ (cf. Theorem 2.9). It is now enough to show that $q_{n|F}$ satisfies the property (P1). We split up $q_{n|F}$ into sub-polynomials based on parallel faces of the Newton polytope

(cf. Theorem 2.10 and Example 2.11). One of these sub-polynomials corresponds to q_{n-1}. For the remaining sub-polynomials, we show that their signed support has a strict separating hyperplane, which implies that they have one negative connected component (cf. Proposition 2.8). To conclude that $q_n|_F$ and consequently, q_n has one negative connected component we will use induction on n and Theorem 2.10. To identify negative edges between parallel faces, as required in Theorem 2.10, we will apply Proposition 2.12 and 2.13.

Combining these arguments shows that the critical polynomial q_n of the reduced networks satisfy properties (P1) and (P2). Consequently, Theorem 2.2 implies that the parameter region of multistationarity is connected for both the reduced and the original network.

3. **Strongly Irreversible Phosphorylation Networks**

In this section, we study the family of reduced strongly irreversible phosphorylation networks $(\mathcal{F}_{1,n})$. First in Section 3.1 we compute their critical polynomial and then in Section 3.2 we use it to show that the parameter region of multistationarity is connected for all $n \geq 2$.

3.1. **Computation of Critical Polynomial.** To define the matrices N_n and A_n for $(\mathcal{F}_{1,n})$, we use the order $K, F, S_0, Y_1, U_1, S_1, \ldots, Y_n, U_n, S_n$ for the species and $\kappa_1, \ldots, \kappa_{4n}$ for the reactions to label the rows and the columns, respectively, of N_n and A_n. We now give the matrix for N_1 and a recursive expression of the stoichiometric matrices $N_n \in \mathbb{R}^{(3n+3)\times 4n}$:

$$N_1 = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \in \mathbb{R}^{6\times 4} \qquad \text{and} \qquad N_n = \begin{pmatrix} N_{n-1} & P_1 \\ 0_{3\times(4n-4)} & P_2 \end{pmatrix} \in \mathbb{R}^{(3n+3)\times 4n}, \tag{14}$$

where the $(3n \times 4)$-matrix P_1 and (3×4)-matrix P_2 are given by

$$P_1 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3n\times 4} \qquad \text{and} \qquad P_2 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4}. \tag{15}$$

Remark 3.1. We point out that matrix P_1 has exactly $3n-3$ rows with only zero entries. Matrix P_2 has full rank and the rank of N_1 is 3. From the recursive relation it is easy to deduce that $\operatorname{rank}(N_n) = 3n$. In particular, $\dim(\operatorname{im}(N_n)) = 3n$ and the dimension of the left kernel of N_n is 3.

A recursive expression for the reactant matrices $A_n \in \mathbb{R}^{(3n+3)\times 4n}$ is given as

$$A_1 = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix} \in \mathbb{R}^{6\times 4} \qquad \text{and} \qquad A_n = \begin{pmatrix} A_{n-1} & Q_1 \\ 0_{3\times(4n-4)} & Q_2 \end{pmatrix} \in \mathbb{R}^{(3n+3)\times 4n}, \tag{16}$$

where the matrices Q_1 and Q_2 are given by

$$Q_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{3n\times 4} \qquad \text{and} \qquad Q_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4}. \tag{17}$$

Same as P_1, matrix Q_1 has exactly $3n-3$ rows with zero entries. With the stoichiometric matrix as in (14), we now determine a conservation law matrix W_n.

Lemma 3.2. *For the stoichiometric matrix N_n in (14), a conservation law matrix W_n is given by:*

$$W_n := \begin{pmatrix} \mathrm{Id}_3 & \underbrace{\overline{W} \quad \cdots \quad \overline{W}}_{n} \end{pmatrix} \in \mathbb{R}^{3\times(3n+3)}, \tag{18}$$

where Id_3 *is the identity matrix of size 3 and*

$$\overline{W} := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 3}.$$

Proof. Since the dimension of the left kernel of N_n is 3 (Remark 3.1) and W_n has rank 3 for all n, it is enough to show that $W_nN_n = 0$. Using $W_0 := \mathrm{Id}_3$ and $W_n = \begin{pmatrix} W_{n-1} & \overline{W} \end{pmatrix}$, for $n \geq 2$ we have

$$W_nN_n = \begin{pmatrix} W_{n-1}N_{n-1} & W_{n-1}P_1 + \overline{W}P_2 \end{pmatrix}.$$

Simple computations give

$$W_{n-1}P_1 = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 4} \qquad \text{and} \qquad \overline{W}P_2 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 1 & 0 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{3\times 4}.$$

Thus, for $n = 1$, we get $W_1N_1 = W_{n-1}P_1 + \overline{W}P_2 = 0$. Let us now assume that the result holds for $k = n-1$ for some $n \geq 1$. By assumption $W_{n-1}N_{n-1} = 0_{3\times(4k-4)}$. Consequently, $W_nN_n = 0_{3\times 4k}$ and therefore, W_n is a conservation law matrix for N_n. □

Proposition 3.3. *An extreme matrix of $(\mathcal{F}_{1,n})$ has the following recursive form*

$$E_n = \left(\begin{array}{c|c} E_{n-1} & 0_{(4n-4)\times 1} \\ \hline 0_{4\times(n-1)} & E_1 \end{array}\right) \in \mathbb{R}^{4n\times n}, \tag{19}$$

where

$$E_1^\top = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{1\times 4}.$$

Moreover, the n column vectors $E_n^{(1)}, \ldots, E_n^{(n)} \in \mathbb{R}^{4n}$ of E_n form a basis of $\ker(N_n)$.

Proof. Using (14), it is easy to check that $E_n^{(1)}, \ldots, E_n^{(n)}$ are contained in $\ker(N_n) \cap \mathbb{R}^{4n}_{\geq 0}$. Since

$$\dim\ker(N_n) = 4n - \dim\mathrm{im}(N_n) = 4n - 3n = n$$

(cf. Remark 3.1) and $E_n^{(1)}, \ldots, E_n^{(n)}$ are linearly independent, they form a basis of $\ker(N_n)$. Moreover, $E_n^{(1)}, \ldots, E_n^{(n)}$ have pairwise disjoint support and hence, by Proposition 2.1 E_n gives an extreme matrix for the network $(\mathcal{F}_{1,n})$. □

The matrix N_n' is obtained by removing the first three rows of N_n. In the remainder of this subsection, we compute a Gale dual matrix $D_n(\lambda)$ of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top \in \mathbb{R}(\lambda)^{3n\times(3n+3)}$ and $\delta_n(\lambda)$ as in Theorem 2.7 for the network $(\mathcal{F}_{1,n})$, considering $\lambda = (\lambda_1, \ldots, \lambda_n)$ as symbolic variables.

In Lemma 3.4, we compute the maximal minor $\det([N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I})$ for $I = \{4, \ldots, 3n+3\}$ and $I^c = \{1, 2, 3\}$.

Lemma 3.4. *Given $n \in \mathbb{N}$, let $s = 3n$, $I = \{4, \ldots, 3n+3\}$. Then,*

$$\det([N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I}) = (-1)^n \prod_{i=1}^{n} \lambda_i^3.$$

In particular, $\ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$ *has dimension* 3.

Proof. Note that the matrix $[N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I}$ is obtained by deleting the first 3 columns of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$. Let A_n' denote the matrix obtained from A_n after deleting the first three rows. Then,

$$[N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I} = N_n' \operatorname{diag}(E_n\lambda)A_n^{'\top}.$$

We define the matrices $C_n \in \mathbb{R}(\lambda)^{3n\times 3n}$ recursively as:

$$C_n := \left(\begin{array}{c|c} \begin{matrix} C_{n-1} \\ 0_{3\times(3n-6)} \end{matrix} & \begin{matrix} 0_{3\times 3} \\ Z_n \end{matrix} \\ V_n & \end{array}\right) \in \mathbb{R}(\lambda)^{3n\times 3n} \quad \text{and} \quad C_1 := Z_1$$

where

$$V_n := \begin{pmatrix} 0 & 0 & \lambda_n \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad Z_n := \begin{pmatrix} -\lambda_n & 0 & 0 \\ 0 & -\lambda_n & \lambda_n \\ \lambda_n & 0 & -\lambda_n \end{pmatrix} = P_2 \operatorname{diag}(E_1\lambda_n)Q_2^\top,$$

and P_2, Q_2 are the matrices from (15) and (17). Therefore, for $n = 1$, we have:

$$N_1' \operatorname{diag}(E_1\lambda)A_1^{'\top} = P_2 \operatorname{diag}(E_1\lambda_n)Q_2^\top = C_1.$$

To prove the result we show that C_n is obtained by performing elementary row operations on $N_n' \operatorname{diag}(E_n\lambda)A_n^{'\top}$. Consider the matrix $X_n \in \mathbb{R}^{3n\times 3n}$ given by

$$(X_n)_{ij} := \begin{cases} 1 \text{ if } i = j \\ 1 \text{ if } k < n,\ i = 3k,\ j > 3k \\ 0 \text{ else.} \end{cases}$$

Furthermore, we set $X_1 := \mathrm{Id}_3$. Multiplying by X_n is the same as adding the rows $3k+1, 3k+2$, and $3k+3$ to the row $3k$ for each $k = 1, \ldots, n-1$. Note that X_n can be written recursively as

$$X_n = \begin{pmatrix} X_{n-1} & X_{n-1}' \\ 0_{3\times(3n-3)} & \mathrm{Id}_3 \end{pmatrix}$$

where $X_n' \in \mathbb{R}^{(3n-3)\times 3}$ and

$$(X_{n-1}')_{ij} := \begin{cases} 1 \text{ if } i = 3k,\ k = 1, \ldots, n-1, \\ 0 \text{ else.} \end{cases}$$

It follows that for all $n > 1$ the matrix X_nN_n' can be written recursively as:

$$X_nN_n' = \begin{pmatrix} X_{n-1}N_{n-1}' & X_{n-1}P_1' + X_{n-1}'P_2 \\ 0_{3\times(4n-4)} & P_2 \end{pmatrix} = \begin{pmatrix} X_{n-1}N_{n-1}' & 0_{(3n-3)\times 4} \\ 0_{3\times(4n-4)} & P_2 \end{pmatrix},$$

where P_1' is obtained by removing the first three rows of P_1 and P_2 is defined as in (15).

Finally, we show inductively that $C_n = X_nN_n' \operatorname{diag}(E_n\lambda)A_n^{'\top}$ for all $n \in \mathbb{N}$. The statement holds for $n = 1$. Assume that for $n \in \mathbb{N}$, $C_k = X_kN_k' \operatorname{diag}(E_k\lambda)A_k^{'\top}$ holds for all $2 \le k < n$. Using the block structure of X_nN_n', $E_n\lambda$, $A_n^{'\top}$ and C_n, we have:

$$\begin{aligned} X_nN_n' \operatorname{diag}(E_n\lambda)A_n^{'\top} &= \begin{pmatrix} X_{n-1}N_{n-1}' \operatorname{diag}(E_{n-1}\hat{\lambda}) & 0_{(3n-3)\times 4} \\ 0_{3\times(3n-3)} & P_2 \operatorname{diag}(E_1\lambda_n) \end{pmatrix} \begin{pmatrix} A_{n-1}^{'\top} & 0_{(3n-3)\times 3} \\ Q_1^{'\top} & Q_2^\top \end{pmatrix} \\ &= \left(\begin{array}{c|c} \begin{matrix} C_{n-1} \\ 0_{3\times(3n-6)} \end{matrix} & \begin{matrix} 0_{3\times 3} \\ Z_n \end{matrix} \\ V_n & \end{array}\right) = C_n, \end{aligned}$$

where $\lambda = (\lambda_1, \ldots, \lambda_n)$, $\hat{\lambda} = (\lambda_1, \ldots, \lambda_{n-1})$, and Q_1, Q_2 are the matrices from (17), with Q_1' being obtained from Q_1 by removing the first three rows.

Since $\det(X_n) = 1$, we get

$$\det(N_n' \operatorname{diag}(E_n\lambda)A_n^{'\top}) = \det(C_n) = \prod_{i=1}^{n} \det(Z_i) = (-1)^n \prod_{i=1}^{n} \lambda_i^3.$$

Since this determinant is non-zero, the matrix $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$ has rank $s = 3n$ for all n, and hence, its kernel has dimension 3. □

Next, we find an explicit representation of a Gale dual matrix of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$ for all n.

Theorem 3.5. *Let N_n', A_n and E_n be the matrices associated with the network $(\mathcal{F}_{1,n})$ and $\lambda = (\lambda_1, \dots, \lambda_n)$. Consider the matrices*

$$D^{(0)} := \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad D^{(i)} := \begin{pmatrix} 0 & 0 & -1 \\ -(i-1) & -(i-1) & -i \\ 1 & 1 & 1 \end{pmatrix}, \quad i = 1, \dots, n, \quad \text{and}$$

$$D_n^\top(\lambda) := \begin{pmatrix} D^{(0)} & \dots & D^{(n)} \end{pmatrix} \in \mathbb{R}(\lambda)^{3\times(3n+3)}.$$

Then $D_n(\lambda)$ is a Gale dual matrix of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top \in \mathbb{R}(\lambda)^{3n\times(3n+3)}$.

Proof. By Lemma 3.4, $\ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$ has dimension 3. The columns of $D_n(\lambda)$ are linearly independent and hence, it is enough to show that they are contained in the kernel of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$. Note that a vector $v \in \mathbb{R}(\lambda)^{3n+3}$ is in this kernel if and only if $\operatorname{diag}(E_n\lambda)A_n^\top v \in \ker(N_n')$. Moreover, $\ker(N_n') = \ker(N_n)$. Since the columns of E_n form a basis of $\ker(N_n)$ by Proposition 3.3, it follows that $v \in \ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$ if and only if there exists $\mu \in \mathbb{R}(\lambda)^n$ such that

$$\operatorname{diag}(E_n\lambda)A_n^\top v = E_n\mu = \begin{bmatrix} u_1 & \dots & u_n \end{bmatrix}^\top$$

where

$$u_i = \begin{bmatrix} \mu_i & \mu_i & \mu_i & \mu_i \end{bmatrix} \in \mathbb{R}(\lambda)^4 \qquad \text{for } i = 1, \dots, n.$$

Using the block form of A_n and the block form of E_n, for each $v \in \mathbb{R}(\lambda)^{3n+3}$ we have

$$\operatorname{diag}(E_n\lambda)A_n^\top v = \begin{bmatrix} \omega_1 & \dots & \omega_n \end{bmatrix}^\top. \tag{20}$$

such that

$$\omega_i = \begin{bmatrix} \lambda_i(v_1 + v_{3i}) & \lambda_i v_{3i+1} & \lambda_i(v_2 + v_{3i+3}) & \lambda_i v_{3i+2} \end{bmatrix} \qquad \text{for } i = 1, \dots n.$$

So, the vector $v \in \mathbb{R}(\lambda)$ is such that the entries $4i+1, 4i+2, 4i+3$, and $4i+4$ are equal in the column vector in (20) for all $i = 1, \dots, n$. In particular, $v_1 + v_{3i} = v_{3i+1} = v_2 + v_{3i+3} = v_{3i+2}$. It is easy to check that the rows of $D_n^\top(\lambda)$ are linearly independent and satisfy these relations. This concludes the proof. □

Note that the Gale dual matrix obtained in Theorem 3.5 is independent of λ. While this is the case for $(\mathcal{F}_{1,n})$, we will see in Section 4.1 that the same does not hold for $(\mathcal{F}_{2,n})$.

The main objective in this section is to obtain the expression of critical polynomial using Theorem 2.7. The final missing ingredient is the expression of $\delta_n(\lambda)$. We will use (11) to find this expression.

Proposition 3.6. *With the choice of Gale dual matrix in Theorem 3.5, it holds*

$$\delta_n(\lambda) = \prod_{i=1}^{n} \lambda_i^3.$$

Proof. Consider the index set $I = \{4, \dots, 3n+3\}$. Then τ_I is the permutation that sends $(1, \dots, 3n+3)$ to $(4, \dots, 3n+3, 1, 2, 3)$. A straightforward computation gives

$$\operatorname{sgn}(\tau_I) = (-1)^n.$$

Moreover, $\det\left([D_n^\top(\lambda)]_{[3],I^c}\right) = \det D^{(0)} = 1$. Using Lemma 3.4 and substituting these values in (11), we obtain the statement. □

3.2. **Connectivity of the Multistationarity Region of $(\mathcal{N}_{1,n})$.** For $h = (h_1, \dots, h_{3n+3}) \in \mathbb{R}^{3n+3}_{>0}$, by Theorem 2.7 and Theorem 3.5, the critical polynomial associated with $(\mathcal{F}_{1,n})$ can be written as $q_n(h, \lambda) = g_n(h)\delta_n(\lambda)$, where

$$g_n(h) = \sum_{\substack{I \subseteq [3n+3] \\ |I|=3}} \det\left([W_n]_{[3],I}\right) \det\left([D_n^\top]_{[3],I}\right) \prod_{i \in I^c} h_i \quad \text{and} \quad \delta_n(\lambda) = \prod_{i=1}^{n} \lambda_i^3. \tag{21}$$

Remark 3.7. The polynomial g_n is independent of λ. Moreover, $\delta_n(\lambda)$ is a positive function if $\lambda \in \mathbb{R}^n_{>0}$ and hence, if $g_n^{-1}(\mathbb{R}_{<0})$ is connected, then so is $q_n^{-1}(\mathbb{R}_{<0})$. Therefore, it is enough to consider the polynomial $g_n(h)$.

Henceforth, we will use the following notation:

$$\alpha_{n,I} := \det\left([W_n]_{[3],I}\right) \det\left([D_n^\top]_{[3],I}\right). \tag{22}$$

In the next remark, we list various cases when $\alpha_{n,I} = 0$.

Remark 3.8. For $n \geq 1$ and $I \subset [3n+3]$ with $|I| = 3$, the following holds:

(i) If $3\ell + k$, $3\ell' + k \in I$ for $\ell, \ell' \in [n]$ and $k \in \{1, 2\}$, then $\det\left([W_n]_{[3],I}\right) = 0$.
(ii) If $3\ell + 3$, $3\ell' + 3 \in I$ for $\ell, \ell' \in [n] \cup \{0\}$, then $\det\left([W_n]_{[3],I}\right) = 0$.
(iii) If $3\ell + 1 \notin I$ or $3\ell' + 2 \notin I$ for $\ell, \ell' \in [n] \cup \{0\}$, then $\det\left([W_n]_{[3],I}\right) = 0$.
(iv) If $3\ell + 1, 3\ell + 2 \in I$ for $\ell \in [n]$, then $\det\left([D_n^\top]_{[3],I}\right) = 0$.

To show that the region in the parameter space for $(\mathcal{N}_{1,n})$ and $(\mathcal{F}_{1,n})$ that enables multistationarity is path connected for all n, we first write g_n as a polynomial in $h_{3n+1}, h_{3n+2}, h_{3n+3}$ with coefficients in $\hat{h} = (h_1, \dots, h_{3n})$:

$$g_n(h) = \sum_{J \subseteq [3]} \left(a_J(\hat{h}) \prod_{i \in J^c} h_{3n+i} \right) = \; b_0(h) + b_1(h),$$

where

$$\begin{aligned} b_0(h) &= a_{\{1\}}(\hat{h}) h_{3n+2} h_{3n+3} + a_{\{1,2\}}(\hat{h}) h_{3n+3} + a_{\{1,3\}}(\hat{h}) h_{3n+2} + a_{\{1,2,3\}}(\hat{h}) \\ b_1(h) &= (a_\emptyset(\hat{h}) h_{3n+2} h_{3n+3} + a_{\{2\}}(\hat{h}) h_{3n+3} + a_{\{3\}}(\hat{h}) h_{3n+2} + a_{\{2,3\}}(\hat{h})) h_{3n+1}. \end{aligned} \tag{23}$$

In the next result, we focus on the polynomial $b_0(h)$.

Proposition 3.9. *For all $n \geq 1$, the polynomial $b_0(h)$ is non-zero and all of its coefficients are positive.*

Proof. First we show that $a_{\{1,2\}}(\hat{h})$ and $a_{\{1,2,3\}}(\hat{h})$ are zero polynomials. To see this, note that their coefficients are obtained from summands in (21) indexed by I such that $\{3n+1, 3n+2\} \subseteq I$. For such an I, we have $\det\left([D_n^\top]_{[3],I}\right) = 0$ by Remark 3.8(iv). Thus, $a_{\{1,2\}}(\hat{h}) = 0$ and $a_{\{1,2,3\}}(\hat{h}) = 0$.

We now show that $a_{\{1,3\}}(\hat{h})$ and $a_{\{1\}}(\hat{h})$ only have positive coefficients. First consider the polynomial $a_{\{1,3\}}(\hat{h})$. Here the coefficients (22) of $a_{\{1,3\}}(\hat{h})$ are computed for $I = \{r, 3n+1, 3n+3\}$ such that $r \in [3n]$. By Remark 3.8, we only need to consider the case when $r = 3\ell+2$ for $\ell = 0, \dots, n-1$. When $\ell = 0$, we have that $\det\left([D_n^\top]_{[3],I}\right) = 0$. When $\ell \in [n-1]$, we have:

$$\alpha_{n,I} = \det \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \det \begin{pmatrix} 0 & 0 & -1 \\ -(\ell-1) & -(n-1) & -n \\ 1 & 1 & 1 \end{pmatrix} = (-1)(\ell - n) = n - \ell > 0.$$

This ensures $a_{\{1,3\}}(\hat{h})$ only has positive coefficients.

We now consider the coefficients of $a_{\{1\}}(\hat{h})$. In this case, $3n+1 \in I$ and $3n+2, 3n+3 \notin I$. We can write $I = \{r_1, r_2, 3n+1\}$ for $r_1, r_2 \in [3n]$ and let $\ell, \ell' \in [n-1]$.

(i) For $(r_1, r_2) = (1, 2)$ and $(r_1, r_2) = (2, 3)$, we have $\alpha_{n,I} = 1$ and $\alpha_{n,I} = n$ respectively.
(ii) Finally, $(r_1, r_2) \in \{(1, 3\ell + 2), (2, 3\ell + 2), (2, 3\ell + 3), (3, 3\ell + 2), (3\ell + 2, 3\ell' + 3)\}$, we get, $\alpha_{n,I} = n - \ell$.

By Remark 3.8, in all the other cases $\alpha_{n,I} = 0$. Since in all the cases above $\alpha_{n,I} \geq 0$, the coefficients of $a_{\{1\}}(\hat{h})$ are positive and hence, the result. □

Following corollary is now a direct consequence of Proposition 3.9.

Corollary 3.10. *For $n \geq 2$, the face $\mathrm{NP}_{e_{3n+1}}(g_n)$ is a proper face of $\mathrm{NP}(g_n)$ and $\sigma_-(g_n) \subseteq \mathrm{NP}_{e_{3n+1}}(g_n)$. In particular, g_n satisfies the closure property and if the polynomial*

$$b_1(h) = \big(a_\emptyset(\hat{h})h_{3n+2}h_{3n+3} + a_{\{2\}}(\hat{h})h_{3n+3} + a_{\{3\}}(\hat{h})h_{3n+2} + a_{\{2,3\}}(\hat{h})\big)h_{3n+1}$$

has one negative connected component, then so does g_n.

Proof. For $I_1 = \{1, 2, 3\}$, $\alpha_{n,I_1} = 1$, and therefore $z_{I_1} \in \sigma(g_n)$. From Proposition 3.9, there exists $I_2 \subseteq [3n + 3]$ with $|I_2| = 3$, $3n + 1 \notin I_2$ and $z_{I_2} \in \sigma(g_n)$. Since $e_{3n+1} \cdot \mu \in \{0, 1\}$ for all $\mu \in \sigma(g_n)$, we have that $\mathrm{NP}_{e_{3n+1}}(g_n)$ is a proper face of $\mathrm{NP}(g_n)$.

Since $\sigma_-(g_n) \subseteq \mathrm{NP}_{e_{3n+1}}(g_n)$ by Proposition 3.9, the second part of the corollary follows from Theorem 2.9. □

By Corollary 3.10, it is enough to focus on $b_1(h)$. Next, we consider the polynomial $a_\emptyset(\hat{h})$. We recall that $z_{\{i,j,k\}} \in \{0, 1\}^{3n+3}$ denotes the vector whose entries indexed by $i, j, k \in [3n + 3]$ are zero and all the other entries are 1. If $i, j, k \in [3n]$, we write $\hat{z}_{\{i,j,k\}} \in \mathbb{R}^{3n}$ for the vector that is obtained from $z_{\{i,j,k\}}$ by removing its last three coordinates.

Proposition 3.11. *For $n \geq 2$, and $J = \emptyset$, $a_J(\hat{h})$ is the polynomial $g_{n-1}(\hat{h})$ associated with the network $(\mathcal{F}_{1,n-1})$. Furthermore, if $\hat{z}_{\{i,j,k\}} \in \sigma_-(g_{n-1})$, then $z_{\{i,j,k\}} \in \sigma_-(g_n)$.*

Proof. The terms in $a_\emptyset(\hat{h})$ are computed from the summands indexed by I in (21) such that $\{3n + 1, 3n + 2, 3n + 3\} \cap I = \emptyset$. Note the following set equaltiy:

$$\big\{I \subseteq [3n + 3] \mid |I| = 3 \text{ and } \{3n + 1, 3n + 2, 3n + 3\} \cap I = \emptyset\big\} = \big\{I \subseteq [3n] \mid |I| = 3\big\}.$$

Using the block structures of W_n and D_n, the first part of the statement follows directly from (21).

For the second part, let I be in the set above. Since $a_\emptyset(\hat{h}) = g_{n-1}(\hat{h})$, by (21), (23), the coefficient of g_n corresponding to z_I is the same as the coefficient of g_{n-1} that corresponds to $\hat{z}_I$. □

Next, we look at the support of $a_{\{2,3\}}(\hat{h})h_{3n+1}$, $\quad a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3}$ and $a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}$. In each of these polynomials, every exponent vector is of the form $z_{\{i,j,k\}}$ by (21) and is a vertex of the Newton polytope.

Proposition 3.12. *For $n \geq 2$, consider the supports $\sigma(a_{\{2,3\}}(\hat{h})h_{3n+1})$ and $\sigma(a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3})$. Each set has exactly one positive exponent vector given by $z_{\{1,3n+2,3n+3\}}$ and $z_{\{1,2,3n+2\}}$, respectively. Moreover, both supports have a strict separating hyperplane, and it holds that*

(i) $\sigma_-(a_{\{2,3\}}(\hat{h})h_{3n+1}) = \big\{z_{\{3\ell+1,3n+2,3n+3\}} \mid \ell \in [n - 1]\big\}$,
(ii) $z_{\{1,3,3n+2\}}, z_{\{1,3\ell+1,3n+2\}} \in \sigma_-\big(a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3}\big)$ *for* $\ell \in [n - 1]$.

Proof. Let us consider the polynomial $a_{\{2,3\}}(\hat{h})h_{3n+1}$. The terms in $a_{\{2,3\}}$ are obtained from the summands in (21) indexed by $I = \{r, 3n + 2, 3n + 3\}$ where $r \in [3n]$. By Remark 3.8, we only need to consider the case when $r = 3\ell + 1$ for $\ell = 0, \ldots, n - 1$. When $\ell = 0$, we have $\alpha_{n,I} = 1 > 0$. When $r = 3\ell + 1$, for $\ell = [n - 1]$, we have:

$$\alpha_{n,I} = \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \det \begin{pmatrix} 0 & 0 & -1 \\ -(\ell - 1) & -(n - 1) & -n \\ 1 & 1 & 1 \end{pmatrix} = \ell - n < 0.$$

This shows that $\sigma_-(a_{\{2,3\}}(\hat{h})h_{3n+1}) = \{z_{\{3\ell+1,3n+2,3n+3\}} \mid \ell \in [n-1]\}$. Additionally, for $I = \{1, 3n+2, 3n+3\}$ we get the unique positive exponent vector.

For $a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3}$, we haven $I = \{r_1, r_2, 3n+2\}$ for some $r_1, r_2 \in [3n]$. Let $\ell, \ell' \in [n-1]$.

(i) For $(r_1, r_2) = (1, 2)$ we obtain $\alpha_{n,I} = 1$.
(ii) For $(r_1, r_2) \in \{(1, 3), (1, 3\ell+1)\}$, we get, $\alpha_{n,I} = -(n-1)$.
(iii) For $(r_1, r_2) \in \{(2, 3\ell+1), (3, 3\ell+1), (3\ell+1, 3\ell'+3)\}$, we have $\alpha_{n,I} = -(n-\ell)$.
(iv) Finally, for $(r_1, r_2) = (1, 3\ell+3)$, we have $\alpha_{n,I} = -(n-\ell-1)$.

By Remark 3.8, in all the other cases $\alpha_{n,I} = 0$. From the computations above we obtain the unique positive exponent vector for $I = \{1, 2, 3n+2\}$.

To see that both $\sigma(a_{\{2,3\}}(\hat{h})h_{3n+1})$ and $\sigma(a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3})$ have a strict separating hyperplane, note that the unique positive exponent vector is a vertex of the coppesponding Newton polytope. Therefore, the unique positive exponent vector can be separated by an affine hyperplane from the other exponent vectors. This concludes the proof. □

Finally, we consider the set of exponent vectors of $a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}$.

Proposition 3.13. *For $n \geq 3$, the set of exponent vectors $\sigma\Big(a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}\Big)$ has a strict separating hyperplane. Moreover, we have $\{z_{\{2,3\ell+1,3n+3\}} \mid \ell \in [n-1]\} \subseteq \sigma_-\big(a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}\big)$.*

Proof. We consider $I = \{r_1, r_2, 3n+3\}$ where $r_1, r_2 \in [3n]$. First, we investigate the case $r_1 = 1$, $r_2 \neq 2$. By Remark 3.8, we only need to consider the case when $r_2 = 3\ell+2$ for $\ell \in [n-1]$. In this case $\alpha_{n,I} = (n-\ell+1) \geq 0$ with equality only if $\ell = n-1$.

Next, we work with the case when $r_1 = 2$, $r_2 \neq 1$. Again by Remark 3.8, we only need to consider the case when $r_2 = 3\ell+1$ for $\ell \in [n-1]$. For this case we find $\alpha_{n,I} = -(n-\ell) < 0$. This shows that the exponent vectors of the form $z_{\{2,3\ell+1,3n+3\}}$ correspond to negative coefficients.

Furthermore,

(i) When $1 \in I$ and $2 \notin I$, we have $(e_1 - e_2) \cdot z_I = -1 < 0$.
(ii) When either $1, 2 \in I$ or $1, 2 \notin I$, we have $(e_1 - e_2) \cdot z_I = 0$.
(iii) Finally, when $2 \in I$ and $1 \notin I$, we have $(e_1 - e_2) \cdot z_I = 1 > 0$.

Thus, the open half-space $\mathcal{H}^{+,\circ}_{e_1-e_2,0}$ contains only negative exponent vectors and $\mathcal{H}^{-,\circ}_{e_1-e_2,0}$ contains only positive exponent vectors. This implies that $\mathcal{H}_{e_1-e_2,0}$ is a strict separating hyperplane of $\sigma\big(a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}\big)$. □

Proposition 3.14. *The polynomial g_2 from* (21) *has one negative connected component.*

Proof. The polynomial g_2 equals

$$
\begin{aligned}
&- h_1h_2h_3h_5h_6h_7 + h_2h_3h_4h_5h_6h_7 + h_1h_2h_3h_4h_6h_8 + 2\,h_2h_3h_4h_6h_7h_8 - h_1h_3h_5h_6h_7h_8\\
&+ h_3h_4h_5h_6h_7h_8 - h_1h_2h_3h_5h_7h_9 - h_1h_2h_5h_6h_7h_9 - h_1h_3h_5h_6h_7h_9 - h_2h_3h_5h_6h_7h_9\\
&- h_2h_4h_5h_6h_7h_9 + h_3h_4h_5h_6h_7h_9 + h_1h_2h_3h_4h_8h_9 + h_1h_3h_4h_5h_8h_9 + h_1h_2h_4h_6h_8h_9\\
&+ h_1h_3h_4h_6h_8h_9 + h_2h_3h_4h_6h_8h_9 + 2\,h_1h_4h_5h_6h_8h_9 + h_3h_4h_5h_6h_8h_9 + h_2h_3h_4h_7h_8h_9\\
&+ h_3h_4h_5h_7h_8h_9 + h_3h_4h_6h_7h_8h_9 + h_1h_5h_6h_7h_8h_9 + h_3h_5h_6h_7h_8h_9 + h_4h_5h_6h_7h_8h_9.
\end{aligned}
$$

An easy computation shows that the hyperplane $\mathcal{H}_{v,4}$ with $v = (1, 1, 1, 0, 1, 0, 1, 0, 1)$ is a strict separating hyperplane of $\sigma(g_2)$. The statement now follows by Proposition 2.8. □

We are now ready to show that the multistationarity region of $(\mathcal{F}_{1,n})$ and $(\mathcal{N}_{1,n})$ are connected for all $n \geq 2$.

Theorem 3.15. *For all $n \geq 2$, the parameter region of multistationarity is path connected for the networks $(\mathcal{F}_{1,n})$ and $(\mathcal{N}_{1,n})$.*

Proof. By Corollary 3.10, it is enough to show that the following polynomial

$$b_1 = \big(a_\emptyset(\hat{h})h_{3n+2}h_{3n+3} + a_{\{2,3\}}(\hat{h}) + a_{\{2\}}(\hat{h})h_{3n+3} + a_{\{3\}}(\hat{h})h_{3n+2}\big)h_{3n+1}$$

has exactly one negative connected component. To show that b_1 has one negative connected component for all $n \geq 2$ we will use an induction argument over n. From Proposition 3.14, we know that g_2 has exactly one negative component, so we assume that it holds for all g_k where $2 \leq k \leq n-1$.

Now, we write $b_1 = b_2 + b_3$ based on the exponent of h_{3n+3}:

$$b_2 := \big(a_\emptyset(\hat{h})h_{3n+1}h_{3n+2} + a_{\{2\}}(\hat{h})h_{3n+1}\big)h_{3n+3} \quad \text{and} \quad b_3 := a_{\{2,3\}}(\hat{h})h_{3n+1} + a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}. \tag{24}$$

Note that $\sigma(b_2) \subseteq \mathrm{NP}_{e_{3n+3}}(b_1)$ and $\sigma(b_3) \subseteq \mathrm{NP}_{-e_{3n+3}}(b_1)$. Moreover, these two faces are parallel and $\sigma(b_1) = \sigma(b_2) \cup \sigma(b_3)$. Since by Proposition 3.12 we have $z_{\{1,4,3n+2\}} \in \sigma_-(b_2)$ and $z_{\{4,3n+2,3n+3\}} \in \sigma_-(b_3)$, using Proposition 2.12 we obtain a negative edge $\mathrm{Conv}(z_{\{1,4,3n+2\}}, z_{\{4,3n+2,3n+3\}})$ of $\mathrm{NP}(b_1)$. Therefore by Theorem 2.10, it is now enough to show that b_2 and b_3 have one negative connected component.

We now consider b_2. The two summands in (24) split b_2 in terms of the exponent of h_{3n+2}. Similar to b_1, we have $\sigma(a_\emptyset(\hat{h})h_{3n+1}h_{3n+2}h_{3n+3}) \subseteq \mathrm{NP}_{e_{3n+2}}(b_2)$ and $\sigma(a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3})$ is contained in $\mathrm{NP}_{-e_{3n+2}}(b_2)$. Moreover, from Proposition 2.12, Proposition 3.11 and Proposition 3.12 it follows that $\mathrm{Conv}(z_{\{1,3,3(n-1)+2\}}, z_{\{1,3,3n+2\}})$ is an edge of the Newton polytope $\mathrm{NP}(b_2)$ between the two negative exponent vectors. Using Theorem 2.10, it is enough to show that the two summands have exactly one negative connected component. By Proposition 3.11, $a_\emptyset(\hat{h}) = g_{n-1}(\hat{h})$, and it has one negative connected component by the inductive assumption. On the other hand, by Proposition 3.12 $a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3}$ has a strict separating hyperplane, which implies that $a_{\{2\}}(\hat{h})h_{3n+1}h_{3n+3}$ has a unique negative connected component by Proposition 2.8. Hence, so does b_2.

To complete the proof, we show that b_3 has one negative connected component. We follow the same argument thread as for b_2. We split the two summands of b_3 in (24) in terms of the exponent of h_{3n+2}. Their exponent vectors then lie on parallel faces of $\mathrm{NP}(b_3)$ with normal vectors $\pm e_{3n+2}$. We observe that $\mathrm{Conv}(z_{\{4,3n+2,3n+3\}}, z_{\{2,4,3n+3\}})$ is an edge of the Newton polytope $\mathrm{NP}(b_3)$ between two negative exponent vectors of $a_{\{2,3\}}(\hat{h})h_{3n+1}$ and $a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}$, respectively. By Theorem 2.10, it is therefore enough to show that the two summands have exactly one negative connected component. By Proposition 3.12 and Proposition 3.13 the supports of $a_{\{2,3\}}(\hat{h})h_{3n+1}$ and $a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}$ have a strict separating hyperplane. Therefore, both $a_{\{2,3\}}(\hat{h})h_{3n+1}$ and $a_{\{3\}}(\hat{h})h_{3n+1}h_{3n+2}$ have one negative connected component and hence, so does b_1. □

4. **Weakly Irreversible Phosphorylation Networks**

We now study the family of reduced weakly irreversible phosphorylation networks $(\mathcal{F}_{2,n})$. This section has several results analogous to Section 3. First, in Section 4.1 we give the basic set up that allow us to compute the critical polynomial of the reduced networks. In Section 4.2 we use this to show that this network has multistationarity for all $n \geq 2$ and the parameter region of multistationarity is connected.

4.1. **Computation of Critical Polynomial.**

The family of networks $(\mathcal{F}_{2,n})$ is obtained from $(\mathcal{N}_{2,n})$ by removing reversible reactions as in Theorem 2.2. We fix the order of the species as $K, F, S_0, Y_1, Y_2, U_1, U_2, S_1, \ldots, Y_{2n-1}, Y_{2n}, U_{2n-1}, U_{2n}, S_n$ to obtain the stoichiometric matrix N_n and the reactant matrix A_n as below. For $n = 1$ these matrices are given by:

$$N_1 = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}, \qquad \text{and} \qquad A_1 = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}. \tag{25}$$

For general $n \geq 2$, we can write these matrices recursively as:

$$N_n = \begin{pmatrix} N_{n-1} & P_1 \\ 0_{5\times(8n-8)} & P_2 \end{pmatrix} \in \mathbb{R}^{(5n+3)\times 8n}, \quad A_n = \begin{pmatrix} A_{n-1} & Q_1 \\ 0_{5\times(8n-8)} & Q_2 \end{pmatrix} \in \mathbb{R}^{(5n+3)\times 8n}, \tag{26}$$

where

$$P_1 = \begin{pmatrix} -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \in \mathbb{R}^{(5n-2)\times 8}, P_2 = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -1 & -1 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{5\times 8}, \tag{27}$$

and

$$Q_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{(5n-2)\times 8}, \quad Q_2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{5\times 8}. \tag{28}$$

Remark 4.1. The matrices P_1 and Q_1 have $5n-5$ rows with all entries zero. Note that the matrix P_2 has full rank and the rank of N_1 is 5. From the recursive relation it follows that $\operatorname{rank}(N_n) = 5n$. Equivalently, the dimension of left kernel of N_n is 3 and $\dim(\operatorname{im}(N_n)) = 5n$.

Using the above matrices, we compute a conservation law matrix and an extreme matrix for $(\mathcal{F}_{2,n})$ in Lemma 4.2 and Proposition 4.3 respectively.

Lemma 4.2. *Given the stoichiometric matrix N_n in (26), a conservation law matrix W_n is given by:*

$$W_n := \begin{pmatrix} \mathrm{Id}_3 & \underbrace{\overline{W} \quad \cdots \quad \overline{W}}_{n} \end{pmatrix} \in \mathbb{R}^{3\times(5n+3)}, \tag{29}$$

where Id_3 is the identity matrix of size 3 and

$$\overline{W} := \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \in \mathbb{R}^{3\times 5}.$$

Proof. Let $W_0 := \mathrm{Id}_3$, then using $W_n = \begin{pmatrix} W_{n-1} & \overline{W} \end{pmatrix}$, for $n \geq 2$ we have

$$W_n N_n = \begin{pmatrix} W_{n-1}N_{n-1} & W_{n-1}P_1 + \overline{W}P_2 \end{pmatrix}.$$

Simple computation gives:

$$W_{n-1}P_1 = \begin{pmatrix} -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix} \quad \text{and} \quad \overline{W}P_2 = \begin{pmatrix} 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}.$$

For $n = 1$, we get $W_1N_1 = W_0P_1 + \overline{W}P_2 = 0$. We now assume that the claim holds for all $k \in \mathbb{N}$ such that $1 \leq k \leq n-1$. By assumption $W_{n-1}N_{n-1}$ is a zero matrix and hence, $W_nN_n = 0_{3\times 8n}$ and since the rank of W_n and the dimension of the left kernel of N_n are both 3 for all n by Remark 4.1, it follows that W_n is a conservation law matrix for N_n. □

Proposition 4.3. *An extreme matrix of $(\mathcal{F}_{2,n})$ has the following recursive form*

$$E_n = \left(\begin{array}{c|c} E_{n-1} & 0_{(8n-8)\times 3} \\ \hline 0_{8\times(3n-3)} & E_1 \end{array}\right) \in \mathbb{R}^{8n\times 3n}, \tag{30}$$

where

$$E_1 = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}^\top \in \mathbb{R}^{8\times 3}.$$

Moreover, the column vectors $E_n^{(1)}, \dots, E_n^{(3n)} \in \mathbb{R}^{8n}$ *of* E_n *form a basis of* $\ker(N_n)$.

Proof. Using (26), a simple computation shows that $E_n^{(1)}, \dots, E_n^{(3n)} \in \ker(N_n) \cap \mathbb{R}^{8n}_{\geq 0}$. Since by Remark 4.1

$$\dim \ker(N_n) = 8n - \dim \operatorname{im}(N_n) = 8n - 5n = 3n$$

and $E_n^{(1)}, \dots, E_n^{(3n)}$ are linearly independent, they form a basis of $\ker(N_n)$.

Furthermore, $E_1^{(1)}, E_1^{(2)}, E_1^{(3)}$ satisfy (5) and since the blocks of E_n have pairwise disjoint supports, $E_n^{(1)}, \dots, E_n^{(3n)}$ also satisfy (5). Thus, E_n is an extreme matrix for $(\mathcal{F}_{2,n})$ by Proposition 2.1. □

The matrix N_n' is obtained by removing the first three rows of N_n. We will now compute a Gale dual $D_n(\lambda)$ (Theorem 4.5) of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$ and $\delta_n(\lambda)$ (Proposition 4.6) for the network $(\mathcal{F}_{2,n})$.

In the following, we view $\lambda = (\lambda_1, \dots, \lambda_{3n})$ as symbolic variables. In Lemma 4.4, we first compute the maximal minor $\det([N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I})$ for $I = \{4, \dots, 3n+3\}$ and $I^c = \{1, 2, 3\}$.

Lemma 4.4. *For a fixed* n*, let* $s = 5n$*,* $I = \{4, \dots, 5n+3\}$*. Then,*

$$\det([N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I}) = (-1)^n \prod_{k=0}^{n-1} (\lambda_{3k+1} + \lambda_{3k+2})(\lambda_{3k+1} + \lambda_{3k+3})\lambda_{3k+1}^3.$$

In particular, $\ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top) \in \mathbb{R}(\lambda)^{5n}$ *has dimension* 3.

Proof. Let A_n' be obtained from A_n after removing the first three rows. Then we have,

$$[N_n' \operatorname{diag}(E_n\lambda)A_n^\top]_{[s],I} = N_n' \operatorname{diag}(E_n\lambda)A_n'^\top.$$

Consider the matrix $X_n \in \mathbb{R}^{5n\times 5n}$, $n \geq 2$ given by

$$(X_n)_{ij} := \begin{cases} 1 \text{ if } i = j, \\ 1 \text{ if } k < n,\ i = 5k,\ j > 5k, \\ 0 \text{ else.} \end{cases}$$

Furthermore, we set $X_1 := \mathrm{Id}_5$. Multiplying by X_n is the same as adding the rows $5k+1, 5k+2, 5k+3, 5k+4$ and $5k+5$ to the row $5k$ for each $k = 1, \dots, n-1$. The matrix X_n can be written recursively as

$$X_n = \begin{pmatrix} X_{n-1} & X_{n-1}' \\ 0_{5\times(5n-5)} & \mathrm{Id}_5 \end{pmatrix} \qquad \text{with} \quad (X_{n-1}')_{ij} := \begin{cases} 1 \text{ if } i = 5k, k = 1, \dots, n-1, \\ 0 \text{ else.} \end{cases}$$

It follows that for all $n > 1$ the matrix X_nN_n' can be written recursively as:

$$X_nN_n' = \begin{pmatrix} X_{n-1}N_{n-1}' & X_{n-1}P_1' + X_{n-1}'P_2 \\ 0_{5\times(8n-8)} & P_2 \end{pmatrix} = \begin{pmatrix} X_{n-1}N_{n-1}' & 0_{(5n-5)\times 8} \\ 0_{5\times(8n-8)} & P_2 \end{pmatrix},$$

where P_1 and P_2 is defined as in (27) and P_1' is obtained by removing the first three rows of P_1. We define the matrices $C_n \in \mathbb{R}(\lambda)^{5n\times 5n}$ recursively as:

$$C_n := \left(\begin{array}{c|c} C_{n-1} & 0_{5\times 5} \\ \hline 0_{5\times(5n-10)} \quad V_n & Z_n \end{array}\right) \in \mathbb{R}(\lambda)^{5n\times 5n} \quad \text{and} \quad C_1 := Z_1$$

where

$$V_n := \begin{pmatrix} 0 & 0 & 0 & 0 & \lambda_{3n-2} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{3n} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad Z_n := \begin{pmatrix} -\lambda_{3n-2} & 0 & 0 & 0 & 0 \\ \lambda_{3n-2} & -\lambda_{3n-2}-\lambda_{3n-1} & 0 & 0 & \lambda_{3n-1} \\ 0 & 0 & -\lambda_{3n-2} & 0 & \lambda_{3n-2} \\ 0 & 0 & \lambda_{3n-2} & -\lambda_{3n-2}-\lambda_{3n} & 0 \\ 0 & \lambda_{3n-2}+\lambda_{3n-1} & 0 & 0 & -\lambda_{3n-2}-\lambda_{3n-1} \end{pmatrix}.$$

Let $\omega = [\lambda_{3n-2}, \lambda_{3n-1}, \lambda_{3n}]^\top$. A simple computation shows that $Z_n = P_2 \operatorname{diag}(E_1\omega) Q_2^\top$ and hence, for $n = 1$, we have $X_1 N_1' \operatorname{diag}(E_1\lambda) A_1'^\top = C_1$. Assume that $X_k N_k' \operatorname{diag}(E_k\lambda) A_k'^\top = C_k$ for all $1 \le k < n$, and let $\hat{\lambda} = (\lambda_1, \ldots, \lambda_{3n-3})$. Then using the block structure of N_n, A_n, and E_n, we get

$$\begin{aligned} X_n N_n' \operatorname{diag}(E_n\lambda) A_n'^\top &= \begin{pmatrix} X_{n-1} N_{n-1}' \operatorname{diag}(E_{n-1}\hat{\lambda}) & 0_{(5n-5)\times 8} \\ 0_{5\times(8n-8)} & P_2 \operatorname{diag}(E_1\omega) \end{pmatrix} \begin{pmatrix} A_{n-1}'^\top & 0_{(8n-8)\times 5} \\ Q_1'^\top & Q_2^\top \end{pmatrix} \\ &= \left(\begin{array}{cc|c} & C_{n-1} & 0_{5\times 5} \\ 0_{5\times(5n-10)} & & V_n \\ & & Z_n \end{array}\right) = C_n, \end{aligned}$$

where Q_1, Q_2 are the matrices from (28) and Q_1' is obtained from Q_1 by removing the first three rows. Since $\det(X_n) = 1$, it follows

$$\begin{aligned} \det(N_n' \operatorname{diag}(E_n\lambda) A_n'^\top) = \det(C_n) &= \prod_{k=1}^{n} \det(Z_k) \\ &= (-1)^n \prod_{k=0}^{n-1} (\lambda_{3k+1}+\lambda_{3k+2})(\lambda_{3k+1}+\lambda_{3k+3}) \lambda_{3k+1}^3. \end{aligned}$$

The second part of the lemma holds since the $5n \times 5n$ maximal minor $\det([N_n' \operatorname{diag}(E_n\lambda) A_n^\top]_{[s],I})$ is non-zero for all n. □

Theorem 4.5. *Let N_n, A_n and E_n be the matrices associated with the network $(\mathcal{F}_{2,n})$ as described in (26) and (30). Let $\lambda = (\lambda_1, \ldots, \lambda_{3n})$ and for $k = 0, \ldots, n-1$ consider the matrices*

$$D^{(k)} := \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ -k & \frac{-k\lambda_{3k+1}-(k+1)\lambda_{3k+2}}{\lambda_{3k+1}+\lambda_{3k+2}} & -k & \frac{-k\lambda_{3k+1}+(1-k)\lambda_{3k+3}}{\lambda_{3k+1}+\lambda_{3k+3}} & -(k+1) \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad D^{(-1)} := \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ and}$$

$$D_n^\top(\lambda) := \begin{pmatrix} D^{(-1)} & D^{(0)} & \ldots & D^{(n-1)} \end{pmatrix} \in \mathbb{R}(\lambda)^{3\times(5n+3)}.$$

Then $D_n(\lambda)$ is a Gale dual matrix of $N_n' \operatorname{diag}(E_n\lambda) A_n^\top$.

Proof. A vector $v \in \mathbb{R}(\lambda)^{5n+3}$ lies in the kernel of $N_n' \operatorname{diag}(E_n\lambda) A_n^\top$ if and only if $\operatorname{diag}(E_n\lambda) A_n^\top v \in \ker(N_n')$. For $v \in \mathbb{R}(\lambda)^{5n+3}$, let

$$\operatorname{diag}(E_n\lambda) A_n^\top v = \begin{bmatrix} \omega_0, \ldots, \omega_{n-1} \end{bmatrix}^\top$$

such that for $i = 0, \ldots, n-1$, $\omega_i \in \mathbb{R}(\lambda)^8$ is given by:

$$\begin{aligned} \Big[&\lambda_{3k+1}(v_1+v_{5k+3}) \quad \lambda_{3k+1}v_{5k+4} \quad (\lambda_{3k+1}+\lambda_{3k+2})v_{5k+5} \quad \lambda_{3k+2}(v_1+v_{5k+8}) \\ &\lambda_{3k+1}(v_2+v_{5k+8}) \quad \lambda_{3k+1}v_{5k+6} \quad (\lambda_{3k+1}+\lambda_{3k+3})v_{5k+7} \quad \lambda_{3k+3}(v_2+v_{5k+3}) \Big]. \end{aligned} \tag{31}$$

Moreover, since the columns of E_n span $\ker(N_n')$, there exists $\mu \in \mathbb{R}(\lambda)^{3n}$ such that $\operatorname{diag}(E_n\lambda) A_n^\top v = E_n\mu$. Let

$$E_n\mu = \begin{bmatrix} u_0, \ldots, u_{n-1} \end{bmatrix}^\top$$

such that for $i = 0, \ldots, n-1$

$$u_i = \begin{bmatrix} \mu_{3i+1} & \mu_{3i+1} & \mu_{3i+1}+\mu_{3i+2} & \mu_{3i+2} & \mu_{3i+1} & \mu_{3i+1} & \mu_{3i+1}+\mu_{3i+3} & \mu_{3i+3} \end{bmatrix} \in \mathbb{R}(\lambda)^8. \tag{32}$$

Comparing equations (31) and (32), we get that $v \in \ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$ if it satisfies the following relations for all $i = 0, \dots, n-1$:

$$\begin{aligned}
&v_1 + v_{5k+3} = v_{5k+4}, \quad v_1 + v_{5k+3} = v_2 + v_{5k+8}, \quad v_1 + v_{5k+3} = v_{5k+6},\\
&(\lambda_{3k+1} + \lambda_{3k+2})v_{5k+5} = \lambda_{3k+1}(v_1 + v_{5k+3}) + \lambda_{3k+2}(v_1 + v_{5k+8}),\\
&(\lambda_{3k+1} + \lambda_{3k+3})v_{5k+7} = \lambda_{3k+1}(v_1 + v_{5k+3}) + \lambda_{3k+3}(v_2 + v_{5k+3}).
\end{aligned}$$

It is easy to check that the columns of $D_n(\lambda)$ are linearly independent and satisfy these relations. Moreover, since $\ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$ has dimension 3 (Lemma 4.4), $D_n(\lambda)$ is a Gale dual matrix of $\ker(N_n' \operatorname{diag}(E_n\lambda)A_n^\top)$. □

Using Lemma 4.4, we compute $\delta_n(\lambda)$, which is the final ingredient needed to calculate the critical polynomial of $(\mathcal{F}_{2,n})$ as in Theorem 2.7.

Proposition 4.6. *With the choice of Gale dual matrix in Theorem 4.5, from* (11) *we get*

$$\delta_n(\lambda) = \prod_{k=0}^{n-1} (\lambda_{3k+1} + \lambda_{3k+2})(\lambda_{3k+1} + \lambda_{3k+3})\lambda_{3k+1}^3.$$

Proof. By Lemma 2.4, $\delta_n(\lambda)$ in (11) is independent of the choice of index set. For $I = \{4, \dots, 5n+3\}$ we have

$$\operatorname{sgn}(\tau_I) = (-1)^n \quad \text{and} \quad \det([D_n^\top(\lambda)]_{[3],I^c}) = 1.$$

Now, using Lemma 4.4 and Corollary 2.6, we obtain the result. □

4.2. **Connectivity of Multistationarity Region of $(\mathcal{N}_{2,n})$.** Using the conservation law matrix W_n (29), the Gale dual matrix $D_n(\lambda)$ of $N_n' \operatorname{diag}(E_n\lambda)A_n^\top$ from Theorem 4.5, and applying Theorem 2.7 we write the critical polynomial of $(\mathcal{F}_{2,n})$ as

$$q_n(h,\lambda) = \sum_{\substack{I \subseteq [5n+3] \\ |I|=3}} \delta_n(\lambda) \det\left([W_n]_{[3],I}\right) \det\left([D_n^\top(\lambda)]_{[3],I}\right) \prod_{i \in I^c} h_i. \tag{33}$$

Unlike for $(\mathcal{F}_{1,n})$, the matrix $D_n^\top(\lambda)$ depends on λ in this case. Since the product

$$\alpha(\lambda)_{n,I} := \det\left([W_n]_{[3],I}\right) \det\left([D_n^\top(\lambda)]_{[3],I}\right) \tag{34}$$

is not independent of λ, we cannot factor q_n in terms of polynomial in h and polynomial in λ like we did in Section 3.2. Here, we study the whole polynomial q_n. Going forward, for ease of notations, we will only write $\alpha_{n,I}$ for $\alpha(\lambda)_{n,I}$.

From $n-1$ to n, the critical polynomial depends on eight new variables $h_{5n-1}, \dots, h_{5n+3}$, and $\lambda_{3n-2}, \lambda_{3n-1}, \lambda_{3n}$. For any subset $J \subseteq [5]$, we define

$$\hat{J} := \{5n-2+j \mid j \in J\} \quad \text{and} \quad \hat{J}^c := \{5n-2+j \mid j \in [5] \setminus J\}, \tag{35}$$

which will be used to index the variables $h_{5n-1}, \dots, h_{5n+3}$. For instance, for $J = [5]$ we have $\{h_{5n-1}, \dots, h_{5n+3}\} = \{h_i \mid i \in \hat{J}\}$. Using this notation, we write q_n as

$$q_n(h,\lambda) = \sum_{\substack{J \subseteq [5] \\ |J| \leq 3}} \Big(a_J(\hat{h},\lambda) \prod_{i \in \hat{J}^c} h_i \Big), \tag{36}$$

where $\hat{h} = (h_1, \dots, h_{5n-2})$. From (33), it follows that the a_J's have the following form

$$a_J(\hat{h},\lambda) = \sum_{\substack{I \subseteq [5n+3], |I|=3 \\ \hat{J} \subseteq I,\, \hat{J}^c \cap I = \emptyset}} \delta_n(\lambda) \det\left([W_n]_{[3],I}\right) \det\left([D_n(\lambda)]_{I,[3]}\right) \prod_{i \in I^c \setminus \hat{J}^c} h_i. \tag{37}$$

First, we relate $a_\emptyset$ to the critical polynomial of $(\mathcal{F}_{2,n-1})$. In what follows, we set $\hat{\lambda} = (\lambda_1, \dots, \lambda_{3n-3})$.

Proposition 4.7. *For $n \geq 2$ and $J = \emptyset$,*

$$a_J(\hat{h}, \lambda) = q_{n-1}(\hat{h}, \hat{\lambda})(\lambda_{3n-2} + \lambda_{3n-1})(\lambda_{3n-2} + \lambda_{3n})\lambda_{3n-2}^3,$$

where $q_{n-1}(\hat{h}, \hat{\lambda})$ denotes the critical polynomial of $(\mathcal{F}_{2,n-1})$.

Proof. The summands in (37) depend on the subsets $I \subseteq [5n-2]$ with $|I| = 3$. Note that the following sets coincide

$$\big\{I \subseteq [5n+3] \mid |I| = 3 \text{ and } \{5n-1, 5n, 5n+1, 5n+2, 5n+3\} \cap I = \emptyset\big\} = \big\{I \subseteq [5(n-1)+3] \mid |I| = 3\big\}.$$

The proof now follows from (33) and (37) using the block structures of W_n and $D_n(\lambda)$. □

In (37) the cardinality of I is 3. Moreover, when $J \neq \emptyset$, the intersection $I \cap \{5n-1, 5n, 5n+1, 5n+2, 5n+3\} \neq \emptyset$. In the next lemma, we list cases when $\alpha_{n,I} = 0$ from (34). The proof of Lemma 4.8 follows directly by considering the columns of the matrices $[W_n]_{[3],I}$ and $[D_n^\top(\lambda)]_{[3],I}$.

Lemma 4.8. *Let $I \subset [5n+3]$ with $|I| = 3$ and $\ell, \ell' \in [n-1] \cup \{0\}$.*

(i) If $5\ell + 4, 5\ell' + 5 \in I$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.
(ii) If $5\ell + 6, 5\ell' + 7 \in I$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.
(iii) If $5\ell + 4, 5\ell + 6 \in I$, then $\det([D_n^\top(\lambda)]_{[3],I}) = 0$ and hence, $\alpha_{n,I} = 0$.
(iv) If $5\ell + k, 5\ell' + k \in I$ for $k \in \{4, 5, 6, 7, 8\}$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.
(v) If $3, 5\ell + 8 \in I$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.
(vi) If $I \in \big\{\{1, 3, 5\ell+4\}, \{1, 3, 5\ell+5\}, \{1, 5\ell+4, 5\ell'+8\}, \{1, 5\ell+5, 5\ell'+8\}$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.
(vii) If $I \in \big\{\{2, 3, 5\ell+6\}, \{2, 3, 5\ell+7\}, \{2, 5\ell+6, 5\ell'+8\}, \{2, 5\ell+7, 5\ell'+8\}$, then $\det\big([W_n]_{[3],I}\big) = 0$ and hence, $\alpha_{n,I} = 0$.

Proposition 4.9. *For all $n \in \mathbb{N}$, if $\{1, 2\} \subseteq J, \{3, 4\} \subseteq J$ or $\{1, 3\} \subseteq J$, then the polynomial a_J is the zero polynomial. Equivalently, $a_{\{1,2\}}, a_{\{1,3\}}, a_{\{3,4\}}, a_{\{1,2,3\}}, a_{\{1,2,4\}}, a_{\{1,2,5\}}, a_{\{1,3,4\}}, a_{\{1,3,5\}}, a_{\{2,3,4\}}$, and $a_{\{3,4,5\}}$ are zero polynomials.*

Proof. The result follows from Lemma 4.8(i),(ii),(iii). □

To determine the signs of the coefficients of q_n and to find vertices of the Newton polytope the following lemmas will be particularly helpful.

Lemma 4.10. *For all $I \subseteq [5n+3]$ with $|I| = 3$, there exist $p_I \in \mathbb{R}[\lambda]$ and g_I in*

$$\mathcal{G} = \Big\{1,\ \lambda_{3k+1} + \lambda_{3k+2},\ \lambda_{3k+1} + \lambda_{3k+3},\ (\lambda_{3k+1} + \lambda_{3k+2})(\lambda_{3k'+1} + \lambda_{3k'+3}) \mid k, k' \in \{0, \dots, n-1\}\Big\}$$

such that $\alpha_{n,I} = \frac{p_I}{g_I}$, and $\sigma(p_I) \subseteq \sigma(g_I)$.

Proof. By Lemma 4.8(iv) if $\alpha_{n,I} \neq 0$, then $[D_n^\top(\lambda)]_{[3],I}$ has at most two entries involving λ. These entries are of the following form:

$$\frac{-k\lambda_{3k+1} - (k+1)\lambda_{3k+2}}{\lambda_{3k+1} + \lambda_{3k+2}} \quad \text{and} \quad \frac{-k\lambda_{3k+1} + (1-k)\lambda_{3k+3}}{\lambda_{3k+1} + \lambda_{3k+3}}. \tag{38}$$

Hence, g_I can be chosen from $\mathcal{G}$. The inclusion $\sigma(p_I) \subseteq \sigma(g_I)$ follows from the observation that for all entries of $[D_n^\top(\lambda)]_{[3],I}$, the support of the numerator is contained in the support of the denominator. □

We recall from Proposition 4.6 that $\delta_n(\lambda)$ is given by

$$\delta_n(\lambda) = \prod_{k=0}^{n-1} (\lambda_{3k+1} + \lambda_{3k+2})(\lambda_{3k+1} + \lambda_{3k+3})\lambda_{3k+1}^3.$$

Lemma 4.11. *For every $n \in \mathbb{N}$, consider the vector*

$$\nu_n := \begin{bmatrix} \nu & \dots & \nu \end{bmatrix}^\top \in \mathbb{R}^{3n}, \quad \text{where} \quad \nu := \begin{bmatrix} 5 & 0 & 0 \end{bmatrix}. \tag{39}$$

Then ν_n is a vertex of $\mathrm{NP}(\delta_n(\lambda))$. Furthermore, for $n \geq 2$ and $I \subseteq [5n-2]$, if $(\hat{z}_I, \nu_{n-1}) \in \sigma_-(q_{n-1})$, then $(z_I, \nu_n) \in \sigma_-(q_n)$, where $\hat{z}_I \in \mathbb{R}^{5n-2}$ is obtained from $z_I \in \mathbb{R}^{5n}$ by removing the last 5 coordinates.

Proof. In each monomial of $\delta_n(\lambda)$, the degree of the variable λ_{3k+1} is at most 5 for each $k = 0, \dots, n-1$. Moreover, there is a unique monomial in $\delta_n(\lambda)$ which is divisible by $\prod_{i=0}^{n-1} \lambda_{3k+1}^5$. This implies that ν_n is a vertex of $\mathrm{NP}(\delta_n(\lambda))$. The second part follows from Proposition 4.7. □

For $J \subseteq [5]$ and $|J| \leq 3$, let a_J and $\hat{J}$ be as defined in (37) and in (35). Let $I \subseteq [5n+3]$, $|I| = 3$ such that $\hat{J} \subseteq I$ and $\hat{J}^c \cap I = \emptyset$. If $\mu \in \sigma\big(\delta_n(\lambda)\alpha_{n,I}\big)$, then from the definition of a_J we get that $(z_I, \mu) \in \sigma(a_J \prod_{i \in \hat{J}^c} h_i)$. Moreover, we have the following result.

Lemma 4.12. *Let J and I be as defined above, and write $\alpha_{n,I} = \frac{p_I}{g_I}$. If p_I has only positive (resp. negative) coefficients, then $(z_I, \mu) \in \sigma_+(a_J \prod_{i \in \hat{J}^c} h_i)$ (resp. $(z_I, \mu) \in \sigma_-(a_J \prod_{i \in \hat{J}^c} h_i)$) for all $\mu \in \sigma\big(\delta_n(\lambda)\alpha_{n,I}\big)$.*

Proof. By Lemma 4.10, g_I divides $\delta_n(\lambda)$. Note that $\frac{\delta_n(\lambda)}{g_I}$ has only positive coefficients. Since p_I has only positive (resp. negative) coefficients, then

$$\delta_n(\lambda)\alpha_{n,I} \prod_{i \in I^c} h_i = \tfrac{\delta_n(\lambda)}{g_I} p_I \prod_{i \in I^c} h_i \tag{40}$$

has only positive (resp. negative) coefficients. Thus, $(z_I, \mu) \in \sigma_+(a_J \prod_{i \in \hat{J}^c} h_i)$ (resp. $(z_I, \mu) \in \sigma_-(a_J \prod_{i \in \hat{J}^c} h_i)$) for all $\mu \in \sigma\big(\delta_n(\lambda)\alpha_{n,I}\big)$. □

The polynomial q_n has $8n+3$ many variables, $h_1, \dots, h_{5n+3}, \lambda_1, \dots, \lambda_{3n}$. Thus, $\sigma(q_n) \subseteq \mathbb{R}^{5n+3} \times \mathbb{R}^{3n}$ Let pr_1 (resp. pr_2) denote the coordinate projection from $\mathbb{R}^{5n+3} \times \mathbb{R}^{3n}$ onto $\mathbb{R}^{5n+3}$ (resp. $\mathbb{R}^{3n}$). The following result finds vertices of the projected Newton polytope of a_J's.

Lemma 4.13. *Let $\mu \in \mathrm{Vert}\big(\mathrm{NP}(\delta_n(\lambda))\big)$, and $\mathcal{I} \subseteq \{I \subseteq [5n+3] \mid |I| = 3\}$. Consider the polynomial*

$$f := \sum_{I \in \mathcal{I}} \delta_n(\lambda)\alpha_{n,I} \prod_{i \in I^c} h_i.$$

If there exists $I_0 \in \mathcal{I}$ such that $(z_{I_0}, \mu) \in \sigma(f)$, then μ is a vertex of $\mathrm{pr}_2(\mathrm{NP}(f))$.

Proof. Let $\alpha_{n,I} = \frac{p_I}{g_I}$. By Lemma 4.10, we have $\sigma(p_I) \subseteq \sigma(g_I)$. Since g_I divides $\delta_n(\lambda)$, it follows that $\sigma(\delta_n(\lambda)\alpha_{n,I}) \subseteq \sigma(\delta_n(\lambda))$ for all $I \in \mathcal{I}$. This implies that

$$\mathrm{pr}_2(\sigma(f)) \subseteq \bigcup_{I \in \mathcal{I}} \sigma(\delta_n(\lambda)\alpha_{n,I}) \subseteq \sigma(\delta_n(\lambda)). \tag{41}$$

Since μ is a vertex of $\mathrm{NP}(\delta_n(\lambda))$, there exists $v \in \mathbb{R}^{3n}$ such that $v \cdot \omega < v \cdot \mu$ for all $\omega \in \mathrm{NP}(\delta_n(\lambda)) \setminus \{\mu\}$. From $(z_{I_0}, \mu) \in \sigma(f)$ follows that $\mu \in \mathrm{pr}_2(\sigma(f))$. Using (41), we conclude that μ is a vertex of $\mathrm{pr}_2(\mathrm{NP}(f))$. □

In the following, we investigate the signed supports of different a_J's.

Proposition 4.14. *For every $n \in \mathbb{N}$, if $1 \in J$ or $2 \in J$, then a_J has only non-negative coefficients. Furthermore, a_J is not the zero polynomial for $J = \{2, 3, 5\}$.*

Proof. If $1 \in J$, write $I = \{r_1, r_2, 5n-1\}$. For $\ell, \ell' \in \{0, \dots, n-1\}$, by Lemma 4.8 $\alpha_{n,I} \neq 0$ in the following cases:

(i) When $(r_1, r_2) = (1, 2)$ and $(r_1, r_2) = (2, 3)$, we get $\alpha_{n,I} = 1$ and $\alpha_{n,I} = n$ respectively.

(ii) When $(r_1, r_2) \in \{(1, 5\ell+6), (2, 5\ell+6), (2, 5\ell+8), (3, 5\ell+6), (5\ell+6, 5\ell'+8)\}$, we have $\alpha_{n,I} = n - \ell - 1$.

(iii) When $(r_1, r_2) \in \{(1, 5\ell+7), (2, 5\ell+7), (3, 5\ell+7), (5\ell+7, 5\ell'+8)\}$, we obtain $\alpha_{n,I} = (n-\ell-1) + \frac{\lambda_{3\ell+3}}{\lambda_{3\ell+1}+\lambda_{3\ell+3}}$.

Since the numerator of all these $\alpha_{n,I}$ has only non-negative coefficients, from Lemma 4.12 it follows that a_J with $1 \in J$ has only non-negative coefficients.

If $2 \in J$, $I = \{r_1, r_2, 5n\}$. For $\ell, \ell' \in \{0, \dots, n-1\}$, by Lemma 4.8 $\alpha_{n,I} \neq 0$ for following cases:

(i) When $(r_1, r_2) = (1, 2)$, we get $\alpha_{n,I} = 1$.

(ii) When $(r_1, r_2) = (2, 3)$, we get $\alpha_{n,I} = n + \frac{\lambda_{3n-1}}{\lambda_{3n-2}+\lambda_{3n-1}}$.

(iii) When $(r_1, r_2) \in \{(1, 5\ell+6), (2, 5\ell+6), (2, 5\ell+8), (3, 5\ell+6), (5\ell+6, 5\ell'+8)\}$, we have $\alpha_{n,I} = (n-\ell-1) + \frac{\lambda_{3n-1}}{\lambda_{3n-2}+\lambda_{3n-1}}$.

(iv) When $(r_1, r_2) \in \{(1, 5\ell+7), (2, 5\ell+7), (3, 5\ell+7), (5\ell+7, 5\ell'+8)\}$, we obtain $\alpha_{n,I} = (n-\ell-1) + \frac{\lambda_{3\ell+3}}{\lambda_{3\ell+1}+\lambda_{3\ell+3}} + \frac{\lambda_{3n-1}}{\lambda_{3n-2}+\lambda_{3n-1}}$.

Since the numerator of $\alpha_{n,I}$ in the above cases has only non-negative coefficients, using again Lemma 4.12, we conclude that a_J has only non-negative coefficients if $2 \in J$.

The second part of the proposition follows from case (iii) listed above and Lemma 4.12. □

Proposition 4.14 implies the following corollary.

Corollary 4.15. *The polynomials* $a_{\{1\}}, a_{\{2\}}, a_{\{1,4\}}, a_{\{1,5\}}, a_{\{2,3\}}, a_{\{2,4\}}, a_{\{2,5\}}, a_{\{1,4,5\}}, a_{\{2,3,5\}}$, *and* $a_{\{2,4,5\}}$ *have non-negative coefficients for all* $n \in \mathbb{N}$.

Next, in Corollary 4.16 we show that q_n satisfies the closure property (P2) and to show connectivity it is enough to consider the following polynomial

$$\begin{aligned} b(h,\lambda) := \Big(&a_{\emptyset}(\hat{h},\lambda)h_{5n+1}h_{5n+2}h_{5n+3} + a_{\{3\}}(\hat{h},\lambda)h_{5n+2}h_{5n+3} + a_{\{4\}}(\hat{h},\lambda)h_{5n+1}h_{5n+3} \\ &+ a_{\{5\}}(\hat{h},\lambda)h_{5n+1}h_{5n+2} + a_{\{3,5\}}(\hat{h},\lambda)h_{5n+2} + a_{\{4,5\}}(\hat{h},\lambda)h_{5n+1}\Big)h_{5n-1}h_{5n}. \end{aligned} \tag{42}$$

Corollary 4.16. *For* $n \geq 2$, q_n *satisfies the closure property (P2). Additionally, if* b *in* (42) *has one negative connected component, then* q_n *also has exactly one negative connected component.*

Proof. To prove this result, we first claim that there exists a proper face of $\mathrm{NP}(q_n)$ that contains all the negative exponent vectors of q_n. Since in all the terms of q_n the variables h_{5n-1}, h_{5n} have exponents either 0 or 1, it follows that

$$(e_{5n-1} + e_{5n}) \cdot \omega \leq 2 \quad \text{for all } \omega \in \sigma(q_n).$$

From Proposition 4.7, it follows that the above inequality is attained for some $\omega' \in \sigma(q_n)$. Since the polynomial $a_{\{2,3,5\}}$ is non-zero by Proposition 4.14, there exists $\omega'' \in \sigma(q_n)$ that is not contained in the face $\mathrm{NP}_{e_{5n-1}+e_{5n}}(q_n)$. We conclude that $\mathrm{NP}_{e_{5n-1}+e_{5n}}(q_n)$ is a proper face of $\mathrm{NP}(q_n)$.

It is easy to check that q_n restricted to $\mathrm{NP}_{e_{5n-1}+e_{5n}}(q_n)$ is exactly the polynomial given by $b(h,\lambda)$. From Proposition 4.9 and Proposition 4.14 it follows that each monomial of q_n with negative coefficient is divisible by $h_{5n-1}h_{5n}$. Therefore, $\sigma_-(q_n) \subseteq \mathrm{NP}_{e_{5n-1}+e_{5n}}(q_n) = \mathrm{NP}(b)$. The corollary now follows from Theorem 2.9. □

Next we will consider different summands of $b(h,\lambda)$.

Proposition 4.17. *For* $n \geq 2$, *there exists a strict separating hyperplane of the signed support for both* $a_{\{3\}}(\hat{h},\lambda)h_{5n-1}h_{5n}h_{5n+2}h_{5n+3}$ *and* $a_{\{3,5\}}(\hat{h},\lambda)h_{5n-1}h_{5n}h_{5n+2}$. *Furthermore,*

(a) $(z_{\{4,8,5n+1\}}, \mu) \in \sigma_-(a_{\{3\}}(\hat{h},\lambda)h_{5n-1}h_{5n}h_{5n+2}h_{5n+3})$ *for all* $\mu \in \sigma(\delta_n(\lambda))$.

(b) $(z_{\{4,5n+1,5n+3\}}, \mu) \in \sigma_-(a_{\{3,5\}}(\hat{h},\lambda)h_{5n-1}h_{5n}h_{5n+2})$ *for all* $\mu \in \sigma(\delta_n(\lambda))$.

Proof. If $J = \{3\}$, let $I = \{r_1, r_2, 5n+1\}$. By Lemma 4.8, $\alpha_{n,I} \neq 0$ in following cases, where $\ell, \ell' \in \{0, \dots, n-2\}$:

(i) When $(r_1, r_2) = (1, 2)$, we have $\alpha_{n,I} = 1$.
(ii) When $(r_1, r_2) = (1, 3)$ and $(r_1, r_2) = (1, 5\ell+8)$, we get $\alpha_{n,I} = -n+1$ and $\alpha_{n,I} = -n+\ell+2$ respectively.
(iii) When $(r_1, r_2) \in \{(1, 5\ell+4), (2, 5\ell+4), (3, 5\ell+4), (5\ell+4, 5\ell'+8)\}$, we have $\alpha_{n,I} = -n+\ell+1$.
(iv) When $(r_1, r_2) \in \{(1, 5\ell+5), (2, 5\ell+5), (3, 5\ell+5), (5\ell+5, 5\ell'+8)\}$, we obtain $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3\ell+1}+(-n+\ell+2)\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}}$.

From Lemma 4.12, it follows

$$\sigma_+(a_{\{3\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+2}h_{5n+3}) = \{(z_{\{1,2,5n+1\}}, \mu) \mid \mu \in \sigma(\delta_n(\lambda))\}.$$

Since $(e_1 + e_2) \cdot (z_I, \mu) \geq 0$ for all $I \subseteq [5n+3]$, $|I| = 3$ and $\mu \in \sigma(\delta_n(\lambda)\alpha_{n,I})$, the hyperplane $\mathcal{H}_{e_1+e_2,0}$ is a strict separating hyperplane of the support of $a_{\{3\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+2}h_{5n+3}$.

Let $J = \{3, 5\}$, and let $I = \{r_1, 5n+1, 5n+3\}$. We only need to consider the following cases,

(i) When $r_1 = 1$, we have $\alpha_{n,I} = 1$.
(ii) When $r_1 = 5\ell + 4$, we have $\alpha_{n,I} = -n + \ell + 1$.
(iii) When $r_1 = 5\ell + 5$, we obtain $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3\ell+1}+(-n+\ell+2)\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}}$.

The only non-negative numerator is given by $I = \{1, 5n+1, 5n+3\}$ and hence, $\mathcal{H}_{e_1,0}$ gives a strict separating hyperplane.

To prove the second part of the proposition, we note that for $I = \{4, 8, 5n+1\}$ and $I = \{4, 5n+1, 5n+3\}$, $\alpha_{n,I} = -n + 1 < 0$. Additionally, since $\alpha_{n,I}$ is an integer, $\sigma(\delta_n(\lambda)\alpha_{n,I}) = \sigma(\delta_n(\lambda))$. Hence, for $I = \{4, 8, 5n+1\}$ (resp. $I = \{4, 5n+1, 5n+3\}$) and for all $\mu \in \sigma(\delta_n(\lambda))$, $(z_I, \mu) \in \sigma_-(a_{\{3\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+2}h_{5n+3})$ (resp. $\sigma_-(a_{\{3,5\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+2})$). □

Corollary 4.18. *For $n \geq 2$, the polynomial*

$$b_0 := \Big(a_{\{3\}}(\hat{h}, \lambda)h_{5n+2}h_{5n+3} + a_{\{3,5\}}(\hat{h}, \lambda)h_{5n+2}\Big)h_{5n-1}h_{5n}$$

has one negative connected component.

Proof. Let $b_{0,1} := a_{\{3\}}(\hat{h}, \lambda)h_{5n+2}h_{5n+3}h_{5n-1}h_{5n}$ and $b_{0,2} := a_{\{3,5\}}(\hat{h}, \lambda)h_{5n+2}h_{5n-1}h_{5n}$. The polynomial $b_{0,1}$ (resp. $b_{0,2}$) is the restriction of b_0 to the face of $\mathrm{NP}_{e_{5n+3}}(b_0)$ (resp. $\mathrm{NP}_{-e_{5n+3}}(b_0)$). Since $\sigma(b_{0,1})$ and $\sigma(b_{0,2})$ have strict separating hyperplanes (Proposition 4.17), $b_{0,1}$ and $b_{0,2}$ have one negative connected component by Proposition 2.8.

Let $\mu \in \sigma(\delta_n(\lambda))$ be a vertex of $\mathrm{NP}\big(\delta_n(\lambda)\big)$. By Proposition 4.17, $(z_{\{4,8,5n+1\}}, \mu) \in \sigma_-(b_{0,1})$ and $(z_{\{4,5n+1,5n+3\}}, \mu) \in \sigma_-(b_{0,2})$. Moreover, by Proposition 2.12, $\mathrm{Conv}(z_{\{4,8,5n+1\}}, z_{\{4,5n+1,5n+3\}})$ is an edge of $\mathrm{pr}_1(\mathrm{NP}(b_0))$ and by Lemma 4.13, μ is a vertex of $\mathrm{pr}_2(\mathrm{NP}(b_0))$. Using Proposition 2.13 we get that $\mathrm{Conv}((z_{\{4,8,5n+1\}}, \mu), (z_{\{4,5n+1,5n+3\}}, \mu))$ is an edge of $\mathrm{NP}(b_0)$. Since $b_{0,1}$ and $b_{0,2}$ have exactly one negative connected component and $\mathrm{NP}(b_0)$ has an edge connecting two negative vertices of $\mathrm{NP}(b_{0,1})$ and $\mathrm{NP}(b_{0,2})$, we conclude by Theorem 2.10 that b_0 has one negative connected component. □

Proposition 4.19. *For $n \geq 2$, the supports $\sigma(a_{\{4\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$, $\sigma(a_{\{5\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+2})$, and $\sigma(a_{\{4,5\}}h_{5n-1}h_{5n}h_{5n+1})$ have strict separating hyperplanes. For $n \geq 3$, we have*

(a) $(z_{\{1,4,5n+2\}}, \mu) \in \sigma_-(a_{\{4\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$ for all $\mu \in \sigma(\delta_n(\lambda))$,
(b) $(z_{\{2,4,5n+3\}}, \mu) \in \sigma_-(a_{\{5\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+2})$ for all $\mu \in \sigma(\delta_n(\lambda))$,
(c) $(z_{\{4,5n+2,5n+3\}}, \mu) \in \sigma_-(a_{\{4,5\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1})$ for all $\mu \in \sigma(\delta_n(\lambda))$.

Let $n = 2$ and ν_2 be as given in (39). Then,

(a') $(z_{\{1,4,12\}}, \nu_2) \in \sigma_-(a_{\{4\}}(\hat{h}, \lambda)h_9h_{10}h_{11}h_{13})$,

(b') $(z_{\{2,4,13\}}, \nu_2) \in \sigma_-(a_{\{5\}}(\hat{h}, \lambda)h_9h_{10}h_{11}h_{12})$,
(c') $(z_{\{4,12,13\}}, \nu_2) \in \sigma_-(a_{\{4,5\}}(\hat{h}, \lambda)h_9h_{10}h_{11})$.

Proof. Let $J = \{4\}$ and $I = \{r_1, r_2, 5n+2\}$. By Lemma 4.8, $\alpha_{n,I} \neq 0$ in the following cases, where $\ell, \ell' \in \{0, \dots, n-2\}$:

(i) When $(r_1, r_2) = (1, 2)$ and $(r_1, r_2) = (1, 3)$, we get $\alpha_{n,I} = 1$ and $\alpha_{n,I} = \frac{(-n+1)\lambda_{3n-2}+(-n+2)\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$ respectively.
(ii) When $(r_1, r_2) \in \{(1, 5\ell+4), (2, 5\ell+4), (3, 5\ell+4), (5\ell+4, 5\ell'+8)\}$, we obtain the expression $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3n-2}+(-n+\ell+2)\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$
(iii) When $(r_1, r_2) \in \{(1, 5\ell+5), (2, 5\ell+5), (3, 5\ell+5), (5\ell+5, 5\ell'+8)\}$, we obtain $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3\ell+1}\lambda_{3n-2}+(-n+\ell+2)\lambda_{3\ell+2}\lambda_{3n-2}+(-n+\ell+2)\lambda_{3\ell+1}\lambda_{3n}+(-n+\ell+3)\lambda_{3\ell+2}\lambda_{3n}}{(\lambda_{3\ell+1}+\lambda_{3\ell+2})(\lambda_{3n-2}+\lambda_{3n})}$.
(iv) Finally, when $(r_1, r_2) = (1, 5\ell+8)$, we get $\alpha_{n,I} = \frac{(-n+\ell+2)\lambda_{3n-2}+(-n+\ell+3)\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$.

For all $(z_I, \mu) \in \sigma(a_{\{4\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$, we have

$$(-e_4) \cdot (z_I, \mu) \geq -1, \tag{43}$$

with strict inequality if and only if $I = \{1, 4, 5n+2\}, \{2, 4, 5n+2\}, \{3, 4, 5n+2\}, \{4, 5\ell+8, 5n+2\}$, $\ell \in \{0, \dots, n-2\}$. Using above cases and Lemma 4.12 if the inequality is strict, $(z_I, \mu) \in \sigma_-(a_{\{4\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$ for all $\mu \in \sigma(\delta_n(\lambda)\alpha_{n,I})$. Hence, $\mathcal{H}_{-e_4,-1}$ is a strict separating hyperplane of $\sigma(a_{\{4\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$.

Let $J = \{5\}$ and $I = \{r_1, r_2, 5n+3\}$. By Lemma 4.8, it is enough to compute $\alpha_{n,I}$ for $\ell, \ell' \in \{0, \dots, n-2\}$ in the following cases:

(i) When $(r_1, r_2) = (1, 2)$, $(r_1, r_2) = (1, 5\ell+6)$, $(r_1, r_2) = (2, 5\ell+4)$, and $(r_1, r_2) = (5\ell+4, 5\ell'+6)$ we get $\alpha_{n,I}$ as $1, n-\ell, -n+\ell+1$, and $\ell-\ell'$, respectively.
(ii) When $(r_1, r_2) = (1, 5\ell+7)$ and $(r_1, r_2) = (2, 5\ell+5)$, we get $\alpha_{n,I} = \frac{(n-\ell)\lambda_{3\ell+1}+(n-\ell+1)\lambda_{3\ell+3}}{\lambda_{3\ell+1}+\lambda_{3\ell+3}}$ and $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3\ell+1}+(-n+\ell+1)\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}}$, respectively.
(iii) When $(r_1, r_2) = (5\ell+4, 5\ell'+7)$ and $(r_1, r_2) = (5\ell+5, 5\ell'+6)$, we obtain the expression $\alpha_{n,I} = (\ell-\ell') + \frac{\lambda_{3\ell'+3}}{\lambda_{3\ell'+1}+\lambda_{3\ell'+3}}$ and $\alpha_{n,I} = (\ell-\ell') + \frac{\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}}$, respectively.
(iv) Finally, when $(r_1, r_2) = (5\ell+5, 5\ell'+7)$, we get $\alpha_{n,I} = (\ell-\ell') + \frac{\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}} + \frac{\lambda_{3\ell'+3}}{\lambda_{3\ell'+1}+\lambda_{3\ell'+3}}$.

To show that $a_{\{5\}}$ has a separating hyperplane, we observe that:

(a) if $1 \in I$ and $2 \notin I$ then the numerator of $\alpha_{n,I}$ has only positive coefficients, and
(b) if $1 \notin I$ and $2 \in I$ then the numerator of $\alpha_{n,I}$ has only negative coefficients.

Hence, $\mathcal{H}_{e_1-e_2,0}$ is a strict separating hyperplane of $a_{\{5\}}h_{5n-1}h_{5n}h_{5n+1}h_{5n+2}$ by Lemma 4.12.

Finally, let $J = \{4, 5\}$ and let $I = \{r_1, 5n+2, 5n+3\}$. Then by Lemma 4.8, $\alpha_{n,I}$ is non-zero in the following cases:

(i) When $r_1 = 1$, we have $\alpha_{n,I} = \frac{\lambda_{3n-2}+2\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$.
(ii) When $r_1 = 5\ell+4$, we have $\alpha_{n,I} = \frac{(-n+\ell+1)\lambda_{3n-2}+(-n+\ell+2)\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$.
(iii) When $r_1 = 5\ell+5$, we obtain $\alpha_{n,I} = (-n+\ell+1)\frac{\lambda_{3\ell+2}}{\lambda_{3\ell+1}+\lambda_{3\ell+2}} + \frac{\lambda_{3n}}{\lambda_{3n-2}+\lambda_{3n}}$,

where $\ell, \ell' \in \{0, \dots n-2\}$.

To show that $\sigma(a_{\{4,5\}})$ has a strict separating hyperplane, we note that:

(a) if $1 \in I$ and $4 \notin I$ then the numerator of $\alpha_{n,I}$ has only positive coefficients. and
(b) if $1 \notin I$ and $4 \in I$ then the numerator of $\alpha_{n,I}$ has only negative coefficients.

Hence, $\mathcal{H}_{e_1-e_4,0}$ is a strict separating hyperplane of $\sigma(a_{\{4,5\}}h_{5n-1}h_{5n}h_{5n+1})$ by Lemma 4.12.

We now focus on the second part of the proposition. We only explicitly prove part (a), the cases (b) and (c) follow the similar argument and observations. The coefficients of the numerator

of $\alpha_{n,I}$ are negative for $I = \{1, 4, 5n+2\}$ and $(z_{\{1,4,5n+2\}}, \mu) \in \sigma(a_{\{4\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+3})$. Moreover, the support of the numerator of $\alpha_{n,I}$ is same as the support of the denominator of $\alpha_{n,I}$ for $n \geq 3$. Hence, $\sigma(\delta_n(\lambda)) = \sigma(\delta_n(\lambda)\alpha_{n,I})$. Now, the statement follows from Lemma 4.12.

To prove (a'),(b'),(c'), we observe that

$$\alpha_{2,I} = \tfrac{-\lambda_4}{\lambda_4+\lambda_6}, \quad \text{if } I = \{1,4,12\} \text{ or } I = \{4,12,13\}, \qquad \text{and} \qquad \alpha_{2,I} = -1, \qquad \text{if } I = \{2,4,13\}.$$

Thus, for $I \in \{\{1,4,12\},\{4,12,13\},\{2,4,13\}\}$, $\lambda_1^5\lambda_4^5$ is a monomial of $\delta_2(\lambda)\alpha_{2,I}$ with negative coefficients. This completes the proof. □

We now present the main result of this section.

Theorem 4.20. *For all $n \geq 2$, both for reduced weakly irreversible phosphorylation network $(\mathcal{F}_{2,n})$ and for the weakly irreversible phosphorylation network $(\mathcal{N}_{2,n})$, the parameter region of multistationarity is non-empty and path connected.*

Proof. To prove that the parameter region of multistationarity is connected, it is enough to show that the polynomial b in (42) has one negative connected component for all n by Corollary 4.16. We relabel the summands of b as follows:

$$\begin{aligned} b_0 &= \big(a_{\{3\}}(\hat{h}, \lambda)h_{5n+2}h_{5n+3} + a_{\{3,5\}}(\hat{h}, \lambda)h_{5n+2})\big)h_{5n-1}h_{5n}, \\ b_1 &:= a_{\emptyset}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+2}h_{5n+3}, \qquad b_2 := a_{\{4\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+3}, \\ b_3 &:= a_{\{5\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}h_{5n+2}, \quad \text{and} \quad b_4 := a_{\{4,5\}}(\hat{h}, \lambda)h_{5n-1}h_{5n}h_{5n+1}. \end{aligned} \tag{44}$$

For all $n \geq 2$, we make the following observations:

(O1) Polynomials b_1 and b_2 are the restrictions of the polynomial $b_1 + b_2$ to the parallel faces $\mathrm{NP}_{e_{5n+2}}(b_1 + b_2)$ and $\mathrm{NP}_{-e_{5n+2}}(b_1 + b_2)$.

(O2) Polynomials b_3 and b_4 are the restrictions of the polynomial $b_3 + b_4$ to the parallel faces $\mathrm{NP}_{e_{5n+2}}(b_3 + b_4)$ and $\mathrm{NP}_{-e_{5n+2}}(b_3 + b_4)$.

(O3) Polynomials $b_1 + b_2$ and $b_3 + b_4$ are the restrictions of $b_1 + b_2 + b_3 + b_4$ to the parallel faces $\mathrm{NP}_{e_{5n+3}}(b_1 + b_2 + b_3 + b_4)$, and $\mathrm{NP}_{-e_{5n+3}}(b_1 + b_2 + b_3 + b_4)$.

(O4) Lastly, polynomials b_0 and $b_1 + b_2 + b_3 + b_4$ are the restrictions of b to the parallel faces $\mathrm{NP}_{e_{5n+1}}(b)$, and $\mathrm{NP}_{-e_{5n+1}}(b)$.

From Proposition 4.19, it follows that the support of b_2, b_3, b_4 have strict separating hyperplanes. Thus, b_2, b_3, b_4 have one negative connected component by Proposition 2.8. Moreover, the polynomial b_0 satisfies (P1) by Corollary 4.18.

We will use these observations to first show that b has one negative connected component for $n = 2$ and then use induction on n.

For $n = 1$, the critical polynomial q_1 only has positive coefficients. For $n = 2$, by Lemma 4.7, we have

$$b_1 = q_1(\hat{h}, \hat{\lambda})(\lambda_4 + \lambda_5)(\lambda_4 + \lambda_6)\lambda_4^3 h_9 h_{10} h_{11} h_{12} h_{13}.$$

Therefore, $\sigma_-(b_1 + b_2) \subseteq \mathrm{NP}_{-e_{12}}(b_1 + b_2)$. Since b_2 has one negative connected component, by observation (O1) above and Theorem 2.9, $b_1 + b_2$ satisfies (P1).

Let $\nu_2 = [5, 0, 0, 5, 0, 0]^\top$. By Lemma 4.11, $\nu_2 \in \mathrm{Vert}(\mathrm{NP}(\delta_2(\lambda)))$. Let $I_0 := \{4, 11, 13\}$, $I_2 := \{1, 4, 12\}$, $I_3 = \{2, 4, 13\}$, $I_4 := \{4, 12, 13\}$. From Proposition 4.17 and Proposition 4.19 it follows that

$$(z_{I_0}, \nu_2) \in \sigma_-(b_0), \qquad (z_{I_2}, \nu_2) \in \sigma_-(b_1 + b_2), \quad (z_{I_3}, \nu_2) \in \sigma_-(b_3), \quad (z_{I_4}, \nu_2) \in \sigma_-(b_4).$$

Using Lemma 4.13, Proposition 2.12, and Proposition 2.13 we have the following edges joining negative vertices of parallel faces in (O2)-(O4):

(i) $\mathrm{Conv}((z_{I_3}, \nu_2), (z_{I_4}, \nu_2))$ is an edge joining negative vertices of $\mathrm{NP}(b_3 + b_4)$,

(ii) $\mathrm{Conv}((z_{I_2}, \nu_2), (z_{I_4}, \nu_2))$ is an edge joining negative vertices of $\mathrm{NP}(b_1 + b_2 + b_3 + b_4)$,

(iii) $\mathrm{Conv}((z_{I_0},\nu_2),(z_{I_4},\nu_2))$ is an edge joining negative vertices of $\mathrm{NP}(b_0+b_1+b_2+b_3+b_4)$.

Since, b_0, b_1+b_2, b_3, b_4 satisfy (P1), the cases above and repeated application of Theorem 2.10 gives that b has one negative connected component for $n=2$.

Let $n\geq 3$ and assume that q_k has one negative connected component for all $2\leq k\leq n-1$. By the inductive assumption q_{n-1} has one negative connected component and thus, by Proposition 4.7 so does b_1.

Let $I_0 := \{4, 5n+1, 5n+3\}$, $I_1 := \{1, 4, 5n-3\}$, $I_2 := \{1, 4, 5n+2\}$, $I_3 = \{2, 4, 5n+3\}$, $I_4 := \{4, 5n+2, 5n+3\}$ and let ν_n as defined in (39). Note that $\nu_n \in \mathrm{Vert}(\mathrm{NP}(\delta_n(\lambda)))$ by Lemma 4.11. From Lemma 4.11, Proposition 4.17 and Proposition 4.19 it follows that

$$(z_{I_0},\nu_n)\in\sigma_-(b_0),\quad (z_{I_1},\nu_n)\in\sigma_-(b_1),\quad (z_{I_2},\nu_n)\in\sigma_-(b_2),\quad (z_{I_3},\nu_n)\in\sigma_-(b_3),\quad (z_{I_4},\nu_n)\in\sigma_-(b_4).$$

Using Lemma 4.13, Proposition 2.12, and Proposition 2.13 we conclude that:

(i) $\mathrm{Conv}((z_{I_1},\nu_n),(z_{I_2},\nu_n))$ is an edge joining negative vertices of $\mathrm{NP}(b_1+b_2)$,
(ii) $\mathrm{Conv}((z_{I_3},\nu_n),(z_{I_4},\nu_n))$ is an edge joining negative vertices of $\mathrm{NP}(b_3+b_4)$,
(iii) $\mathrm{Conv}((z_{I_2},\nu_n),(z_{I_4},\nu_n))$ is an edge joining negative vertices of $\mathrm{NP}(b_1+b_2+b_3+b_4)$,
(iv) $\mathrm{Conv}((z_{I_0},\nu_n),(z_{I_4},\nu_n))$ is an edge joining negative vertices of $\mathrm{NP}(b_0+b_1+b_2+b_3+b_4)$.

Using observations (O1)-(O4) and Theorem 2.10, we conclude that b has one negative connected component.

Since q_n attains negative values, from [4, Theorem 1] it follows that the multistationarity region is non-empty for the reduced network $(\mathcal{F}_{2,n})$. From [12, Theorem 5.1], it follows that $(\mathcal{N}_{2,n})$ is multistationary for some choice of the parameters. □

Acknowledgements. The authors thank Elisenda Feliu for useful comments on the manuscript, which significantly improved its readability. NK was supported by Independent Research Fund of Denmark. MLT was supported by the European Union under the Grant Agreement no. 101044561, POSALG. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or European Research Council (ERC). Neither the European Union nor ERC can be held responsible for them.

References

[1] A. Ben-Israel. Notes on linear inequalities, I: The intersection of the nonnegative orthant with complementary orthogonal subspaces. *J. Math. Anal. Appl.*, 9(2):303–314, 1964.
[2] R. Bradford, J. H. Davenport, M. England, H. Errami, V. Gerdt, D. Grigoriev, C. Hoyt, K. Košta, O. Radulescu, T. Sturm, and A. Weber. Identifying the parametric occurrence of multiple steady states for some biological networks. *J. Symb. Comput.*, 98:84–119, 2020. Special Issue on Symbolic and Algebraic Computation: ISSAC 2017.
[3] B. L. Clarke. *Stability of Complex Reaction Networks*, pages 1–215. John Wiley and Sons, Ltd, 1980.
[4] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *Plos. Comput. Biol.*, 13(10):1–25, 2017.
[5] A. Dennis and A. Shiu. On the connectedness of multistationarity regions of small reaction networks. *arXiv*, 2303.03960, 2023.
[6] A. Dickenstein, M.P. Millán, A. Shiu, and X. Tang. Multistationarity in structured reaction networks. *Bull. Math. Biol.*, 81:1527–1581, 2019.
[7] M. Feinberg. *Foundations of Chemical Reaction Network Theory*, volume 202. Springer, 2019.
[8] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürück. The kinetic space of multistationarity in dual phosphorylation. *J. Dyn. Differ. Equ.*, 34:825–852, 2022.
[9] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. Parameter region for multistationarity in $\boldsymbol{n}-$site phosphorylation networks. *SIAM J. Appl. Dyn. Syst.*, 22(3):2024–2053, 2023.
[10] E. Feliu and A. Sadeghimanesh. Kac-Rice formulas and the number of solutions of parametrized systems of polynomial equations. *Math. Comput.*, 91:2739–2769, 2022.
[11] E. Feliu and M. L. Telek. On generalizing Descartes' rule of signs to hypersurfaces. *Adv. Math.*, 408(A), 2022.

[12] E. Feliu and C. Wiuf. Simplifying biochemical models with intermediate species. *J. R. Soc. Interface*, 10(87):20130484, 2013.
[13] J. Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophys. J.*, 93:3828–3834, 2007.
[14] H. A. Harrington, D. Mehta, H. M. Byrne, and J. D. Hauenstein. Decomposing the parameter space of biological networks via a numerical discriminant approach. In J. Gerhard and I. Kotsireas, editors, *Maple in Mathematics Education and Research*, pages 114–131. Springer International Publishing, 2020.
[15] J. Hell and A. D. Rendall. Dynamical features of the map kinase cascade. In F. Graw, F. Matthäus, and J. Pahle, editors, *Modeling Cellular Systems*, volume 11. Springer, 2017.
[16] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Springer, 2013.
[17] M. Laurent and N. Kellershohn. Multistability: a major means of differentiation and evolution in biological systems. *Trends. Biochem. Sci.*, 24(11):418–22, 1999.
[18] M. Marcondes de Freitas, E. Feliu, and C. Wiuf. Intermediates, catalysts, persistence, and boundary steady states. *J. Math. Biol.*, 74:887–932, 2017.
[19] S. Müller and G. Regensburger. Elementary vectors and conformal sums in polyhedral geometry and their relevance for metabolic pathway analysis. *Front. Genet.*, 7:90, 2016.
[20] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16:69–97, 2016.
[21] K. M. Nam, B. M. Gyori, S. V. Amethyst, D. J. Bates, and J. Gunawardena. Robustness and parameter geography in post-translational modification systems. *Plos. Comput. Biol.*, 16(5):1–50, 05 2020.
[22] E. M. Ozbudak, M. Thattai, H. N. Lim, B. I. Shraiman, and A. van Oudenaarden. Multistability in the lactose utilization network of escherichia coli. *Nature.*, 427:737–740, 2004.
[23] V.V. Prasolov and S. Ivanov. *Problems and Theorems in Linear Algebra*. History of Mathematics. American Mathematical Society, 1994.
[24] R. T. Rockafellar. *Convex analysis*. Princeton University Press, 1972.
[25] M. L. Telek. Geometry of the signed support of a multivariate polynomial and Descartes' rule of signs. *arXiv*, 2310.05466, 2023.
[26] M. L. Telek and E. Feliu. Topological descriptors of the parameter region of multistationarity: Deciding upon connectivity. *Plos. Comput. Biol.*, 19(3):1–38, 03 2023.

Authors' addresses:

Nidhi Kaihnsa, University of Copenhagen nidhi@math.ku.dk
Máté L. Telek, University of Copenhagen mlt@math.ku.dk

III

On generalizing Descartes' rule of signs to hypersurfaces

Elisenda Feliu
Department of Mathematical Sciences
University of Copenhagen

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Publication details

Published in Advances in Mathematics 408(A), 2022
DOI: https://doi.org/10.1016/j.aim.2022.108582

ON GENERALIZING DESCARTES' RULE OF SIGNS TO HYPERSURFACES

ELISENDA FELIU AND MÁTÉ L. TELEK

Abstract. We give partial generalizations of the classical Descartes' rule of signs to multivariate polynomials (with real exponents), in the sense that we provide upper bounds on the number of connected components of the complement of a hypersurface in the positive orthant. In particular, we give conditions based on the geometrical configuration of the exponents and the sign of the coefficients that guarantee that the number of connected components where the polynomial attains a negative value is at most one or two. Our results fully cover the cases where such an upper bound provided by the univariate Descartes' rule of signs is one. This approach opens a new route to generalize Descartes' rule of signs to the multivariate case, differing from previous works that aim at counting the number of positive solutions of a system of multivariate polynomial equations.

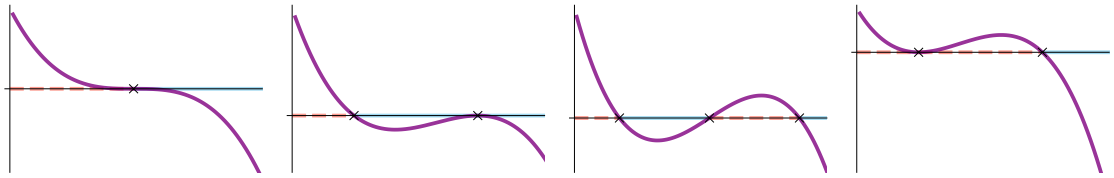
Keywords: semi-algebraic set; signomial; Newton polytope; connectivity; convex function

1. Introduction

Descartes' rule of signs, established by René Descartes in his book *La Géométrie* in 1637, provides an easily computable upper bound for the number of positive real roots of a univariate polynomial with real coefficients. Specifically, it states that the polynomial cannot have more positive real roots than the number of sign changes in its coefficient sequence (excluding zero coefficients). In 1828, Gauss improved the rule by showing that the number of positive real roots, counted with multiplicity, and the number of sign changes in the coefficients sequence, have the same parity [22]. Since then, several different proofs were published e.g. [1, 17, 43], and several generalizations were made in several directions. In 1918, Curtiss gave a proof that works for real exponents and even for some infinite series [17]. In 1999, Grabiner showed that Descartes' bound is sharp, that is, for every given sign sequence, one can always find compatible coefficients such that the polynomial has the maximum possible number of positive roots provided by Descartes' bound [25]. Generalizations of the Descartes' rule to other types of functions in one variable are also available [28, 42].

Efforts to generalize Descartes' rule of signs to the multivariate case have focused on systems of n multivariate polynomial equations in n variables, and on bounding the number of solutions in the positive orthant using sign properties of the coefficients of the system. The first conjecture for such a bound was published in 1996 by Itenberg and Roy [29]. They were able to show their conjecture for some special cases. The first non-trivial example supporting the conjecture was presented by Lagarias and Richardson [32] in 1997. Almost at the same time, Li and Wang gave a counterexample to the Itenberg-Roy conjecture [33]. The first generalization was given recently and identifies systems with at most one solution in the positive orthant [35], see also [15]. Afterwards, a sharp upper bound was given for systems of polynomials supported on circuits [10, 11]. In these works, the bound is given in terms of the sign variation of a sequence associated both with the exponents and the coefficients of the system. To the best of our knowledge, these are the only known generalizations of Descartes' rule of signs to the multivariate case.

Descartes' rule of signs allows however for a "dual" presentation: it gives an upper bound on the number of connected components of $\mathbb{R}_{>0}$ minus the zero set of the polynomial, and if the sign of the highest degree term is fixed, then it also gives an upper bound on the number of connected components where the polynomial evaluates positively or negatively. Specifically, if we write $f(x) = a_0 + a_1x + \cdots + a_nx^n$ with $a_n \neq 0$, and let ρ be the Descartes' bound on the number of positive roots, then there are at most $\rho + 1$ connected components. If ρ is odd, the



(a) (b) (c) (d)

FIGURE 1. Graphs of polynomials p of degree three with coefficient sign sequence $(+-+-)$. In each figure, the connected components of $\mathbb{R}_{>0}$ minus the zero set of p, where p evaluates positively or negatively, are shown in red and blue respectively. (a) $8-12x+6x^2-x^3$. (b) $9-15x+7x^2-x^3$. (c) $15-23x+9x^2-x^3$. (d) $3-7x+5x^2-x^3$.

upper bounds for the number of components where f is positive or negative agree, while if ρ is even, then there are at most $\frac{\rho}{2}+1$ connected components where f attains the sign of a_n. For example, if after ignoring zero coefficients, the sign sequence of the coefficients is $(++--)$, then there is one connected component where the polynomial evaluates positively and one where it evaluates negatively. If the sequence is $(+-+-)$, then there at most two connected components where the polynomial evaluates positively and at most two where it evaluates negatively, see Fig. 1.

With this presentation, Descartes' rule of signs may be generalized to hypersurfaces in the following sense. Let $f\colon \mathbb{R}^n_{>0}\to\mathbb{R}$ be a *signomial* (a multivariate generalized polynomial, where we allow real exponents, restricted to the positive orthant), and consider the sets

$$V_{>0}(f) := \{x\in\mathbb{R}^n_{>0} \mid f(x)=0\}, \qquad V^c_{>0}(f) := \mathbb{R}^n_{>0}\setminus V_{>0}(f). \tag{1}$$

We aim at bounding the number of connected components of $V^c_{>0}(f)$ in terms of the relative position of the exponent vectors of each monomial of f in $\mathbb{R}^n$, and the sign of the coefficients. This leads to the formulation of the following problem for the *generalization of Descartes' rule of signs to hypersurfaces.*

Problem 1.1. *Consider a signomial $f\colon \mathbb{R}^n_{>0}\to\mathbb{R}$ with $f(x)=\sum_{\mu\in\sigma(f)} c_\mu x^\mu$, and $\sigma(f)\subseteq\mathbb{R}^n$ a finite set. Find a (sharp) upper bound on the number of connected components of $V^c_{>0}(f)$, where f takes negative (resp. positive) values, based on the sign of the coefficients and the geometry of $\sigma(f)$.*

In this paper we address Problem 1.1 for generic n in some scenarios, which, in particular, include the univariate Descartes' rule of signs when the upper bound on the number of connected components where f is negative is one, that is, when the sign sequence is one of $(+\cdots+-\cdots-)$, $(-\cdots-+\cdots+)$, or $(+\cdots+-\cdots-+\cdots+)$.

Specifically, we show that $V^c_{>0}(f)$ has at most one connected component where f is negative if f has only one negative coefficient (Theorem 3.4). The same holds if there exists a hyperplane separating the exponents with positive coefficients from those with negative coefficients (Theorem 3.6), or if the exponents with negative coefficient lie on a simplex such that the exponents with positive coefficient lie outside the simplex in a certain way (Theorem 4.6). A detailed account of our results is given in Section 5. We focus on finding upper bounds for the number of negative connected components, as statements about the number of positive connected components of $V^c_{>0}(f)$ follows by studying $-f$.

If f is a polynomial, that is, $\sigma(f)\subseteq\mathbb{Z}^n_{\geq 0}$, the set $V^c_{>0}(f)$ is semi-algebraic and hence it has a finite number of connected components [7, Theorem 5.22]. Computing topological invariants of semi-algebraic sets, such as the number of connected components, has been heavily studied in real algebraic geometry. Upper bounds of the sum of the Betti numbers of a semi-algebraic set in terms of the number of variables, the degree and the number of the defining polynomials can be found for example in [4, Theorem 1], [21, Theorem 6.2], and [8, Theorems 1.8 and 2.7]. For

the number of connected components of a semi-algebraic set, that is, the 0-th Betti number, an upper bound was given in [6, Theorem 1], [3, Theorem 1.1].

There exist several algorithms to compute the number of connected components of a semi-algebraic set. One algorithm is provided by Cylindrical Algebraic Decomposition, but it has double exponential complexity (see [7, Remark 11.19]). A more efficient way to compute connected components is using so-called road maps. In this way, one has an algorithm with single exponential complexity. For more details about this algorithm, see [5, Section 3].

The Descartes' rule of signs is of special importance in applications where positive solutions to polynomial systems are the object of study. This is the case in models in biology and (bio)chemistry where variables are concentrations or abundances. It is precisely in this setting, namely the theory of biochemical reaction networks, where our motivation to consider Problem 1.1 comes from. In an upcoming paper, we show that the connectivity of the set of parameters that give rise to multistationarity in a reaction network [14, 16, 30] relies on the number of connected components of the complementary of a hypersurface. The hypersurface of interest is large for realistic networks, with many monomials and variables, and hence not manageable by algorithms from semi-algebraic geometry. The advantage of the techniques presented here is that they rely on linear optimization problems, and can handle this application.

The paper is organized as follows. In Section 2, we provide the notation and basic results on signomials. In Section 3, we give bounds answering Problem 1.1 using separating hyperplanes (Theorem 3.6, 3.8), while in Section 4 bounds are found by providing conditions that guarantee that the signomial can be transformed into a convex function, while preserving the number of connected components of $V^c_{>0}(f)$ (Theorem 4.6). In Section 5, we compare the two approaches. Throughout we illustrate our results with examples and figures, worked out using SageMath [41].

Notation. $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{<0}$ refer to the sets of non-negative, positive and negative real numbers respectively. We denote the Euclidean scalar product of two vectors $v, w \in \mathbb{R}^n$ by $v \cdot w$. For a set $\sigma \subseteq \mathbb{R}^m$, a matrix $M \in \mathbb{R}^{n\times m}$ and a vector $v \in \mathbb{R}^n$ we write $M\sigma + v$ for the set $\{Ms + v \mid s \in \sigma\}$. For two sets $A, B \subseteq \mathbb{R}^n$, the set $A + B = \{a + b \mid a \in A, b \in B\}$ is the Minkowski sum of A and B. We let $\mathrm{Conv}(A)$ denote the convex hull of A. For $a_1, \ldots, a_m \in \mathbb{R}^n$, we write $\mathrm{Conv}(a_1, \ldots, a_m) := \mathrm{Conv}(\{a_1, \ldots, a_m\})$. By convention, the maximum over an empty set is $-\infty$, and the minimum over an empty set is ∞. The symbol $\#S$ denotes the cardinality of a finite set S.

2. Preliminaries

The central object of study is a function

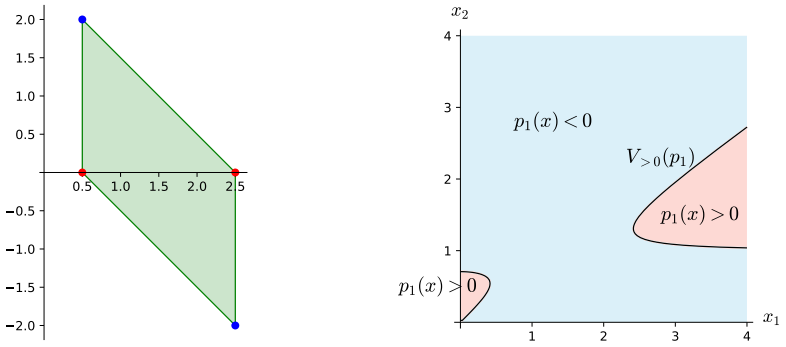
$$f\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \qquad f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu, \quad \text{with } c_\mu \in \mathbb{R}\setminus\{0\}, \tag{2}$$

where $\sigma(f) \subseteq \mathbb{R}^n$ is a finite set, called the *support* of f. Here x^μ is the usual short notation for $x_1^{\mu_1} \ldots x_n^{\mu_n}$. To emphasize that we restrict the domain of f to the positive orthant, we call f a *signomial*. That is, a signomial is a generalized polynomial on the positive orthant. The term signomial was introduced by Duffin and Peterson in the early 1970s [19]. Since then, it is commonly used in geometric programming [12, 39].

Given a signomial f as in (2) and a set $S \subseteq \sigma(f)$, we define *the restriction of f to S* by considering the monomials with exponent vectors in S:

$$f_{|S}(x) = \sum_{\mu\in S} c_\mu x^\mu. \tag{3}$$

With the notation in (1) and by continuity, the signomial f has constant sign in each connected component of $V^c_{>0}(f)$.



(a) (b)

FIGURE 2. (a) Newton polytope of $p_1(x_1, x_2)$ from Example 2.2. Blue points are negative and red points are positive. (b) The positive and negative connected components of $V^c_{>0}(p_1)$.

Definition 2.1. Let f be a signomial in n variables.

- A connected component U of $V^c_{>0}(f)$ is said to be *positive* if $f(x) > 0$ for every $x \in U$. We say U is *negative*, if $f(x) < 0$ for every $x \in U$.
- The convex hull of $\sigma(f)$ is called the *Newton polytope* of f and denoted by $\mathrm{N}(f)$.
- A point $\alpha \in \sigma(f)$ is called positive, resp. negative, if the coefficient c_α is positive, resp. negative. The set $\sigma(f)$ is partitioned into the set of positive points and the set of negative points:

$$\sigma_+(f) := \{\alpha \in \sigma(f) \mid c_\alpha > 0\} \quad \text{and} \quad \sigma_-(f) := \{\beta \in \sigma(f) \mid c_\beta < 0\}.$$

Example 2.2. The support of the signomial

$$p_1(x_1, x_2) = x_1^{2.5} - 2x_1^{0.5}x_2^2 + x_1^{0.5} - x_1^{2.5}x_2^{-2}$$

is $\sigma(p_1) = \{(2.5, 0), (0.5, 2), (0.5, 0), (2.5, -2)\}$. The points $(2.5, 0)$, $(0.5, 0)$ are positive, while the points $(0.5, 2)$, $(2.5, -2)$ are negative. The Newton polytope of p_1 and the positive and negative connected components of $V^c_{>0}(p_1)$ are displayed in Fig. 2.

In what follows, it will be convenient to consider transformations of the support that do not change the number of negative (resp. positive) connected components. Any invertible matrix $M \in \mathrm{GL}_n(\mathbb{R})$ induces a function

$$h_M \colon \mathbb{R}^n_{>0} \to \mathbb{R}^n_{>0}, \qquad x \mapsto x^M := (x^{M_1}, \dots, x^{M_n}) \tag{4}$$

where $M_1, \dots, M_n$ denote the columns of M. The function h_M is called a *monomial change of variables* and it is a homeomorphism.

Lemma 2.3. *For $M \in \mathrm{GL}_n(\mathbb{R})$, $v \in \mathbb{R}^n$, and a signomial f on $\mathbb{R}^n_{>0}$, define the signomial*

$$F_{M,v,f} \colon \mathbb{R}^n_{>0} \to \mathbb{R}, \qquad F_{M,v,f}(x) = x^v f(h_M(x)).$$

There is a homeomorphism between the positive (resp. negative) connected components of $V^c_{>0}(f)$ and $V^c_{>0}(F_{M,v,f})$. Furthermore,

$$\sigma_+\big(F_{M,v,f}\big) = M\sigma_+(f) + v \quad \text{and} \quad \sigma_-\big(F_{M,v,f}\big) = M\sigma_-(f) + v.$$

Proof. If $f(x) = \sum_{\mu \in \sigma(f)} c_\mu x^\mu$, we have

$$F_{M,v,f}(x) = x^v f(h_M(x)) = \sum_{\mu \in \sigma(f)} c_\mu x^v (x^M)^\mu = \sum_{\mu \in \sigma(f)} c_\mu x^{M\mu + v}.$$

From this, the second part of the lemma follows.

For the first part, clearly, the identity map induces a sign-preserving homeomorphism between $V^c_{>0}(F_{M,v,f})$ and $V^c_{>0}(f \circ h_M)$, and the map h_M induces a homeomorphism between $V^c_{>0}(f \circ h_M)$ and $V^c_{>0}(f)$, which also preserves the sign of each connected component. □

In view of Lemma 2.3, we can for example assume that all exponent vectors belong to $\mathbb{R}^n_{>0}$ if necessary. Moreover, if $\sigma(f) \subseteq \mathbb{Q}^n$, then f can be replaced by a polynomial and the number of negative (resp. positive) connected components of $V^c_{>0}(f)$ remains unchanged.

Example 2.4. The matrix $M = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0 \end{pmatrix}$ and the vector $v = (-0.25, -0.25)$ transform the signomial p_1 from Example 2.2 to the polynomial $F_{M,v,p_1}(x_1, x_2) = x_1x_2 - 2x_2 + 1 - x_1$.

3. Paths on logarithmic scale

In this section, we provide the first results towards Problem 1.1. The idea behind the proofs in this section relies on reducing the multivariate signomial to a univariate signomial, and applying Descartes' rule of signs. To this end, given $v \in \mathbb{R}^n$ and $x \in \mathbb{R}^n_{>0}$, we consider continuous paths

$$\gamma_{v,x} \colon [1, \infty) \to \mathbb{R}^n_{>0}, \qquad t \mapsto (t^{v_1}x_1, \ldots, t^{v_n}x_n). \tag{5}$$

In logarithmic scale, applying the coordinate-wise natural logarithm map

$$\mathrm{Log} \colon \mathbb{R}^n_{>0} \to \mathbb{R}^n, \qquad (x_1, \ldots, x_n) \mapsto (\log(x_1), \ldots, \log(x_n)), \tag{6}$$

each path $\gamma_{v,x}$ is transformed into a half-line $\tau_{v,\mathrm{Log}(x)} \colon [0, \infty) \to \mathbb{R}^n$, $s \mapsto s\,v + \mathrm{Log}(x)$, with start point $\mathrm{Log}(x)$ and direction vector v. Specifically,

$$\mathrm{Log} \circ \gamma_{v,x} = \tau_{v,\mathrm{Log}(x)} \circ \log, \qquad \text{in } [1, \infty). \tag{7}$$

Since the logarithm map Log is a homeomorphism, the topological properties of $f^{-1}(\mathbb{R}_{<0})$ and of its image under Log are the same. This observation gives us an easy geometric way to think about paths $\gamma_{v,x}$.

Given a signomial f, each $v \in \mathbb{R}^n$ and $x \in \mathbb{R}^n_{>0}$ induce a signomial function in one variable:

$$f_{v,x} \colon \mathbb{R}_{>0} \to \mathbb{R}, \qquad t \mapsto \sum_{\mu \in \sigma(f)} (c_\mu x^\mu) t^{v \cdot \mu}. \tag{8}$$

Note that $f_{v,x}(1) = f(x)$. Since the restriction of $f_{v,x}$ to $[1, \infty)$ is the composition $f \circ \gamma_{v,x}$, understanding the properties of $f_{v,x}$ allows us to determine whether the path $\gamma_{v,x}$ is in the pre-image of the negative real line under f. This motivates the study of signomials in one variable. The following lemma will be used repeatedly in what follows. Its proof is a direct application of Descartes' rule of signs.

Lemma 3.1. *Let $g \colon \mathbb{R}_{>0} \to \mathbb{R}$, $g(t) = \sum_{\nu \in \sigma(g)} a_\nu t^\nu$, be a signomial in one variable such that $g(1) < 0$.*

(i) *If the sign sequence of the coefficients of g has at most two sign changes, and the leading coefficient is positive, then there is unique $\rho \in (1, \infty)$ such that $g(\rho) = 0$, and it holds that $g(t) < 0$ for all $t \in [1, \rho)$ and $g(t) > 0$ for all $t \in (\rho, \infty)$.*

(ii) *If the sign sequence of the coefficients of g has at most one sign change, and the leading coefficient is negative, then $g(t) < 0$ for all $t \in [1, \infty)$.*

Following the notation of [31, Section 2.3.1] and [26, Section 1.1], for every $v \in \mathbb{R}^n$ and $a \in \mathbb{R}$, we define a *hyperplane* $\mathcal{H}_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu = a\}$, and two *half-spaces*

$$\mathcal{H}^+_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \geq a\} \quad \text{and} \quad \mathcal{H}^-_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \leq a\}.$$

We let $\mathcal{H}^{+,\circ}_{v,a}, \mathcal{H}^{-,\circ}_{v,a}$ denote the interior of $\mathcal{H}^+_{v,a}$, and $\mathcal{H}^-_{v,a}$ respectively. Although $\mathcal{H}_{v,a} = \mathcal{H}_{-v,-a}$, the choice of sign determines which half-space is positive and which one is negative.

As we will see in Lemma 3.3, the relative position of a hyperplane $\mathcal{H}_{v,a}$ and the points in $\sigma(f)$ gives valuable information about the behavior of the function $f_{v,x}$ in (8). To this end, we introduce the following types of vectors v.

Definition 3.2. Let $v \in \mathbb{R}^n$.

(i) We say that v is a *separating vector* of $\sigma(f)$ if for some $a \in \mathbb{R}$ it holds

$$\sigma_-(f) \subseteq \mathcal{H}^+_{v,a}, \qquad \sigma_+(f) \subseteq \mathcal{H}^-_{v,a}.$$

The separating vector v is *strict* if $\sigma_-(f) \cap \mathcal{H}^{+,\circ}_{v,a} \neq \emptyset$, and *very strict* if additionally $\sigma_-(f) \cap \mathcal{H}_{v,a} = \emptyset$ for some $a \in \mathbb{R}$. Let $\mathcal{S}^-(f)$ denote the set of separating vectors of $\sigma(f)$.

(ii) We say that v is an *enclosing vector* of $\sigma(f)$ if for some $a, b \in \mathbb{R}$, $a \leq b$, it holds

$$\sigma_-(f) \subseteq \mathcal{H}^+_{v,a} \cap \mathcal{H}^-_{v,b}, \qquad \sigma_+(f) \subseteq \mathbb{R}^n \setminus (\mathcal{H}^{+,\circ}_{v,a} \cap \mathcal{H}^{-,\circ}_{v,b}).$$

We say that v is a *strict enclosing vector* of $\sigma(f)$ if additionally $\sigma_+(f) \cap \mathcal{H}^{-,\circ}_{v,a} \neq \emptyset$ and $\sigma_+(f) \cap \mathcal{H}^{-,\circ}_{v,b} \neq \emptyset$. We denote by $\mathcal{E}^-(f)$ the set of enclosing vectors of $\sigma(f)$.

The sets of separating and enclosing vectors can be described algebraically as

$$\mathcal{S}^-(f) = \Big\{ v \in \mathbb{R}^n \mid \max_{\alpha \in \sigma_+(f)} v \cdot \alpha \leq \min_{\beta \in \sigma_-(f)} v \cdot \beta \Big\}, \tag{9}$$

$$\mathcal{E}^-(f) = \{ v \in \mathbb{R}^n \mid \forall \alpha \in \sigma_+(f) \colon v \cdot \alpha \leq \min_{\beta \in \sigma_-(f)} v \cdot \beta \text{ or } \max_{\beta \in \sigma_-(f)} v \cdot \beta \leq v \cdot \alpha \}. \tag{10}$$

For $v \in \mathcal{S}^-(f)$, setting $a := \max_{\alpha \in \sigma_+(f)} v \cdot \alpha$, Definition 3.2(i) holds. For $v \in \mathcal{E}^-(f)$, we let $a := \min_{\beta \in \sigma_-(f)} v \cdot \beta$ and $b := \max_{\beta \in \sigma_-(f)} v \cdot \beta$ and Definition 3.2(ii) holds.

Note that a separating vector is in particular an enclosing vector, that is, $\mathcal{S}^-(f) \subseteq \mathcal{E}^-(f)$. Using the algebraic description of $\mathcal{S}^-(f)$ from (9), one can easily show that $\mathcal{S}^-(f)$ is a convex cone, i.e. it is closed under addition and multiplication by a nonnegative scalar [44, Ch. 1].

For a separating vector v to be strict, there must be a negative point in $\sigma(f)$ in $\mathcal{H}^+_{v,a}$ that is not in the hyperplane $\mathcal{H}_{v,a}$. That is, there exists $\beta_0 \in \sigma_-(f)$ such that $\max_{\alpha \in \sigma_+(f)} v \cdot \alpha < v \cdot \beta_0$. For it to be very strict, no negative point of $\sigma(f)$ lies on the hyperplane, or equivalently, the inequality defining $\mathcal{S}^-(f)$ in (9) is strict. Fig. 3(a) shows a strict separating vector.

Enclosing vectors *enclose* all negative points of $\sigma(f)$ between two parallel hyperplanes separated from the positive points, but points of both signs are allowed to be in the two hyperplanes. For an enclosing vector v to be strict, there must be positive points on the side of the hyperplanes not containing the negative points, that is, there exist $\alpha_1, \alpha_2 \in \sigma_+(f)$ such that the inequalities in (10) are strict for that v respectively. See Fig. 4(a).

Enclosing and separating vectors order the exponents of $f_{v,x}$ in (8), such that the negative and positive coefficients are grouped. This has the following consequences.

Lemma 3.3. *Let $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial and $x \in \mathbb{R}^n_{>0}$.*

(i) *If $v \in \mathcal{E}^-(f)$, then there are at most two sign changes in the coefficient sign sequence of the signomial $f_{v,x}$. If v is additionally strict, then both the leading coefficient and the coefficient of smallest degree of $f_{v,x}$ are positive.*

(ii) *If $v \in \mathcal{S}^-(f)$, then there is at most one sign change in the coefficient sign sequence of the signomial $f_{v,x}$. If v is strict, then the leading coefficient of $f_{v,x}$ is negative.*

Additionally if $f(x) < 0$, then the following statements hold:

(i') *If $v \in \mathcal{E}^-(f)$, then there is a unique $\rho \in (1, \infty]$ such that $f_{v,x}(t) < 0$ for all $t \in [1, \rho)$ and $f_{v,x}(t) > 0$ for all $t > \rho$ (note that ρ might be ∞).*

(ii") *If $v \in \mathcal{S}^-(f)$, then $f_{v,x}(t) < 0$ for all $t \in [1, \infty)$.*

Proof. (i) and (i'). For $v \in \mathcal{E}^-(f)$, v orders the exponents $v \cdot \mu$ such that the sign sequence is $(+ \cdots + - \cdots - + \cdots +)$, with potentially one or more of the three blocks of repeated signs not present. The positive blocks are present if v is strict by definition, showing (i).

For $f(x) < 0$, if the leading coefficient of $f_{v,x}$ is positive, then Lemma 3.1(i) gives the existence of a unique $\rho \in (1, \infty)$ satisfying (i') in the statement. If the leading coefficient of $f_{v,x}$ is negative, then $v \in \mathcal{S}^-(f)$ and this case is covered next, and gives $\rho = \infty$.

(ii) and (ii'). From $v \in \mathcal{S}^-(f)$, it follows that the signomial $f_{v,x}$ has at most one sign change in its coefficient sequence, as $\max_{\alpha \in \sigma_+(f)} v \cdot \alpha \leq \min_{\beta \in \sigma_-(f)} v \cdot \beta$. If v is strict, then for at least one $\beta_0 \in \sigma_-(f)$ we have $\max_{\alpha \in \sigma_+(f)} v \cdot \alpha < v \cdot \beta_0$, and hence the leading term is negative,

showing (ii). If $f_{v,x}(1) = f(x) < 0$, $f_{v,x}$ must have some negative coefficient. Using $v \in \mathcal{S}^-(f)$, we conclude that the leading coefficient is negative and v is strict. Lemma 3.1(ii) gives now statement (ii'). □

Theorem 3.4. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If at most one coefficient of f is negative, then $f^{-1}(\mathbb{R}_{<0})$ is a logarithmically convex set. In particular, $V^c_{>0}(f)$ has at most one negative connected component.*

Proof. Let $x, y \in f^{-1}(\mathbb{R}_{<0})$, define $v := \mathrm{Log}(y) - \mathrm{Log}(x)$, and let e denote Euler's number. Since f has at most one negative coefficient, v is an enclosing vector, c.f. Definition 3.2(ii). Since $f_{v,x}(1) = f(x) < 0$ and $f_{v,x}(e) = f(y) < 0$, Lemma 3.3(i') implies that $f_{v,x}(t) < 0$ for all $t \in [1, e]$ and hence $\gamma_{v,x}(t) \in f^{-1}(\mathbb{R}_{<0})$ for $t \in [1, e]$. Applying Log, equality (7) gives that $\tau_{v,\mathrm{Log}(x)}(s) \in \mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$ for all $s \in [0, 1]$. As $\tau_{v,\mathrm{Log}(x)}$ in the interval $[0, 1]$ is simply the line segment joining $\mathrm{Log}(x)$ and $\mathrm{Log}(y)$, $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$ is convex. This concludes the proof. □

We will now show that the existence of one strict separating vector implies that $V^c_{>0}(f)$ has at most one negative connected component, which in addition is contractible. To this end, we need an auxiliary proposition, that states that the existence of one very strict separating vector is enough to guarantee that there is a basis of very strict separating vectors. The idea is simply that the property of being a very strict separating vector is robust under small perturbations.

For a finite collection of vectors $w_1, \ldots, w_k \in \mathbb{R}^n$ we write

$$\mathrm{Cone}(w_1, \ldots, w_k) := \Big\{ \sum_{i=1}^{k} \lambda_i w_i \mid \lambda_1, \ldots, \lambda_k \in \mathbb{R}_{\geq 0} \Big\} \tag{11}$$

for the convex cone generated by $w_1, \ldots, w_k$. If $w_1, \ldots, w_k$ are linearly independent, then the relative interior of $\mathrm{Cone}(w_1, \ldots w_k)$ is given by

$$\mathrm{Cone}^\circ(w_1, \ldots, w_k) = \Big\{ \sum_{i=1}^{k} \lambda_i w_i \mid \lambda_1, \ldots, \lambda_k \in \mathbb{R}_{>0} \Big\}. \tag{12}$$

Proposition 3.5. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial and $v \in \mathbb{R}^n$ a very strict separating vector of $\sigma(f)$. Then there exists a basis $\{w_1, \ldots, w_n\}$ of $\mathbb{R}^n$ consisting of very strict separating vectors, and a constant $c \in \mathbb{R}$ such that*

$$\sigma_-(f) \subseteq \mathcal{H}^+_{w_i,c}, \qquad \sigma_+(f) \subseteq \mathcal{H}^-_{w_i,c} \qquad \text{for every} \quad i \in \{1, \ldots, n\}, \tag{13}$$

$$v \in \mathrm{Cone}^\circ(w_1, \ldots, w_n). \tag{14}$$

Proof. Define

$$a := \max_{\alpha \in \sigma_+(f)} v \cdot \alpha, \qquad b := \min_{\beta \in \sigma_-(f)} v \cdot \beta, \qquad c := \tfrac{a+b}{2}.$$

As $v \in \mathcal{S}^-(f)$, $\sigma_-(f) \subseteq \mathcal{H}^+_{v,c}$ and $\sigma_+(f) \subseteq \mathcal{H}^-_{v,c}$ by (9). Since v is very strict, we have $b > c > a$.

Choose a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{R}^n$ such that $v \in \mathrm{Cone}^\circ(v_1, \ldots, v_n)$. By (12) this is equivalent to the existence of $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_{>0}$ such that $v = \sum_{i=1}^n \lambda_i v_i$. For this basis, we define

$$K := \min_{i=1,\ldots,n} \min_{\mu \in \sigma(f)} v_i \cdot \mu, \qquad L := \max_{i=1,\ldots,n} \max_{\mu \in \sigma(f)} v_i \cdot \mu.$$

In the following, we show that it is possible to choose $\epsilon_i > 0$ such that the vectors $w_i := v + \epsilon_i v_i$, for $i = 1, \ldots, n$, with the given c satisfy (13). For $\beta \in \sigma_-(f)$ and $i \in \{1, \ldots, n\}$, using that $v_i \cdot \beta \geq K$ and $v \cdot \beta \geq b$, it holds that

$$w_i \cdot \beta = v \cdot \beta + \epsilon_i (v_i \cdot \beta) \geq b + \epsilon_i K \begin{cases} \geq b > c & \text{if } K \geq 0 \text{ and for } \epsilon_i > 0, \\ > b + \frac{a-b}{2K} K = c & \text{if } K < 0 \text{ and for } 0 < \epsilon_i < \frac{a-b}{2K}. \end{cases} \tag{15}$$

Similarly, for every $\alpha \in \sigma_+(f)$ and $i \in \{1, \ldots, n\}$, it follows that

$$w_i \cdot \alpha = v \cdot \alpha + \epsilon_i (v_i \cdot \alpha) \leq a + \epsilon_i L \begin{cases} \leq a < c & \text{if } L \leq 0 \text{ and for } \epsilon_i > 0, \\ < a + \frac{b-a}{2L} L = c & \text{if } L > 0 \text{ and for } 0 < \epsilon_i < \frac{b-a}{2L}. \end{cases} \tag{16}$$

Therefore, there exists an $\epsilon > 0$ such that w_i satisfies (15) and (16) for all $0 < \epsilon_i < \epsilon$ and $i \in \{1, \dots, n\}$. Hence for sufficiently small $\epsilon_1, \dots, \epsilon_n$ the vectors $w_1, \dots, w_n$ are very strict separating vectors satisfying (13).

To obtain (14), we specify a choice of $\epsilon_1, \dots, \epsilon_n$. For each $i \in \{1, \dots, n\}$, choose $p_i > 0$ such that $\epsilon_i := \frac{\lambda_i}{p_i} < \epsilon$ and define $P := \sum_{i=1}^n p_i$. By construction, we have that

$$\sum_{i=1}^n \tfrac{p_i}{P+1} w_i = \sum_{i=1}^n \tfrac{p_i}{P+1}(v + \tfrac{\lambda_i}{p_i} v_i) = \tfrac{P}{P+1} v + \tfrac{1}{P+1} \sum_{i=1}^n \lambda_i v_i = v,$$

which gives that $v \in \mathrm{Cone}^\circ(w_1, \dots, w_n)$.

Finally, since v is a positive linear combination of $v_1, \dots, v_n$ and $\epsilon_1, \dots, \epsilon_n$ are positive, an easy linear algebra argument shows that $w_1, \dots, w_n$ form a basis of $\mathbb{R}^n$. □

Theorem 3.6. *Let $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If there exists a strict separating vector of $\sigma(f)$, then*

(i) *$f^{-1}(\mathbb{R}_{<0})$ is non-empty and contractible.*
(ii) *The closure of $f^{-1}(\mathbb{R}_{<0})$ equals $f^{-1}(\mathbb{R}_{\leq 0})$.*

In particular, $V^c_{>0}(f)$ has at most one negative connected component.

Proof. Let $v \in \mathcal{S}^-(f)$ be a strict separating vector. Define

$$a := \max_{\alpha \in \sigma_+(f)} v \cdot \alpha, \quad \text{and} \quad M := \{\beta \in \sigma_-(f) \mid v \cdot \beta = a\} = \sigma_-(f) \cap \mathcal{H}_{v,a}.$$

Since v is a strict separating vector, $\sigma_-(f) \setminus M \neq \emptyset$. Consider the restriction of f to $\sigma(f) \setminus M$, c.f. (3):

$$\tilde{f} := f_{|\sigma(f) \setminus M}.$$

As $\tilde{f}$ is obtained from f only by removing monomials with negative coefficients, $f(x) \leq \tilde{f}(x)$ for all $x \in \mathbb{R}^n_{>0}$ and hence $\tilde{f}^{-1}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0})$. By construction $\sigma_-(\tilde{f}) \neq \emptyset$, and v is also a strict separating vector of $\sigma(\tilde{f})$, which additionally satisfies

$$\max_{\alpha \in \sigma_+(\tilde{f})} v \cdot \alpha < \min_{\beta \in \sigma_-(\tilde{f})} v \cdot \beta.$$

Hence, v is a very strict separating vector of $\sigma(\tilde{f})$. Note that for any $x \in \mathbb{R}^n_{>0}$, the leading coefficient of $\tilde{f}_{v,x}$ is negative by Lemma 3.3(ii), and hence $\tilde{f}^{-1}(\mathbb{R}_{<0}) \neq \emptyset$. It follows that $f^{-1}(\mathbb{R}_{<0}) \neq \emptyset$ as well.

We show that $f^{-1}(\mathbb{R}_{<0})$ is contractible, by showing that this is the case for $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. First, note that by Proposition 3.5, there exists a basis $\{w_1, \dots, w_n\}$ of $\mathbb{R}^n$, consisting of very strict separating vectors of $\sigma(\tilde{f})$ such that v can be written as

$$v = \sum_{i=1}^n \lambda_i w_i \qquad \text{for some} \quad \lambda = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n_{>0}. \tag{17}$$

To show that $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$ is contractible, we will show that for any $\xi \in \mathrm{Log}(\tilde{f}^{-1}(\mathbb{R}_{<0}))$, it holds that $\xi + \mathrm{Cone}(w_1, \dots, w_n)$ is a strong deformation retract of $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. As $\xi + \mathrm{Cone}(w_1, \dots w_n)$ is contractible, this will conclude the proof of (i), c.f. [27].

To this end, fix $x \in \tilde{f}^{-1}(\mathbb{R}_{<0})$ and let $\xi = \mathrm{Log}(x)$. For $w \in \mathcal{S}^-(\tilde{f})$, the path $\gamma_{w,x}$ is contained in $\tilde{f}^{-1}(\mathbb{R}_{<0})$ by Lemma 3.3(ii'). Hence, by equality (7), the path $\tau_{w,\xi}$ is contained in $\mathrm{Log}(\tilde{f}^{-1}(\mathbb{R}_{<0}))$. In particular, it holds that $\xi + w \in \mathrm{Log}(\tilde{f}^{-1}(\mathbb{R}_{<0}))$ for all $w \in \mathcal{S}^-(\tilde{f})$. As $\mathcal{S}^-(\tilde{f})$ is a convex cone and contains $w_1, \dots, w_n$, we have $\mathrm{Cone}(w_1, \dots, w_n) \subseteq \mathcal{S}^-(\tilde{f})$ [44, Ch. 1]. It follows that $\xi + \mathrm{Cone}(w_1, \dots, w_n) \subseteq \mathrm{Log}(\tilde{f}^{-1}(\mathbb{R}_{<0})) \subseteq \mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$.

We now construct a homotopy map giving that $\xi + \mathrm{Cone}(w_1, \dots, w_n)$ is a strong deformation retract of $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. To this end, we consider the map $s^* \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ defined by

$$s^*(\zeta) = \min\{s \in \mathbb{R}_{\geq 0} \mid \zeta + s\, v \ \in \xi + \mathrm{Cone}(w_1, \dots, w_n)\}.$$

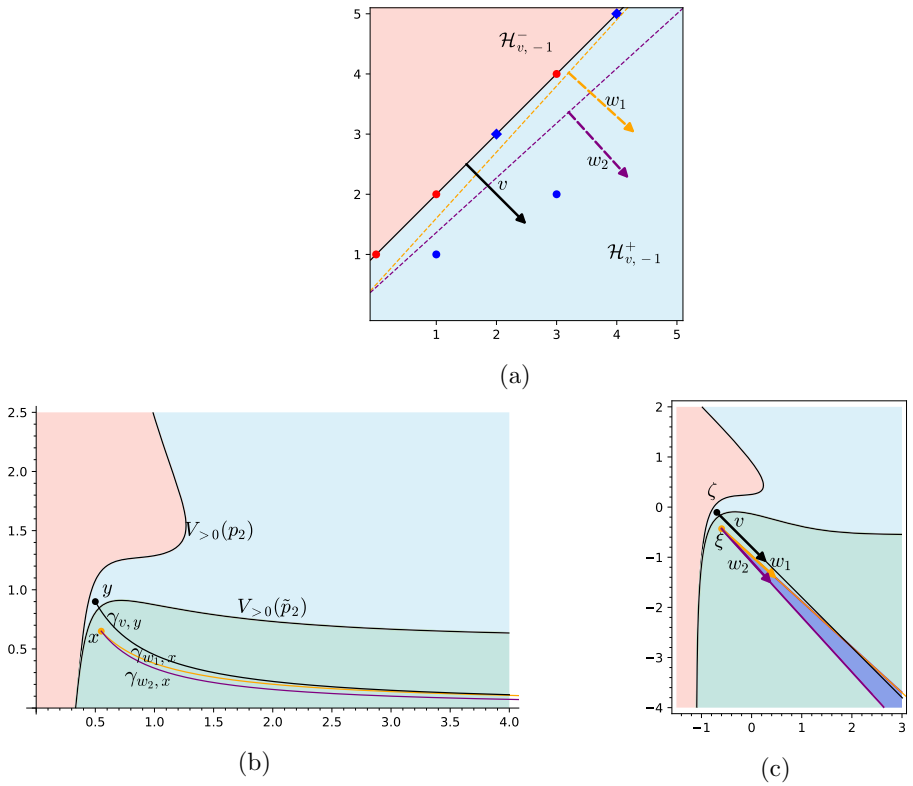


FIGURE 3. Graphical representation of Example 3.7. (a) $v = (1,-1) \in \mathcal{S}^-(p_2)$ is a strict separating vector, the vectors $w_1 = (1.1,-1)$ and $w_2 = (1,-1.1)$ are very strict separating vectors of the support of $\tilde{p}_2(x_1,x_2)$ and form a basis of $\mathbb{R}^2$. (b) $p_2^{-1}(\mathbb{R}_{<0})$ shown in blue and its subset $\tilde{p}_2^{-1}(\mathbb{R}_{<0})$ shown in green. (c) The half-line $\mathrm{Log}(\gamma_{v,y})$ intersects the cone generated by w_1, w_2 with apex $\xi = \mathrm{Log}(x)$.

To see that s^* is well defined and continuous, we note that

$$s^*(\zeta) = \max\left\{0, -\tfrac{(W^{-1}(\zeta-\xi))_1}{\lambda_1}, \ldots, -\tfrac{(W^{-1}(\zeta-\xi))_n}{\lambda_n}\right\},$$

where $W \in \mathbb{R}^{n\times n}$ is the matrix of the linear isomorphism that sends the i-th standard basis vector of $\mathbb{R}^n$ to w_i, and $\lambda_1, \ldots, \lambda_n > 0$ are from (17).

Consider the following continuous map

$$\rho\colon [0,1] \times \mathrm{Log}(f^{-1}(\mathbb{R}_{<0})) \to \mathrm{Log}(f^{-1}(\mathbb{R}_{<0})), \quad (t,\zeta) \mapsto \zeta + t\, s^*(\zeta)\, v. \tag{18}$$

Since v is a strict separating vector of $\sigma(f)$, from Lemma 3.3(ii') follows that $\rho(t,\zeta) \in \mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$ for all $(t,\zeta) \in [0,1] \times \mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. Clearly, $\rho(0,\cdot)$ is the identity map, and by definition of s^*, $\rho(1,\zeta) \in \xi + \mathrm{Cone}(w_1,\ldots,w_n)$ for all $\zeta \in \mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. Furthermore, if $\zeta \in \xi + \mathrm{Cone}(w_1,\ldots,w_n)$, then $s^*(\zeta) = 0$ and $\rho(t,\zeta) = \zeta$ for all $t \in [0,1]$.

We conclude that ρ is a homotopy showing that $\xi + \mathrm{Cone}(w_1,\ldots,w_n)$ is a strong deformation retract of $\mathrm{Log}(f^{-1}(\mathbb{R}_{<0}))$. This implies (i).

Finally, we show statement (ii). Let $x \in f^{-1}(\{0\})$. Since $v \in \mathcal{S}^-(f)$ and strict, Lemma 3.3(ii) gives that $f_{v,x}(t) < 0$ for all $t > 1$. Thus the sequence $\left(\gamma_{v,x}(1+\frac{1}{n})\right)_{n\in\mathbb{N}}$ belongs to $f^{-1}(\mathbb{R}_{<0})$. As $\gamma_{v,x}$ is continuous and $\gamma_{v,x}(1) = x$, the sequence $\left(\gamma_{v,x}(1+\frac{1}{n})\right)_{n\in\mathbb{N}}$ converges to x. So each $x \in f^{-1}(\mathbb{R}_{\leq 0})$ is the limit of a convergent sequence in $f^{-1}(\mathbb{R}_{<0})$. Hence $f^{-1}(\mathbb{R}_{\leq 0}) \subseteq \overline{f^{-1}(\mathbb{R}_{<0})}$. The other inclusion is clear by the continuity of f. □

Example 3.7. Consider the signomial

$$p_2(x_1,x_2) = -x_1^4x_2^5 + 3x_1^3x_2^4 - x_1^3x_2^2 - x_1^2x_2^3 + x_1x_2^2 - 3x_1x_2 + x_2.$$

Then $v = (1,-1) \in \mathcal{S}^-(p_2)$ is strict, see Fig. 3(a), and by Theorem 3.6, $V^c_{>0}(p_2)$ has one negative connected component which is a contractible set.

Fig. 3 displays the idea of the proof of Theorem 3.6. First, one considers the signomial obtained by removing the negative monomials on the separating hyperplane $\mathcal{H}_{v,-1}$ from Fig. 3(a):

$$\tilde{p}_2(x_1,x_2) = 3x_1^3x_2^4 - x_1^3x_2^2 + x_1x_2^2 - 3x_1x_2 + x_2.$$

Using Proposition 3.5, one can find strict separating vectors $w_1 = (1.1,-1)$ and $w_2 = (1,-1.1)$ of $\sigma(\tilde{p}_2)$ such that $v \in \mathrm{Cone}(w_1,w_2)$. For a fixed $x \in \tilde{p}_2^{-1}(\mathbb{R}_{<0})$, the paths $\gamma_{w_1,x}, \gamma_{w_2,x}$ turn into half-lines with start point $\xi = \mathrm{Log}(x)$ under the coordinate-wise logarithm map (see Fig. 3 (b,c)). For each point $\zeta = \mathrm{Log}(y) \in \mathrm{Log}(p_2^{-1}(\mathbb{R}_{<0}))$, the half-line with start point ζ and direction vector v intersects $\mathrm{Cone}(w_1,w_2)$. By sending ζ to the first such intersection point, we obtain that $\mathrm{Cone}(w_1,w_2)$ is a strong deformation retract of $\mathrm{Log}(p_2^{-1}(\mathbb{R}_{<0}))$.

The results provided so far guarantee that $V^c_{>0}(f)$ has at most one negative connected component. With analogous techniques, the existence of strict enclosing vectors of $\sigma(-f)$ gives that $V^c_{>0}(f)$ has at most two negative connected components. Note that a strict enclosing vector of $\sigma(-f)$ defines two parallel hyperplanes such that the positive points of $\sigma(f)$ are between them, and the negative points of $\sigma(f)$ are on the other side of these hyperplanes.

Theorem 3.8. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If there exists a strict enclosing vector of $\sigma(-f)$, then $V^c_{>0}(f)$ has at most two negative connected components.*

Proof. Let $v \in \mathcal{E}^-(-f)$ be a strict enclosing vector. Then for $\beta \in \sigma_+(-f) = \sigma_-(f)$, it holds that either

$$v\cdot\beta \leq \min_{\alpha\in\sigma_+(f)} v\cdot\alpha \quad \text{or} \quad \max_{\alpha\in\sigma_+(f)} v\cdot\alpha \leq v\cdot\beta.$$

As v is strict, the following sets are non-empty:

$$M := \{\beta \in \sigma_-(f) \mid \max_{\alpha\in\sigma_+(f)} v\cdot\alpha < v\cdot\beta\}, \qquad N := \{\beta \in \sigma_-(f) \mid v\cdot\beta < \min_{\alpha\in\sigma_+(f)} v\cdot\alpha\}.$$

Consider the restriction of f to the sets $M \cup \sigma_+(f)$ and $N \cup \sigma_+(f)$:

$$\tilde{f}_M := f_{|M\cup\sigma_+(f)} \qquad\qquad \tilde{f}_N := f_{|N\cup\sigma_+(f)}.$$

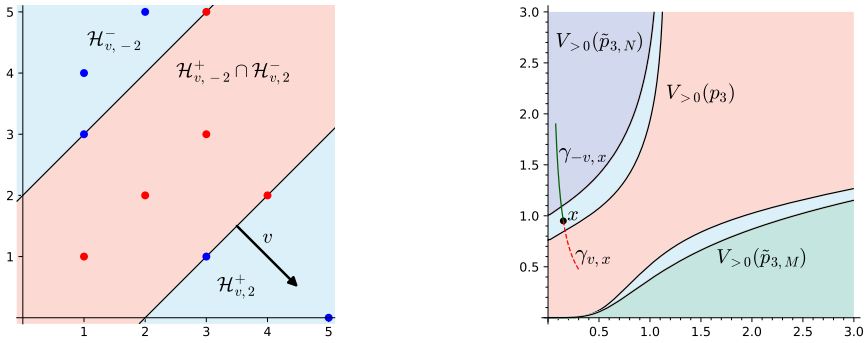
By construction, see (9), v and $-v$ are strict separating vectors of $\sigma(\tilde{f}_M)$ and $\sigma(\tilde{f}_N)$ respectively. Hence $\tilde{f}_N^{-1}(\mathbb{R}_{<0})$ and $\tilde{f}_M^{-1}(\mathbb{R}_{<0})$ are path connected by Theorem 3.6. Additionally, as the sets of negative points in $\sigma(\tilde{f}_M)$ and $\sigma(\tilde{f}_N)$ are included in $\sigma_-(f)$, it holds $f(x) \leq \tilde{f}_N(x)$ and $f(x) \leq \tilde{f}_M(x)$ for all $x \in \mathbb{R}^n_{>0}$ and hence

$$\tilde{f}_M^{-1}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0}), \qquad \tilde{f}_N^{-1}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0}).$$

With this in place, if we show that for every $x \in f^{-1}(\mathbb{R}_{<0})$ there is a continuous path to a point in $\tilde{f}_M^{-1}(\mathbb{R}_{<0})$ or to a point in $\tilde{f}_N^{-1}(\mathbb{R}_{<0})$ and this path is contained in $f^{-1}(\mathbb{R}_{<0})$, then the number of connected components of $f^{-1}(\mathbb{R}_{<0})$ is at most 2.

Fix $x \in f^{-1}(\mathbb{R}_{<0})$. As v is a strict separating vector of $\sigma(\tilde{f}_M)$ and $-v$ of $\sigma(\tilde{f}_N)$, there exist $t_x, d_x > 1$ such that $\gamma_{v,x}(t_x) \in \tilde{f}_M^{-1}(\mathbb{R}_{<0})$ and $\gamma_{-v,x}(d_x) \in \tilde{f}_N^{-1}(\mathbb{R}_{<0})$ by Lemma 3.3(ii).

By Lemma 3.3(i), $f_{v,x}$ has negative leading and smallest degree coefficients, and the coefficient sign sequence has at most two sign changes. Hence either $f_{v,x}(t) < 0$ for all $t \geq 1$ or $f_{v,x}(t) < 0$ for all $t \leq 1$. If $f_{v,x}(t) = f(\gamma_{v,x}(t)) < 0$ for all $t \geq 1$, then the path $\gamma_{v,x}$ connects x to a point in $\tilde{f}_M^{-1}(\mathbb{R}_{<0})$. If $f_{v,x}(t) < 0$ for all $t \leq 1$, then $f_{-v,x}(t) = f_{v,x}(t^{-1}) < 0$ for all $t \geq 1$. Hence the path $\gamma_{-v,x}$ connects x to a point in $\tilde{f}_N^{-1}(\mathbb{R}_{<0})$. This concludes the proof. □



(a) (b)

FIGURE 4. Illustration of Example 3.9. (a) A strict enclosing vector for $-p_3$ is shown. (b) The positive connected component of $V^c_{>0}(p_3)$ is shown in red, the negative connected components of $V^c_{>0}(p_3)$ are shown in blue, the subset $\tilde{p}^{-1}_{3,M}(\mathbb{R}_{<0})$ is shown in green, and the subset $\tilde{p}^{-1}_{3,N}(\mathbb{R}_{<0})$ is shown in purple. The path $\gamma_{v,x}$ from $x = (0.15, 0.95)$ to $\tilde{p}^{-1}_{3,M}(\mathbb{R}_{<0})$, shown dashed in red, is not contained in $p_3^{-1}(\mathbb{R}_{<0})$. The path $\gamma_{-v,x}$, shown in solid green, connects x with $\tilde{p}^{-1}_{3,N}(\mathbb{R}_{<0})$ and does not leave $p_3^{-1}(\mathbb{R}_{<0})$.

Example 3.9. Consider the signomial

$$p_3(x_1, x_2) = x_1^3x_2^5 - x_1^2x_2^5 + x_1^4x_2^2 + x_1^3x_2^3 - x_1^5 - x_1x_2^4 - x_1^3x_2 + 3x_1^2x_2^2 - x_1x_2^3 + x_1x_2.$$

The vector $v = (1, -1)$ is a strict enclosing vector of $-p_3$, see Fig. 4(a). Hence, the number of negative connected components of $V^c_{>0}(p_3)$ is at most two by Theorem 3.8.

In Fig. 4(b), the idea of the proof of Theorem 3.8 is illustrated. The following two signomials are considered

$$\tilde{p}_{3,M}(x_1, x_2) = x_1^3x_2^5 + x_1^4x_2^2 + x_1^3x_2^3 - x_1^5 + 3x_1^2x_2^2 + x_1x_2,$$
$$\tilde{p}_{3,N}(x_1, x_2) = x_1^3x_2^5 - x_1^2x_2^5 + x_1^4x_2^2 + x_1^3x_2^3 - x_1x_2^4 + 3x_1^2x_2^2 + x_1x_2.$$

For each of these signomials, the pre-image of $\mathbb{R}_{<0}$ is path connected and contained in $p_3^{-1}(\mathbb{R}_{<0})$. Using the paths $\gamma_{v,x}$ or $\gamma_{-v,x}$, any point $x \in p_3^{-1}(\mathbb{R}_{<0})$ is connected to one of these two connected sets.

Remark 3.10. The conditions of Theorems 3.6 and 3.8 can be checked computationally using linear programming. Finding a separating vector of $\sigma(f)$ corresponds to finding a solution of the linear inequality system

$$v \cdot \alpha \leq a, \quad \alpha \in \sigma_+(f), \qquad v \cdot \beta \geq a, \qquad \beta \in \sigma_-(f), \tag{19}$$

where $v \in \mathbb{R}^n, a \in \mathbb{R}$ are treated as unknown variables. Existing software like SageMath [41], Polymake [23] and other linear programming software can find a solution to (19) even for large number of variables and of inequalities.

Finding an enclosing hyperplane as in Theorem 3.8 can be more demanding computationally. A naive approach is to consider all partitions of $\sigma_-(f)$ into two sets $\sigma_{-,1}(f), \sigma_{-,2}(f)$ and for each partition decide the feasibility of the system of linear inequalities

$$v \cdot \beta \leq a, \quad \beta \in \sigma_{-,1}(f) \qquad a \leq v \cdot \alpha \leq b, \quad \alpha \in \sigma_+(f), \qquad v \cdot \beta \leq b, \quad \beta \in \sigma_{-,2}(f).$$

Remark 3.11. One might be tempted to believe that in the situation of Theorem 3.8, $V^c_{>0}(f)$ has at most one positive connected component. However, Example 2.2 gives a counter example, as $V^c_{>0}(p_1)$ has two positive connected components, and the vector $v = (0, 1)$ satisfies the hypotheses of Theorem 3.8, see Fig. 2.

A direct consequence of Theorems 3.6 and 3.8 applies to the case where the positive points of $\sigma(f)$ belong to a hyperplane that does not contain all the negative points of $\sigma(f)$.

Corollary 3.12. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If for some $v \in \mathbb{R}^n$ and $a \in \mathbb{R}$*

$$\sigma_+(f) \subseteq \mathcal{H}_{v,a} \qquad \text{and} \qquad \sigma_-(f) \nsubseteq \mathcal{H}_{v,a},$$

then $V^c_{>0}(f)$ has at most two negative connected components.

Proof. The conditions imply that either v is a strict enclosing vector of $\sigma(-f)$, or either v or $-v$ is a strict separating vector of $\sigma(-f)$. The statement then follows from Theorem 3.8 or Theorem 3.6. □

Corollary 3.13. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If*

$$\#\sigma_+(f) \leq \dim \mathrm{N}(f),$$

then $V^c_{>0}(f)$ has at most two negative connected components.

Proof. Since $\#\sigma_+(f) \leq \dim \mathrm{N}(f) \leq n$, the points $\sigma_+(f)$ lie on an affine subspace of dimension at most $\dim \mathrm{N}(f) - 1$. Necessarily, this subspace cannot contain all points of $\sigma(f)$. Hence, there exists an affine hyperplane $\mathcal{H}_{v,a}$ containing $\sigma_+(f)$ and not containing $\sigma_-(f)$. Now, the statement follows from Corollary 3.12. □

Remark 3.14. The techniques used in this section rely on the observation that the paths (5) become half-lines at the logarithmic scale. Studying images of algebraic sets under the coordinate-wise logarithm map has a rich history. In 1994, Gelfand et al. [24] introduced the *amoeba* of a Laurent polynomial $f \in \mathbb{C}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ which is the image of the set $\{z \in (\mathbb{C}^*)^n \mid f(z) = 0\}$ under the map $(\mathbb{C}^*)^n \to \mathbb{R}^n, (z_1, \ldots, z_n) \mapsto (\log(|z_1|), \ldots, \log(|z_n|))$. Since then, many results have been proved about the structure of the connected components of the complement of the amoeba. It is known that these connected components are convex [24, Corollary 1.6], their number is at least equal to the number of vertices of the Newton polytope $\mathrm{N}(f)$ and at most equal to the total number of integer points in $\mathrm{N}(f) \cap \mathbb{Z}^n$ [20, Theorem 2.8]. Furthermore, if the polynomial is maximally sparse (i.e. every exponent of f is a vertex of $\mathrm{N}(f)$), then the number of connected components of the complement of the amoeba is equal to the number of vertices of $\mathrm{N}(f)$ [36], and each of these components is unbounded [24, Corollary 1.8].

The logarithmic image of $V_{>0}(f)$ can be seen as the "positive real part" of the amoeba of f. Therefore, one might hope that statements about amoebas can be translated directly to answer Problem 1.1. However, logarithmic images of $V_{>0}(f)$ have been studied in [2], where the author concluded that, in general, it is not possible to use properties of the amoeba to understand the logarithmic image of $V_{>0}(f)$ [2, Section 5.1]. To illustrate that the amoeba of f and the logarithmic image can behave differently, we recall the following example [38, Example 2.6]. Consider the maximally sparse polynomial $f = 1 - x_1 - x_2 + \frac{6}{5}x_1^4 x_2 + \frac{6}{5}x_1 x_2^4$. The complement of the amoeba of f has 5 connected components, which are convex and unbounded. However, it is easy to see that the complement of $\mathrm{Log}(V_{>0}(f))$ has a bounded connected component, which is contained in the amoeba of f.

4. Convexification of signomials

In Section 3, we used continuous paths (5), which are half-lines on logarithmic scale, to derive bounds for the number of negative connected components of $V^c_{>0}(f)$, where f is a signomial function. In this section, we take a different approach to bound the number of negative connected components of $V^c_{>0}(f)$. We use the almost trivial observation that every sublevel set of a convex function is a convex set (see e.g. [37, Theorem 4.6.]). Therefore, $V^c_{>0}(f)$ has at most one negative connected component, if f is a convex function. With this in mind, we investigate what signomials can be transformed into a convex function using Lemma 2.3.

From [34, Theorem 7], one can easily derive a sufficient condition for convexity of signomials.

Lemma 4.1. *A signomial $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ is a convex function if the following holds:*

(a) *For each* $\alpha \in \sigma_+(f)$*, it holds that*
  (i) $\alpha_i \leq 0$ *for all* $i = 1, \ldots, n$*, or*
  (ii) *there exists* $j \in \{1, \ldots, n\}$ *such that* $\alpha_i \leq 0$ *for all* $i \neq j$ *and* $(1, \ldots, 1) \cdot \alpha \geq 1$*,*
(b) *For each* $\beta \in \sigma_-(f)$*, it holds that* $\beta_i \geq 0$ *for all* $i = 1, \ldots, n$ *and* $(1, \ldots, 1) \cdot \beta \leq 1$*.*

Proof. By [34, Theorem 7], hypotheses (a) and (b) imply that each term $c_\alpha x^\alpha$, $\alpha \in \sigma_+(f)$ and $c_\beta x^\beta$, $\beta \in \sigma_-(f)$ is convex. The result follows from the fact that the sum of convex functions is convex. □

We proceed to interpret the conditions in Lemma 4.1 geometrically.

Definition 4.2. Given an n-simplex $P \subseteq \mathbb{R}^n$ with vertices $\mu_0, \ldots, \mu_n$, we define for $k \in \{0, \ldots, n\}$ the *negative vertex cone* at the vertex μ_k as

$$\begin{aligned} P^{-,k} &:= \mu_k + \operatorname{Cone}(\mu_k - \mu_0, \ldots, \mu_k - \mu_n) \\ &= \left\{ \sum_{i=0}^n \lambda_i \mu_i \mid \sum_{i=0}^n \lambda_i = 1,\ \lambda_i \leq 0 \text{ for all } i \neq k \right\}. \end{aligned}$$

We write $P^- = \bigcup_{k=0}^n P^{-,k}$.

Note that it follows that $\lambda_k > 0$ in the definition of $P^{-,k}$. The name 'negative vertex cone' comes from [9, 13], where the authors refer to the *vertex cone* as the pointed convex cone with apex μ_k and generators the edge directions pointing out of μ_k. Fig. 5(a) shows an example of the negative vertex cones in the plane.

The next proposition provides another geometric interpretation of negative vertex cones. First recall that every n-simplex $P \subseteq \mathbb{R}^n$ has $n+1$ facets, each facet F is supported on a hyperplane $\mathcal{H}_{v_F, a_F}$, and it holds that $P = \bigcap_{F \subseteq P \text{ facet}} \mathcal{H}^-_{v_F, a_F}$ [26, Section 4.1].

Proposition 4.3. *Let* $P = \operatorname{Conv}(\mu_0, \ldots, \mu_n) \subseteq \mathbb{R}^n$ *be an* n*-simplex. A point* $\alpha \in \mathbb{R}^n$ *belongs to* $P^{-,k}$ *for* $k \in \{0, \ldots, n\}$*, if and only if* $\alpha \in \mathcal{H}^+_{v_F, a_F}$ *for all facets* F *of* P *containing* μ_k*. In that case, it holds* $\alpha \in \mathcal{H}^-_{v_F, a_F}$ *for the facet* F *not containing* μ_k*.*

Proof. Denote by F_i the facet of P that does not contain μ_i and $\mathcal{H}_{v_i, a_i}$ a supporting hyperplane. In particular it holds that

$$v_j \cdot \mu_i = a_j \quad \text{for} \quad i \neq j \qquad \text{and} \qquad v_i \cdot \mu_i < a_i, \qquad \text{for} \quad i = 0, \ldots, n. \tag{20}$$

The condition in the statement is equivalent to the existence of $k \in \{0, \ldots, n\}$ such that

$$v_i \cdot \alpha \geq a_i \quad \text{for} \quad i \neq k. \tag{21}$$

Write $\alpha = \sum_{j=0}^n \lambda_j \mu_j$ for $\lambda_0, \ldots, \lambda_n \in \mathbb{R}$ such that $\sum_{j=0}^n \lambda_j = 1$. Then

$$\begin{aligned} v_i \cdot \alpha &= \sum_{j=0}^n \lambda_j (v_i \cdot \mu_j) = \lambda_i (v_i \cdot \mu_i) + \sum_{j=0, j \neq i}^n \lambda_j a_i \\ &= \lambda_i (v_i \cdot \mu_i) + (1 - \lambda_i) a_i = a_i + \lambda_i (v_i \cdot \mu_i - a_i). \end{aligned} \tag{22}$$

Using this, condition (21) holds for some k if and only if

$$\lambda_i (v_i \cdot \mu_i - a_i) \geq 0 \quad \text{for} \quad i \neq k.$$

By (20), this holds if and only if $\lambda_i \leq 0$ for $i \neq k$, that is, if and only if $\alpha \in P^{-,k} \subseteq P^-$. As then, $\lambda_k \geq 0$, (22) gives that $v_k \cdot \alpha < a_k$ and hence $\alpha \in \mathcal{H}^-_{v_k, a_k}$. □

We write $\Delta_n := \operatorname{Conv}(e_0, e_1, \ldots, e_n)$ for the standard n-simplex in $\mathbb{R}^n$, where $e_1, \ldots, e_n$ are the standard basis vectors of $\mathbb{R}^n$ and e_0 denotes the zero vector.

Lemma 4.4. *Let* $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ *be a signomial. If* $\sigma_-(f) \subseteq \Delta_n$ *and* $\sigma_+(f) \subseteq \Delta_n^-$*, then* f *is a convex function.*

Proof. We show that the conditions in Lemma 4.1 are equivalent to $\sigma_-(f) \subseteq \Delta_n$ and $\sigma_+(f) \subseteq \Delta_n^-$. For $\beta \in \mathbb{R}^n$, find the unique $\lambda_0, \dots, \lambda_n \in \mathbb{R}$ such that $\sum_{i=0}^n \lambda_i e_i = \beta$ and $\sum_{i=0}^n \lambda_i = 1$. Note that $(1, \dots, 1) \cdot \beta = \sum_{i=1}^n \lambda_i = 1 - \lambda_0$, which is at most 1 if and only if $\lambda_0 \geq 0$.

Lemma 4.1(b) holds if and only if $\lambda_i \geq 0$ for all $i = 1, \dots, n$ and $\sum_{i=1}^n \lambda_i \leq 1$. Equivalently, $\lambda_i \geq 0$ for all $i = 1, \dots, n$ and $\lambda_0 \geq 0$, that is, $\beta \in \Delta_n$.

We show now that $\beta \in \Delta_n^-$ if and only if Lemma 4.1(a) holds. By definition, $\beta \in \Delta_n^-$ if and only if for some k,

$$\lambda_i \leq 0 \quad \text{for} \quad i \neq k. \tag{23}$$

For $k = 0$, (23) holds if and only if $\beta_i \leq 0$ for all i, thus Lemma 4.1(a,i) holds. For $k > 0$, (23) holds, if and only if all but the k-th coordinate of β are non-positive, and $\lambda_0 \leq 0$, equivalently $(1, \dots, 1) \cdot \beta \geq 1$, which is Lemma 4.1(a,ii). This concludes the proof. □

We next look into what signomials can be transformed into a convex signomial using the transformations from Lemma 2.3. It is well known that any two n-simplices are affinely isomorphic [44]. The next lemma shows that the negative vertex cones are preserved under such an affine transformation.

Lemma 4.5. *Let $P, Q \subseteq \mathbb{R}^n$ be n-simplices. For every $B \subseteq P$ and $A \subseteq P^-$, there exist an invertible matrix $M \in \mathrm{GL}_n(\mathbb{R})$ and a vector $v \in \mathbb{R}^n$ such that $MB + v \subseteq Q$ and $MA + v \subseteq Q^-$.*

Proof. Denote by $\{p_0, \dots, p_n\}$ and $\{q_0, \dots, q_n\}$ the vertex sets of P and Q respectively. Since P and Q are simplices, there is an invertible matrix $M \in \mathrm{GL}_n(\mathbb{R})$ such that $M(p_i - p_0) = q_i - q_0$ for $i = 1, \dots, n$. Define $v := -Mp_0 + q_0$. By construction, it holds that $Mp_i + v = q_i$ for every $i = 0, \dots, n$.

For each $\mu \in \mathbb{R}^n$, write $\mu = \sum_{i=0}^n \lambda_i p_i$ with $\sum_{i=0}^n \lambda_i = 1$. It holds that

$$M\mu + v = \sum_{i=0}^n \lambda_i M p_i + \sum_{i=0}^n \lambda_i v = \sum_{i=0}^n \lambda_i (M p_i + v) = \sum_{i=0}^n \lambda_i q_i.$$

That is, the coordinates of μ according to P and those of $M\mu + v$ according to Q are the same. From this the statement follows. □

Theorem 4.6. *Let $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If there exists an n-simplex P such that*

$$\sigma_-(f) \subseteq P, \qquad \text{and} \qquad \sigma_+(f) \subseteq P^-,$$

then $f^{-1}(\mathbb{R}_{<0})$ is either empty or contractible. In particular, $V^c_{>0}(f)$ has at most one negative connected component.

Proof. By Lemma 4.5 with $B = \sigma_-(f)$ and $A = \sigma_+(f)$, there exists $M \in \mathrm{GL}_n(\mathbb{R})$ and $v \in \mathbb{R}^n$ such that $M\sigma_-(f) + v \subseteq \Delta_n$ and $M\sigma_+(f) + v \subseteq \Delta_n^-$. By Lemma 2.3, $\sigma_+(F_{M,v,f}) = M\sigma_+(f) + v$ and $\sigma_-(F_{M,v,f}) = M\sigma_-(f) + v$. Hence by Lemma 4.4, $F_{M,v,f}$ is a convex function and thus $F^{-1}_{M,v,f}(\mathbb{R}_{<0})$ is either empty or contractible. By Lemma 2.3 again, $f^{-1}(\mathbb{R}_{<0})$ is homeomorphic to $F^{-1}_{M,v,f}(\mathbb{R}_{<0})$, and the statement of the theorem follows. □

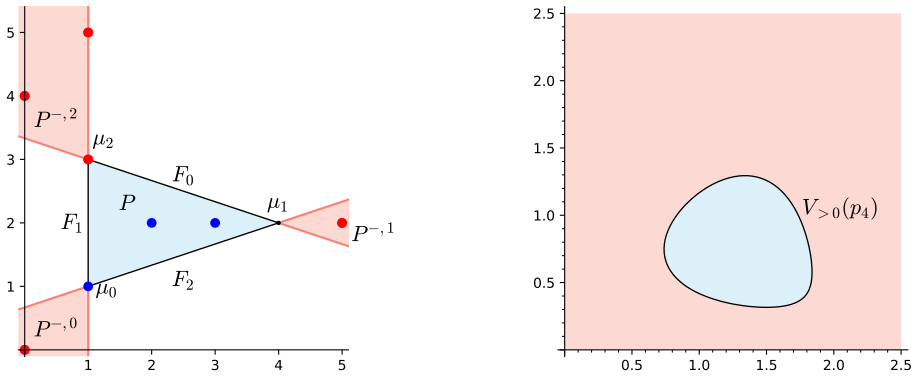
In view of Theorem 4.6, understanding P^- for a simplex P allows us to determine whether f can be transformed to a convex function.

Example 4.7. Consider the signomial

$$p_4(x_1, x_2) = x_1^5 x_2^2 + x_1 x_2^5 - 2x_1^3 x_2^2 - 3x_1^2 x_2^2 + x_1 x_2^3 + x_2^4 - x_1 x_2 + 1$$

and the simplex $P = \mathrm{Conv}((1,1),(4,2),(1,3))$. We have $\sigma_-(p_4) \subseteq P$ and $\sigma_+(p_4) \subseteq P^-$, see Fig. 5. By Theorem 4.6, the set $p_4^{-1}(\mathbb{R}_{<0})$ is contractible, since $p_4(1,1) = -1$.

A direct consequence of Theorem 4.6 states that if all positive points of $\sigma(f)$ are vertices of the Newton polytope and this is a simplex, then $f^{-1}(\mathbb{R}_{<0})$ is either empty or contractible. Let $\mathrm{Vert}(\mathrm{N}(f))$ denote the set of vertices of $\mathrm{N}(f)$.



(a) (b)

FIGURE 5. Illustration of Example 4.7. (a) A 2-simplex P, its negative cones P^- and the support of $p_4(x_1, x_2)$. (b) The set $p_4^{-1}(\mathbb{R}_{<0})$ is shown in blue.

Corollary 4.8. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If $\sigma_+(f) \subseteq \mathrm{Vert}(\mathrm{N}(f))$ and $\mathrm{N}(f)$ is a simplex, then $f^{-1}(\mathbb{R}_{<0})$ is either empty or contractible.*

Proof. Let $d := \dim \mathrm{N}(f)$ and denote by $e_1, \dots, e_d$ the first d standard basis vectors of $\mathbb{R}^n$. Without loss of generality, we can assume that $\sigma(f)$ belongs to the linear subspace generated by $e_1, \dots, e_d$ in $\mathbb{R}^n$, as this can be achieved via a change of variables as in Lemma 2.3. Hence f depends only on the variables $x_1, \dots, x_d$, and can be seen as a signomial in $\mathbb{R}^d_{>0}$ with full dimensional Newton polytope. Viewing $V^c_{>0}(f)$ in $\mathbb{R}^d_{>0}$, the statement follows from Theorem 4.6, since $\sigma_+(f) \subseteq \mathrm{Vert}(\mathrm{N}(f)) \subseteq \mathrm{N}(f)^-$ and $\sigma_-(f) \subseteq \mathrm{N}(f)$.

The proof is completed noticing that the pre-image of a contractible subset of $\mathbb{R}^d_{>0}$ under the projection map $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_d)$ is contractible in $\mathbb{R}^n_{>0}$. □

Remark 4.9. Finding a simplex P that satisfies the conditions of Theorem 4.6 might be challenging even in low dimensions. For a partition of $\sigma_+(f)$ into $n+1$ sets $\sigma_{+,0}(f), \dots, \sigma_{+,n}(f)$, Proposition 4.3 give rise to a system of linear inequalities that the normal vectors of the facets of P need to satisfy to ensure that $\sigma_-(f) \subseteq P$ and $\sigma_{+,i}(f) \subseteq P^{-,i}$ for $i = 0, \dots, n$. To verify that a solution of this system gives indeed an n-simplex, one can employ Lemma 4.10 below, whose proof is given for completeness.

Using these observations, the existence of a simplex P satisfying the conditions of Theorem 4.6 can be established by verifying the feasibility of a system of polynomial inequalities. This can be for example achieved using quantifier elimination [18]; see [40] for an implementation.

Lemma 4.10. *Let $\{\mathcal{H}_{w_0,a_0}, \dots, \mathcal{H}_{w_n,a_n}\}$ be a set of hyperplanes of $\mathbb{R}^n$ such that:*

(i) Every proper subset of $\{w_0, \dots, w_n\}$ is linearly independent.

(ii) For every $i \in \{0, \dots, n\}$ it holds that $\bigcap_{j=0, j\neq i}^n \mathcal{H}_{w_j,a_j} \subseteq \mathcal{H}^{-,\circ}_{w_i,a_i}$.

Then $\bigcap_{j=0}^n \mathcal{H}^-_{w_j,a_j}$ is an n-simplex.

Proof. First, note that (ii) implies

$$(ii') \qquad \bigcap_{j=0}^{n} \mathcal{H}_{w_j,a_j} = \emptyset.$$

As a finite intersection of closed half-spaces, $P := \bigcap_{j=0}^n \mathcal{H}^-_{w_j,a_j}$ is a convex polyhedron. Each face of P has the form

$$P_I = P \cap H_I, \qquad H_I = \bigcap_{i\in I} \mathcal{H}_{w_i,a_i},$$

for some non-empty subset $I \subseteq \{0, \dots, n\}$. By (i) and (ii'), H_I is zero dimensional if and only if I has n elements. By (ii), for $I = \{0, \dots, n\} \setminus \{i\}$, $P_I \neq \emptyset$ and hence P_I is a vertex of P, denoted

by μ_i. Furthermore, the points $\mu_0, \dots, \mu_n$ are affinely independent. This follows from (ii'), as for each k, $\mu_i \in \mathcal{H}_{w_k,a_k}$ for $i \neq k$ and $\mu_k \notin \mathcal{H}_{w_k,a_k}$. Hence $\mathrm{Conv}(\mu_0, \dots, \mu_n)$ is an n-simplex. Finally, $P = \mathrm{Conv}(\mu_0, \dots, \mu_n)$ as $\mathcal{H}^{-,\circ}_{w_k,a_k}$ contains a vertex for each k. □

We conclude the section with Proposition 4.11, which states that if there are $n-1$ linearly independent non-strict separating vectors and the convex hull of the negative points does not contain positive points, then a simplex satisfying the conditions of Theorem 4.6 exists. This case, together with the scenario with one negative point in Theorem 3.4 or the existence of a strict separating vector in Theorem 3.6, conform the situations where one can effectively conclude that $V^c_{>0}(f)$ has at most one negative connected component.

Proposition 4.11. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial, such that $\sigma(f)$ has at least two negative points. Assume that there exist $n-1$ linearly independent separating vectors of $\sigma(f)$, which are not strict and that $\mathrm{Conv}(\sigma_-(f)) \cap \sigma_+(f) = \emptyset$. Then there exists an n-simplex P such that $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$.*

Proof. Let $w_1, \dots, w_{n-1}$ be non-strict separating vectors. Then with $a_i := \max\{w_i \cdot \alpha \mid \alpha \in \sigma_+(f)\}$, it holds

$$\sigma_+(f) \subseteq \bigcap_{i=1}^{n-1} \mathcal{H}^-_{w_i,a_i} \quad \text{and} \quad \sigma_-(f) \subseteq L \qquad \text{with} \quad L := \bigcap_{i=1}^{n-1} \mathcal{H}_{w_i,a_i}. \tag{24}$$

If $\sigma_-(f) \subseteq L$, then any simplex P having as edge $\mathrm{Conv}(\sigma_-(f))$ satisfies the statement. Hence, we assume that this is not the case. We prove the proposition by applying Lemma 4.10. We introduce the following:

$$v := \sum_{i=1}^{n-1} w_i \in \mathbb{R}^{n-1}, \qquad d := \sum_{i=1}^{n-1} a_i \in \mathbb{R}, \qquad K := \max\{v \cdot \alpha \mid \alpha \in \sigma_+(f),\ v \cdot \alpha \neq d\} \in \mathbb{R}.$$

By assumption, $\epsilon := d - K > 0$ and we have $\sigma_-(f) \subseteq \mathcal{H}_{v,d}$. Let $z \in \mathbb{R}^n$ such that $z, w_1, \dots, w_{n-1}$ are linearly independent, and denote by β_0, β_1 the vertices of $\mathrm{Conv}(\sigma_-(f))$ where the linear form induced by z attains its minimum and its maximum respectively. These vertices are different, otherwise each $\beta \in \mathrm{Conv}(\sigma_-(f))$ would be the unique solution of $z \cdot \beta = z \cdot \beta_0$, $w_i \cdot \beta = a_i$, $i = 1, \dots, n-1$. This would be a contradiction, since $\sigma_-(f)$ contains at least two points.

We let $M := \max\{z \cdot \alpha \mid \alpha \in \sigma_+(f)\}$, choose $\lambda > \mu$ positive real numbers such that

$$\lambda(M - z \cdot \beta_0) \leq \epsilon = d - K, \qquad \mu(M - z \cdot \beta_1) \leq \epsilon = d - K, \tag{25}$$

and define $w_0 := v + \lambda z$, $w_n := -v - \mu z$, $a_0 := d + \lambda(z \cdot \beta_0)$, and $a_n := -d - \mu(z \cdot \beta_1)$. By construction, $\beta_0 \in \mathcal{H}_{-w_0,-a_0}$ and $\beta_1 \in \mathcal{H}_{-w_n,-a_n}$.

We show that $P := \bigcap_{i=0}^{n} \mathcal{H}^-_{-w_i,-a_i}$ is an n-simplex using Lemma 4.10, and satisfies the hypotheses of the statement. Lemma 4.10(i) holds by construction. To show Lemma 4.10(ii), we consider first $i \in \{0, n\}$. As

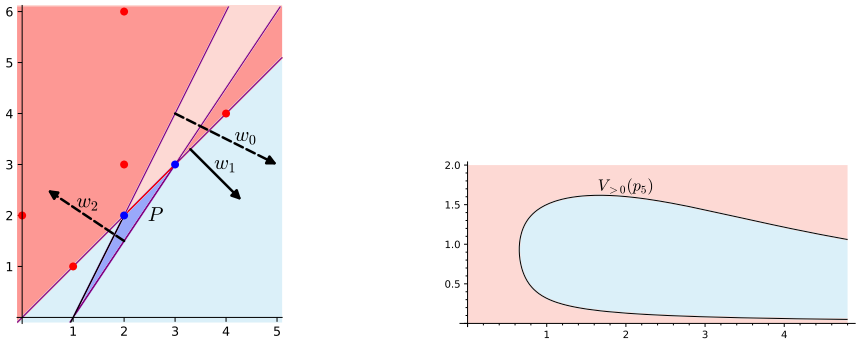
$$\bigcap_{j=0}^{n-1} \mathcal{H}_{-w_j,-a_j} = \{\beta_0\}, \qquad \bigcap_{j=1}^{n} \mathcal{H}_{-w_j,-a_j} = \{\beta_1\}, \tag{26}$$

it suffices to show that $\beta_0 \in \mathcal{H}^{-,\circ}_{-w_n,-a_n}$ and $\beta_1 \in \mathcal{H}^{-,\circ}_{-w_0,-a_0}$. For each $\beta \in \sigma_-(f)$, it holds that

$$w_n \cdot \beta = -v \cdot \beta - \mu(z \cdot \beta) \geq -d - \mu(z \cdot \beta_1) = a_n, \quad \text{and} \tag{27}$$

$$w_0 \cdot \beta = v \cdot \beta + \lambda(z \cdot \beta) \geq d + \lambda(z \cdot \beta_0) = a_0 \tag{28}$$

as z attains its minimum resp. its maximum on $\mathrm{Conv}(\sigma_-(f))$ at β_0 resp. at β_1 and $\lambda, \mu > 0$. From these we get that $\beta_0 \in \mathcal{H}^{-,\circ}_{-w_n,-a_n}$ and $\beta_1 \in \mathcal{H}^{-,\circ}_{-w_0,-a_0}$, since $z \cdot \beta_1 > z \cdot \beta_0$ and hence the inequalities in (27) and (28) are strict.



(a) (b)

FIGURE 6. Illustration of Example 4.12. (a) Shows $\sigma(p_5)$ with blue indicating negative points and red positive points. The vector $w_1 = (1, -1)$ is a non-strict separating vector of the support of p_5. (b) The negative connected component of $V^c_{>0}(p_5)$ is shown in blue.

Consider now $i \in \{1, \dots, n-1\}$ and $x \in \bigcap_{j=0, j\neq i}^{n} \mathcal{H}_{-w_i,-a_i}$. In particular, $x \in \mathcal{H}_{w_0,a_0} \cap \mathcal{H}_{w_n,a_n}$. Solving the linear system $w_0 \cdot x = v \cdot x + \lambda(z \cdot x) = a_0$ and $w_n \cdot x = -v \cdot x - \mu(z \cdot x) = a_n$ for $v \cdot x$ and $z \cdot x$ and using the definition of a_0, a_n, we obtain

$$z \cdot x = \tfrac{a_0+a_n}{\lambda-\mu}, \qquad v \cdot x = a_0 - \lambda \cdot \tfrac{a_0+a_n}{\lambda-\mu} = d + \tfrac{\lambda\mu}{\lambda-\mu}(z \cdot \beta_1 - z \cdot \beta_0) > d,$$

as $\lambda, \mu, \lambda - \mu, z \cdot \beta_1 - z \cdot \beta_0 > 0$. Hence

$$\sum_{j=1}^{n-1} w_j \cdot x = v \cdot x > d = \sum_{j=1}^{n-1} a_j.$$

From this follows that $w_i \cdot x > a_i$, since $w_j \cdot x = a_j$ for $j \neq i$. Therefore $x \in \mathcal{H}^{-,\circ}_{-w_i,-a_i}$ and Lemma 4.10(ii) holds. We conclude that P is an n-simplex.

Finally, we show that $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$. The inclusion $\sigma_-(f) \subseteq P$ follows from (24), (27) and (28).

Let $\alpha \in \sigma_+(f)$ and assume that $v \cdot \alpha < d$. By (25),

$$w_0 \cdot \alpha = v \cdot \alpha + \lambda(z \cdot \alpha) \leq K + \lambda M = d - \epsilon + \lambda M \leq d + \lambda(z \cdot \beta_0) = a_0,$$

which implies $\alpha \in \mathcal{H}^+_{-w_0,-a_0}$. This together with (24) imply that $\alpha \in P^-$ by Proposition 4.3.

Now, consider the case $v \cdot \alpha = d$. In this case, (24) implies that $w_i \cdot \alpha = a_i$ for each $i = 1, \dots, n-1$. Thus, $\alpha \in L$ and recall $\alpha \notin \mathrm{Conv}(\sigma_-(f))$. Hence $\alpha \in L \setminus \mathrm{Conv}(\sigma_-(f)) \subseteq P^-$, where the last inclusion follows from the fact that the supporting hyperplanes of each cone $P^{-,k}$ are supporting hyperplanes of P. □

Example 4.12. Consider the signomial

$$p_5(x_1, x_2) = x_1^4 x_2^4 + x_1^2 x_2^6 + x^2 y^3 - 5x_1^3 x_2^3 - 3x_1^2 x_2^2 + x_1 x_2 + x_2^2,$$

with $\sigma(p_5)$ depicted in Fig. 6(a). The vector $w_1 = (1, -1)$ is a separating vector of $\sigma(p_5)$. The convex hull of $\sigma_-(p_5)$ does not intersect $\sigma_+(p_5)$ as we can see from Fig. 6(a). Hence, we can use Proposition 4.11 to conclude that there exists a simplex P such that $\sigma_-(p_5) \subset P$ and $\sigma_+(p_5) \subset P^-$. In fact, the proof Proposition 4.11 is constructive, the corresponding P is depicted also in Fig. 6(a). Now, we can apply Theorem 4.6 to conclude that $f^{-1}(\mathbb{R}_{<0})$ is contractible.

5. Comparing the different approaches

Theorems 3.4, 3.6, 3.8, 4.6 cover some cases of a generalization of Descartes' rule of signs to hypersurfaces. In particular, we have shown that $f^{-1}(\mathbb{R}_{<0})$ is contractible in the following relevant cases:

- f has at most one negative point in $\sigma(f)$.
- There exists a strict separating vector of $\sigma(f)$.
- There exists a simplex P such that negative points of $\sigma(f)$ belong to P and positive points to P^-; in particular if all positive points are vertices of the Newton polytope and this is a simplex, or if there are $n-1$ linearly independent non-strict separating vectors and the convex hull of the negative points does not contain positive points.

The techniques to study the case where $f^{-1}(\mathbb{R}_{<0})$ is path connected could also be used to derive a condition for $f^{-1}(\mathbb{R}_{<0})$ having at most two connected components:

- There exists a strict enclosing vector of $\sigma(-f)$; in particular if the positive points belong to a hyperplane that does not contain all negative points, or if the number of positive points is smaller than $\dim \mathrm{N}(f)$.

Theorem 4.6 covers all the cases where the classical Descartes' rule guarantees that the number of negative connected components of $V^c_{>0}(f)$ is at most one. These are the cases when the coefficients of the one-variable signomial f has one of the following sign patterns:

$$(-\cdots-+\cdots+) \qquad (+\cdots+-\cdots-) \qquad (+\cdots+-\cdots-+\cdots+)\,.$$

Although Theorem 3.6, and 4.6 build apparently on different techniques, we show in this section that they are equivalent in some situations. Computationally, checking whether Theorem 3.6 applies is less demanding than to verifying that the conditions of Theorem 4.6 hold.

We start by noting that Theorem 4.6 applies for the signomial p_4 in Example 4.7, but $\sigma(p_4)$ does not have any separating vector. However, under some assumptions, the existence of an n-simplex as in Theorem 4.6 implies the existence of a separating vector.

Proposition 5.1. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial and let $P \subseteq \mathbb{R}^n$ be an n-simplex such that $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$. If there exists $k \in \{0,\dots,n\}$ such that $P^{-,k} \cap \sigma_+(f) = \emptyset$, then $\sigma(f)$ has a separating vector. Moreover, there is a strict separating vector if there is a negative point in $P \setminus F_k$, where F_k denotes the facet of P opposite to $P^{-,k}$.*

Proof. Let $\mathcal{H}_{v_k,a_k}$ be a supporting hyperplane for the facet F_k. By hypothesis and from Proposition 4.3 we obtain $\sigma_+(f) \subseteq \mathcal{H}^+_{v_k,a_k}$. By hypothesis we also have that $\sigma_-(f) \subseteq P \subseteq \mathcal{H}^-_{v_k,a_k}$. Therefore, $-v_k$ is a separating vector of $\sigma(f)$. If there is a negative point $\beta \notin F_k$, then $v_k \cdot \beta < a_k$ giving that $-v_k$ is strict. □

We inspect now whether or when Theorem 3.6 follows from Theorem 4.6, in which case we obtain the additional information that f can be transformed into a convex signomial. The existence of a strict separating vector does not imply the existence of an n-simplex satisfying the condition in Theorem 4.6. To see this, we consider the signomial p_2 in Example 3.7. The positive point $(3,4)$ lies in $\mathrm{Conv}(\sigma_-(p_2))$, and is not a vertex. Therefore, there is no n-simplex P such that $\sigma_-(p_2) \subseteq P$ and $(3,4) \in P^-$.

However, if there exists a very strict separating vector, then there is an n-simplex satisfying the conditions in Theorem 4.6 and Theorem 3.6 follows from it. For an example, see Fig. 7.

Proposition 5.2. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial. If there is a very strict separating vector $v \in \mathbb{R}^n$ of $\sigma(f)$, then there exists an n-simplex P such that $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$.*

Proof. By Proposition 3.5 there exist n linearly independent very strict separating vectors $-w_1,\dots,-w_n$, and $c \in \mathbb{R}^n$ such that

$$\sigma_-(f) \subseteq \bigcap_{i=1}^n \mathcal{H}^-_{w_i,c} \quad \text{and} \quad \sigma_+(f) \subseteq \bigcap_{i=1}^n \mathcal{H}^+_{w_i,c}. \tag{29}$$

We consider minus the basis in Proposition 3.5, as separating vectors leave the negative points on the positive side of the hyperplane, while the simplex P leaves them on the negative side of the defining hyperplanes.

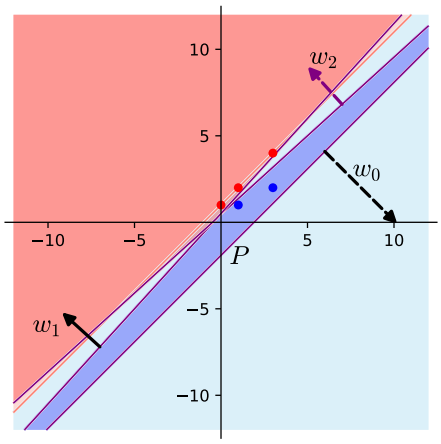


FIGURE 7. The support of the signomial $\tilde{p}_2$ in Example 3.7 has a very strict separating vector as in Proposition 5.2, namely $v = (1, -1)$. The 2-simplex P shown in blue is constructed following the proof of Proposition 5.2 with the choice $v_1, = (1, 0), v_2 = (0, -1)$, $a_0 = 4$.

We define $w_0 := -\sum_{i=1}^n w_i$, choose $a_0 \in \mathbb{R}$ such that $a_0 > \max_{\mu\in\sigma(f)} w_0 \cdot \mu$ and define

$$P := \mathcal{H}^-_{w_0,a_0} \cap \bigcap_{i=1}^n \mathcal{H}^-_{w_i,c}.$$

It then holds that $\sigma_-(f)$ and $\sigma_+(f)$ belong to $\mathcal{H}^-_{w_0,a_0}$. Thus, $\sigma_-(f) \subseteq P$, and $\sigma_+(f) \subseteq P^-$ by Proposition 4.3.

All that is left is to show that P is an n-simplex. To this end, we apply Lemma 4.10. It is clear that every subset of $\{w_0, \dots, w_n\}$ with n elements is linearly independent, so Lemma 4.10(i) holds. From (29) follows that

$$n\,(-c) \le \max_{\beta\in\sigma_-(f)} \sum_{i=1}^n -w_i \cdot \beta = \max_{\beta\in\sigma_-(f)} w_0 \cdot \beta \le \max_{\mu\in\sigma(f)} w_0 \cdot \mu < a_0. \tag{30}$$

For $x \in \bigcap_{j=1}^n \mathcal{H}_{w_j,c}$, we obtain $w_0 \cdot x = -n\,c < a_0$, so $x \in \mathcal{H}^{-,\circ}_{w_0,a_0}$. If $x \in \mathcal{H}_{w_0,a_0} \cap \bigcap_{j=1,j\neq i}^n \mathcal{H}_{w_j,c}$, again by (30) we have that

$$w_i \cdot x = -w_0 \cdot x - \sum_{j=1,j\neq i}^n w_j \cdot x = -a_0 - (n-1)c < n\,c - (n-1)\,c = c.$$

Hence $x \in \mathcal{H}^{-,\circ}_{w_i,c}$ for each $i \in \{1, \dots, n\}$. We conclude that Lemma 4.10(ii) holds, so P is an n-simplex and this completes the proof. □

In the scenario where f has exactly one negative point neither the existence of a separating hyperplane nor the existence of a simplex satisfying the conditions of Theorem 4.6 are guaranteed. In fact, if f has one negative point, then a strict separating hyperplane exists if and only if the negative point is a vertex of the Newton polytope of f. The following example illustrates a scenario where a simplex as in Theorem 4.6 does not exist, and f has only one negative point.

Example 5.3. Let $f\colon \mathbb{R}^2_{>0} \to \mathbb{R}$ be a signomial with only one negative point $\beta_0 \in \sigma(f)$. If $\sigma_+(f)$ is equal to the vertex set of a regular m-gon for some $m \ge 7$ with circumcenter β_0, then there does not exist a simplex P such that $\sigma_-(f) \subseteq P$ and $\sigma_+(f) \subseteq P^-$.

To see this, assume that such a simplex exists and write $P = \mathcal{H}^-_{w_0,b_0} \cap \mathcal{H}^-_{w_1,b_1} \cap \mathcal{H}^-_{w_2,b_2}$, with $w_0, w_1, w_2 \in \mathbb{R}^2$, and $b_0, b_1, b_2 \in \mathbb{R}$. For $a_i := w_i \cdot \beta_0$, $i = 0, 1, 2$, the three lines $\mathcal{H}_{w_0,a_0}$, $\mathcal{H}_{w_1,a_1}$, and $\mathcal{H}_{w_2,a_2}$, intersect each other at β_0 and divide the circumsphere of the m-gon into 6 regions.

Let $\gamma_0, \gamma_1, \gamma_2 \in [0, \pi]$ be the angles of the regions cut out by $\mathcal{H}_{w_0,a_0}$ and $\mathcal{H}_{w_1,a_1}$, by $\mathcal{H}_{w_1,a_1}$ and $\mathcal{H}_{w_2,a_2}$, and by $\mathcal{H}_{w_2,a_2}$ and $\mathcal{H}_{w_0,a_0}$ respectively. Note that $\gamma_0 + \gamma_1 + \gamma_2 = \pi$. Since $\sigma_+(f) \subseteq P^-$, the positive points are in alternating regions. Therefore one of the two regions cut out by $\mathcal{H}_{w_0,a_0}$ and $\mathcal{H}_{w_1,a_1}$ with angle γ_0 cannot contain any positive point. Since $\sigma_+(f)$ is the vertex set of a regular m-gon, for each pair of consecutive positive point α_i, α_{i+1} (counted counterclockwise), the angle $\measuredangle \alpha_i \beta_0 \alpha_{i+1}$ equals $\frac{2\pi}{m}$. From this follows that $\gamma_0 \leq \frac{2\pi}{m}$. A similar argument shows that $\gamma_1 \leq \frac{2\pi}{m}$, $\gamma_2 \leq \frac{2\pi}{m}$. We conclude that $\gamma_0 + \gamma_1 + \gamma_2 \leq \frac{6\pi}{m}$. Since $m \geq 7$, this contradicts $\gamma_0 + \gamma_1 + \gamma_2 = \pi$. Therefore, such a simplex P does not exist.

References

[1] A. A. Albert. An inductive proof of Descartes' rule of signs. *Am. Math. Mon.*, 50(3):178–180, 1943.

[2] D. Alessandrini. Logarithmic limit sets of real semi-algebraic sets. *Adv. Geom.*, 13(1):155–190, 2013.

[3] S. Barone and S. Basu. Refined bounds on the number of connected components of sign conditions on a variety. *Discrete. Comput. Geom.*, 47:577–597, 2012.

[4] S. Basu. On bounding the Betti numbers and computing the Euler characteristic of semi-algebraic sets. *Discrete. Comput. Geom.*, 22:1–18, 1999.

[5] S. Basu. Algorithms in real algebraic geometry: A survey. *Panor. Synthèses*, 51:107–153, 2017.

[6] S. Basu, R. Pollack, and M. Roy. On the number of cells defined by a family of polynomials on a variety. *Mathematika.*, 43(1):120–126, 1996.

[7] S. Basu, R. Pollack, and M. F. Roy. *Algorithms in Real Algebraic Geometry (Algorithms and Computation in Mathematics)*. Springer-Verlag, 2006.

[8] S. Basu and A. Rizzie. Multi-degree bounds on the Betti numbers of real varieties and semi-algebraic sets and applications. *Discrete. Comput. Geom.*, 59:553–620, 2018.

[9] M. Beck, C. Haase, and F. Sottile. Formulas of Brion, Lawrence, and Varchenko on rational generating functions for cones. *Math Intell.*, 31:9–17, 01 2009.

[10] F. Bihan and A. Dickenstein. Descartes' rule of signs for polynomial systems supported on circuits. *Int. Math. Res. Notices.*, 39(22):6867–6893, 2017.

[11] F. Bihan, A. Dickenstein, and J. Forsgård. Optimal Descartes' rule of signs for systems supported on circuits. *Math. Ann.*, 2021.

[12] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optim. Eng.*, 8:67–127, 2007.

[13] M. Brion. Points entiers dans les polyèdres convexes. *Ann. Sci. Ecole. Norm. S.*, 21(4):653–663, 1988.

[14] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *PLoS Comput. Biol.*, 13(10):e1005751, 2017.

[15] G. Craciun, L. Garcia-Puente, and F. Sottile. Some geometrical aspects of control points for toric patches. In M Dæhlen, M S Floater, T Lyche, J-L Merrien, K Morken, and L L Schumaker, editors, *Mathematical Methods for Curves and Surfaces*, volume 5862 of *Lecture Notes in Comput. Sci.*, pages 111–135, Heidelberg, 2010. Springer.

[16] G. Craciun, Y. Tang, and M. Feinberg. Understanding bistability in complex enzyme-driven reaction networks. *Proc. Natl. Acad. Sci. U.S.A.*, 103:8697–8702, 2006.

[17] D. R. Curtiss. Recent extentions of Descartes' rule of signs. *Ann. Math.*, 19(4):251–278, 1918.

[18] A. Dolzmann, T. Sturm, and V. Weispfenning. *Real Quantifier Elimination in Practice*. 01 1999.

[19] R. J. Duffin and E. L. Peterson. Geometric programming with signomials. *J. Optimiz. Theory. App.*, 11(1):3–35, 1973.

[20] M. Forsberg, M. Passare, and A. Tsikh. Laurent determinants and arrangements of hyperplane amoebas. *Adv. Math.*, 151:45–70, 2000.

[21] A. Gabrielov and N. Vorobjov. Approximation of definable sets by compact families, and upper bounds on homotopy and homology. *J. London. Math. Soc.*, 80:35–54, 2009.

[22] C. F. Gauß. Beweis eines algebraischen lehrsatzes. *J. Reine. Angew. Math.*, 3:1–4, 1828.

[23] E. Gawrilow and M. Joswig. `polymake`: a framework for analyzing convex polytopes. In *Polytopes—combinatorics and computation (Oberwolfach, 1997)*, volume 29 of *DMV Sem.*, pages 43–73. Birkhäuser, Basel, 2000.

[24] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants*. Mathematics (Boston, Mass.). Birkhäuser, 1994.

[25] D. J. Grabiner. Descartes' rule of signs: Another construction. *Am. Math. Mon.*, 106(9):854–856, 1999.
[26] B. Grünbaum, V. Kaibel, V. Klee, and G. M. Ziegler. *Convex Polytopes*. Graduate Texts in Mathematics. Springer, 2003.
[27] A. Hatcher. *Algebraic Topology*. Cambridge University Press, 2001.
[28] P. Haukkanen and T. Tossavainen. A generalization of Descartes' rule of signs and fundamental theorem of algebra. *Appl. Math. Comput.*, 218:1203–1207, 2011.
[29] I. Itenberg and M. F. Roy. Multivariate Descartes' rule. *Beitr. Algebra. Geom.*, 37(2):337–346, 1996.
[30] B. Joshi and A. Shiu. A survey of methods for deciding whether a reaction network is multistationary. *Mathematical Modelling of Natural Phenomena*, 10(5):47–67, 2015.
[31] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Universitext. Springer London, 2013.
[32] J. C. Lagarias and T. J. Richardson. Multivariate Descartes rule of signs and Sturmfels's challenge problem. *Math. Intell.*, 19:9–15, 1997.
[33] T. Y. Li and X. Wang. On multivariate Descartes' rule - a counterexample. *Beitr. Algebra. Geom.*, 39(1):1–5, 1998.
[34] C. D. Maranas and C. A. Floudas. All solutions of nonlinear constrained systems of equations. *J. Global. Optim.*, 7:143–182, 1995.
[35] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16:69–97, 2016.
[36] M. Nisse. Maximally sparse polynomials have solid amoebas. *arXiv*, (0704.2216), 2008.
[37] R. T. Rockafellar. *Convex analysis*. Princeton University Press, 1972.
[38] J. M. Rojas and K. Rusek. A-discriminants for complex exponents, and counting real isotopy types. *arXiv*, (1612.03458), 2017.
[39] I. Sahidul and A. M. Wasim. *Fuzzy Geometric Programming Techniques and Applications*. Springer, 2019.
[40] T. Sturm. Redlog online resources for applied quantifier elimination. *Act. Acad. Ab.*, 67:177–191, 02 2007.
[41] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.2)*, 2021. `https://www.sagemath.org`.
[42] D. V. Tokarev. A generalisation of Descartes' rule of signs. *J. Aust. Math. Soc.*, 91(3):415–420, 2011.
[43] X. Wang. A simple proof of Descartes's rule of signs. *Am. Math. Mon.*, 111:525–526, 2004.
[44] G. M. Ziegler. *Lectures on Polytopes*. Springer, 2007.

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: `efeliu@math.ku.dk`

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: `mlt@math.ku.dk`

IV

Geometry of the signed support of a multivariate polynomial and Descartes' rule of signs

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Publication details

Submitted (2023)
Available on arXiv: https://doi.org/10.48550/arXiv.2310.05466

GEOMETRY OF THE SIGNED SUPPORT OF A MULTIVARIATE POLYNOMIAL AND DESCARTES' RULE OF SIGNS

MÁTÉ L. TELEK

Abstract. We describe conditions on the signed support, that is, on the set of the exponent vectors and on the signs of the coefficients, of a multivariate polynomial f ensuring that the semi-algebraic set $\{f < 0\}$ defined in the positive orthant has at most one connected component. These results generalize Descartes' rule of signs in the sense that they provide a bound which is independent of the values of the coefficients and the degree of the polynomial. Based on how the exponent vectors lie on the faces of the Newton polytope, we give a recursive algorithm that verifies a sufficient condition for the set $\{f < 0\}$ to have one connected component. We apply the algorithm to reaction networks in order to prove that the parameter region of multistationarity of a ubiquitous network comprising phosphorylation cycles is connected.

Keywords: semi-algebraic set, connected component, Newton polytope, reaction network

1. Introduction

Descartes' rule of signs is a classical theorem in real algebraic geometry that provides an upper bound on the number of positive real roots of a univariate real polynomial. The bound is given by the number of sign changes in the coefficient sequence of the polynomial, therefore it is easy to compute. Since Descartes' bound is independent from the degree of the polynomial, it shows a crucial difference between real and complex roots.

Since Descartes published his result in 1637, a lot of effort has been made to improve and generalize his statement. It is known that the result is valid for polynomials with real exponents [11], and the number of positive roots has the same parity as the number of sign changes in the coefficient sequence [22]. Moreover, Descartes's bound is sharp, that is, for every given sign sequence there exists a polynomial matching the sign sequence that has as many positive roots as provided by Descartes's bound [25].

The research question of multivariate generalizations is still rather open. In his seminal book *Fewnomials* [30], Khovanskii gave an upper bound on the number of positive solutions of a polynomial system given by n real polynomials in n variables that depends only on n and the number of monomials appearing in the polynomials. Khovanskii's bound has been improved in [7, 9]. For some specific systems there are also better bounds available [1, 2, 5, 31, 32]. These works generalize Descartes' rule of signs in the sense that they provide upper bounds which are independent of the degree, however the signs of the coefficients are not taken into account. Recently in [3, 4], for systems whose polynomials are supported on a circuit, a sharp upper bound was given that depends on the number of sign changes of a given sequence associated both with the exponents and the coefficients of the polynomials.

Descartes' rule of signs allows also different types of generalizations. In a notorious one, instead of focusing on a system of polynomials, one considers a single polynomial in n variables and bounds topological invariants of the hypersurface given by the positive real zero set of the polynomial. In [8], the authors provided upper bounds on the sum of the Betti numbers of the hypersurface. Bounds on the connected component of the hypersurface were given in [6, 21]. These bounds depend on the number of variables n and the number of monomials of the polynomial.

In this work, we aim to bound connected components but in a slightly different setting. As in [19], we consider connected components of the complement of the hypersurface. To make it precise, let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial (a generalized polynomial whose exponent vectors are real) and let $f^{-1}(\mathbb{R}_{<0})$ be the set of points in $\mathbb{R}^n_{>0}$ where f takes negative values. In

[19], the authors phrased the following problem as *generalization of Descartes' rule of signs to hypersurfaces*.

Problem 1.1. *Consider a signomial* $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ *with* $f(x) = \sum_{\mu\in\sigma(f)} c_\mu x^\mu$, *and* $\sigma(f) \subseteq \mathbb{R}^n$ *a finite set. Find a (sharp) upper bound on the number of connected components of* $f^{-1}(\mathbb{R}_{<0})$ *based on the sign of the coefficients and the geometry of* $\sigma(f)$.

To avoid wordy sentences, we might call connected components of $f^{-1}(\mathbb{R}_{<0})$ *negative connected components of* f. In Section 2, we give new conditions on the set of exponent vectors $\sigma(f)$ that ensure that f has at most one negative connected component. For instance, the existence of two parallel hyperplanes, which enclose in a certain way the exponent vectors of f with positive coefficients, implies that f has at most one negative connected component (Theorem 2.11). In case f is multivariate ($n \geq 2$), we show that the number of negative connected components is one if f has only one positive coefficient (Corollary 2.13), or the exponent vectors are separated by a simplex in a specific manner (Corollary 2.15).

In Section 3, we show that the problem of finding the number negative connected components can be reduced to the same problem for a signomial in fewer monomials, if all the exponent vectors of f with negative coefficients are contained in a face of the Newton polytope (Theorem 3.1). A similar reduction is possible if the Newton polytope of f has two parallel faces containing all the exponent vectors of f (Theorem 3.6). These statements lead to a recursive algorithm that can verify connectivity of $f^{-1}(\mathbb{R}_{<0})$ (Algorithm 1). Since the algorithm is based on polyhedral geometry computations, its running time remains reasonable even for polynomials with many variables and many monomials.

The motivation to consider Problem 1.1 came from chemical reaction network theory. In [39], the authors associated with a reaction network (satisfying some technical conditions) a polynomial function $q\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ such that connectivity of $q^{-1}(\mathbb{R}_{<0})$ implies that the parameter region of multistationarity of the reaction network is connected. Using the results from [19], for several reaction networks it was verified that the associated polynomial q has one negative connected component, so the parameter region of multistationarity is connected.

However, there were some biologically relevant reaction networks, where the results from [19] did not suffice. In particular, one of these networks was the weakly irreversible phosphorylation system with two binding sites [34]. Using numerical methods, the authors in [34] showed that its parameter region of multistationarity is connected, however they did not give a rigorous proof.

Another, important family of reaction networks, where the connectivity of the parameter region of multistationarity has been investigated, is the sequential and distributive phosphorylation cycles with m-binding sites, $m \in \mathbb{N}$. In [39], it has been showed that the parameter region of multistationarity is connected for $m = 2, 3$. Furthermore, it is known that the projection of the parameter region of multistationarity to a subset of the parameters (reaction rate constants) is connected for all m [16, 17]. In Section 4, we revisit these networks. We use Algorithm 1 to show connectivity of the parameter region of multistationarity for the weakly irreversible phosphorylation system and for phosphorylation cycles with $m = 4, 5, 6, 7$ binding sites.

Notation. $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{<0}$ refer to the sets of non-negative, positive and negative real numbers respectively. For monomials, we use the notation $x^\mu = x_1^{\mu_1} \cdots x_n^{\mu_n}$, where $x \in \mathbb{R}^n_{>0}, \mu \in \mathbb{R}^n$. For two vectors $v, w \in \mathbb{R}^n$, $v \cdot w$ denotes the Euclidean scalar product, and $v * w$ denotes the coordinate-wise product of v and w. We denote the Euclidean interior of a set $X \subseteq \mathbb{R}^n$ by $\operatorname{int}(X)$. If $X \subseteq \mathbb{R}^n$ is a polyhedron, $\operatorname{relint}(X)$ denotes the relative interior of X, that is the Euclidean interior of X in its affine hull. The symbol $\#S$ denotes the cardinality of the finite set S. We write $S \sqcup T$ for the disjoint union of two sets S, T.

2. Separating and enclosing hyperplanes

2.1. **Background and definitions.** In this section, we investigate *signomials* and their *negative connected components* using certain affine hyperplanes that partition the exponent vectors of the signomial. Recall that a signomial is a multivariate generalized polynomial with real

exponents whose domain is restricted to the positive orthant $\mathbb{R}^n_{>0}$ [14, 36]. In other words, a signomial is a function of the form:

$$f\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu,$$

where $\sigma(f) \subseteq \mathbb{R}^n$ is a finite set, called the *support* of f, and the *coefficients* c_μ are non-zero real numbers. The negative connected components of f are connected components of the set:

$$f^{-1}(\mathbb{R}_{<0}) = \{x \in \mathbb{R}^n_{>0} \mid f(x) < 0\}.$$

We write $\mathcal{B}_0^-(f)$ for the set of negative connected components, and $b_0(f^{-1}(\mathbb{R}_{<0}))$ for the cardinality of $\mathcal{B}_0^-(f)$.

If $c_\mu > 0$ (resp. $c_\mu < 0$), we call μ a *positive* (resp. *negative*) *exponent vector* of f. We write

$$\sigma_+(f) := \{\mu \in \sigma(f) \mid c_\mu > 0\} \quad \text{and} \quad \sigma_-(f) := \{\mu \in \sigma(f) \mid c_\mu < 0\}$$

for the set of positive and negative exponent vectors respectively. The convex hull of the support

$$\mathrm{N}(f) := \mathrm{Conv}(\sigma(f))$$

is called the *Newton polytope* of f. For a set $S \subseteq \mathbb{R}^n$, we define the *restriction* of f to S as

$$f_{|S}\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad x \mapsto f_{|S}(x) := \sum_{\mu\in\sigma(f)\cap S} c_\mu x^\mu.$$

Similarly to [19], our arguments will benefit from the existence of *separating* and *enclosing hyperplanes* of the support of f. We briefly recall these objects. Each $v \in \mathbb{R}^n \setminus \{0\}$ and $a \in \mathbb{R}$ define a *hyperplane*

$$\mathcal{H}_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu = a\},$$

and two *half-spaces*

$$\mathcal{H}^+_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \geq a\}, \qquad \mathcal{H}^-_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \leq a\}.$$

The interiors of these half-spaces are denoted by $\mathcal{H}^{+,\circ}_{v,a}$ and $\mathcal{H}^{-,\circ}_{v,a}$. If the support of a signomial f satisfies

$$\sigma_-(f) \subseteq \mathcal{H}^+_{v,a} \qquad \text{and} \qquad \sigma_+(f) \subseteq \mathcal{H}^-_{v,a}$$

then we call $\mathcal{H}_{v,a}$ a *separating hyperplane* of $\sigma(f)$ and v is called a *separating vector*. A separating hyperplane is *strict*, if

$$\sigma_-(f) \cap \mathcal{H}^{+,\circ}_{v,a} \neq \emptyset,$$

meaning that the negative exponent vectors are not all contained in the hyperplane. We call a vector $v \in \mathbb{R}^n$ an *enclosing vector* of $\sigma_+(f)$ if there exist parallel hyperplanes $\mathcal{H}_{v,a}$, $\mathcal{H}_{v,b}$, $a \geq b$ such that

$$\sigma_+(f) \subseteq \mathcal{H}^-_{v,a} \cap \mathcal{H}^+_{v,b}, \qquad \text{and} \qquad \sigma_-(f) \subseteq \mathbb{R}^n \setminus \left(\mathcal{H}^{-,\circ}_{v,a} \cap \mathcal{H}^{+,\circ}_{v,b}\right).$$

In that case, we call the pair $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ a *pair of enclosing hyperplanes* of $\sigma_+(f)$. A pair of enclosing hyperplanes is strict, if

$$\sigma_-(f) \cap \mathcal{H}^{+,\circ}_{v,a} \neq \emptyset, \quad \text{and} \quad \sigma_-(f) \cap \mathcal{H}^{-,\circ}_{v,b} \neq \emptyset.$$

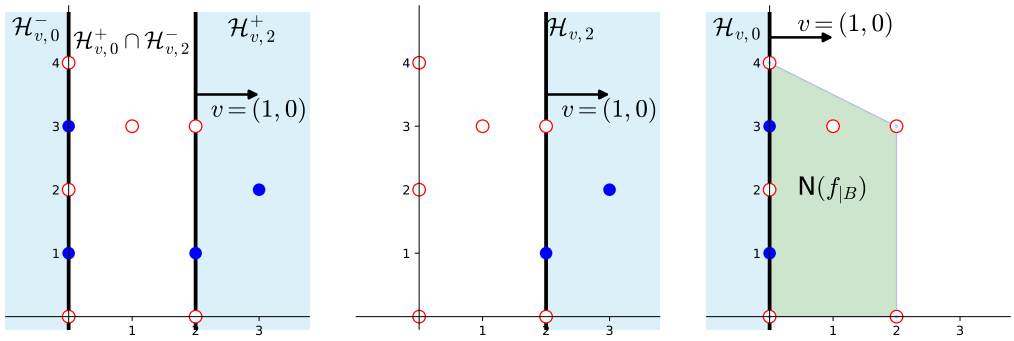
Example 2.1. To illustrate the above definitions, consider the polynomial

$$f(x,y) = -101x^3y^2 + 50x^2y^3 + xy^3 + y^4 - x^2y - 9.5y^3 + 51x^2 + 30.5y^2 - 37y + 12. \tag{1}$$

The set of positive and negative exponent vectors are given by:

$$\sigma_+(f) = \{(2,3),(1,3),(0,4),(2,0),(0,2),(0,0)\}, \quad \sigma_-(f) = \{(3,2),(2,1),(0,3),(0,1)\}.$$

The pair $(\mathcal{H}_{v,2}, \mathcal{H}_{v,0})$, $v = (1,0)$ is a pair of enclosing hyperplanes of $\sigma_+(f)$, see Figure 1(a).



(a) (b) (c)

FIGURE 1. Exponent vectors of $f, f_{|A}, f_{|B}$ from Example 2.1. Positive exponent vectors are marked by red circles and negative exponent vectors by blue dots. (a) A pair of enclosing hyperplanes of $\sigma_+(f)$. (b) A strict separating hyperplane of $\sigma(f_{|A})$. (c) The Newton polytope of $f_{|B}$ and a non-strict separating hyperplane of $\sigma(f_{|B})$.

If we remove the two negative exponent vectors $(0,3),(0,1)$ which are contained in $\mathcal{H}^-_{v,0}$, or with other words we restrict f to $A = (\mathcal{H}^+_{v,2} \cap \sigma_-(f)) \cup \sigma_+(f)$, then the support of

$$f_{|A}(x,y) = -101x^3y^2 + 50x^2y^3 + xy^3 + y^4 - x^2y + 51x^2 + 30.5y^2 + 12 \tag{2}$$

has a strict separating hyperplane given by $\mathcal{H}_{v,2}$, $v = (1,0)$, see Figure 1(b).

By restricting f to $B = (\mathcal{H}^-_{v,0} \cap \sigma_-(f)) \cup \sigma_+(f)$, we have

$$f_{|B}(x,y) = 50x^2y^3 + xy^3 + y^4 - 9.5y^3 + 51x^2 + 30.5y^2 - 37y + 12. \tag{3}$$

The hyperplane $\mathcal{H}_{v,0}$ is a non-strict separating hyperplane of $\sigma(f)$, see Figure 1(c). The face of $\mathrm{N}(f_{|B})$ with inner normal vector $v = (1,0)$ equals $\mathrm{N}(f_{|B}) \cap \mathcal{H}_{v,0}$ and contains all the negative exponent vectors of $f_{|B}$.

The relevance of separating and enclosing vectors arises from the following observation. Each $v \in \mathbb{R}^n$ and $x \in \mathbb{R}^n_{>0}$ induce a univariate signomial

$$\mathbb{R}_{>0} \to \mathbb{R}, \quad t \mapsto f(t^v * x) = \sum_{\mu \in \sigma(f)} c_\mu x^\mu t^{v\cdot\mu}. \tag{4}$$

If v is a separating (resp. enclosing) vector, then $f(t^v * x)$, viewed as a signomial in t, has at most one (resp. two) sign changes in its coefficient sequence.

Recall that a coefficient of a univariate signomial g is called the *leading coefficient* $\mathrm{LC}(g)$ (resp. *trailing coefficient* $\mathrm{TC}(g)$) if the corresponding exponent is the largest (resp. smallest). If $\mathrm{LC}(g) < 0$ (resp. $\mathrm{TC}(g) < 0$), then g attains negative values for large (resp. small) enough $t \in \mathbb{R}_{>0}$. This simple observation and the univariate Descartes' rule of signs give the following statement.

Lemma 2.2. *Let $g\colon \mathbb{R}_{>0} \to \mathbb{R}$, $g(t) = \sum_{i=1}^d a_i t^{\nu_i}$ be a univariate signomial such that $g(1) < 0$.*

(i) *If $\mathrm{LC}(g) < 0$ and g has at most one sign change in its coefficient sign sequence, then $g(t) < 0$ for all $t \geq 1$.*
(ii) *If $\mathrm{TC}(g) < 0$ and g has at most one sign change in its coefficient sign sequence, then $g(t) < 0$ for all $t \leq 1$.*
(iii) *If g has at most two sign changes in its coefficient sign sequence and $\mathrm{LC}(g) < 0$ or $\mathrm{TC}(g) < 0$, then $g(t) < 0$ for all $t \leq 1$ or $g(t) < 0$ for all $t \geq 1$.*

Lemma 2.2 played a crucial role in the proof of the following theorems in [19].

Theorem 2.3. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial.*

(i) *If there exists a strict separating hyperplane of $\sigma(f)$, then $f^{-1}(\mathbb{R}_{<0})$ is non-empty and contractible [19, Theorem 3.6].*
(ii) *If there exists a pair of strict enclosing hyperplanes of $\sigma_+(f)$, then $b_0(f^{-1}(\mathbb{R}_{<0})) \leq 2$ [19, Theorem 3.8].*
(iii) *If f has at most one negative coefficient, then $f^{-1}(\mathbb{R}_{<0})$ is either empty or logarithmically convex [19, Theorem 3.4].*

In addition to Theorem 2.3, another condition on the signed support of the signomial f implying that f has at most one negative connected component, is that the negative and positive exponent vectors of f are separated by a simplex and its negative vertex cones [19, Theorem 4.6]. We postpone recalling this result to Section 2.3, and continue with investigating separating and enclosing hyperplanes.

2.2. **Non-strict separating and non-strict enclosing hyperplanes.** In the following propositions, we investigate what happens when $\sigma(f)$ (resp. $\sigma_+(f)$) has a separating (resp. enclosing) hyperplane that is not necessarily strict. In these cases, a bound on $b_0(f^{-1}(\mathbb{R}_{<0}))$ is given by the number of negative connected components of restrictions of f to certain subsets of $\sigma(f)$. These technical statements serve as the core part of the proofs in Section 2.3 and Section 3.1.

The first such statement considers subsets of the support containing all negative exponent vectors and those positive exponent vectors which lie on the separating hyperplane.

Proposition 2.4. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial whose support has a separating hyperplane $\mathcal{H}_{v,a}$. For any subset $R \subseteq \sigma(f)$ such that $\sigma_-(f) \subseteq R$ and $\mathcal{H}_{v,a}\cap\sigma_+(f) \subseteq R$, we have:*

(i) *For all $U \in \mathcal{B}_0^-(f_{|R})$, there exists a unique $V \in \mathcal{B}_0^-(f)$ such that $U \cap f^{-1}(\mathbb{R}_{<0}) \subseteq V$.*
(ii) *The map*

$$\phi\colon \mathcal{B}_0^-(f_{|R}) \to \mathcal{B}_0^-(f), \quad U \mapsto \begin{array}{l}\text{connected component of } f^{-1}(\mathbb{R}_{<0})\\ \text{that contains } U \cap f^{-1}(\mathbb{R}_{<0})\end{array}$$

is well defined and bijective. In particular, it holds

$$b_0(f_{|R}^{-1}(\mathbb{R}_{<0})) = b_0(f^{-1}(\mathbb{R}_{<0})).$$

Proof. First observe that

$$f^{-1}(\mathbb{R}_{<0}) \subseteq f_{|R}^{-1}(\mathbb{R}_{<0}), \tag{5}$$

because $f_{|R}(x) \leq f(x)$ for all $x \in \mathbb{R}^n_{>0}$.

If $\sigma_+(f) \subseteq R$, then $R = \sigma(f)$ and we are done. So we assume that there exists $\alpha_0 \in \sigma_+(f)$ such that $\alpha_0 \notin \mathcal{H}_{v,a}$. Since $\mathcal{H}_{v,a}$ is a separating hyperplane of $\sigma(f)$, we have

$$v \cdot \beta \geq a \geq v \cdot \alpha \qquad \text{for all } \beta \in \sigma_-(f),\ \alpha \in \sigma_+(f), \qquad \text{and} \quad a > v \cdot \alpha_0. \tag{6}$$

For $x \in \mathbb{R}^n_{>0}$, consider the path

$$\gamma_{v,x}\colon [1,\infty) \to \mathbb{R}^n_{>0}, \quad t \mapsto t^v * x$$

and the univariate signomial $f(t^v * x)$ as in (4). Note that $f(\gamma_{v,x}(t)) = f(t^v * x)$ for all $t \geq 1$.

If v is a strict separating vector of $\sigma(f)$, then $f(t^v * x)$ has a negative leading coefficient. If v is not strict, then

$$\mathrm{LC}(f(t^v * x)) = f_{|\mathcal{H}_{v,a}}(x), \quad \text{and} \quad f_{|R}(x) = f_{|\mathcal{H}_{v,a}}(x) + \sum_{\alpha\in\sigma_+(f)\cap R\cap\mathcal{H}_{v,a}^{-;\circ}} c_\alpha x^\alpha.$$

Thus, if $f_{|R}(x) < 0$, then $\mathrm{LC}(f(t^v * x)) < 0$.

By (6), $f(t^v * x)$ and $f_{|R}(t^v * x)$ have at most one sign change in their coefficient sign sequence. By Lemma 2.2(i), we have:

$$\operatorname{im}\gamma_{v,x} \subseteq f^{-1}(\mathbb{R}_{<0}), \qquad \text{for all } x \in f^{-1}(\mathbb{R}_{<0}), \tag{7}$$

$$\operatorname{im}\gamma_{v,x} \subseteq f_{|R}^{-1}(\mathbb{R}_{<0}) \quad \text{and} \quad \gamma_{v,x}(t) \in f^{-1}(\mathbb{R}_{<0}), \qquad \text{for all } x \in f_{|R}^{-1}(\mathbb{R}_{<0}) < 0,\ t \gg 1. \tag{8}$$

This gives that each connected component of $f_{|R}^{-1}(\mathbb{R}_{<0})$ has a non-empty intersection with $f^{-1}(\mathbb{R}_{<0})$.

To prove (i), let $U \in \mathcal{B}_0^-(f_{|R})$, $x, y \in U \cap f^{-1}(\mathbb{R}_{<0})$ and consider a continuous path

$$\gamma\colon [0,1] \to \mathbb{R}^n_{>0}$$

such that $\gamma(0) = x$, $\gamma(1) = y$ and $\gamma(s) \subseteq f_{|R}^{-1}(\mathbb{R}_{<0})$ for all $s \in [0,1]$.

We construct now a path between x and y that is contained in $f^{-1}(\mathbb{R}_{<0})$. For a fixed $s \in [0,1]$, since $f_{|R}(\gamma(s)) < 0$, and the signomial $f(t^v * \gamma(s))$ has exactly one sign change in its coefficient sign sequence, Descartes' rule of signs implies that $f(t^v * \gamma(s))$ has exactly one positive real root $\tau(s)$, which is simple. Now, the Implicit Function Theorem [15] implies that the function

$$\tau\colon [0,1] \to \mathbb{R}_{>0}, \quad s \mapsto \tau(s)$$

is continuous. So $T := \max_{s\in[0,1]} \tau(s)$ exists, and we have:

$$f(t_0^v * \gamma(s)) < 0 \qquad \text{for some } t_0 > \max\{1, T\} \text{ and all } s \in [0,1].$$

Thus, the path

$$[0,1] \to \mathbb{R}^n_{>0}, \quad s \to t_0^v * \gamma(s)$$

connects $t_0^v * x$ and $t_0^v * y$, and is contained in $f^{-1}(\mathbb{R}_{<0})$. By (7), the paths $\gamma_{v,x}, \gamma_{v,y}$ are contained in $f^{-1}(\mathbb{R}_{<0})$ and join x and $t_0 * x$, resp. y and $t_0 * y$. Thus, if x, y are in the same connected component of $f_{|R}^{-1}(\mathbb{R}_{<0})$, then they are also in the same connected component of $f^{-1}(\mathbb{R}_{<0})$, which gives (i).

By (i), the map ϕ is well defined. Now we show that ϕ is bijective. From (5), it follows that every $W \in \mathcal{B}_0^-(f)$ lies in a connected component of $f_{|R}^{-1}(\mathbb{R}_{<0})$, which is a preimage of W under ϕ. Thus, ϕ is surjective.

To show injectivity of ϕ, consider $U_1, U_2 \in \mathcal{B}_0^-(f_{|R})$ such that $\phi(U_1) = \phi(U_2)$ and let $x_1 \in U_1, x_2 \in U_2$. By (8)

$$\operatorname{im}\gamma_{v,x_1} \subseteq f_{|R}^{-1}(\mathbb{R}_{<0}), \quad \operatorname{im}\gamma_{v,x_2} \subseteq f_{|R}^{-1}(\mathbb{R}_{<0}), \quad \text{and}$$
$$\gamma_{v,x_1}(t_0) \in f^{-1}(\mathbb{R}_{<0}), \quad \gamma_{v,x_2}(t_0) \in f^{-1}(\mathbb{R}_{<0}) \text{ for some } t_0 \gg 1.$$

Since $U_1 \cap f^{-1}(\mathbb{R}_{<0})$ and $U_2 \cap f^{-1}(\mathbb{R}_{<0})$ are in the same connected component of $f^{-1}(\mathbb{R}_{<0})$, there exists a continuous path γ_3 between $\gamma_{v,x_1}(t_0)$ and $\gamma_{v,x_2}(t_0)$ such that $\operatorname{im}\gamma_3 \subseteq f^{-1}(\mathbb{R}_{<0})$. By (5), we have $\operatorname{im}\gamma_3 \subseteq f_{|R}^{-1}(\mathbb{R}_{<0})$. Thus, γ_{v,x_1}, γ_3 and γ_{v,x_2} give a continuous path between x_1 and x_2 contained in $f_{|R}^{-1}(\mathbb{R}_{<0})$. Thus, x, y lie in the same connected component of $f_{|R}^{-1}(\mathbb{R}_{<0})$, which implies that $U_1 = U_2$. □

In the following, we generalize the proof of Theorem 2.3(ii) to the case, where the enclosing hyperplanes are not necessarily strict. To ease the reading of the proofs, we discuss first the notation we are going to use. Given a pair of enclosing hyperplanes $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ of $\sigma_+(f)$, we define

$$A = (\mathcal{H}_{v,a}^+ \cap \sigma_-(f)) \cup \sigma_+(f), \qquad B = (\mathcal{H}_{v,b}^- \cap \sigma_-(f)) \cup \sigma_+(f). \tag{9}$$

That is, both A and B contain all the positive exponent vectors of f. The set A (resp. B) contains the negative exponent vectors from one (resp. the other) side of the area enclosed by the hyperplanes $\mathcal{H}_{v,a}, \mathcal{H}_{v,b}$. Note that by definition, $\mathcal{H}_{v,a}$ (resp $\mathcal{H}_{-v,b}$) is a separating hyperplane of the support of the restricted signomials $f_{|A}$ (resp $f_{|B}$). If $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of strict enclosing hyperplanes, then $\mathcal{H}_{v,a}$ and $\mathcal{H}_{-v,b}$ are strict separating hyperplanes of $\sigma(f_{|A})$ and $\sigma(f_{|B})$ respectively. In this case, by Theorem 2.3(i) we have:

$$b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) = 1, \qquad b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big) = 1. \tag{10}$$

For example, for the signomial (1) from Example 2.1, the sets A, B are given by

$$A = (\mathcal{H}^+_{v,2} \cap \sigma_-(f)) \cup \sigma_+(f) = \{(3,2),(2,1),(2,3),(1,3),(0,4),(2,0),(0,2),(0,0)\} \tag{11}$$

$$B = (\mathcal{H}^-_{v,0} \cap \sigma_-(f)) \cup \sigma_+(f) = \{(0,3),(0,1),(2,3),(1,3),(0,4),(2,0),(0,2),(0,0)\}. \tag{12}$$

The supports of the restricted signomials $f_{|A}$ and $f_{|B}$ are depicted in Figure 1(b),(c). For fixed $x \in \mathbb{R}^n_{>0}$, the induced univariate signomials as in (4) equal

$$\begin{aligned} f(t^1x, t^0y) &= (-101x^3y^2)t^3 + (51x^2 - x^2y + 50x^2y^3)t^2 + xy^3t^1 + y^4 - 9.5y^3 + 30.5y^2 - 37y + 12,\\ f_{|A}(t^1x, t^0y) &= (-101x^3y^2)t^3 + (51x^2 - x^2y + 50x^2y^3)t^2 + xy^3t^1 + y^4 + 30.5y^2 + 12,\\ f_{|B}(t^{-1}x, t^0y) &= y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 + xy^3t^{-1} + (51x^2 + 50x^2y^3)t^{-2}. \end{aligned}$$

The leading coefficient of $f_{|A}(t^1x, t^0y)$ is $-101x^3y^2$, which is negative for all $(x,y) \in \mathbb{R}^2_{>0}$. This phenomenon always happens. If v is a strict separating vector of $\sigma(f)$, then the leading coefficient of the induced signomial in (4) is negative. On the contrary, this might not be true for non-strict separating vectors. For the above example, we have

$$\begin{aligned} &\mathrm{LC}(f_{|B}(t^{-1}x, t^0y)) = y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 < 0, && \text{if } y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 < 0,\\ &\mathrm{LC}(f_{|B}(t^{-1}x, t^0y)) = xy^3 > 0, && \text{if } y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 = 0,\\ &\mathrm{LC}(f_{|B}(t^{-1}x, t^0y)) = y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 > 0, && \text{if } y^4 - 9.5y^3 + 30.5y^2 - 37y + 12 > 0. \end{aligned}$$

In the next proposition, we consider subsets $R \subseteq A$, $S \subseteq B$ such that $f_{|A}$ and R (resp. $f_{|B}$ and S) satisfy the hypothesis of Proposition 2.4.

Proposition 2.5. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial and $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ be a pair of enclosing hyperplanes of $\sigma_+(f)$. Let $A, B \subseteq \sigma(f)$ be as in (9) and let $R \subseteq A$, $S \subseteq B$ such that*

- $R \cap \sigma_-(f) = A \cap \sigma_-(f)$, $\quad \mathcal{H}_{v,a} \cap \sigma_+(f) \subseteq R$,
- $S \cap \sigma_-(f) = B \cap \sigma_-(f)$, $\quad \mathcal{H}_{v,b} \cap \sigma_+(f) \subseteq S$.

Then the map

$$\varphi\colon \mathcal{B}^-_0(f_{|R}) \sqcup \mathcal{B}^-_0(f_{|S}) \to \mathcal{B}^-_0(f),$$
$$U \mapsto \begin{cases} \text{connected component of } f^{-1}(\mathbb{R}_{<0}) \text{ that contains } U \cap f^{-1}_{|A}(\mathbb{R}_{<0}), & \text{if } U \in \mathcal{B}^-_0(f_{|R}),\\ \text{connected component of } f^{-1}(\mathbb{R}_{<0}) \text{ that contains } U \cap f^{-1}_{|B}(\mathbb{R}_{<0}), & \text{if } U \in \mathcal{B}^-_0(f_{|S}). \end{cases}$$

is well defined and surjective. In particular, we have

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \le b_0\big(f^{-1}_{|R}(\mathbb{R}_{<0})\big) + b_0\big(f^{-1}_{|S}(\mathbb{R}_{<0})\big),$$

Proof. The idea of the proof follows closely the arguments in [19, Theorem 3.8]. Since $f_{|A}$ and $f_{|B}$ are obtained by removing some of the monomials with negative coefficient from f, $f(x) \le f_{|A}(x)$ and $f(x) \le f_{|B}(x)$ for all $x \in \mathbb{R}^n_{>0}$. This implies:

$$f^{-1}_{|A}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0}), \qquad \text{and} \qquad f^{-1}_{|B}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0}).$$

Thus, every connected component of $f^{-1}_{|A}(\mathbb{R}_{<0})$ and $f^{-1}_{|B}(\mathbb{R}_{<0})$ is contained in a unique connected component of $f^{-1}(\mathbb{R}_{<0})$. By this observation, the map

$$\psi\colon \mathcal{B}^-_0(f_{|A}) \sqcup \mathcal{B}^-_0(f_{|B}) \to \mathcal{B}^-_0(f), \quad V \mapsto \text{connected component of } f^{-1}(\mathbb{R}_{<0}) \text{ that contains } V$$

is well defined.

From Proposition 2.4 applied to $f_{|A}$ and R or $f_{|B}$ and S, it follows that the map

$$\phi\colon \mathcal{B}^-_0(f_{|R}) \sqcup \mathcal{B}^-_0(f_{|S}) \to \mathcal{B}^-_0(f_{|A}) \sqcup \mathcal{B}^-_0(f_{|B}),$$
$$U \mapsto \begin{cases} \text{connected component of } f^{-1}_{|A}(\mathbb{R}_{<0}) \text{ that contains } U \cap f^{-1}_{|A}(\mathbb{R}_{<0}), & \text{if } U \in \mathcal{B}^-_0(f_{|R}),\\ \text{connected component of } f^{-1}_{|B}(\mathbb{R}_{<0}) \text{ that contains } U \cap f^{-1}_{|B}(\mathbb{R}_{<0}), & \text{if } U \in \mathcal{B}^-_0(f_{|S}). \end{cases}$$

is well defined and bijective. Since φ is the composition of ϕ and ψ, it is enough to show that ψ is surjective.

To prove that ψ is surjective, for $W \in \mathcal{B}_0^-(f)$ and any $x \in W$, we show that one of the two paths

$$\gamma_v \colon [1,\infty) \to \mathbb{R}^n_{>0}, \quad t \to t^v * x, \qquad \gamma_{-v} \colon (0,1] \to \mathbb{R}^n_{>0}, \quad t \to t^{-v} * x$$

(a) connects x to a point $y \in f_{|A}^{-1}(\mathbb{R}_{<0}) \cup f_{|B}^{-1}(\mathbb{R}_{<0})$, and
(b) the image of the path is contained in $f^{-1}(\mathbb{R}_{<0})$.

Then, any connected component $V \in \mathcal{B}_0^-(f_{|A}) \sqcup \mathcal{B}_0^-(f_{|B})$ that contains y will be a preimage of W under ψ. To see this, we define

$$S_{x,a}(t) := \sum_{\mu \in \sigma(f),\, a \le v\cdot\mu} c_\mu x^\mu t^{v\cdot\mu}, \qquad S_{x,b}(t) := \begin{cases} \sum\limits_{\mu \in \sigma(f),\, v\cdot\mu \le b} c_\mu x^\mu t^{v\cdot\mu}, \text{ if } a \neq b \\ \sum\limits_{\mu \in \sigma(f),\, v\cdot\mu < b} c_\mu x^\mu t^{v\cdot\mu}, \text{ if } a = b \end{cases}.$$

and consider the signomials

$$\tilde{f}_A(t) := f_{|A}(t^v * x) = S_{x,a}(t) + \sum_{\mu \in \sigma_+(f), v\cdot\mu < a} c_\mu x^\mu t^{v\cdot\mu},$$
$$\tilde{f}_B(t) := f_{|B}(t^v * x) = \sum_{\mu \in \sigma_+(f), b < v\cdot\mu} c_\mu x^\mu t^{v\cdot\mu} + S_{x,b}(t).$$

A simple argument shows that

$$\begin{aligned} &\text{If } \mathrm{LC}(\tilde{f}_A) < 0, \text{ then } \tilde{f}_A(t_A) < 0 \text{ for some } t_A > 1 \text{ and } \gamma_v \text{ connects } x \text{ to } f_{|A}^{-1}(\mathbb{R}_{<0}). \\ &\text{If } \mathrm{TC}(\tilde{f}_B) < 0, \text{ then } \tilde{f}_B(t_B) < 0 \text{ for some } t_B < 1 \text{ and } \gamma_{-v} \text{ connects } x \text{ to } f_{|B}^{-1}(\mathbb{R}_{<0}). \end{aligned} \tag{13}$$

Since $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$, the univariate signomial

$$\tilde{f}(t) := f(t^v * x) = S_{x,a}(t) + \sum_{\substack{\mu \in \sigma_+(f), \\ b < v\cdot\mu < a}} c_\mu x^\mu t^{v\cdot\mu} + S_{x,b}(t)$$

has at most two sign changes in its coefficient sign sequence. We denote the number of sign changes by $\#\mathrm{signvar}(\tilde{f})$. By Lemma 2.2, we have:

$$\begin{aligned} &\mathrm{im}(\gamma_v) \subseteq f^{-1}(\mathbb{R}_{<0}) \text{ or } \mathrm{im}(\gamma_{-v}) \subseteq f^{-1}(\mathbb{R}_{<0}), && \text{if } \mathrm{LC}(\tilde{f}) < 0 \text{ and } \mathrm{TC}(\tilde{f}) < 0, \\ &\mathrm{im}(\gamma_v) \subseteq f^{-1}(\mathbb{R}_{<0}), && \text{if } \mathrm{LC}(\tilde{f}) < 0 \text{ and } \#\mathrm{signvar}(\tilde{f}) \le 1, \\ &\mathrm{im}(\gamma_{-v}) \subseteq f^{-1}(\mathbb{R}_{<0}), && \text{if } \mathrm{TC}(\tilde{f}) < 0 \text{ and } \#\mathrm{signvar}(\tilde{f}) \le 1. \end{aligned} \tag{14}$$

In view of (13) and (14), to obtain (a) and (b) all we need is to show that one of the following holds:

(I) $\mathrm{LC}(\tilde{f}) = \mathrm{LC}(\tilde{f}_A) < 0$ and $\mathrm{TC}(\tilde{f}) = \mathrm{TC}(\tilde{f}_B) < 0$,
(II) $\mathrm{LC}(\tilde{f}) = \mathrm{LC}(\tilde{f}_A) < 0$ and $\#\mathrm{signvar}(\tilde{f}) \le 1$,
(III) $\mathrm{TC}(\tilde{f}) = \mathrm{TC}(\tilde{f}_B) < 0$ and $\#\mathrm{signvar}(\tilde{f}) \le 1$.

The leading and trailing coefficients $\mathrm{LC}(\tilde{f}), \mathrm{LC}(\tilde{f}_A), \mathrm{TC}(\tilde{f})$ and $\mathrm{TC}(\tilde{f}_B)$ depend on the signs of $S_{x,a}(t)$, $S_{x,b}(t)$. If $S_{x,a}(t)$ (resp. $S_{x,b}(t)$) is not the zero polynomial, then $\mathrm{LC}(\tilde{f}) = \mathrm{LC}(\tilde{f}_A) = \mathrm{LC}(S_{x,a})$ (resp. $\mathrm{TC}(\tilde{f}) = \mathrm{TC}(\tilde{f}_B) = \mathrm{TC}(S_{x,b})$). As $\tilde{f}(1) < 0$, we have that at least one of $S_{x,a}(t)$ and $S_{x,b}(t)$ is not the zero polynomial, furthermore $\mathrm{LC}(S_{x,a}) < 0$ or $\mathrm{TC}(S_{x,b}) < 0$. We have the following cases:

- If $\mathrm{LC}(S_{x,a}) < 0$ and $\mathrm{TC}(S_{x,b}) < 0$, then $S_{x,a} \not\equiv 0$, $S_{x,b} \not\equiv 0$ and (I) is satisfied.
- If $\mathrm{LC}(S_{x,a}) < 0$ and $\mathrm{TC}(S_{x,b}) > 0$ or $S_{x,b} \equiv 0$, then (II) holds.
- If $\mathrm{TC}(S_{x,b}) < 0$ and $\mathrm{LC}(S_{x,a}) > 0$ or $S_{x,a} \equiv 0$, then (III) holds.

□

If $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$, then by Proposition 2.4(iii), for every $R, S \subseteq \sigma(f)$ as in Proposition 2.5 the number of negative connected components of $f_{|R}$ respectively $f_{|S}$ is the same. So to reduce the computational cost of finding $b_0(f_{|R}^{-1}(\mathbb{R}_{<0}))$ and $b_0(f_{|S}^{-1}(\mathbb{R}_{<0}))$ one might choose the sets R, S as small as possible, that is $R = \mathcal{H}_{v,a}^+ \cap \sigma(f)$ and $S = \mathcal{H}_{v,b}^- \cap \sigma(f)$. However, if R and S are as large as possible, that is $R = A$ and $S = B$, then one can refine the bound from Proposition 2.5 by identifying negative connected components of $f_{|A}$ and $f_{|B}$ that intersect. To make this precise, consider the bipartite graph with vertex set and edges defined as

$$\begin{aligned} &\mathcal{B}_{A,B} := \mathcal{B}_0^-(f_{|A}) \sqcup \mathcal{B}_0^-(f_{|B}), \\ &\mathcal{E}_{A,B} := \{(U,V) \mid U \in \mathcal{B}_0^-(f_{|A}), V \in \mathcal{B}_0^-(f_{|B}) \colon U \cap V \neq \emptyset\}. \end{aligned} \tag{15}$$

Proposition 2.6. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial, $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ a pair of enclosing hyperplanes of $\sigma_+(f)$, and let $A, B \subseteq \sigma(f)$ be as in (9). Then*

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \leq C \leq b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) + b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big),$$

where C denotes the number of connected components of the graph $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$ from (15).

Proof. The inequality $C \leq b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) + b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big)$ follows from the fact that a graph cannot have more connected components than the number of its vertices.

Let φ be the surjective map from Proposition 2.5 with $R = A, S = B$. If the intersection of two connected sets is non-empty, then their union is connected. This gives that if $U, V \in \mathcal{B}_{A,B}$ lie in the same connected component of the graph $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$, then U and V are contained in the same connected component of $f^{-1}(\mathbb{R}_{<0})$. This yields a well defined map

$$\tilde{\varphi}\colon \mathcal{B}_{A,B}/\!\sim \to \mathcal{B}_0^-(f), \quad [U] \mapsto \varphi(U) = \text{ connected component of } f^{-1}(\mathbb{R}_{<0}) \text{ that contains } U.$$

where $\sim$ denotes the equivalence relation that identifies two elements of $\mathcal{B}_{A,B}$ if they lie in the same connected component of the graph $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$, and $[U]$ denotes the equivalence class of U under this relation. Since φ is surjective, so is $\tilde{\varphi}$, which gives $b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \leq C$. □

Remark 2.7. If $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of strict enclosing hyperplanes of $\sigma_+(f)$, then from (10) and Proposition 2.6 follows that

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \leq 2,$$

which recovers the bound from Theorem 2.3(ii).

Example 2.8. We revisit the signomial f in (1) from Example 2.1, whose support and a pair of enclosing hyperplanes $(\mathcal{H}_{v,2}, \mathcal{H}_{v,0})$, $v = (1,0)$ of $\sigma_+(f)$ are depicted in Figure 1(a). We determined the corresponding sets A, B in (11), (12).

Using the `Maple` [33] function `IsEmpty()`, one can check that $f_{|A}^{-1}(\mathbb{R}_{<0}) \cap f_{|B}^{-1}(\mathbb{R}_{<0}) = \emptyset$. Thus, the graph $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$ does not have any edges and the bound C on $b_0\big(f^{-1}(\mathbb{R}_{<0})\big)$ provided by Proposition 2.6 equals

$$C = b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) + b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big).$$

Since $\sigma(f_{|A})$ has a strict separating hyperplane, $b_0\big(f_{|A}^{-1}(\mathbb{R}_{<0})\big) = 1$ by Theorem 2.3(i). In Example 3.2, we show that $b_0\big(f_{|B}^{-1}(\mathbb{R}_{<0})\big) = 2$. Thus, Proposition 2.6 gives the bound $C = 3$. In this case, C equals the number of negative connected components of f. In Example 3.5, we will consider a signomial where the bound C is not sharp.

The sets $f^{-1}(\mathbb{R}_{<0}), f_{|A}^{-1}(\mathbb{R}_{<0})$ and $f_{|B}^{-1}(\mathbb{R}_{<0})$ are depicted in Figure 2(a),(b).

One might wonder if, for other choices of R and S, the bound from Proposition 2.5 could be refined using a similar idea as in the proof of Proposition 2.6. We will show that such a refinement is possible in the special case where all the exponent vectors are contained in the enclosing hyperplanes, or with other words the exponent vectors lie on parallel faces of the Newton polytope (Proposition 3.3). The next example shows that Proposition 2.6 might fail for other choices of R and S.

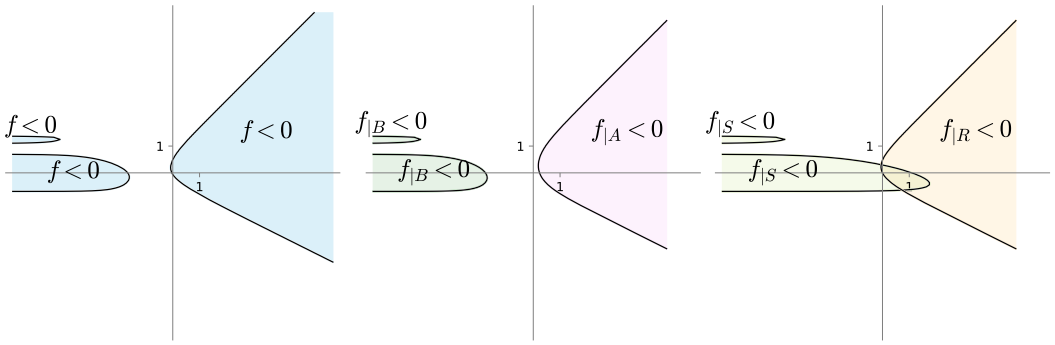


FIGURE 2. Negative connected components of $f, f_{|A}, f_{|B}, f_{|R}, f_{|S}$ from Example 2.8 and Example 2.9. For better visibility, the figures show the images of these sets under the coordinate-wise natural logarithm map $\mathbb{R}^2_{>0} \to \mathbb{R}^2, (x, y) \mapsto (\log(x), \log(y))$.

Example 2.9. Consider again the signomial f in (1) from Example 2.1 and choose

$$R = \{(3,2),(2,1),(2,3),(1,3),(2,0)\}, \quad S = \{(0,3),(0,1),(1,3),(0,4),(0,2),(0,0)\}.$$

Similarly to (15), we investigate the graph whose vertices are

$$\mathcal{B}_{R,S} := \mathcal{B}_0^-(f_{|R}) \sqcup \mathcal{B}_0^-(f_{|S}), \text{ and edges}$$
$$\mathcal{E}_{R,S} := \{(U,V) \mid U \in \mathcal{B}_0^-(f_{|R}), V \in \mathcal{B}_0^-(f_{|S}) \colon U \cap V \neq \emptyset\}.$$

Since $\sigma(f_{|R})$ has a strict separating hyperplane, $f_{|R}$ has one negative connected component. From Theorem 3.1, it will follow that $f_{|S}$ has two negative connected components. Thus, the graph $(\mathcal{B}_{R,S}, \mathcal{E}_{R,S})$ has three vertices. One can check that $(2,1) \in f_{|R}^{-1}(\mathbb{R}_{<0}) \cap f_{|S}^{-1}(\mathbb{R}_{<0}) \neq \emptyset$. This implies that the graph has at most two connected components. Since $b_0(f^{-1}(\mathbb{R}_{<0})) = 3$, the number of connected components of the graph $(\mathcal{B}_{R,S}, \mathcal{E}_{R,S})$ cannot be an upper bound on $b_0(f^{-1}(\mathbb{R}_{<0}))$. In Figure 2(c), we depicted the sets $f_{|R}^{-1}(\mathbb{R}_{<0}), f_{|S}^{-1}(\mathbb{R}_{<0})$.

To compute the bound C in Proposition 2.6, one should check whether the negative connected components of two signomials intersect. We finish this subsection with a criterion on the signed supports of two signomials guaranteeing that the intersection of their negative connected components is non-empty.

Proposition 2.10. *Let $f, g \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be signomials. If there exist negative exponent vectors $\beta_1 \in \sigma_-(f)$, $\beta_2 \in \sigma_-(g)$ such that*

$$\operatorname{Conv}\big(\{\beta_1, \beta_2\}\big) \cap \operatorname{Conv}\big(\sigma_+(f) \cup \sigma_+(g)\big) = \emptyset,$$

then $f^{-1}(\mathbb{R}_{<0}) \cap g^{-1}(\mathbb{R}_{<0}) \neq \emptyset$.

Proof. We start by observing

$$f(x) \leq f_{|\sigma_+(f)\cup\{\beta_1\}}(x), \quad \text{and} \quad g(x) \leq g_{|\sigma_+(g)\cup\{\beta_2\}}(x) \quad \text{for all } x \in \mathbb{R}^n_{>0}, \tag{16}$$

since $f, f_{|\sigma_+(f)\cup\{\beta_1\}}$ (resp. $g, g_{|\sigma_+(g)\cup\{\beta_2\}}$) differ only in monomials with negative coefficients.

As non-intersecting closed convex sets can be separated by an affine hyperplane see e.g. [26, Section 2.2, Theorem 1], from $\operatorname{Conv}(\{\beta_1, \beta_2\}) \cap \operatorname{Conv}(\sigma_+(f) \cup \sigma_+(g)) = \emptyset$ it follows that there exists $w \in \mathbb{R}^n$ such that

$$w \cdot \beta_1 > w \cdot \alpha, \quad w \cdot \beta_2 > w \cdot \alpha, \quad \text{for all } \alpha \in \sigma_+(f) \cup \sigma_+(g).$$

For a fixed $x \in \mathbb{R}^n_{>0}$, both univariate signomials

$$f_{|\sigma_+(f)\cup\{\beta_1\}}(t^w * x), \qquad g_{|\sigma_+(g)\cup\{\beta_2\}}(t^w * x)$$

have negative leading coefficients. Thus, there exists $t_0 \gg 0$ such that

$$f_{|\sigma_+(f)\cup\{\beta_1\}}(t_0^w * x) < 0 \quad \text{and} \quad g_{|\sigma_+(g)\cup\{\beta_2\}}(t_0^w * x) < 0.$$

By (16), we have $t_0^w * x \in f^{-1}_{|\sigma_+(f)\cup\{\beta_1\}}(\mathbb{R}_{<0}) \cap g^{-1}_{|\sigma_+(g)\cup\{\beta_2\}}(\mathbb{R}_{<0}) \subseteq f^{-1}(\mathbb{R}_{<0}) \cap g^{-1}(\mathbb{R}_{<0})$, which completes the proof. □

2.3. **One negative connected component.** Building on the results of Section 2.2, we describe conditions on the signs of the coefficients of f and its support $\sigma(f)$ that guarantee that f has one negative connected component. Similarly to Theorem 2.3(i), for polynomials satisfying these conditions the number of negative connected components does not depend on the values of the coefficients but only on their signs.

Theorem 2.11. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial such that $\sigma_+(f)$ has a pair of strict enclosing hyperplanes $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$. Assume that there exist negative exponent vectors $\beta_1, \beta_2 \in \sigma_-(f)$ such that*

- *$\beta_1 \in \mathcal{H}^+_{v,a}$, $\beta_2 \in \mathcal{H}^-_{v,b}$ and*
- *$\mathrm{Conv}(\{\beta_1, \beta_2\}) \cap \mathrm{Conv}(\sigma_+(f)) = \emptyset$.*

Then $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected.

Proof. Let $A, B \subseteq \sigma(f)$ as in (9). Since $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of strict enclosing hyperplanes of $\sigma_+(f)$, by (10) we have:

$$b_0\big(f^{-1}_{|A}(\mathbb{R}_{<0})\big) = 1, \qquad b_0\big(f^{-1}_{|B}(\mathbb{R}_{<0})\big) = 1.$$

The assumptions of the theorem are equivalent to $\beta_1 \in \sigma_-(f_{|A})$, $\beta_2 \in \sigma_-(f_{|B})$ and $\mathrm{Conv}(\{\beta_1, \beta_2\}) \cap \mathrm{Conv}(\sigma_+(f_{|A}) \cup \sigma_+(f_{|B})) = \emptyset$. Thus, by Proposition 2.10, $f^{-1}_{|A}(\mathbb{R}_{<0}) \cap f^{-1}_{|B}(\mathbb{R}_{<0}) \neq \emptyset$. Now, from Proposition 2.6 it follows that $b_0(f^{-1}(\mathbb{R}_{<0})) = 1$. □

Example 2.12. Consider the signomial $f = -x^4y^4 + 10\,x^3y^3 - 10\,x^4 - 10\,y^4 + 7\,xy + 5\,x - 1$. Figure 3(a) displays the exponent vectors of f. The pair $(\mathcal{H}_{v,3.5}, \mathcal{H}_{v,0.5})$, $v = (1, 0)$ is a pair of strict enclosing hyperplanes of $\sigma_+(f)$, furthermore

$$\mathrm{Conv}(\{(0, 4), (4, 4)\}) \cap \mathrm{Conv}(\sigma_+(f)) = \emptyset. \tag{17}$$

Thus, $f^{-1}(\mathbb{R}_{<0})$ is connected by Theorem 2.11. Figure 3(b),(c) show $f^{-1}(\mathbb{R}_{<0})$, $f^{-1}_{|A}(\mathbb{R}_{<0})$ and $f^{-1}_{|B}(\mathbb{R}_{<0})$, where $f_{|A} = -x^4y^4 + 10\,x^3y^3 - 10\,x^4 + 7\,xy + 5\,x$ and $f_{|B} = 10\,x^3y^3 - 10\,y^4 + 7\,xy + 5\,x - 1$.

By Theorem 2.3(iii), $f^{-1}(\mathbb{R}_{<0})$ is either empty or logarithmically convex, if the signomial f has at most one negative coefficient. In particular, in that case f has one negative connected component. A natural question to ask is what happens for signomials with at most one positive coefficient. For a univariate signomial f, it is easy to construct examples where f has one positive coefficient and $f^{-1}(\mathbb{R}_{<0})$ is disconnected, take for example $f = -x^2 + 3x - 1$.

If $n \geq 2$ and the positive exponent vector lies in the interior of the Newton polytope, then $f^{-1}(\mathbb{R}_{<0})$ is homeomorphic to the complement of a bounded convex set [38, Corollary 3.5] [19, Theorem 3.4], therefore $f^{-1}(\mathbb{R}_{<0})$ is connected. This argument does not work if the positive exponent vector lies on the boundary of $\mathrm{N}(f)$, but the conclusion is still true.

Corollary 2.13. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $\quad x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial. If $\dim \mathrm{N}(f) \geq 2$ and $\#\sigma_+(f) = 1$, then $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected.*

Proof. Write $\{\alpha\} = \sigma_+(f)$. Since $\dim \mathrm{N}(f) \geq 2$, there exist $\beta_1, \beta_2 \in \sigma_-(f)$ such that β_1, β_2 and α do not lie on a line. In that case,

$$\mathrm{Conv}(\{\beta_1, \beta_2\}) \cap \mathrm{Conv}(\{\alpha\}) = \emptyset.$$

Pick a hyperplane $\mathcal{H}_{v,a}$ that contains α such that β_1, β_2 lie in different open half-spaces determined by $\mathcal{H}_{v,a}$. Thus, $(\mathcal{H}_{v,a}, \mathcal{H}_{v,a})$ is a pair of strict enclosing hyperplanes of $\sigma_+(f)$.

Using Theorem 2.11, we conclude that $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected. □

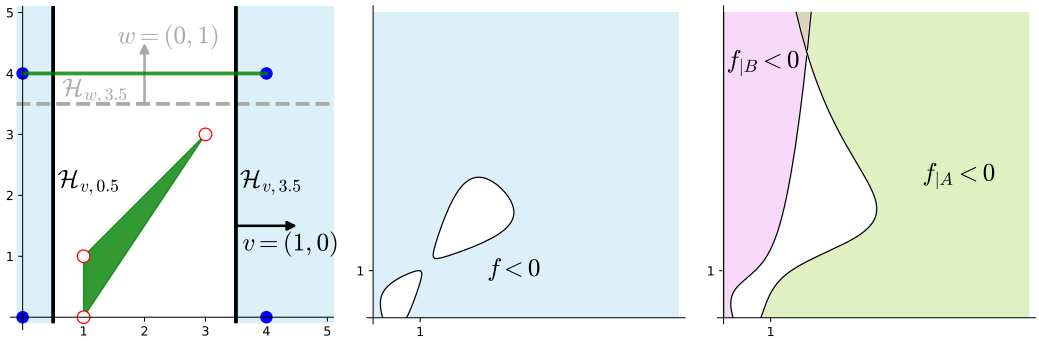


FIGURE 3. Illustration of Example 2.12 (a) Negative and positive exponent vectors of $f=-x^4y^4+10\,x^3y^3-10\,x^4-10\,y^4+7\,xy+5\,x-1$, blue dots are negative, red circles are positive. The black solid lines are strict enclosing hyperplanes of $\sigma_+(f)$. The gray dashed line separates $\mathrm{Conv}((0,4),(4,4))$ from $\mathrm{Conv}(\sigma_+(f))$. (b) Negative connected component of f. (c) Negative connected component of $f_{|A}=-x^4y^4+10\,x^3y^3-10\,x^4+7\,xy+5\,x$ and $f_{|B}=10\,x^3y^3-10\,y^4+7\,xy+5\,x-1$.

Corollary 2.13 shows that one can flip the signs in Theorem 2.3(iii) if $\dim \mathrm{N}(f)\geq 2$, i.e. if f has one negative coefficient then both $f^{-1}(\mathbb{R}_{<0})$ and $(-f)^{-1}(\mathbb{R}_{<0})$ are (possibly empty) connected sets. As discussed at the end of Section 2.1, a signomial has at most one negative connected component if its positive and negative exponent vectors are separated by a simplex and its negative vertex cones [19, Theorem 4.6]. We proceed by recalling this statement, and show that it is possible the flip the signs also in that case under some mild assumptions.

First we recall the definition of the *negative vertex cone* of a simplex. For an n-simplex $P\subseteq\mathbb{R}^n$ with vertices $\mu_0,\dots,\mu_n$, the negative vertex cone of P at the vertex μ_k is defined as

$$P^{-,k}:=\mu_k+\mathrm{Cone}(\mu_k-\mu_0,\dots,\mu_k-\mu_n).$$

Thus, $P^{-,k}$ is the cone with apex at μ_k which is generated by the edges pointing into μ_k. For the union of the negative vertex cones $P^{-,0},\dots,P^{-,n}$, we write P^-.

Theorem 2.14. *[19, Theorem 4.6] Let $f\colon\mathbb{R}^n_{>0}\to\mathbb{R}$, $x\mapsto\sum_{\mu\in\sigma(f)}c_\mu x^\mu$ be a signomial. If there exists an n-simplex $P\subseteq\mathbb{R}^n$ such that*

$$\sigma_-(f)\subseteq P,\quad \text{and}\quad \sigma_+(f)\subseteq P^-,$$

then $f^{-1}(\mathbb{R}_{<0})$ is either empty or contractible.

As another consequence of Theorem 2.11 we have the following result.

Corollary 2.15. *Let $n\geq 2$ and let $f\colon\mathbb{R}^n_{>0}\to\mathbb{R}$, $x\mapsto\sum_{\mu\in\sigma(f)}c_\mu x^\mu$ be a signomial. Assume that there exists an n-simplex $P\subseteq\mathbb{R}^n$ such that*

$$\sigma_+(f)\subseteq P\quad \text{and}\quad \sigma_-(f)\subseteq P^-.$$

If $\sigma_-(f)\cap\mathrm{int}(P^-)\neq\emptyset$, then $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected.

Proof. Denote $\mu_0,\dots,\mu_n$ the vertices of P and write P as

$$P=\bigcap_{j=0}^{n}\mathcal{H}^-_{v_j,a_j} \tag{18}$$

for a choice of normal vectors $v_0,\dots,v_n\in\mathbb{R}^n$ and scalars $a_0,\dots,a_n\in\mathbb{R}$ such that μ_k equals the intersection point of the hyperplanes $\mathcal{H}_{v_j,a_j}, j\in\{0,\dots,n\}\setminus\{k\}$. By [19, Proposition 4.3],

each negative vertex cone has the form:

$$P^{-,k} = \bigcap_{j=0, j\neq k}^{n} \mathcal{H}^{+}_{v_j,a_j} \qquad k = 0, \dots, n.$$

Note that if $\beta \in P^{-,k}$, then β is contained in $\mathcal{H}^{-,\circ}_{v_k,a_k}$. Since $\sigma_+(f) \subseteq P$ and $\sigma_-(f) \subseteq P^-$, we have

$$v_k \cdot \beta \geq a_k \geq v_k \cdot \alpha \geq v \cdot \mu_k \geq v_k \cdot \beta', \quad \text{and } v_k \cdot \alpha > v_k \cdot \beta' \tag{19}$$

for all $\beta \in \sigma_-(f) \setminus P^{-,k}$, $\alpha \in \sigma_+(f)$, $\beta' \in \sigma_-(f) \cap P^{-,k}$, $k = 0, \dots, n$.

Let $\beta_0 \in \sigma_-(f) \cap \operatorname{int}(P^-)$ and assume without loss of generality that $\beta_0 \in \operatorname{int}(P^{-,0})$, so

$$a_0 > v_0 \cdot \beta_0 \qquad \text{and} \qquad v_j \cdot \beta_0 > a_j, \quad \text{for } j = 1, \dots, n. \tag{20}$$

If $\sigma_-(f) \cap P^{-,1} = \emptyset$, then $\sigma_-(f) \subseteq \mathcal{H}^{+}_{v_1,a_1}$. Since $\beta_0 \in \mathcal{H}^{+,\circ}_{v_1,a_1}$, $\mathcal{H}_{v_1,a_1}$ is a strict separating hyperplane of $\sigma(f)$, which implies that $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected by Theorem 2.3(i).

Assume now that $\sigma_-(f) \cap P^{-,1} \neq \emptyset$, let $\beta_1 \in \sigma_-(f) \cap P^{-,1}$ and $b_1 \in \mathbb{R}$ such that

$$v_1 \cdot \alpha > b_1 > v_1 \cdot \beta_1 \qquad \text{for all } \alpha \in \sigma_+(f).$$

From (19), it follows that $(\mathcal{H}_{v_1,a_1}, \mathcal{H}_{v_1,b_1})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$. By (20), we have $\beta_0 \in \mathcal{H}^{+,\circ}_{v_1,a_1}$. Thus, $(\mathcal{H}_{v_1,a_1}, \mathcal{H}_{v_1,b_1})$ is a pair of strict enclosing hyperplanes.

Every element $\mu \in \operatorname{Conv}(\beta_0, \beta_1) \setminus \{\beta_1\}$ has the form $\mu = t\beta_0 + (1-t)\beta_1$ for some $t \in (0, 1]$. By (19) and (20), we have

$$v_2 \cdot \mu = t(v_2 \cdot \beta_0) + (1-t)(v_2 \cdot \beta_1) > a_2.$$

Thus, $\operatorname{Conv}(\beta_0, \beta_1) \setminus \{\beta_1\} \subseteq \mathcal{H}^{+,\circ}_{v_2,a_2}$, which implies that

$$\operatorname{Conv}(\{\beta_0, \beta_1\}) \cap \operatorname{Conv}(\sigma_+(f)) = \emptyset.$$

From Theorem 2.11, it follows that $f^{-1}(\mathbb{R}_{<0})$ is non-empty and connected. □

Example 2.16. (a) Consider the signomial

$$f = -x^5 y^{\frac{7}{3}} - x^5 y^2 + x^2 y^2 - xy^3 - y^4 + 2\,x^2 y^{\frac{4}{3}} + 2\,xy^2 - xy.$$

The simplex $P = \operatorname{Conv}((1,1),(4,2),(1,3))$ contains $\sigma_+(f)$, and the negative exponent vectors are contained in the union of the negative vertex cones P^-. By Theorem 2.14, we have that $(-f)^{-1}(\mathbb{R}_{<0})$ is connected. From Corollary 2.15, it follows that $f^{-1}(\mathbb{R}_{<0})$ is connected as well. The exponent vectors of f, the simplex P and its negative vertex cones are depicted in Figure 4 (a),(b).

To write P as an intersection of half-spaces as in (18) in the proof of Corollary 2.15, one can choose $v_0 = (-1, 0)$, $v_1 = (0.5, 1.5)$, $v_2 = (0.5, -1.5)$. With this choice we have:

$$P = \mathcal{H}^{-}_{v_0,-1} \cap \mathcal{H}^{-}_{v_1,5} \cap \mathcal{H}^{-}_{v_2,-1}.$$

(b) The following example shows that the assumption $\sigma_-(f) \cap \operatorname{int}(P^-) \neq \emptyset$ in Corollary 2.15 is necessary. If we remove the exponent vectors $(0,4), (5,2) \in \operatorname{relint}(P^-)$ from the support of f, the signomial

$$g = -x^5 y^{\frac{7}{3}} + x^2 y^2 - xy^3 + 2\,x^2 y^{\frac{4}{3}} + 2\,xy^2 - xy$$

satisfies that $\sigma_+(g) \subseteq P$ and $\sigma_-(g) \subseteq P^-$, however $g^{-1}(\mathbb{R}_{<0})$ has two connected components. Figure 4(c),(d) displays $\sigma(g)$ and $g^{-1}(\mathbb{R}_{<0})$.

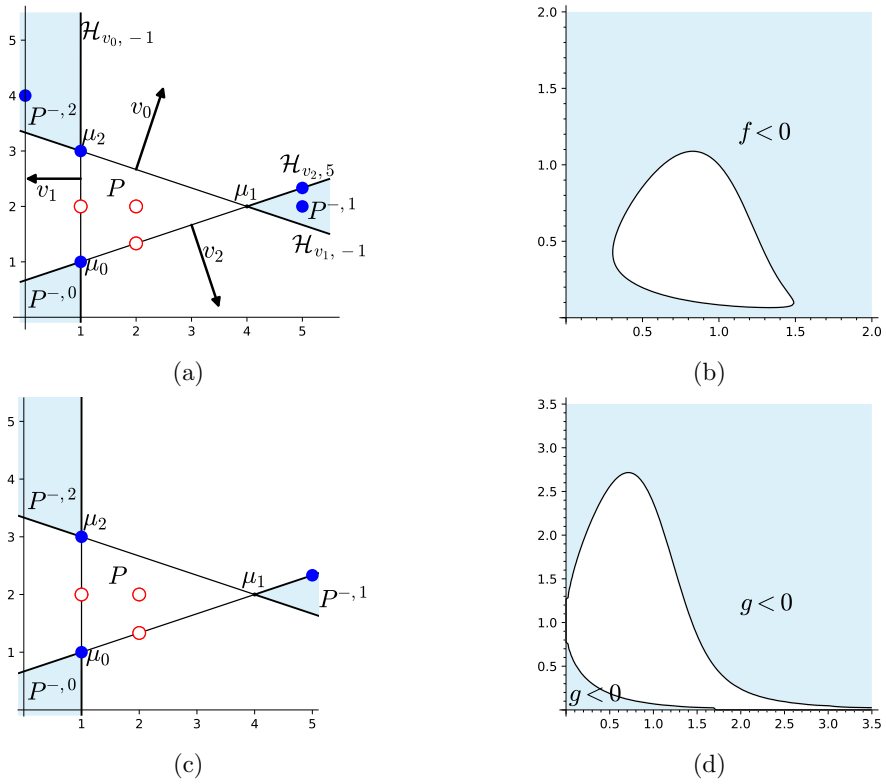


FIGURE 4. Illustration of Example 2.16 (a) A simplex P, its negative vertex cones, and the support of $f = -x^5y^{\frac{7}{3}} - x^5y^2 + x^2y^2 - xy^3 - y^4 + 2\,x^2y^{\frac{4}{3}} + 2\,xy^2 - xy$ (c) The support of $g = -x^5y^{\frac{7}{3}} + x^2y^2 - xy^3 + 2\,x^2y^{\frac{4}{3}} + 2\,xy^2 - xy$

3. Reduction to faces of the Newton polytope

3.1. **Negative and parallel faces.** In this section, we present two criteria to reduce the problem of finding the number of negative connected components of a signomial to the same problem for a signomial in less variables and monomials. The approach is based on how the exponent vectors of the signomial lie on the faces of the Newton polytope. A *face* of the Newton polytope of a signomial f is a set of the form

$$\mathrm{N}(f)_v := \{\omega \in \mathrm{N}(f) \mid v \cdot \omega = \max_{\mu \in \mathrm{N}(f)} v \cdot \mu\}$$

for some $v \in \mathbb{R}^n$. The vector v is called the *outer normal vector* of the face $\mathrm{N}(f)_v$. A polytope has finitely many faces [29, Theorem 3.46]. For a fixed face $F \subseteq \mathrm{N}(f)$, the set of vectors $v \in \mathbb{R}^n$ such that $\mathrm{N}(f)_v = F$ is called the *outer normal cone* of F. For more details about polytopes and their faces, we refer to the books [28, 29, 43]. To compute faces and outer normal cones of polytopes, one can use e.g. `Polymake` [23] or `SageMath` [40].

Theorem 3.1. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu \in \sigma(f)} c_\mu x^\mu$ be a signomial. If there exists a face $F \subseteq \mathrm{N}(f)$ such that $\sigma_-(f) \subseteq F$, then*

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) = b_0\big(f_{|F}^{-1}(\mathbb{R}_{<0})\big).$$

Proof. It follows directly from Proposition 2.4 with $R = \sigma(f) \cap F$. □

Example 3.2. Consider the signomial

$$f_{|B}(x,y) = 50x^2y^3 + xy^3 + y^4 - 9.5y^3 + 51x^2 + 30.5y^2 - 37y + 12$$

from Example 2.1 whose Newton polytope is shown in Figure 1(c).

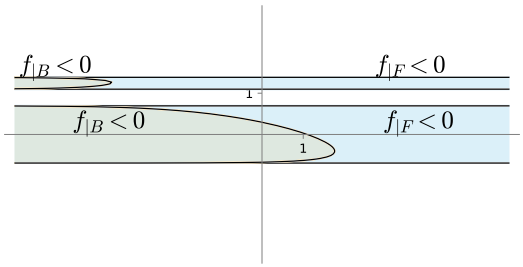


FIGURE 5. Images of the negative connected components of $f_{|B}$ and $f_{|F}$ from Example 3.2 under the coordinate-wise natural logarithm map.

The face $F = \operatorname{Conv}\big((0,0),(0,4)\big) \subseteq \mathrm{N}(f_{|B})$ contains all the negative exponent vectors of $f_{|B}$. Since $f_{|F}$ is univariate, it is easy to conclude that $f_{|F}^{-1}(\mathbb{R}_{<0})$ has two connected components. By Theorem 3.1, $f_{|B}^{-1}(\mathbb{R}_{<0})$ has also two connected components.

The next proposition refines the bound from Proposition 2.5 in the special case where the pair of enclosing hyperplanes $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ of $\sigma_+(f)$ is non-strict and all the exponent vectors of f lie on the union of the two hyperplanes. In this case, the hyperplanes $\mathcal{H}_{v,a}, \mathcal{H}_{v,b}$ cut out two parallel faces of the Newton polytope of f, i.e.:

$$\mathrm{N}(f)_v = \mathrm{N}(f) \cap \mathcal{H}_{v,a}, \quad \mathrm{N}(f)_{-v} = \mathrm{N}(f) \cap \mathcal{H}_{v,b}.$$

Similarly to (15), we define a bipartite graph whose set of vertices and edges are

$$\begin{aligned} &\mathcal{B}_v := \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_v}) \sqcup \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_{-v}}), \\ &\mathcal{E}_v := \{(U,V) \mid U \in \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_v}), V \in \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_{-v}}) \colon U \cap V \neq \emptyset\}. \end{aligned} \tag{21}$$

Proposition 3.3. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial. Assume that there exists $v \in \mathbb{R}^n$ such that $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$. Then*

$$b_0\big(f^{-1}(\mathbb{R}_{<0})\big) \leq D \leq b_0\big(f_{|N(f)_v}^{-1}(\mathbb{R}_{<0})\big) + b_0\big(f_{|N(f)_{-v}}^{-1}(\mathbb{R}_{<0})\big),$$

where D denotes the number of the connected components of the graph $(\mathcal{B}_v, \mathcal{E}_v)$ from (21).

Proof. Let $a := \max_{\mu\in\sigma(f)} v\cdot\mu$ and $b := \min_{\mu\in\sigma(f)} v\cdot\mu$. Since $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$, the pair $(\mathcal{H}_{v,a}, \mathcal{H}_{v,b})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$. Let $A, B \subseteq \sigma(f)$ be as in (9) and let φ be the map from Proposition 2.5 with $R = \sigma(f) \cap \mathrm{N}(f)_v$ and $S = \sigma(f) \cap \mathrm{N}(f)_{-v}$. Recall that $A = R \cup \sigma_+(f)$, $B = S \cup \sigma_+(f)$.

We show that the map

$$\tilde{\varphi}\colon \Big(\mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_v}) \sqcup \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_{-v}})\Big)/{\sim} \longrightarrow \mathcal{B}_0^-(f), \quad [U] \mapsto \varphi(U)$$

is well defined and surjective. Here $\sim$ denotes the equivalence relation that identifies two elements in $\mathcal{B}_v$ if they lie in the same connected component of the graph $(\mathcal{B}_v, \mathcal{E}_v)$.

Let U and V be in the same connected component of the bipartite graph $(\mathcal{B}_v, \mathcal{E}_v)$ and let $U_1 = U, V_1, U_2, \ldots, V_{m-1}, U_m, V_m = V$ be a path in the graph such that $U_i \in \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_v})$, $V_i \in \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_{-v}})$, $U_i \cap V_i \neq \emptyset$ for $i = 1, \ldots, m$ and $V_i \cap U_{i+1} \neq \emptyset$ for $i = 1, \ldots, m-1$. It is enough to show that, for all $i = 1, \ldots, m$ and $j = i, i+1$ it holds that $U_i \cap f_{|A}^{-1}(\mathbb{R}_{<0})$ and $V_j \cap f_{|B}^{-1}(\mathbb{R}_{<0})$ lie in the same connected component of $f^{-1}(\mathbb{R}_{<0})$. We show this by constructing a path γ between $U_i \cap f_{|A}^{-1}(\mathbb{R}_{<0})$ and $V_j \cap f_{|B}^{-1}(\mathbb{R}_{<0})$ such that $\operatorname{im}(\gamma) \subseteq f^{-1}(\mathbb{R}_{<0})$.

For a fixed $x \in U_i \cap V_j$, consider the univariate signomial

$$f(t^v * x) = \Big(\sum_{\mu\in\sigma(f)\cap\mathrm{N}(f)_v} c_\mu x^\mu\Big)t^a + \Big(\sum_{\mu\in\sigma(f)\cap\mathrm{N}(f)_{-v}} c_\mu x^\mu\Big)t^b = f_{|N(f)_v}(x)t^a + f_{|N(f)_{-v}}(x)t^b.$$

Since $f_{|N(f)_v}(x) < 0$ and $f_{|N(f)_{-v}}(x) < 0$, we have $f(t^v * x) < 0$, $f_{|N(f)_v}(t^v * x) < 0$ and $f_{|N(f)_{-v}}(t^v * x) < 0$ for all $t > 0$, and therefore the image of the path

$$\gamma\colon \mathbb{R}_{>0} \to \mathbb{R}^n_{>0}, \quad t \mapsto t^v * x$$

is contained in $U_i \cap V_j \cap f^{-1}(\mathbb{R}_{<0})$.

Since $\mathrm{LC}(f_{|A}(t^v * x)) = f_{|N(f)_v}(x) < 0$ and $\mathrm{TC}(f_{|B}(t^v * x)) = f_{|N(f)_{-v}}(x) < 0$, there exist $t_1 \gg 1$ and $0 < t_2 \ll 1$ such that $\gamma(t_1) \in f^{-1}_{|A}(\mathbb{R}_{<0})$ and $\gamma(t_2) \in f^{-1}_{|B}(\mathbb{R}_{<0})$. Since $\mathrm{im}(\gamma) \subseteq U_i$ and $\mathrm{im}(\gamma) \subseteq V_j$, it follows that $\gamma(t_1) \in U_i \cap f^{-1}_{|A}(\mathbb{R}_{<0})$ and $\gamma(t_2) \in V_j \cap f^{-1}_{|B}(\mathbb{R}_{<0})$. Since $x, \gamma(t_1), \gamma(t_2)$ lie in the same connected component of $f^{-1}(\mathbb{R}_{<0})$, it follows that $U_i \cap f^{-1}_{|A}(\mathbb{R}_{<0})$ and $V_j \cap f^{-1}_{|B}(\mathbb{R}_{<0})$ lie in the same connected component of $f^{-1}(\mathbb{R}_{<0})$. Thus, the map $\tilde{\varphi}$ is well defined.

From the surjectivity of φ follows that $\tilde{\varphi}$ is surjective, which gives $b_0(f^{-1}(\mathbb{R}_{<0})) \leq D$. □

Remark 3.4. Let f be a signomial satisfying

$$\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$$

for some $v \in \mathbb{R}$ as in Proposition 3.3 and let A, B the sets as in (9). By Theorem 3.1, there is a bijection between the sets

$$\mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_v}) \sqcup \mathcal{B}_0^-(f_{|\,\mathrm{N}(f)_{-v}}) \longleftrightarrow \mathcal{B}_0^-(f_{|A}) \sqcup \mathcal{B}_0^-(f_B).$$

Thus, the graphs $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$ from (15) and $(\mathcal{B}_v, \mathcal{E}_v)$ from (21) have the same number of vertices. Since

$$f^{-1}_{|A}(\mathbb{R}_{<0}) \subseteq f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0}) \quad \text{and} \quad f^{-1}_{|B}(\mathbb{R}_{<0}) \subseteq f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0}),$$

every edge of the graph $(\mathcal{B}_{A,B}, \mathcal{E}_{A,B})$ corresponds to an edge of the graph $(\mathcal{B}_v, \mathcal{E}_v)$. Thus, the bound provided in Proposition 3.3 is always smaller or equal than the bound given in Proposition 2.6.

Example 3.5. The bound on $b_0(f^{-1}(\mathbb{R}_{<0}))$ in Proposition 2.6 and in Proposition 3.3 can be different, even though the two statements look similar.

To demonstrate this, consider the signomial $f = 73\,x - 55\,x^2 - x^4 + y - 20\,xy + x^4y$. The Newton polytope of f is shown in Figure 6(a). We have that $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$, for $v = (0, 1)$. The restrictions of f to these two faces are given by:

$$f_{\mathrm{N}(f)_v} = 73\,x - 55\,x^2 - x^4, \qquad f_{\mathrm{N}(f)_{-v}} = y - 20\,xy + x^4y.$$

The pair $(\mathcal{H}_{v,1}, \mathcal{H}_{v,0})$ is a pair of enclosing hyperplanes of $\sigma_+(f)$. Let $A, B \subseteq \sigma(f)$ be as defined in (9):

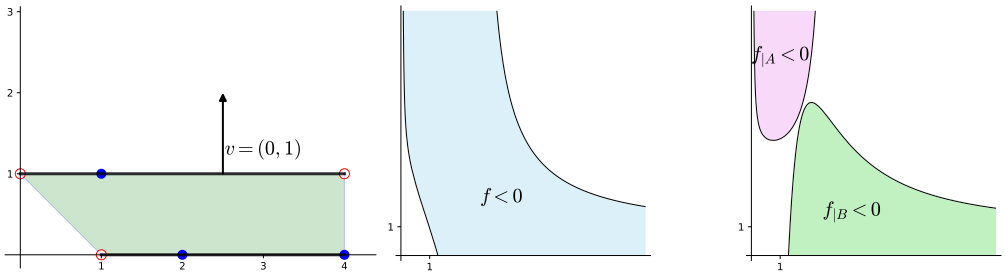
$$A = (\mathcal{H}^+_{v,1} \cap \sigma_-(f)) \cup \sigma_+(f) = \{(1,1), (1,0), (0,1), (4,1)\},$$
$$B = (\mathcal{H}^-_{v,0} \cap \sigma_-(f)) \cup \sigma_+(f) = \{(2,0), (4,0), (1,0), (0,1), (4,1)\}.$$

Since $f_{|\,\mathrm{N}(f)_v}$ and $f_{|A}$ only have one negative coefficient, the sets $f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0})$ and $f^{-1}_{|A}(\mathbb{R}_{<0})$ are connected by Theorem 2.3(iii). Since the supports of $f_{|\,\mathrm{N}(f)_{-v}}$ and $f_{|B}$ have strict separating hyperplanes, Theorem 2.3(i) implies that $f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0})$, $f^{-1}_{|B}(\mathbb{R}_{<0})$ are connected.

One can also verify that $f^{-1}_{|A}(\mathbb{R}_{<0}) \cap f^{-1}_{|B}(\mathbb{R}_{<0}) = \emptyset$, e.g. using the `Maple` [33] function `IsEmpty()`. Thus, the bound on the number of connected components of $f^{-1}(\mathbb{R}_{<0})$ provided by Proposition 2.6 is two. The negative connected components of $f_{|A}$, $f_{|B}$ are shown in Figure 6(c).

On the other hand, $(1,1) \in f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0}) \cap f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0})$. Thus, Proposition 3.3 gives that $f^{-1}(\mathbb{R}_{<0})$ is connected.

From Proposition 3.3, we derive a criterion to ensure that a polynomial has at most one negative connected component. This criterion can be interpreted as a version of Theorem 2.11 where we replace strict separating hyperplanes by non-strict ones.



(a) (b) (c)

FIGURE 6. Illustration of Example 3.5 (a) Newton polytope of $f = 73\,x - 55\,x^2 - x^4 + y - 20\,xy + x^4y$. (b) The negative connected component of f (c)Negative connected components of $f_{|A}$ (purple) and $f_{|B}$ (green), where $A = \{(1,1),(1,0),(0,1),(4,1)\}$, $B = \{(2,0),(4,0),(1,0),(0,1),(4,1)\}$.

Theorem 3.6. *Let $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ be a signomial. Assume that there exists $v \in \mathbb{R}^n$ such that $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$ and*

$$b_0\big(f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0})\big) = b_0\big(f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0})\big) = 1.$$

If there exist negative exponent vectors $\beta_1 \in \mathrm{N}(f)_v$ and $\beta_2 \in \mathrm{N}(f)_{-v}$ such that $\mathrm{Conv}(\beta_1,\beta_2)$ is an edge of $\mathrm{N}(f)$, then $b_0\big(f^{-1}(\mathbb{R}_{<0})\big) = 1$.

Proof. Throughout the proof we assume that $\mathrm{N}(f)_v \neq \mathrm{N}(f)_{-v}$, otherwise the statement is obvious. First, we show that $\mathrm{Conv}(\beta_1,\beta_2) \cap \sigma(f) = \{\beta_1,\beta_2\}$. For $\alpha \in \mathrm{Conv}(\beta_1,\beta_2)\cap\sigma(f)$, if $\alpha \notin \{\beta_1,\beta_2\}$, then there exists $t \in (0,1)$ such that $\alpha = t\beta_1 + (1-t)\beta_2$. Let $\mathcal{H}_{v,a}$ and $\mathcal{H}_{v,b}$ be the supporting hyperplanes of $\mathrm{N}(f)_v$ and $\mathrm{N}(f)_{-v}$ respectively. As $\mathrm{N}(f)_v \neq \mathrm{N}(f)_{-v}$, $a > b$. Thus, we have

$$v\cdot\alpha = t(v\cdot\beta_1) + (1-t)(v\cdot\beta_2) = ta + (1-t)b \neq a,b \qquad \text{if } t\in(0,1).$$

which contradicts $\alpha \in \mathcal{H}_{v,a}\cup\mathcal{H}_{v,b}$.

Proposition 2.10 implies that $f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0}) \cap f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0}) \neq \emptyset$. Thus, by Proposition 3.3 we have $b_0\big(f^{-1}(\mathbb{R}_{<0})\big) = 1$. □

Remark 3.7. In the last step of the proof of Theorem 3.6, instead of Proposition 3.3 one could use Theorem 3.1 and Proposition 2.6 as well to conclude that $b_0\big(f^{-1}(\mathbb{R}_{<0})\big) = 1$.

Example 3.8. For each $c_1,\dots,c_9 \in \mathbb{R}_{>0}$, the Newton polytope of

$$f(x,y,z) = c_1x + c_2xy - c_3y - c_4 + c_5yz - c_6z - c_7xz - c_9xyz$$

is the cube depicted in Figure 7. The top and the bottom faces of the cube are parallel to each other and contain all the exponent vectors of f. Choosing $v = (0,0,1)$, we have that $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$. Since both $\sigma(f_{|\,\mathrm{N}(f)_v})$ and $\sigma(f_{|\,\mathrm{N}(f)_{-v}})$ have a strict separating hyperplane, $b_0\big(f^{-1}_{|\,\mathrm{N}(f)_v}(\mathbb{R}_{<0})\big) = b_0\big(f^{-1}_{|\,\mathrm{N}(f)_{-v}}(\mathbb{R}_{<0})\big) = 1$ by Theorem 2.3(i).

The vertices of the edge $\mathrm{Conv}\big((0,0,0),(0,0,1)\big)$ correspond to negative exponent vectors, thus $f^{-1}(\mathbb{R}_{<0})$ is connected by Theorem 3.6.

3.2. **Algorithm for connectivity.** Based on Theorem 3.1 and Theorem 3.6, we give a recursive algorithm that checks a sufficient condition for having a signomial f one negative connected component. Using Theorem 3.1, we reduce f to a face of its Newton polytope that contains all the negative exponent vectors, if possible. Using Theorem 3.6, we split up f to parallel faces of its Newton polytope whose union contains all exponent vectors of f. We repeat this reduction until the polynomials are simple enough, and we can apply one of the following criterion:

(i) f has one negative coefficient, Theorem 2.3(iii),
(ii) f has one positive coefficient and $n \geq 2$, Corollary 2.13,

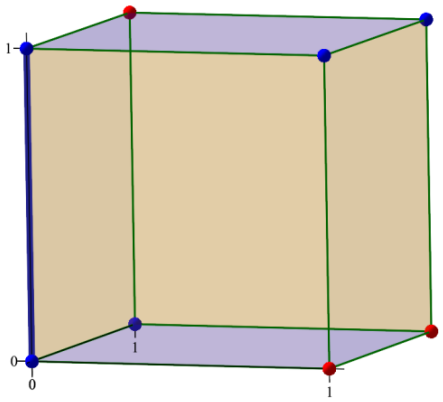


FIGURE 7. Newton polytope of $f(x,y,z) = c_1x+c_2xy-c_3y-c_4+c_5yz-c_6z-c_7xz-c_9xyz$ from Example 3.8. Red (resp. blue) dots correspond to positive (resp. negative) exponent vectors. The blue squares are parallel faces of $\mathrm{N}(f)$ whose union contains $\sigma(f)$. The blue thick edge $\mathrm{Conv}((0,0,0),(0,0,1))$ joins the two blue parallel faces and its vertices are negative exponent vectors.

(iii) the support of f has a strict separating hyperplane, Theorem 2.3(i),
(iv) the exponent vectors of f lie in and outside of a simplex as in Theorem 2.14 or in Corollary 2.15.

We define a submethod `CheckConnectivity()` that checks these conditions. If one of (i)-(iv) holds, `CheckConnectivity(`f`)` returns true, and we know that $f^{-1}(\mathbb{R}_{<0})$ is connected. If none of the conditions (i)-(iv) is true, `CheckConnectivity(`f`)` returns false. Checking (i) and (ii) is simple. Deciding whether $\sigma(f)$ has a strict separating hyperplane can be done using linear programming. Therefore, condition (iii) can be checked even for signomials in many variables and many monomials. Checking condition (iv) is a significantly harder problem, and in practice we might avoid it.

The submethod `IntersectionNonempty()` checks whether there exists an edge of the Newton polytope between two parallel faces such that both vertices of the edge correspond to a negative exponent vector.

To compute the smallest face $F \subseteq \mathrm{N}(f)$ such that $\sigma_-(f) \subseteq F$, one proceeds as follows. First, one finds all the facets of $\mathrm{N}(f)$ (using `Polymake` [23] or `SageMath` [40]). If $\mathrm{N}(f)$ does not have any facet containing $\sigma_-(f)$, then $\mathrm{N}(f)$ is the smallest face that contains $\sigma_-(f)$. Otherwise, the intersection of the facets containing $\sigma_-(f)$ give the smallest face that contains $\sigma_-(f)$.

One possible way to compute all proper faces $\mathrm{N}(f)_v \subseteq \mathrm{N}(f)$ such that $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$ is the following:

(1) Compute the outer normal fan $\mathcal{F}$ of $\mathrm{N}(f)$ and the common refinement $\mathcal{F} \bigwedge -\mathcal{F}$ of $\mathcal{F}$ and $-\mathcal{F}$ ([43, Definition 7.6]). Here, $-\mathcal{F}$ is the inner normal fan of $\mathrm{N}(f)$, i.e. the fan obtained by taking the negative of each cone in $\mathcal{F}$.
(2) Collect a vector from the relative interior of each cone in $\mathcal{F} \bigwedge -\mathcal{F}$. These vectors are the normal vectors of the parallel faces of $\mathrm{N}(f)$.
(3) Consider all the parallel faces and check whether their union contains $\sigma(f)$. This can be done by taking each vector v from step (2) and computing their scalar product with the exponent vectors. The vector v gives a pair of parallel faces containing $\sigma(f)$ if and only if $\{v \cdot \mu \mid \mu \in \sigma(f)\}$ has exactly two elements.

If the Newton polytope has many faces, computing $\mathcal{F} \bigwedge -\mathcal{F}$ might become too expensive. In our implementation, we use the following simplification. For a facet $F \subset \mathrm{N}(f)$, there exist a unique $v \in \mathbb{R}^n$ such that $F = \mathrm{N}(f)_v$. We consider all the facets $\mathrm{N}(f)_v$ and check whether $\sigma(f) \subseteq \mathrm{N}(f)_v \cup \mathrm{N}(f)_{-v}$. Thus, our code runs through only a sublist of the list in the for-loop in line 8 in Algorithm 1.

Algorithm 1 `CheckConnectivityRecursive`

Input: a signomial f
Output: **true** if $f^{-1}(\mathbb{R}_{<0})$ is connected, **false** if the method is inconclusive

```
1: if CheckConnectivity(f) = true then
2:   return true
3: end if
4: F ← smallest face of N(f) that contains σ_(f)
5: if F ⊆ N(f) is a proper face then
6:   return CheckConnectivityRecursive( f_|F)
7: end if
8: for every proper face N(f)_v ⊆ N(f) such that σ(f) ⊆ N(f)_v ∪ N(f)_{-v} do
9:   if IntersectionNonempty( f_|N(f)_v, f_|N(f)_{-v} ) then
10:    v_connected ← CheckConnectivityRecursive( f_|N(f)_v)
11:    vminus_connected ← CheckConnectivityRecursive( f_|N(f)_{-v})
12:    if v_connected and vminus_connected then
13:      return true
14:    end if
15:  end if
16: end for
17: return false
```

Using `OSCAR` [12, 35] and `Polymake` [23], we implemented Algorithm 1 in `Julia`. The code can be found at the Github repository [37].

Example 3.9. To demonstrate how Algorithm 1 works, consider the polynomial

$$f(x,y,z,w) = c_1x + c_2xy - c_3y - c_4 + c_5yz - c_6z - c_7xz - c_9xyz + c_{10}w^3 + c_{11}xw,$$

where $c_1, \dots, c_{11} \in \mathbb{R}_{>0}$. Since $\#\sigma_-(f) \geq 2$, $\#\sigma_+(f) \geq 2$ and $\sigma(f)$ does not have a strict separating hyperplane `CheckConnectivity(`f`)` returns false. Following the algorithm, we compute the smallest face $F \subseteq \mathrm{N}(f)$ containing $\sigma_-(f)$. This is a 3-dimensional face with normal vector $v = (0,0,0,-1)$. Then the algorithm calls `CheckConnectivityRecursive(`$f_{|F}$`)`, where

$$f_{|F} = c_1x + c_2xy - c_3y - c_4 + c_5yz - c_6z - c_7xz - c_9xyz,$$

which is the same polynomial as in Example 3.8.

Since $f_{|F}$ have more than one positive and more than one negative exponent vectors and $\sigma(f_{|F})$ does not have a strict separating hyperplane, `CheckConnectivity(`$f_{|F}$`)` returns false and the algorithm continues with computing the smallest face of $\mathrm{N}(f_{|F}) = F$ containing $\sigma_-(f_{|F})$. Since this smallest face is F itself, the algorithm proceeds with the for-loop in line 8.

The faces

$$G_1 = F_{v_1}, \quad v_1 = (0,0,1,-1), \qquad G_2 = F_{v_2}, \quad v_2 = (0,0,-1,-1)$$

are parallel and contains all the exponent vectors of $f_{|F}$. As discussed in Example 3.8, there is an edge with negative vertices between G_1 and G_2, thus `IntersectionNonEmpty(`$f_{|G_1}, f_{|G_2}$`)` returns true and the method `CheckConnectivityRecursive` is called again for $f_{|G_1}$ and f_{G_2}. Since $\sigma(f_{|G_1})$ and $\sigma(f_{|G_2})$ have strict separating hyperplanes, both `CheckConnectivity(`$f_{|G_1}$`)` and `CheckConnectivity(`$f_{|G_2}$`)` return true, and the algorithm terminates with the result that $f^{-1}(\mathbb{R}_{<0})$ is connected. The steps taken by the algorithm can be found in Figure 8.

3.3. **The closure property.** We finish the section with a statement that will allow us to apply Algorithm 1 for reaction networks in Section 4. We say that a signomial f satisfies the *closure property*, if the closure of $f^{-1}(\mathbb{R}_{<0})$ equals $f^{-1}(\mathbb{R}_{\leq 0})$.

Proposition 3.10. *A signomial $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$, $x \mapsto \sum_{\mu\in\sigma(f)} c_\mu x^\mu$ satisfies the closure property, if one of the following holds:*

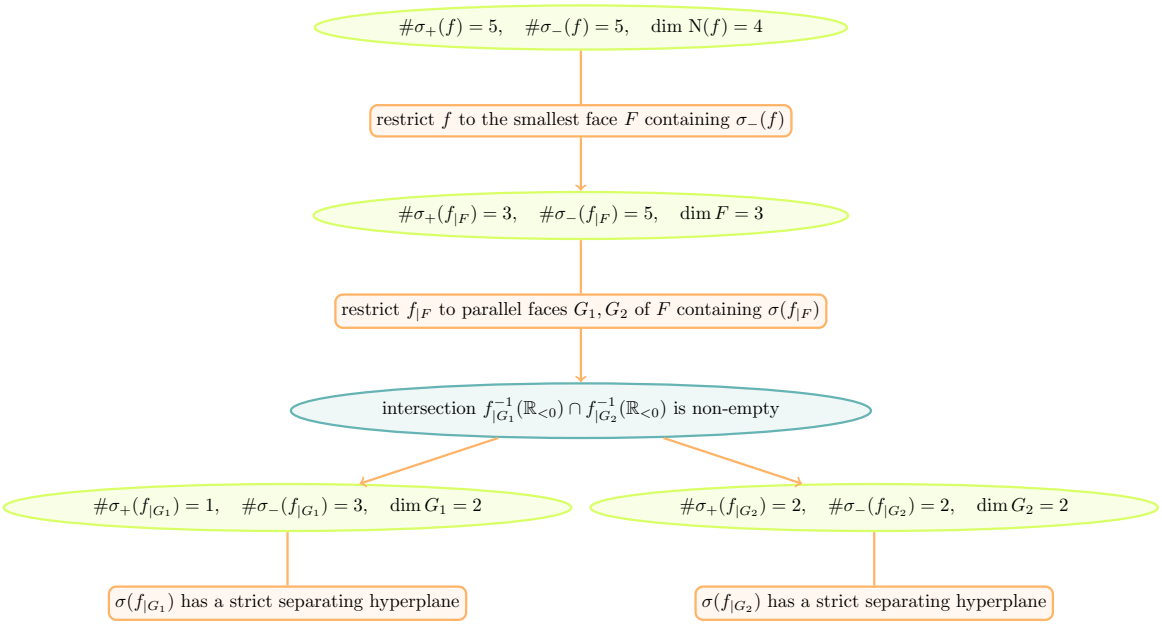


FIGURE 8. Depiction of the steps taken by Algorithm 1 as in Example 3.9.

(i) *$\sigma(f)$ has a strict separating hyperplane.*
(i) *$\sigma_-(f) \subseteq F$, for a proper face $F \subseteq \mathrm{N}(f)$.*

Proof. (i) has been showed in [19, Theorem 3.6]. (ii) follows by almost the same argument. We recall it for the sake of completeness. Since $f^{-1}(\mathbb{R}_{\leq 0})$ is closed, it contains the closure of $f^{-1}(\mathbb{R}_{<0})$. To show the reverse inclusion, we pick a point $x \in f^{-1}(\{0\})$ and show that it can be written as a limit of elements from $f^{-1}(\mathbb{R}_{<0})$.

Let $v \in \mathbb{R}^n \setminus \{0\}$ such that $\mathrm{N}(f)_v = F$. Consider the univariate signomial

$$\mathbb{R}_{>0} \to \mathbb{R}, \quad t \mapsto f(t^v * x) = \Big(\sum_{\mu \in F \cap \sigma(f)} c_\mu x^\mu \Big) t^d + \sum_{\mu \in \sigma(f) \setminus F} c_\mu x^\mu t^{v \cdot \mu},$$

where $d = \max_{\mu \in \sigma(f)} v \cdot \mu$. The leading coefficient of $f(t^v * x)$ is negative since otherwise $f(x) > 0$. Since F is a proper face, $\sigma(f) \setminus F \neq \emptyset$, therefore the trailing coefficient of $f(t^v * x)$ is positive. By Descartes' rule of signs, 1 is a positive real root of $f(t^v * x)$ and $(1 + \frac{1}{n})^v * x \in f^{-1}(\mathbb{R}_{<0})$ for all $n \in \mathbb{N}$. Furthermore, the sequence $\{(1 + \frac{1}{n})^v * x\}_{n \in \mathbb{N}}$ converges to x. □

4. Application to reaction networks

A reaction network is a collection of reactions between (bio)chemical species. The species are represented by formal variables and the reactions by arrows between non-negative integer linear combination of the species. The goal of reaction network theory is to study the evolution of the concentration of the species in time, which is usually modeled by an ordinary differential equation system (ODE). For an introduction to reaction network theory, we refer to [13]. Here, we consider some specific reaction networks, where Algorithm 1 can be applied to verify that the *parameter region of multistationarity* is path connected.

Remark 4.1. An open subset of an Euclidean space is connected if and only if it is path connected. Thus, for negative connected components of a signomial these two notions of connectivity coincide. As the parameter region of multistationarity might not be an open set, path-connectivity is a stronger property than connectivity.

4.1. **Weakly irreversible phosphorylation cycle.** First, we consider the reaction network that was studied in [34]:

$$\begin{aligned} &\mathrm{S}+\mathrm{E} \underset{\kappa_2}{\overset{\kappa_1}{\rightleftharpoons}} Y_1 \xrightarrow{\kappa_3} Y_2 \underset{\kappa_5}{\overset{\kappa_4}{\rightleftharpoons}} \mathrm{S_p}+\mathrm{E} \underset{\kappa_7}{\overset{\kappa_6}{\rightleftharpoons}} Y_3 \xrightarrow{\kappa_8} Y_4 \underset{\kappa_{10}}{\overset{\kappa_9}{\rightleftharpoons}} \mathrm{S_{pp}}+\mathrm{E} \\ &\mathrm{S_{pp}}+\mathrm{F} \underset{\kappa_{12}}{\overset{\kappa_{11}}{\rightleftharpoons}} \mathrm{Y}_5 \xrightarrow{\kappa_{13}} \mathrm{Y}_6 \underset{\kappa_{15}}{\overset{\kappa_{14}}{\rightleftharpoons}} \mathrm{S_p}+\mathrm{F} \underset{\kappa_{17}}{\overset{\kappa_{16}}{\rightleftharpoons}} \mathrm{Y}_7 \xrightarrow{\kappa_{18}} \mathrm{Y}_8 \underset{\kappa_{20}}{\overset{\kappa_{19}}{\rightleftharpoons}} \mathrm{S}+\mathrm{F}, \end{aligned} \tag{22}$$

which represents the two-site phosphorylation cycle of a substrate S, where both the phosphorylation and dephosphorylation processes follow a weakly irreversible mechanism. The reaction network has 13 species and 20 reactions. Under the assumption of mass action kinetics, the evolution of the concentration of the species is modeled by an ODE system of the form:

$$\dot{x} = f_\kappa(x),$$

where $\kappa = (\kappa_1, \ldots, \kappa_{20})$ are positive parameters, called *reaction rate constants*, and f_κ is a polynomial in 13 variables. One can find the exact form of f_κ in the accompanying `Jupyter` notebook [37].

The set of all positive steady states of the ODE system equals the variety

$$V_\kappa := \{x \in \mathbb{R}^{13}_{>0} \mid f_\kappa(x) = 0\}.$$

For a given initial condition, the trajectories of the ODE are contained in *stoichiometric compatibility classes* that are defined as

$$\mathcal{P}_c := \{x \in \mathbb{R}^{13}_{>0} \mid Wx = c\},$$

where $c \in \mathbb{R}^3$ is the so-called *total concentration parameter* and $W \in \mathbb{R}^{3\times 13}$ is a matrix whose rows give the *conservation laws* of the network. A pair of parameters (κ, c) *enables multistationarity* if $V_\kappa \cap \mathcal{P}_c$ contains at least two points. The *parameter region of multistationarity* is the set of all pairs (κ, c) enabling multistationarity.

By sampling parameters and using a connectivity graph, connectivity of the parameter region of multistationarity had been studied for the weakly irreversible phosphorylation cycle (22) in [34]. Based on that numerical observation, it was conjectured that the region is path connected.

In [39, Algorithm 2.5], the authors gave an algorithm that certifies that the parameter region of multistationarity is path connected for reaction networks satisfying some technical conditions. The algorithm is based on the following result.

Proposition 4.2. *[39, Theorem 2.4] For a conservative reaction network without relevant boundary steady states, there exists a polynomial*

$$q\colon \mathbb{R}^{n+\ell}_{>0} \to \mathbb{R}$$

such that if q satisfies the closure property and $q^{-1}(\mathbb{R}_{<0})$ is path connected, then the parameter region of multistationarity is path connected.

In [10, 39], it was described how to check whether the network is conservative, does not have any relevant boundary states and how to compute the polynomial q. To check if q satisfies the closure property and $q^{-1}(\mathbb{R}_{<0})$ is path connected, the authors in [39] used strict separating hyperplanes [19, Theorem 3.6]. Using this strategy, it was verified for a large number of reaction networks that the parameter region of multistationarity is path connected.

The weakly irreversible phosphorylation cycle (22) satisfies the technical conditions of Proposition 4.2, i.e. it is conservative and does not have any relevant boundary steady states. The polynomial q from Proposition 4.2 associated with the network (22) is a large polynomial with 1248 monomials.

The method in [39, Algorithm 2.5] was inconclusive, because $\sigma(q)$ does not have a strict separating hyperplane. In the following, we go through Algorithm 1 step-by-step and show that q satisfies the closure property and $q^{-1}(\mathbb{R}_{<0})$ is path connected, which implies by Proposition 4.2 that the parameter region of multistationarity for the weakly irreversible phosphorylation cycle (22) is path connected. The computations were done using `OSCAR` [12, 35] and `Polymake` [23]. The code can be found at the Github repository [37].

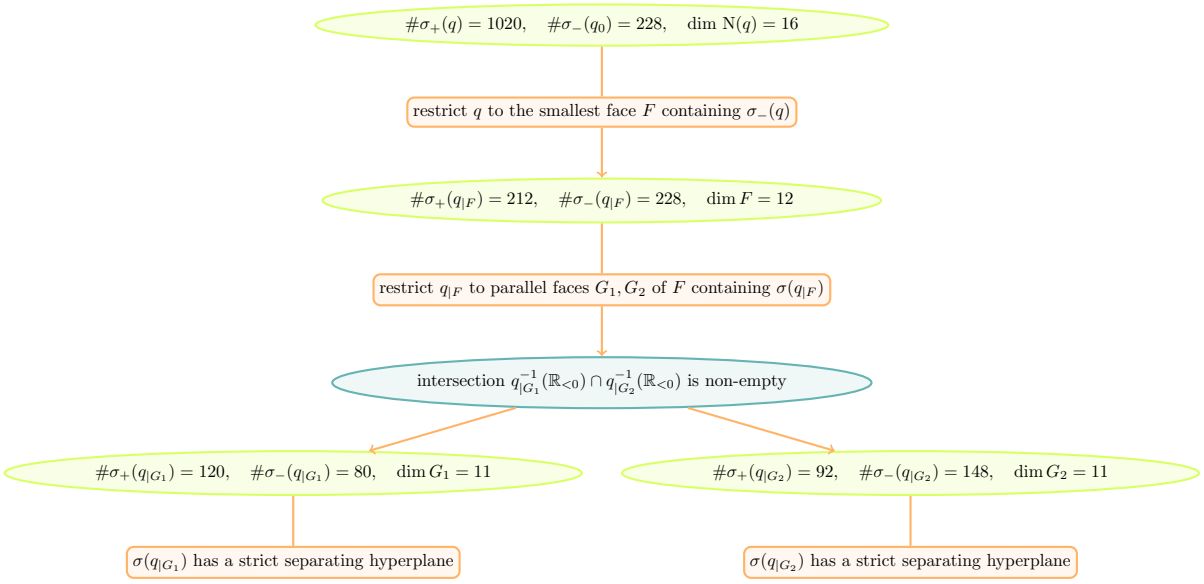


FIGURE 9. Steps taken by Algorithm 1 with input polynomial q from Proposition 4.2 associated with the weakly irreversible phosphorylation cycle (22).

The Newton polytope of q has dimension 16, 1020 out of its 1248 monomials are positive, and 228 are negative. The smallest face $F \subseteq \mathrm{N}(q)$ that contains $\sigma_-(q)$ has dimension 12 and contains 212 positive exponent vectors of q. Using Proposition 3.10, we conclude that q satisfies the closure property.

Following Algorithm 1, we study the restricted polynomial $q_{|F}$. The Newton polytope $\mathrm{N}(q_{|F})$ has 37 facets. We choose the first facet $G_1 = \mathrm{N}(q)_v$ that is provided by `Polymake`. For the face $G_2 = \mathrm{N}(f)_{-v}$ it holds that $\sigma(q_{|F}) \subseteq G_1 \cup G_2$. Furthermore, there exists a pair of negative exponent vectors β_1, β_2 such that $\beta_1 \in G_1$ and $\beta_2 \in G_2$ and $\mathrm{Conv}(\beta_1, \beta_2)$ is an edge of $\mathrm{N}(q_{|F})$. Thus, by Theorem 3.6, it is enough to show that $q_{|G_1}^{-1}(\mathbb{R}_{<0})$ and $q_{|G_2}^{-1}(\mathbb{R}_{<0})$ are connected. This holds by Theorem 2.3(i), since the support of $q_{|G_1}$ and $q_{|G_2}$ have strict separating hyperplanes. For an overview of the steps taken by Algorithm 1, we refer to Figure 9.

Theorem 4.3. *The parameter region of multistationarity of the weakly irreversible phosphorylation system with two binding sites* (22) *is path connected.*

4.2. **Strongly irreversible phosphorylation cycles.** A well-studied family of reaction networks is the family of m-site phosphorylation cycles, where it is assumed that each phosphorylation step follows the Michaelis-Menten mechanism. This is also referred to as *strong irreversibility* in [41]. For a fixed $m \in \mathbb{N}$, the m-site phosphorylation cycle is given by the reactions:

$$\mathrm{S}_i + \mathrm{E} \underset{\kappa_{6i+2}}{\overset{\kappa_{6i+1}}{\rightleftarrows}} \mathrm{ES}_i \xrightarrow{\kappa_{6i+3}} \mathrm{S}_{i+1} + \mathrm{E}, \quad \mathrm{S}_{i+1} + \mathrm{F} \underset{\kappa_{6i+5}}{\overset{\kappa_{6i+4}}{\rightleftarrows}} \mathrm{FS}_{i+1} \xrightarrow{\kappa_{6i+6}} \mathrm{S}_{i+6} + \mathrm{F}, \quad i = 0, \dots, n-1.$$

Questions about the existence of multistationarity and the number of steady states are well understood [18, 20, 24, 27, 42] It is known that the projection of the parameter region of multistationarity onto the space of reaction rate constants (the κ's) is path connected for all $m \geq 2$ [17]. In [39], it was shown that the full parameter region (the region in the κ's and c's) of multistationarity is path connected for $m = 2, 3$, which led to the conjecture that this might be true for every $m \geq 2$.

Let q_m denote the polynomial from Proposition 4.2 associated with the m-site phosphorylation cycle. Algorithm 1 can be successfully applied to show that $q_m^{-1}(\mathbb{R}_{<0})$ is path connected for $m = 4, 5, 6, 7$. The computation can be found in the accompanying `Jupyter` notebook [37]. Since

in these cases, the negative monomials are contained in a proper face of $\mathrm{N}(q_m)$, the polynomial q_m satisfies the closure property by Proposition 3.10. We conclude that the parameter region of multistationarity for the m-site phosphorylation cycle is path connected for $m = 4, 5, 6, 7$.

The steps taken by Algorithm 1 are similar for each $m = 4, 5, 6, 7$, thus Algorithm 1 might provide a strategy to prove the conjecture about the connectivity of the parameter region of multistationarity for all $m \geq 2$. This will be explored in future work.

Acknowledgments

The author thanks Elisenda Feliu and Nidhi Kaihnsa for useful discussions and comments on the manuscript. Funded by the European Union under the Grant Agreement no. 101044561, POSALG. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or European Research Council (ERC). Neither the European Union nor ERC can be held responsible for them.

References

[1] M. Avendaño. The number of roots of a lacunary bivariate polynomial on a line. *J. Symb. Comput.*, 44(9):1280–1284, 2009. Effective Methods in Algebraic Geometry.

[2] F. Bihan. Polynomial systems supported on circuits and dessins d'enfants. *J. Lond. Math. Soc. (2)*, 75(1):116–132, 2007.

[3] F. Bihan and A. Dickenstein. Descartes' rule of signs for polynomial systems supported on circuits. *Int. Math. Res. Notices.*, 39(22):6867–6893, 2017.

[4] F. Bihan, A. Dickenstein, and J. Forsgård. Optimal Descartes' rule of signs for systems supported on circuits. *Math. Ann.*, 2021.

[5] F. Bihan and B. El Hilany. A sharp bound on the number of real intersection points of a sparse plane curve with a line. *J. Symbolic Comput.*, 81:88–96, 2017.

[6] F. Bihan, T. Humbert, and Tavenas S. New bounds for the number of connected components of fewnomial hypersurfaces. *arXiv*, 2208.04590, 2022.

[7] F. Bihan and F. Sottile. New fewnomial upper bounds from Gale dual polynomial systems. *Mosc. Math. J.*, 7(3):387–407, 573, 2007.

[8] F. Bihan and F. Sottile. Betti number bounds for fewnomial hypersurfaces via stratified morse theory. *Proc. Am. Math. Soc.*, 137(9):2825–2833, 2009.

[9] F. Bihan and F. Sottile. Fewnomial bounds for completely mixed polynomial systems. *Adv. Geom.*, 11(3):541–556, 2011.

[10] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *PLoS Comput. Biol.*, 13(10):e1005751, 2017.

[11] D. R. Curtiss. Recent extentions of Descartes' rule of signs. *Ann. Math.*, 19(4):251–278, 1918.

[12] W. Decker, C. Eder, C. Fieker, M. Horn, and M. Joswig, editors. *The OSCAR book*. 2024.

[13] A. Dickenstein. Algebraic geometry tools in systems biology. *Not. Am. Math. Soc.*, 67:1, 12 2020.

[14] R. J. Duffin and E. L. Peterson. Geometric programming with signomials. *J. Optimiz. Theory. App.*, 11(1):3–35, 1973.

[15] C. H. Edwards. *Advanced Calculus of Several Variables*. Academic Press, 1973.

[16] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürück. The kinetic space of multistationarity in dual phosphorylation. *J. Dyn. Differ. Equ.*, 34:825–852, 2022.

[17] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. Parameter region for multistationarity in $\boldsymbol{n}-$site phosphorylation networks. *SIAM J. Appl. Dyn. Syst.*, 22(3):2024–2053, 2023.

[18] E. Feliu, A. D. Rendall, and C. Wiuf. A proof of unlimited multistability for phosphorylation cycles. *Nonlinearity*, 33(11):5629–5658, 2020.

[19] E. Feliu and M. L. Telek. On generalizing Descartes' rule of signs to hypersurfaces. *Adv. Math.*, 408(A), 2022.

[20] D. Flockerzi, K. Holstein, and C. Conradi. N-site Phosphorylation Systems with 2N-1 Steady States. *Bull. Math. Biol.*, 76(8):1892–1916, 2014.

[21] J. Forsgard, M. Nisse, and J. M Rojas. New subexponential fewnomial hypersurface bounds. *arXiv*, 1710.00481, 2017.

[22] C. F. Gauß. Beweis eines algebraischen lehrsatzes. *J. Reine. Angew. Math.*, 3:1–4, 1828.

[23] E. Gawrilow and M. Joswig. *polymake: a Framework for Analyzing Convex Polytope*, pages 43–73. Birkhäuser Basel, Basel, 2000".

[24] M. Giaroli, R. Rischter, M. Pérez Millán, and A. Dickenstein. Parameter regions that give rise to 2[n/2] +1 positive steady states in the n-site phosphorylation system. *Math. Biosci. Eng.*, 16(6):7589–7615, 2019.
[25] D. J. Grabiner. Descartes' rule of signs: Another construction. *Am. Math. Mon.*, 106(9):854–856, 1999.
[26] B. Grünbaum, V. Kaibel, V. Klee, and G. M. Ziegler. *Convex Polytopes*. Graduate Texts in Mathematics. Springer, 2003.
[27] J. Gunawardena. Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. *Biophys. J.*, 93:3828–3834, 2007.
[28] M. Henk, J. Richter-Gebert, and G. M. Ziegler. Basic properties of convex polytopes. In *Handbook of Discrete and Computational Geometry, 3rd Ed.*, 2017.
[29] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Springer, 01 2013.
[30] A. G. Khovanskii. *Fewnomials*, volume 88 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1991.
[31] P. Koiran, N. Portier, and S. Tavenas. A wronskian approach to the real τ-conjecture. *J. Symb. Comput.*, 68:195–214, 2015. Effective Methods in Algebraic Geometry.
[32] T.-Y. Li, J. M. Rojas, and X. Wang. Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces. *Discrete Comput. Geom.*, 30(3):379–414, 2003.
[33] Maplesoft, a division of Waterloo Maple Inc. `Maple`. https://www.maplesoft.com, 2021.
[34] K. M. Nam, B. M. Gyori, S. V. Amethyst, D. J. Bates, and J. Gunawardena. Robustness and parameter geography in post-translational modification systems. *Plos. Comput. Biol.*, 16(5):1–50, 05 2020.
[35] OSCAR – Open Source Computer Algebra Research system, version 0.12.2-dev, 2023.
[36] I. Sahidul and A. M. Wasim. *Fuzzy Geometric Programming Techniques and Applications*. Springer, 2019.
[37] M. L. Telek. CheckConnectivityRecursive, version 1.0.0. available online at https://doi.org/10.5281/zenodo.8419983, 2023.
[38] M. L. Telek. Real tropicalization and negative faces of the Newton polytope. *arXiv*, 2306.05154, 2023.
[39] M. L. Telek and E. Feliu. Topological descriptors of the parameter region of multistationarity: Deciding upon connectivity. *Plos. Comput. Biol.*, 19(3):1–38, 03 2023.
[40] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.2)*, 2021. `https://www.sagemath.org`.
[41] M. Thomson and J. Gunawardena. Unlimited multistability in multisite phosphorylation systems. *Nature*, 406:274–277, 2009.
[42] L. Wang and E. D. Sontag. On the number of steady states in a multiple futile cycle. *J. Math. Biol.*, 57:29–52, 2008.
[43] G. M. Ziegler. *Lectures on Polytopes*. Springer, 2007.

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: `mlt@math.ku.dk`

V

Real tropicalization and negative faces of the Newton polytope

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Publication details

Published in the Journal of Pure and Applied Algebra 228(6), 2024
DOI: https://doi.org/10.1016/j.jpaa.2023.107564

REAL TROPICALIZATION AND NEGATIVE FACES OF THE NEWTON POLYTOPE

MÁTÉ L. TELEK

ABSTRACT. In this work, we explore the relation between the tropicalization of a real semi-algebraic set $S = \{f_1 < 0, \dots, f_k < 0\}$ defined in the positive orthant and the combinatorial properties of the defining polynomials $f_1, \dots, f_k$. We describe a cone that depends only on the face structure of the Newton polytopes of $f_1, \dots, f_k$ and the signs attained by these polynomials. This cone provides an inner approximation of the real tropicalization, and it coincides with the real tropicalization if $S = \{f < 0\}$ and the polynomial f has generic coefficients. Furthermore, we show that for a maximally sparse polynomial f the real tropicalization of $S = \{f < 0\}$ is determined by the outer normal cones of the Newton polytope of f and the signs of its coefficients. Our arguments are valid also for signomials, that is, polynomials with real exponents defined in the positive orthant.

Keywords: logarithmic limit set, signomial, semi-algebraic set, signed support

1. INTRODUCTION

Real algebraic varieties, or more generally, semi-algebraic sets, play a central role in applications, e.g. in chemical reaction network theory [7] or in robotics [29]. In practice, the polynomials defining these sets involve many variables and many monomials, which makes using standard techniques from real (semi-)algebraic geometry infeasible.

Tropicalization is the process of associating a polyhedral complex to an algebraic variety. This can be used to answer questions about the variety, such as computing its dimension [20, Structure Theorem], or to compute Gromov-Witten invariants, that is, to count the number of curves of given degree and genus passing through a given finite number of points in the complex projective plane [21]. Moreover, using tropicalizations one can find the generic number of solutions of a parametric polynomial equation system [12].

One of the roots of tropical geometry goes back to 1971 when Bergman [3] studied the logarithmic limit of an algebraic variety V in the complex torus $(\mathbb{C}^*)^n$. The logarithmic limit of V is defined as the limit in the Hausdorff distance of the images of V under the coordinate-wise logarithm map with base $t > 0$

$$\mathrm{Log}_t \colon (\mathbb{C}^*)^n \mapsto \mathbb{R}^n, \quad z \mapsto (\log_t(|z_1|), \dots, \log_t(|z_n|)$$

as $t \to \infty$. If V is a hypersurface, i.e. it is defined by one polynomial f, the logarithmic limit of V equals the $(n-1)$-skeleton of the outer normal fan of the Newton polytope of f [22, Corollary 6.4].

The logarithmic limit of a real semi-algebraic set $S \subseteq \mathbb{R}^n_{>0}$ was first studied by Alessandrini [1] in 2013. The author showed that the logarithmic limit of S is a closed polyhedral complex of dimension at most the dimension of S [1, Theorem 3.11]. In [4, 5], the authors called the logarithmic limit of S the *real tropicalization* of S. In this manuscript, we follow this notation and write

$$\mathrm{Trop}(S) := \lim_{t \to \infty} \mathrm{Log}_t(S).$$

Computing the real tropicalization of a semi-algebraic set $S \subseteq \mathbb{R}^n_{>0}$ is known to be hard. The Fundamental Theorem [17, Theorem 6.9] implies that $\mathrm{Trop}(S)$ is described as an intersection of outer normal cones as follows: For a polynomial $f \in \mathbb{R}[x_1, \dots, x_n]$, we call an exponent vector negative if the coefficient of the corresponding monomial is negative. Similarly, a face F of the Newton polytope $\mathrm{NP}(f)$ is a *negative face* if F contains a negative exponent vector. The union

of the outer normal cones of $\mathrm{NP}(f)$ that correspond to negative faces is denoted by $\mathcal{N}_f^-$. Using this notation, we have:

$$\mathrm{Trop}(S) = \bigcap_{f\leq 0 \text{ on } S} \mathcal{N}_f^-, \tag{1}$$

where the intersection is taken over all polynomials $f \in \mathbb{R}[x_1, \dots, x_n]$ which are nonpositive on S.

In [17, Theorem 6.9], it was also shown that the intersection in (1) can be taken over finitely many such polynomials, but to the best of our knowledge, no algorithm is known to find such a finite set of polynomials. However, in certain cases, there exist good approximations of the real tropicalization. From [17, Proposition 6.12], it follows that the real tropicalization of a semi-algebraic set S of the form $S = \bigcap_{i=1}^k f_i^{-1}(\mathbb{R}_{<0})$ satisfies

$$\mathrm{int}\big(\bigcap_{i=1}^k \mathcal{N}_{f_i}^-\big) \subseteq \mathrm{Trop}(S) \subseteq \bigcap_{i=1}^k \mathcal{N}_{f_i}^-. \tag{2}$$

If the set on the right-hand side in (2) is *regular* (it is equal to the closure of its interior), then the real tropicalization of S equals the intersection of the cones $\mathcal{N}_{f_i}^-, i = 1, \dots, k$ [2, Corollary 4.8]. In Proposition 3.2, we characterize when regularity happens. As a consequence of this, we show that if S is defined by one polynomial f, i.e. $S = f^{-1}(\mathbb{R}_{<0})$, then $\mathcal{N}_f^-$ is regular if and only if the faces of $\mathrm{NP}(f)$ that contain a negative exponent vector have also a negative vertex. This seems to be a quite restrictive assumption, but it is automatically satisfied for polynomials whose set of exponent vectors equals the set of vertices of the Newton polytope. Such polynomials are called *maximally sparse* in the literature [23]. Thus, if f is maximally sparse, then the real tropicalization of $S = f^{-1}(\mathbb{R}_{<0})$ is easy to compute as it equals $\mathcal{N}_f^-$ (Corollary 3.5).

Our main result is Theorem 3.3, where we give an elementary proof of the inclusions in (2), while we also refine them. We associate a cone $\Sigma(f_1, \dots, f_k)$, called the *actual negative normal cone* of $f_1, \dots, f_k$ (see Section 2) that provides a better inner approximation of $\mathrm{Trop}(S)$:

$$\mathrm{int}\big(\bigcap_{i=1}^k \mathcal{N}_{f_i}^-\big) \subseteq \Sigma(f_1, \dots, f_k) \subseteq \mathrm{Trop}(S) \subseteq \bigcap_{i=1}^k \mathcal{N}_{f_i}^-. \tag{3}$$

The proof of Theorem 3.3 is valid also in the case where the defining polynomials of the semi-algebraic set have real exponents.

Theorem 3.3 has two main consequences. In the case when S is defined by one polynomial, i.e. $S = f^{-1}(\mathbb{R}_{<0})$, we show that $\Sigma(f)$ equals the real tropicalization of S when the coefficients of f are generic (Corollary 3.10).

Specifically, to compute $\Sigma(f)$, one needs to determine all faces of the Newton polytope of f for which the restriction of f to the face takes negative values over $\mathbb{R}^n_{>0}$. Deciding if a polynomial is nonnegative is an NP-hard problem [19]. To address this difficulty, several sufficient criteria implying nonnegativity have been introduced. For instance, one approach, going back to Hilbert [15], is to express the polynomial as a sum of squares (SOS); this decomposition can be found by semidefinite programming, see [19] and the references therein. Another approach is to write the polynomial as sums of nonnegative circuits (SONC) [8, 16], see also [6, 25].

The set of all nonnegative polynomials is a closed convex cone, called the nonnegativity cone. The sets of SOS and SONC polynomials are also convex cones, clearly contained in the nonnegativity cone. The genericity assumption in Corollary 3.10 is that neither f nor its restrictions to the faces of the Newton polytope lie on the boundary of the nonnegativity cone. To the best of our knowledge, there are no complete descriptions of the boundary neither of the nonnegativity cone nor of the SOS cone. On the contrary, the boundary of the SONC cone has been characterized in [11]. This characterization together with Corollary 3.10 provide a method to compute the real tropicalization of semi-algebraic sets defined by one polynomial.

In [14], the authors introduced the DSONC cone, which is a cone contained in the interior of the SONC cone. A remarkable property of the DSONC cone is that it is possible to check

membership using linear programming. This might give a more efficient method to compute $\Sigma(f)$ and to verify the genericity condition in Corollary 3.10.

Finally, an additional consequence of Theorem 3.3, which is an interesting result on its own, is that any logarithmic image of $f^{-1}(\mathbb{R}_{<0})$ is a bounded set if the boundary of $\mathrm{NP}(f)$ does not contain any negative exponent vector (Corollary 3.6).

Notation. $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$ and $\mathbb{R}_{<0}$ refer to the sets of nonnegative, positive and negative real numbers respectively. For two vectors $v, w \in \mathbb{R}^n$, $v \cdot w$ denotes the Euclidean scalar product, and $v * w$ denotes the coordinate-wise product of v and w. We denote the interior of a set $X \subseteq \mathbb{R}^n$ by $\mathrm{int}(X)$. If $X \subseteq \mathbb{R}^n$ is a polyhedron, $\mathrm{relint}(X)$ denotes the relative interior of X. The symbol $\#S$ denotes the cardinality of a finite set S.

2. Negative faces of the Newton polytope

Following [9, 27], a *signomial* is a function

$$f\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \qquad x \mapsto f(x) = \sum_{\mu \in \sigma(f)} c_\mu x^\mu,$$

where $\sigma(f) \subseteq \mathbb{R}^n$ is a finite set, called the *support* of f, and the *coefficients* $c_\mu \in \mathbb{R}$ are non-zero. We use the notation x^μ for $x_1^{\mu_1} \cdots x_n^{\mu_n}$. Thus, a signomial is a polynomial with real exponents whose domain is restricted to the positive orthant. For a set $S \subseteq \mathbb{R}^n$, we define the *restriction* of f to S as

$$f_{|S}(x) := \sum_{\mu \in \sigma(f) \cap S} c_\mu x^\mu.$$

Furthermore, we divide the support of f into the set of *positive and negative exponent vectors*:

$$\sigma_+(f) := \{\mu \in \sigma(f) \mid c_\mu > 0\}, \qquad \sigma_-(f) := \{\mu \in \sigma(f) \mid c_\mu < 0\}.$$

Thus, a positive exponent vector μ is an exponent vector of a monomial x^μ of f such that the corresponding coefficient c_μ is positive.

The *Newton polytope* of f is the convex hull of the support

$$\mathrm{NP}(f) := \mathrm{Conv}\big(\sigma(f)\big).$$

We refer to Figure 1(a) for an illustration of the Newton polytope of $f = x_1^2 - x_1 + 1 - x_2^2$. The set of negative exponent vectors is $\sigma_-(f) = \{(1,0),(0,2)\}$ and the set of positive exponent vectors is $\sigma_+(f) = \{(0,0),(2,0)\}$.

We recall some basic notions in polyhedral geometry that are relevant to the study of real tropicalizations (see [13, 18, 30] for details). Each vector $v \in \mathbb{R}^n$ cuts out the *face* of $\mathrm{NP}(f)$ [13, Section 15.1.1] given by

$$\mathrm{NP}(f)_v := \{\mu \in \mathrm{NP}(f) \mid v \cdot \mu = \max_{\nu \in \mathrm{NP}(f)} v \cdot \nu\}. \tag{4}$$

The vector v is called an *outer normal vector* of $\mathrm{NP}(f)_v$, or normal vector for short. A face is *proper* if it is not the entire polytope.

For a face $F \subseteq \mathrm{NP}(f)$, the set of all vectors v such that $F \subseteq \mathrm{NP}(f)_v$ form a closed convex polyhedral cone, called the *normal cone* of F [13, Section 15.2.2]:

$$\mathcal{N}_f(F) = \{v \in \mathbb{R}^n \mid F \subseteq \mathrm{NP}(f)_v\}. \tag{5}$$

Furthermore, for $v \in \mathcal{N}_f(F)$, it holds:

$$\mathrm{NP}(f)_v = F \quad \Longleftrightarrow \quad v \in \mathrm{relint}\,\mathcal{N}_f(F). \tag{6}$$

The collection of the normal cones of all faces forms a *complete fan* $\mathcal{N}_f$ [30, Example 7.3].

Taking normal cones induces an inclusion reversing correspondence between faces and their normal cones, which behaves well with dimensions. Hence, for any two non-empty faces $F, G \subseteq \mathrm{NP}(f)$, it holds that:

$$F \subseteq G \quad \Longleftrightarrow \quad \mathcal{N}_f(F) \supseteq \mathcal{N}_f(G), \tag{7}$$

$$\dim F = n - \dim \mathcal{N}_f(F). \tag{8}$$

To study several signomials $f_1, \dots, f_k$ at the same time, it will be beneficial to consider the *common refinement* of the normal fans $\mathcal{F}_{f_1}, \dots, \mathcal{F}_{f_k}$[30, Definition 7.6], which is the fan defined as

$$\bigwedge_{i=1}^{k} \mathcal{N}_{f_i} := \Big\{ \bigcap_{i=1}^{k} C_i \mid C_i \in \mathcal{N}_{f_i} \Big\}. \tag{9}$$

It is easy to see from (9) that full-dimensional cones in the common refinement are intersections of full-dimensional cones. Note that for every pair of vectors v, v' in the relative interior of a cone in the common refinement, it holds that

$$\mathrm{NP}(f_i)_v = \mathrm{NP}(f_i)_{v'} \qquad \text{for all } i = 1, \dots, k. \tag{10}$$

Several arguments in this work rely on the following easy observation. Each $v \in \mathbb{R}^n$ and $x \in \mathbb{R}^n_{>0}$ induce a signomial in one variable

$$\mathbb{R}_{>0} \to \mathbb{R}, \qquad t \mapsto f(t^v * x) = \sum_{\mu \in \sigma(f)} c_\mu x^\mu t^{v \cdot \mu}, \tag{11}$$

where t^v is short for the vector $(t^{v_1}, \dots, t^{v_n})$. We call a coefficient of a univariate signomial the *leading coefficient* (LC) if the exponent of the accompanying monomial is the largest. From (4) it follows that for fixed $x \in \mathbb{R}^n_{>0}$

$$\mathrm{LC}(f(t^v * x)) = \sum_{\mu \in \sigma(f) \cap \mathrm{NP}(f)_v} c_\mu x^\mu,$$

if we view $f(t^v * x)$ as a univariate signomial in the variable t. This implies the following well-known fact (see e.g. [10, Proposition 2.3])

Lemma 2.1. *Let $f \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ be a signomial and $v \in \mathbb{R}^n$. If $f_{|N(f)_v}(x) < 0$ for some $x \in \mathbb{R}^n_{>0}$, then there exists $T \in \mathbb{R}_{>0}$ such that for all $t > T$:*

$$f(t^v * x) < 0.$$

To find all vectors $v \in \mathbb{R}^n$ for which Lemma 2.1 applies, we have to decide for which faces $F \subseteq \mathrm{NP}(f)$, the signomial $f_{|F}$ takes negative values. A necessary condition for $f_{|F}$ to take negative values is that the face F contains negative exponent vectors of f. Motivated by this observation, we call a face $F \subseteq \mathrm{NP}(f)$ a *negative face* if $F \cap \sigma_-(f) \neq \emptyset$. Furthermore, we define the *negative normal cone* of $\mathrm{NP}(f)$ as the union of the normal cones corresponding to the negative faces:

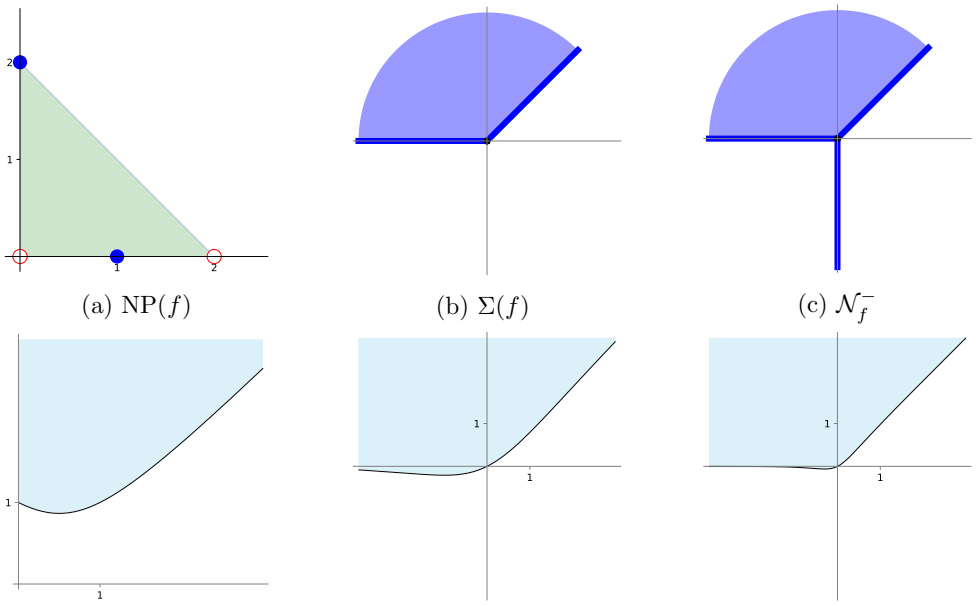
$$\mathcal{N}_f^- = \bigcup_{\substack{F \subseteq \mathrm{NP}(f) \\ \text{negative face}}} \mathcal{N}_f(F) = \{v \in \mathbb{R}^n \mid \mathrm{NP}(f)_v \cap \sigma_-(f) \neq \emptyset\}.$$

Note that $\mathcal{N}_f^-$ is a cone, i.e. it is closed under multiplication by positive scalars, but $\mathcal{N}_f^-$ does not need to be convex.

Even if the face F contains negative exponent vectors of f there is no guarantee that $f_{|F}$ actually takes negative values. For an example, take the signomial $f = x_1^2 - x_1 + 1 - x_2^2$ from Figure 1 and the face $F = \mathrm{Conv}\big((0,0),(2,0)\big)$.

For a signomial f, we define the *actual negative normal cone* as

$$\Sigma(f) := \Big\{ v \in \mathbb{R}^n \mid f^{-1}_{|\mathrm{NP}(f)_v}(\mathbb{R}_{<0}) \neq \emptyset \Big\} \subseteq \mathcal{N}_f^-.$$



(a) NP(f) (b) $\Sigma(f)$ (c) $\mathcal{N}_f^-$

(d) $f^{-1}(\mathbb{R}_{<0})$ (e) $\mathrm{Log}_2\left(f^{-1}(\mathbb{R}_{<0})\right)$ (f) $\mathrm{Log}_{10}\left(f^{-1}(\mathbb{R}_{<0})\right)$

FIGURE 1. (a) Newton polytope of $f = x_1^2 - x_1 + 1 - x_2^2$. Blue dots correspond to negative exponents vectors, red circles correspond to positive exponent vectors. (b),(c) actual negative normal cone and negative normal cone of f, (d) semi-algebraic set defined by f, (e),(f) logarithmic images of $f^{-1}(\mathbb{R}_{<0})$.

In Proposition 2.2, we show that $\Sigma(f)$ is the union of normal cones corresponding to faces $F \subseteq \mathrm{NP}(f)$ such that $f_{|F}$ takes negative values. The actual negative normal cone and the negative normal cone of $f = x_1^2 - x_1 + 1 - x_2^2$ are depicted in Figure 1(b)(c).

Define the *actual negative normal cone* of a collection of signomials $f_1, \ldots, f_k$ as:

$$\Sigma(f_1, \ldots, f_k) := \left\{ v \in \mathbb{R}^n \mid \bigcap_{i=1}^{k} f^{-1}_{i|\,\mathrm{NP}(f_i)_v}(\mathbb{R}_{<0}) \neq \emptyset \right\}. \tag{12}$$

In the proof of Theorem 3.3, we need that the negative normal cones $\mathcal{N}_{f_i}^-, i = 1, \ldots, k$ are closed subsets of $\mathbb{R}^n$. This property follows easily from the fact that a polytope has finitely many faces [18, Theorem 3.46] and the cones $\mathcal{N}_{f_i}(F)$, F is a face of $\mathrm{NP}(f_i)$, are closed. The actual negative normal cone $\Sigma(f_1, \ldots, f_k)$ is also closed, as the following argument shows.

Proposition 2.2. *Let $f_1, \ldots, f_k$ be signomials on $\mathbb{R}^n_{>0}$, let C be a cone in the common refinement $\bigwedge_{i=1}^k \mathcal{N}_{f_i}$ and $v \in \mathrm{relint}(C)$. If $v \in \Sigma(f_1, \ldots, f_k)$, then $C \subseteq \Sigma(f_1, \ldots, f_k)$. In particular, $\Sigma(f_1, \ldots, f_k)$ is a closed subset of $\mathbb{R}^n$.*

Proof. Let $C_i \in \mathcal{N}_{f_i}$, $i = 1, \ldots, k$ such that $C = \cap_{i=1}^k C_i$. Since $v \in \mathrm{relint}(C)$, from [26, Theorem 6.5.] follows that $v \in \mathrm{relint}(C_i)$ for all $i = 1, \ldots, k$.

Let $F_i := \mathrm{NP}(f_i)_v, i = 1, \ldots, k$. Since $v \in \mathrm{relint}(C)$, for any $w \in C$ we have that

$$F_i = \left(\mathrm{NP}(f_i)_w\right)_v \subseteq \mathrm{NP}(f_i)_w$$

for all $i = 1, \ldots, k$ by (5) and (6).

Consider $x \in \bigcap_{i=1}^k f^{-1}_{i|F_i}(\mathbb{R}_{<0})$, which exists since $v \in \Sigma(f_1, \ldots, f_k)$. By Lemma 2.1, there exists $T > 0$ such that

$$t^v * x \in \bigcap_{i=1}^{k} f^{-1}_{i|\,\mathrm{NP}(f_i)_w}(\mathbb{R}_{<0}), \qquad \text{for } t > T,$$

which implies that $w \in \Sigma(f_1, \dots, f_k)$. Hence, $C \subseteq \Sigma(f_1, \dots, f_k)$.

This argument shows that $\Sigma(f_1, \dots, f_k)$ is the union of the cones in $\bigwedge_{i=1}^{k} \mathcal{N}_{f_i}$, whose relative interior intersects $\Sigma(f_1, \dots, f_k)$. Since there exist only finitely many cones in $\bigwedge_{i=1}^{k} \mathcal{N}_{f_i}$ and they are closed, it follows that $\Sigma(f_1, \dots, f_k)$ is a closed set. □

Proposition 2.2 ensures that to compute the actual negative normal cone $\Sigma(f_1, \dots, f_k)$, it is enough to check for finitely many $v \in \mathbb{R}^n$ whether

$$\bigcap_{i=1}^{k} f_{i|\operatorname{NP}(f_i)_v}^{-1}(\mathbb{R}_{<0}) \neq \emptyset. \tag{13}$$

To be precise, for each cone $C \in \bigwedge_{i=1}^{k} \mathcal{N}_{f_i}$, pick $v \in \operatorname{relint}(C)$ and check whether (13) is empty. This observation provides a way to compute $\Sigma(f_1, \dots, f_k)$. Note that, computing the negative normal cone $\mathcal{N}_{f_i}^{-}$ is significantly simpler and in some cases still gives enough information about the real tropicalization (Theorem 3.3, Corollary 3.5).

We conclude the section with a bound on the roots of a univariate signomial that will be used in the proofs of Theorem 3.3 and Theorem 3.9. The statement is a generalization of Cauchy's bound [24, Theorem 8.1.7] to signomials. Although the idea of the proof is almost the same as in the polynomial case, we recall it for the sake of completeness.

Lemma 2.3. *Let $g\colon \mathbb{R}_{>0} \to \mathbb{R}$, $g(t) := \sum_{i=1}^{d} a_i t^{\nu_i}$ be a univariate signomial such that $\nu_1 < \dots < \nu_d$. If there exist $p \in \{1, \dots, d\}$ and $\epsilon > 0$ such that $(a_d - \epsilon)t^{\nu_d} + \sum_{i=p}^{d-1} a_i t^{\nu_i} > 0$ for all $t > 0$, then for all $0 < \delta \leq \nu_p - \nu_{p-1}$ it holds:*

$$\max\{t_0 \in \mathbb{R}_{>0} \mid g(t_0) = 0\} \leq \max\left\{1, \left(\tfrac{1}{\epsilon}\sum_{i=1}^{p-1}|a_i|\right)^{\frac{1}{\delta}}\right\}.$$

Proof. Let $t_0 > 1$ be such that $g(t_0) = 0$. Then,

$$\epsilon\, t_0^{\nu_d} < \sum_{i=p}^{d} a_i t_0^{\nu_i} = -\sum_{i=1}^{p-1} a_i t_0^{\nu_i} \leq \sum_{i=1}^{p-1} |a_i| t_0^{\nu_i} \leq \sum_{i=1}^{p-1} |a_i| t_0^{\nu_d - \delta}.$$

The first inequality follows from $(a_d - \epsilon)t_0^{\nu_d} + \sum_{i=p}^{d-1} a_i t_0^{\nu_i} > 0$. The last inequality follows from $\nu_d - \delta \geq \nu_i$ for all $i = 1, \dots, p-1$. Dividing both sides by $\epsilon\, t_0^{\nu_d - \delta}$, the statement follows. □

Remark 2.4. In Section 3, Lemma 2.3 will be used to ensure that for a signomial f and a convergent sequence $w(m) \to w$ in $\mathbb{R}^n$, the roots of $f(t^{w(m)})$ converge to the roots of $f(t^w)$.

This statement might fail for some signomials not satisfying the assumptions of Lemma 2.3 as the following example shows. Consider the signomial $f(x_1, x_2) = x_1^2 - x_1 + 1 - x_2^2$ from Figure 1 and the sequence $w(m) = (1, 1 - \frac{1}{m}) \to w = (1, 1)$. The induced functions are given by:

$$f(t^w) = -t + 1, \qquad f(t^{w(m)}) = t^2 - t^{2-\frac{2}{m}} - t + 1.$$

The function $f(t^w)$ has a unique root $t = 1$. The function $f(t^{w(m)})$ has two positive real roots, one of which converges to 1 and the other goes to infinity as $m \to \infty$.

3. Real tropicalization

The goal of this section is to relate the real tropicalization of the set

$$S(f_1, \dots, f_k) := \bigcap_{i=1}^{k} f_i^{-1}(\mathbb{R}_{<0}), \tag{14}$$

where $f_1, \dots, f_k$ are signomials on $\mathbb{R}^n_{>0}$, to the negative normal cones $\mathcal{N}_{f_1}^{-}, \dots, \mathcal{N}_{f_k}^{-}$, and to the actual negative normal cone $\Sigma(f_1, \dots, f_k)$. In the case that (14) is described by a single polynomial f, we show that the real tropicalization and the actual negative normal cone coincide

for generic coefficients of f. Furthermore, $\mathrm{Trop}(S(f))$ equals the negative normal cone of f if f is maximally sparse (all exponent vectors of f are vertices of the Newton polytope).

As indicated in the Introduction, following [4, 5], the *real tropicalization* of $S \subseteq \mathbb{R}^n_{>0}$ is defined to be

$$\mathrm{Trop}(S) := \lim_{t\to\infty} \mathrm{Log}_t(S) = \lim_{t\to\infty} \tfrac{1}{\log_e(t)} \mathrm{Log}_e(S), \tag{15}$$

where Log_t denotes the point-wise logarithm with base t, and e is Euler's number. This construction is also called the *logarithmic limit set* of S. For a precise definition and for some basic properties, we refer to [1, Section 2]. Since it will be used frequently, we recall the following results from [1]:

Proposition 3.1. *[1, Proposition 2.2] For any set $S \subseteq \mathbb{R}^n_{>0}$, it holds:*

(i) $\mathrm{Trop}(S)$ *is a closed subset of* $\mathbb{R}^n$.
(ii) *For* $w \in \mathbb{R}^n$, $w \in \mathrm{Trop}(S)$ *if and only if there exist sequences* $\big\{x(m)\big\}_{m\in\mathbb{N}} \subseteq S$ *and* $\big\{t(m)\big\}_{m\in\mathbb{N}} \subseteq (1,\infty)$ *such that* $t(m) \to \infty$ *and* $\mathrm{Log}_{t(m)}(x(m)) \to w$.
(iii) $\mathrm{Trop}(S) \subset \{0\}$ *if and only if* $\mathrm{Log}_t(S)$ *is bounded for all* $t > 0$.

As discussed in the Introduction, if $f_1, \ldots, f_k$ are polynomials, that is their exponents are nonnegative integer vectors, it is known that

$$\mathrm{Trop}\left(S(f_1,\ldots,f_k)\right) = \bigcap_{i=1}^{k} \mathcal{N}^-_{f_i} \tag{16}$$

if the set on the right is a regular set (it is equal to the closure of its interior) [2, Corollary 4.8]. The following proposition characterizes the cases when the intersection of the negative normal cones in (16) is regular.

Proposition 3.2. *Let $f_1, \ldots, f_k$ be signomials on $\mathbb{R}^n_{>0}$. A point $w \in \bigcap_{i=0}^k \mathcal{N}^-_{f_i}$ is in the closure of $\mathrm{int}\left(\bigcap_{i=0}^k \mathcal{N}^-_{f_i}\right)$ if and only if there exists $v \in \mathbb{R}^n$ such that $\mathrm{NP}(f_i)_v$ is a negative vertex of $\mathrm{NP}(f_i)_w$ for all $i = 1, \ldots, k$.*

Proof. First, we show the if part. For each $i = 1, \ldots, k$, let $\beta_i := \mathrm{NP}(f_i)_v$ be the negative vertex of $\mathrm{NP}(f_i)_w$. It follows that w is contained in

$$\bigcap_{i=1}^{k} \mathcal{N}_{f_i}(\beta_i). \tag{17}$$

In the following, we show that the interior of (17) is non-empty. As vertices correspond to full-dimensional normal cones by (8), the interior and the relative interior of $\mathcal{N}_{f_i}(\beta_i)$ coincide. By (6), $v \in \mathrm{relint}(\mathcal{N}_{f_i}(\beta_i)) = \mathrm{int}(\mathcal{N}_{f_i}(\beta_i))$, which implies that there exists an $\epsilon > 0$ such that the ball with radius ϵ and center v is contained in the interior of $\mathcal{N}_{f_i}(\beta_i)$ for all $i = 1, \ldots, k$. Thus, the interior of (17) is non-empty.

As the finite intersection of interiors equals the interior of the intersection [28, Section 1.1], we have

$$\mathrm{int}\Big(\bigcap_{i=1}^{k} \mathcal{N}_{f_i}(\beta_i)\Big) = \bigcap_{i=1}^{k} \mathrm{int}\left(\mathcal{N}_{f_i}(\beta_i)\right) \subseteq \bigcap_{i=1}^{k} \mathrm{int}\left(\mathcal{N}^-_{f_i}\right) = \mathrm{int}\Big(\bigcap_{i=1}^{k} \mathcal{N}^-_{f_i}\Big), \tag{18}$$

where the inclusion in the middle holds since β_i is a negative vertex of $\mathrm{NP}(f_i)$ for all $i = 1, \ldots, k$.

Since (17) is a polyhedral cone with non-empty interior, it is regular and w is contained in the closure of the interior of (17). Using (18), one concludes that w is contained in the closure of $\mathrm{int}\left(\bigcap_{i=1}^k \mathcal{N}^-_{f_i}\right)$.

To show the reverse implication, we assume that w is in the closure of $\mathrm{int}\left(\bigcap_{i=0}^k \mathcal{N}^-_{f_i}\right)$. In that case, for every $\epsilon > 0$ we have that $B_\epsilon(w) \cap \mathrm{int}\left(\bigcap_{i=0}^k \mathcal{N}^-_{f_i}\right) \neq \emptyset$, where $B_\epsilon(w)$ denotes the ball

with radius ϵ and center w. We choose ϵ small enough such that $B_\epsilon(w)$ does not intersect the cones of the $(n-1)$-skeleton of $\bigwedge_{i=1}^k \mathcal{N}_{f_i}$ that do not contain w.

The $(n-1)$-skeleton of $\bigwedge_{i=1}^k \mathcal{N}_{f_i}$ is not full-dimensional, so there exists $v \in B_\epsilon(w) \cap \operatorname{int}\left(\bigcap_{i=0}^k \mathcal{N}_{f_i}^-\right)$ which is not contained in the $(n-1)$-skeleton.

Let C be the smallest cone in $\bigwedge_{i=1}^k \mathcal{N}_{f_i}$ that contains v in its interior, and let $C_i \in \mathcal{N}_{f_i}$, $i = 1, \dots, k$ such that $C = \cap_{i=1}^k C_i$. From $v \in \operatorname{int}(C)$ follows that $v \in \operatorname{int}(C_i)$ for all $i = 1, \dots, k$. From (6) and (8) it follows that $\operatorname{NP}(f_i)_v$ is a vertex for all $i = 1, \dots, k$. Moreover, $\operatorname{NP}(f_i)_v$ is a negative vertex, since $v \in \mathcal{N}_{f_i}^-$.

Since $B_\epsilon(w)$ intersects only the cones of the $(n-1)$-skeleton of $\bigwedge_{i=1}^k \mathcal{N}_{f_i}$ that contain w, it follows that w lies in C. Using (7), we conclude that $\operatorname{NP}(f_i)_v \subseteq \operatorname{NP}(f_i)_w$. □

For an example of a signomial whose negative normal cone is not regular, we refer to Figure 1. The negative normal cone of $f(x_1, x_2) = x_1^2 - x_1 + 1 - x_2^2$ has a ray in southern direction, but the closure of $\operatorname{int}(\mathcal{N}_f^-)$ does not contain this ray. One can also use Proposition 3.2 to conclude that $\mathcal{N}_f^-$ is not regular: The face $\operatorname{Conv}\big((0,0),(2,0)\big) \subseteq \operatorname{NP}(f)$ is negative, but does not contain any negative vertex.

In [2, 17], to prove (2) the authors worked with polynomials over the field of real Puiseux series. In the following theorem, we give an elementary proof of (2) that applies to signomials and extend (2) by relating the real tropicalization to the actual negative normal cone as defined in (12).

Theorem 3.3. *For signomials $f_1, \dots, f_k$ on $\mathbb{R}^n_{>0}$, it holds:*

$$\operatorname{int}\Big(\bigcap_{i=1}^k \mathcal{N}_{f_i}^-\Big) \subseteq \Sigma(f_1, \dots, f_k) \subseteq \operatorname{Trop}\big(S(f_1, \dots, f_k)\big) \subseteq \bigcap_{i=1}^k \mathcal{N}_{f_i}^-.$$

Proof. Let $S = S(f_1, \dots, f_k)$. To show the first inclusion, let $w \in \operatorname{int}\left(\bigcap_{i=1}^k \mathcal{N}_{f_i}^-\right)$. By Proposition 3.2, there exists $v \in \mathbb{R}^n$ such that $\operatorname{NP}(f_i)_v$ is a negative vertex of $\operatorname{NP}(f_i)_w$ for all $i = 1, \dots, k$. For any fixed $x \in \mathbb{R}^n_{>0}$, we have

$$f_{i|\operatorname{NP}(f_i)_w}(t^v * x) < 0 \qquad \text{for} \quad t \gg 0$$

by Lemma 2.1. Thus, $w \in \Sigma(f_1, \dots, f_k)$.

For the proof of the second inclusion, let $w \in \Sigma(f_1, \dots, f_k)$, and let $x \in \bigcap_{i=1}^k f_{i|\operatorname{NP}(f_i)_w}^{-1}(\mathbb{R}_{<0})$. By Lemma 2.1, $f_i(t^w * x) < 0$ for all $i = 1, \dots, k$ and $t \gg 0$. Therefore, $t^w * x \in S$ for $t \gg 0$. Choose a sequence $(t(m))_{m\in\mathbb{N}}$ such that $t(m) \to \infty$ and $t(m)^w * x \in S$ for all $m \in \mathbb{N}$. Then

$$\operatorname{Log}_{t(m)}(t(m)^w * x) = \log_{t(m)}(t(m))w + \operatorname{Log}_{t(m)}(x) = w + \tfrac{1}{\log_e(t(m))} \operatorname{Log}_e(x) \to w.$$

Proposition 3.1 implies that $w \in \operatorname{Trop}(S)$.

The third inclusion remains to be shown. For that, let $w \in \operatorname{Trop}(S)$. By Proposition 3.1, there exist sequences $\{x(m)\}_{m\in\mathbb{N}} \subseteq S$ and $\{t(m)\}_{m\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$ such that $t(m) \to \infty$ and $w(m) := \operatorname{Log}_{t(m)}(x(m)) \to w$. Note that with this notation we have

$$t(m)^{w(m)} = x(m), \qquad \text{for all } m \in \mathbb{N}.$$

If $w \notin \bigcap_{i=1}^k \mathcal{N}_{f_i}^-$, then there exists $i \in \{1, \dots, k\}$ such that $w \in \mathbb{R}^n \setminus \mathcal{N}_{f_i}^-$. In the following, we show that it is possible to choose a subsequence $\{\tilde{w}(m)\}_{m\in\mathbb{N}}$ of $\{w(m)\}_{m\in\mathbb{N}}$ such that

$$f_i(t^{\tilde{w}(m)}) > 0 \qquad \text{for all} \quad t > T,$$

where $T > 0$ can be chosen independently from m. This yields to a contradiction, since for large m, $t(m) > T$ and $f_i(t(m)^{\tilde{w}(m)}) < 0$.

Since the relative interiors of the cones in $\mathcal{N}_{f_i}$ form a partition of $\mathbb{R}^n$ and there are only finitely many such cones, there exists a face $F \subseteq \operatorname{NP}(f_i)$ such that $\operatorname{relint}(\mathcal{N}_{f_i}(F))$ contains infinitely many elements of $\{w(m)\}_{m\in\mathbb{N}}$. Thus, we can pass to a subsequence $\{\tilde{w}(m)\}_{m\in\mathbb{N}}$ such that

$F = \mathrm{NP}(f_i)_{\tilde{w}(m)}$ for all $m \in \mathbb{N}$. Since $\tilde{w}(m) \to w$ and $\mathcal{N}_{f_i}(F)$ is closed, $w \in \mathcal{N}_{f_i}(F)$. Therefore $F \subseteq \mathrm{NP}(f_i)_w$. Note that equality holds if and only if $w \in \operatorname{relint}\big(\mathcal{N}_{f_i}(F)\big)$ by (6).

Since $w \in \mathbb{R}^n \setminus \mathcal{N}_{f_i}^-$ and this set is open in $\mathbb{R}^n$, it follows that $\tilde{w}(m) \in \mathbb{R}^n \setminus \mathcal{N}_{f_i}^-$ for all $m \gg 1$. Since $\mathrm{NP}(f_i)_{\tilde{w}(m)} = F$ for all $m \in \mathbb{N}$, it follows that $\tilde{w}(m) \in \mathbb{R}^n \setminus \mathcal{N}_{f_i}^-$ for all $m \in \mathbb{N}$ and the face F does not contain any negative exponent vector of f_i. This implies:

$$f_{i|N(f_i)_{\tilde{w}(m)}}(1,\dots,1) = \sum_{\mu \in \sigma(f_i) \cap F} c_\mu = \sum_{\mu \in \sigma_+(f_i) \cap F} c_\mu > 0, \tag{19}$$

where c_μ for $\mu \in \sigma(f_i)$ are the coefficients of f_i. Thus, by Lemma 2.1 for each $\tilde{w}(m)$ there exists $T(m) > 0$ such that

$$f_i(t^{\tilde{w}(m)}) > 0, \quad \text{for all} \quad t > T(m).$$

In the following, we show that $T(m)$ can be chosen independently from m. The leading coefficient of $f_i(t^{\tilde{w}(m)})$ is as given in (19), which is positive, so there exists $\epsilon > 0$ such that $f_{i|N(f_i)_{\tilde{w}(m)}}(1,\dots,1) - \epsilon > 0$. Since $w \in \mathbb{R}^n \setminus \mathcal{N}_{f_i}^-$, all the exponent vectors on the face $\mathrm{NP}(f_i)_w$ are positive, i.e.

$$c_\mu > 0, \qquad \text{for all } \mu \in \mathrm{NP}(f_i)_w \cap \sigma(f_i).$$

Thus, the expression

$$\big(f_{i|N(f_i)_{\tilde{w}(m)}}(1,\dots,1) - \epsilon\big)t^d + \sum_{\mu \in (\mathrm{NP}(f_i)_w \setminus F) \cap \sigma(f_i)} c_\mu t^{\mu \cdot \tilde{w}(m)},$$

where $d = \max_{\mu \in \mathrm{NP}(f_i)} \tilde{w}(m) \cdot \mu$, is positive for all $t > 0$ and m.

Note that there exists $\delta > 0$ such that

$$w \cdot \mu - w \cdot \nu > \delta \qquad \text{for all} \quad \mu \in \sigma(f_i) \cap \mathrm{NP}(f_i)_w, \quad \nu \in \sigma(f_i) \setminus \mathrm{NP}(f_i)_w.$$

Since $\tilde{w}(m) \to w$ and the scalar product is continuous,

$$\tilde{w}(m) \cdot \mu - \tilde{w}(m) \cdot \nu > \delta \qquad \text{for all} \quad \mu \in \sigma(f_i) \cap \mathrm{NP}(f_i)_w, \quad \nu \in \sigma(f_i) \setminus \mathrm{NP}(f_i)_w$$

holds for m large enough. By passing to a subsequence if necessary, we may assume that the above inequality holds for all $\tilde{w}(m)$. Lemma 2.3 gives a bound $T \geq 1$ on the positive roots of $f_i(t^{\tilde{w}(m)})$ that depends only on ϵ, δ and c_μ, $\mu \in \sigma(f_i)$. In particular, this bound is independent of m. Since the leading coefficient of $f_i(t^{\tilde{w}(m)})$ is positive for all $m \in \mathbb{N}$ by (19), we conclude that

$$f_i(t^{\tilde{w}(m)}) > 0.$$

for all $m \in \mathbb{N}$ and $t > T$. □

Remark 3.4. Some of the inclusions in Theorem 3.3 can be generalized to semi-algebraic sets over a real closed field $\mathcal{R}$ with a compatible non-trivial non-Archimedean valuation $\mathrm{val}\colon \mathcal{R}^* \to \mathbb{R}$, for instance the field of real Puiseux series $\mathbb{R}\{\{t\}\}$.

Let $f_1, \dots, f_k \in \mathcal{R}[x_1, \dots, x_n]$ be polynomials and consider the semi-algebraic set

$$S_{\mathcal{R}}(f_1, \dots, f_k) = \{z \in \mathcal{R}^n_{>0} \mid f_1(z) < 0, \dots, f_k(z) < 0\}.$$

Here, $>$ denotes the unique ordering on the real closed field $\mathcal{R}$. The real tropicalization of a semi-algebraic set is defined as

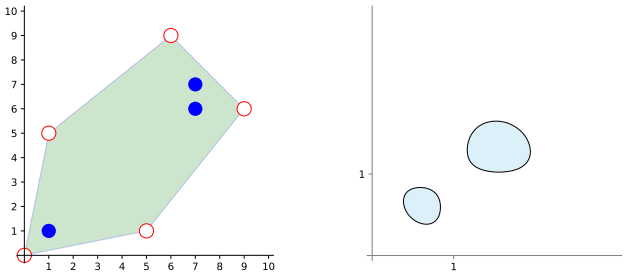
$$\mathrm{Trop}\big(S_{\mathcal{R}}(f_1, \dots, f_k)\big) := \overline{\big\{(-\operatorname{val}(z), \dots, -\operatorname{val}(z)) \mid z \in S_{\mathcal{R}}(f_1, \dots, f_k)\big\}},$$

where the closure is taken in the Euclidean topology of $\mathbb{R}^n$.

This construction generalizes the definition of real tropicalization as a logarithmic limit in the following sense. If $\mathcal{R}$ is a non-Archimedean real closed field of rank one extending $\mathbb{R}$ (e.g. $\mathcal{R} = \mathbb{R}\{\{t\}\}$) and the coefficients of $f_1, \dots, f_k \in \mathcal{R}[x_1, \dots, x_n]$ are real numbers, then by [1, Corollary 4.6] we have,

$$\mathrm{Trop}\big(S(f_1, \dots, f_k)\big) = \mathrm{Trop}\big(S_{\mathcal{R}}(f_1, \dots, f_k)\big),$$

where $S(f_1, \dots, f_k)$ is defined as in (14) and $\mathrm{Trop}(S(f_1, \dots, f_k))$ denotes the logarithmic limit of $S(f_1, \dots, f_k)$ as introduced in (15).



(a) NP(f) (b) $S(f)$

(c) $\mathrm{Log}_2\big(S(f)\big)$ (d) $\mathrm{Log}_{10}\big(S(f)\big)$

FIGURE 2. (a) The Newton polytope of $f = x_1^9x_2^6 + x_1^6x_2^9 - x_1^7x_2^7 - 4x_1^7x_2^6 + 5x_1^5x_2 + 5x_1x_2^5 - 5x_1x_2 + 1$ from Example 3.7 (b) The set $S(f) = f^{-1}(\mathbb{R}_{<0})$ is bounded. (c)(d) Logarithmic images of $S(f)$ for $t = 2$ and $t = 10$.

By replacing the negative normal cones $\mathcal{N}_{f_i}^-$ with a "signed part" of the tropical hypersurface defined by f_i (see [17, Section 5.2] or [5, Section 2.1] for a precise definition), the inclusions in (2) remain true [17, Proposition 6.12].

To the best of our knowledge, there is no known generalization of the actual negative normal cone $\Sigma(f_1, \dots, f_k)$ in the non-trivial valuation case, such that the generalized objects would satisfy similar inclusions as in Theorem 3.3. The techniques used in the current paper do not allow immediately to find such a generalization. However, it is an interesting problem that might be addressed using alternative approaches.

Corollary 3.5. *For a maximally sparse signomial f, it holds:*

$$\mathrm{Trop}(S(f)) = \mathcal{N}_f^-.$$

Proof. If all the exponent vectors of f are vertices of NP(f), then every negative face of NP(f) must contain a negative vertex. Thus, $\mathcal{N}_f^-$ equals the closure of its interior by Proposition 3.2. Using that $\mathrm{Trop}(S(f))$ is closed and Theorem 3.3, we conclude that $\mathrm{Trop}(S(f)) = \mathcal{N}_f^-$. □

For a signomial f, it might be worth to know when $S(f)$ is bounded from the coordinate planes in $\mathbb{R}^n$ and from infinity. Note that this happens if and only if $\mathrm{Log}_t(S(f))$ is bounded for $t > 0$, as the map $\mathrm{Log}_t \colon \mathbb{R}^n_{>0} \to \mathbb{R}$ is a homeomorphism.

Corollary 3.6. *Let f be a signomial on $\mathbb{R}^n_{>0}$. If $\sigma_-(f) \subseteq \mathrm{int}(\mathrm{NP}(f))$, then $\mathrm{Log}_t(S(f))$ is bounded for all $t > 0$.*

Proof. If the boundary of NP(f) does not contain negative monomials, then $\mathcal{N}_f^- \subseteq \{0\}$. Theorem 3.3 implies that $\mathrm{Trop}(S(f)) \subseteq \{0\}$. Now, the statement follows from Proposition 3.1(iii). □

Example 3.7. The boundary of the Newton polytope of $f = x_1^9x_2^6 + x_1^6x_2^9 - x_1^7x_2^7 - 4x_1^7x_2^6 + 5x_1^5x_2 + 5x_1x_2^5 - 5x_1x_2 + 1$ does not contain any negative exponent vector of f. By Corollary 3.6, $\mathrm{Log}_t\big(S(f)\big)$ is bounded for all $t > 0$. For an illustration, we refer to Figure 2.

To conclude, we show that for a semi-algebraic set of the form $S(f)$ the second inclusion in Theorem 3.3 is an equality if the coefficients of f are generic. First, we clarify what we mean by generic coefficients. For a fixed finite set $A \subseteq \mathbb{R}^n$, we consider the space of signomials whose support is contained in A:

$$\mathbb{R}^A = \big\{ f = \sum_{\mu \in \sigma(f)} c_\mu x^\mu \mid \sigma(f) \subseteq A \big\}.$$

Interpreting the coefficients of signomials as vectors, one has an isomorphism $\mathbb{R}^A \cong \mathbb{R}^{\#A}$.

The set of nonnegative signomials in $\mathbb{R}^A$,

$$\mathcal{P}_A^+ = \big\{ f \in \mathbb{R}^A \mid f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n_{>0} \big\},$$

is a full-dimensional closed convex cone, called the *nonnegativity cone* [11, 14]. Using this terminology, the genericity condition we will assume is that for all faces $F \subseteq \mathrm{NP}(f)$, the signomial $f_{|F}$ does not lie on the boundary of the nonnegativity cone, i.e.:

$$f_{|F} \in \mathcal{P}^+_{\sigma(f)\cap F} \quad \Longrightarrow \quad f_{|F} \in \mathrm{int}(\mathcal{P}^+_{\sigma(f)\cap F}). \tag{NB}$$

As a consequence of the non-boundary assumption (NB), we will be able to perturb the coefficients of a nonnegative signomial while preserving nonnegativity.

Remark 3.8. If a signomial, which is nonnegative over the positive real orthant, has a positive real root, then this root is degenerate and the signomial lies on the boundary of the nonnegativity cone. However, there exist signomials contained in the boundary of the nonnegativity cone that do not have any positive real roots as the following well-known example shows.

The signomial $g = x_1^2x_2^2 - 2x_1x_2 + 1 + x_1^2 = (x_1x_2 - 1)^2 + x_1^2$ is positive for all $(x_1, x_2) \in \mathbb{R}^2_{>0}$, but for any $\varepsilon > 0$ the signomial $g_\varepsilon = x_1^2x_2^2 - (2+\varepsilon)x_1x_2 + 1 + x_1^2$ takes negative values in $\mathbb{R}^2_{>0}$, e.g. $g_\varepsilon(\frac{\sqrt{\varepsilon}}{2}, \frac{2}{\sqrt{\varepsilon}}) = \frac{3}{4}\varepsilon < 0$. Thus, even if a nonnegative signomial does not have any roots, i.e. it is strictly positive, we might not be able to perturb its coefficients while preserving nonnegativity.

Theorem 3.9. *Let $f_1, \ldots, f_k$ be signomials on $\mathbb{R}^n_{>0}$ such that condition* (NB) *holds for all f_i and all faces $F \subseteq \mathrm{NP}(f_i)$, $i = 1, \ldots, k$. Then:*

$$\mathrm{Trop}\big(S(f_1, \ldots, f_k)\big) \subseteq \bigcap_{i=1}^k \Sigma(f_i).$$

Proof. We use a similar argument to the proof of Theorem 3.3. Let $w \in \mathrm{Trop}(S)$. By Proposition 3.1, there exist sequences $\{x(m)\}_{m\in\mathbb{N}} \subseteq S$, and $\{t(m)\}_{m\in\mathbb{N}} \subseteq \mathbb{R}_{>0}$ such that $t(m) \to \infty$ and $w(m) := \mathrm{Log}_{t(m)}(x(m)) \to w$.

Assume that $w \notin \bigcap_{j=1}^k \Sigma(f_j)$, and let $i \in \{1, \ldots, k\}$ such that $w \notin \Sigma(f_i)$. In the following, we write c_μ, $\mu \in \sigma(f_i)$ for the coefficients of f_i and $G := \mathrm{NP}(f_i)_w$ for the face of $\mathrm{NP}(f_i)$ which is cut out by w. By an argument as in the proof of Theorem 3.3, we may pass to a subsequence if necessary and assume that there exists a face $F \subseteq G$ of $\mathrm{NP}(f_i)$ such that $F = \mathrm{NP}(f_i)_{w(m)}$ for all $m \in \mathbb{N}$.

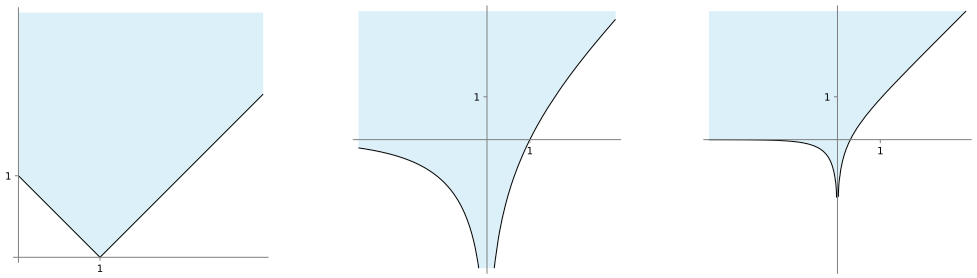
Since w cuts out the face G, there exists $\delta > 0$ such that

$$w \cdot \mu - \delta > w \cdot \nu, \qquad \text{for all } \mu \in G \cap \sigma(f_i), \quad \nu \in \sigma(f_i) \setminus G. \tag{20}$$

By continuity of the scalar product, (20) holds also for $w(m)$ for large enough m. Thus, we can pass to a subsequence again to assume that (20) holds for all $w(m)$.

Since $w \notin \Sigma(f_i)$, the signomial $f_{i|G}$ is in the nonnegativity cone $\mathcal{P}^+_{\sigma(f_i)\cap F}$. By the condition (NB) $f_{i|G}$ is in the interior and hence there exists $\tilde{\epsilon} > 0$ such that

$$\sum_{\mu \in F \cap \sigma(f_i)} (c_\mu - \tilde{\epsilon}) x^\mu + \sum_{\mu \in (G \setminus F) \cap \sigma(f_i)} c_\mu x^\mu > 0 \quad \text{for all} \quad x \in \mathbb{R}^n_{>0}. \tag{21}$$



(a) $S(g)$ (b) $\mathrm{Log}_2\big(S(g)\big)$ (c) $\mathrm{Log}_{10}\big(S(g)\big)$

FIGURE 3. The semi-algebraic set $S(g)$ defined by $g = x_1^2 - 2x_1 + 1 - x_2^2$ (a), and its logarithmic images for $t = 2$ (b) and $t = 10$ (c).

In particular, it holds:

$$\sum_{\mu\in F\cap\sigma(f_i)} (c_\mu - \tilde{\epsilon})t^{w(m)\cdot\mu} + \sum_{\mu\in(G\setminus F)\cap\sigma(f_i)} c_\mu t^{w(m)\cdot\mu} > 0 \quad \text{for all } t\in\mathbb{R}_{>0} \text{ and } m\in\mathbb{N}. \tag{22}$$

From (20) and (22) follows that for each $m \in \mathbb{N}$ the univariate signomial

$$f(t^{w(m)}) = \Big(\sum_{\mu\in F\cap\sigma(f_i)} c_\mu\Big)t^d + \sum_{\mu\in(G\setminus F)\cap\sigma(f_i)} c_\mu t^{w(m)\cdot\mu} + \sum_{\mu\in\sigma(f_i)\setminus G} c_\mu t^{w(m)\cdot\mu},$$

where $d = \max_{\mu\in\sigma(f_i)} w(m)\cdot\mu$, satisfies the conditions in Lemma 2.3 with $\epsilon = \sum_{\mu\in F\cap\sigma(f_i)} \tilde{\epsilon} > 0$ and δ as in (20). Thus, there exists $T > 1$ which is independent of m such that

$$f_i(t^{w(m)}) > 0, \qquad \text{for all } t > T.$$

Since $t(m) \to \infty$, $t(m) > T$ for $m \gg 1$. Thus,

$$0 < f_i(t(m)^{w(m)}) = f_i(x(m)) < 0.$$

which is a contradiction. The last inequality holds as $x(m) \in S(f_1, \dots, f_k)$. □

Corollary 3.10. *Let f be a signomial on $\mathbb{R}^n_{>0}$. If all faces $F \subseteq \mathrm{NP}(f)$ satisfy* (NB) *then:*

$$\mathrm{Trop}\big(S(f)\big) = \Sigma(f).$$

Proof. By Theorem 3.9, $\mathrm{Trop}\big(S(f)\big) \subseteq \Sigma(f)$. The reverse inclusion follows from Theorem 3.3. □

Example 3.11. Consider the signomials $f = x_1^2 - x_1 + 1 - x_2^2$ from Figure 1 and $g = x_1^2 - 2x_1 + 1 - x_2^2$. Note that f and g have the same negative/positive exponent vectors and $\mathrm{NP}(f) = \mathrm{NP}(g)$. The only difference between f and g is the coefficient of x_1. Furthermore, we have:

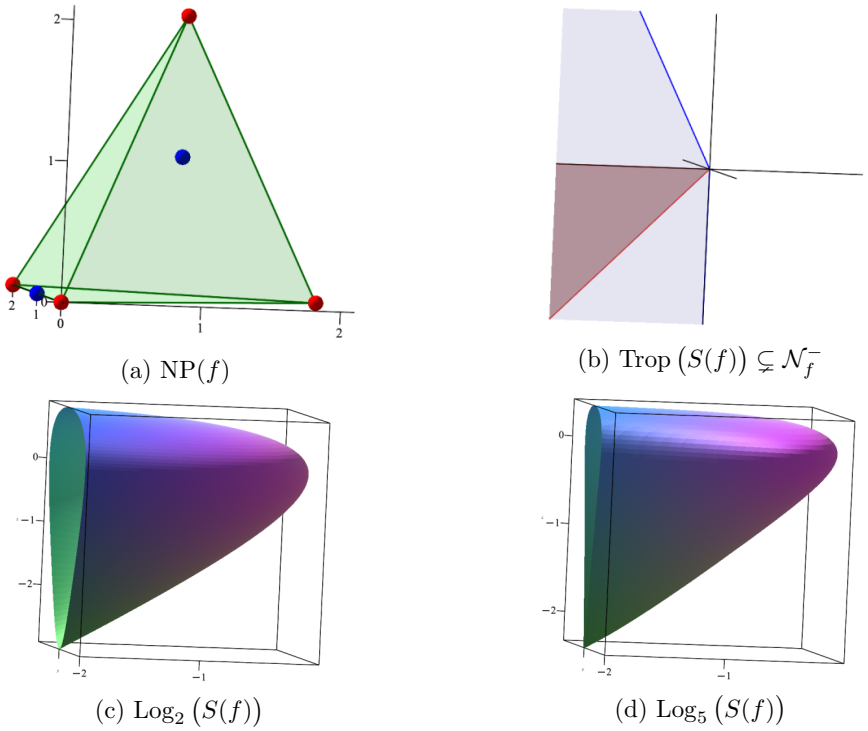
$$\Sigma(f) = \Sigma(g), \qquad \mathcal{N}_f^- = \mathcal{N}_g^-.$$

For the faces $F_1 = \mathrm{Conv}((2,0),(0,2))$, $F_2 = \mathrm{Conv}((0,0),(0,2))$ of $\mathrm{NP}(f)$, the restrictions $f_{|F_i}$ and $g_{|F_i}$, $i = 1, 2$ take negative values by Lemma 2.1. Thus, $f_{|F_i}$ and $g_{|F_i}$, $i = 1, 2$ are not in the nonnegativity cone.

For the face $F = \mathrm{Conv}((0,0),(2,0))$, $f_{|F} = x_1^2 - x_1 + 1$ is contained in the interior of the nonnegativity cone. So the coefficients of f are generic in the sense of Corollary 3.10, thus:

$$\mathrm{Trop}(S(f)) = \Sigma(f).$$

The sets $S(f), \Sigma(f)$ and $\mathrm{Log}_t(S(f))$ for $t = 2, 10$ are shown in Figure 1. The negative normal cone $\mathcal{N}_f^-$ (see Figure1(c)) is strictly larger than $\mathrm{Trop}(S(f))$ as $\mathcal{N}_f^-$ has a ray in southern direction that $\mathrm{Trop}(S(f))$ does not have. This example illustrates that the negative normal cone might not coincide with the real tropicalization.



(a) NP(f) (b) Trop $\big(S(f)\big) \subsetneq \mathcal{N}_f^-$

(c) $\mathrm{Log}_2\big(S(f)\big)$ (d) $\mathrm{Log}_5\big(S(f)\big)$

FIGURE 4. Illustration of Example 3.12 (a) The Newton polytope of $f = x_2^2 - 2x_2 + 1 - 2x_1x_2x_3 + x_1x_2x_3^2 + x_1^2x_2$. The exponent vector $(1,1,1)$ lies in the interior of NP(f). (b) Negative normal cone of f (blue shaded area) and the real tropicalization of $S(f)$ (pink shaded area) (c),(d) Logarithmic images of $S(f)$ for $t = 2$ and $t = 5$ respectively.

As $g_{|F} = x_1^2 - 2x_1 + 1$ is nonnegative but has a positive real root, $g_{|F}$ lies on the boundary of the nonnegativity cone. Thus (NB) is not satisfied, and we cannot apply Corollary 3.10. In fact, it holds that

$$\mathrm{Trop}(S(g)) = \mathcal{N}_g^-,$$

which is strictly larger than the actual negative normal cone $\Sigma(g)$, see Figure 1 (b),(c). This illustrates that if the condition (NB) is not satisfied, Corollary 3.10 might not hold. Logarithmic images of $S(g)$ with base $t = 2, 10$ are shown in Figure 3.

Example 3.12. Consider the signomial

$$f = x_2^2 - 2x_2 + 1 - 2x_1x_2x_3 + x_1x_2x_3^2 + x_1^2x_2.$$

The Newton polytope of f is shown in Figure 4(a). The only negative exponent vector of f that lies on the boundary of NP(f) is $(0,1,0)$. This negative exponent vector is contained in the 1-dimensional face $F = \mathrm{Conv}\big((0,0,0),(0,2,0)\big)$. The negative normal cone of f is 2-dimensional, and it is spanned by the vectors $(0,0,-1)$ and $(-1,0,2)$, see Figure4(b). Since $\mathcal{N}_f^-$ is not full dimensional, we have that $\mathrm{int}(\mathcal{N}_f^-) = \emptyset$.

The restricted signomial $f_{|F} = x_2^2 - 2x_2 + 1$ is non-negative but it has a positive real zero. Thus, $f_{|F}$ does not satisfy the condition (NB). Since $f(0.5,1,0.5) = -0.125 < 0$ and $f_{|G}$ is non-negative for all proper faces $G \subsetneq \mathrm{NP}(f)$, we have $\Sigma(f) = \{0\}$.

The real tropicalization of $S(f)$ is a 2-dimensional cone, spanned by the vectors $(-1,0,0)$ and $(-1,0,-1)$. This example shows that $\mathrm{Trop}(S(f))$ is not always a subfan of the outer normal fan of NP(f), and that all the inclusions in Theorem 3.3 might be strict:

$$\mathrm{int}\left(\mathcal{N}_f^-\right) \subsetneq \Sigma(f) \subsetneq \mathrm{Trop}\left(S(f)\right) \subsetneq \mathcal{N}_f^-.$$

Acknowledgments

The author thanks Elisenda Feliu for useful discussions and comments on the manuscript. Funded by the European Union under the Grant Agreement no. 101044561, POSALG. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or European Research Council (ERC). Neither the European Union nor ERC can be held responsible for them

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. Alessandrini. Logarithmic limit sets of real semi-algebraic sets. *Adv. Geom.*, 13(1):155–190, 2013.
[2] X. Allamigeon, S. Gaubert, and M. Skomra. Tropical spectrahedra. *Discrete. Comput. Geom.*, 63(3):507–548, 2020.
[3] G. M. Bergman. The logarithmic limit-set of an algebraic variety. *Trans. Am. Math. Soc.*, 157:459–469, 1971.
[4] G. Blekherman and A. Raymond. A path forward: Tropicalization in extremal combinatorics. *Adv. Math.*, 407:108561, 2022.
[5] G. Blekherman, F. Rincón, R. Sinn, C. Vinzant, and J. Yu. Moments, sums of squares, and tropicalization. *arXiv*, (2203.06291), 2022.
[6] V. Chandrasekaran and M. P. Shah. Relative entropy relaxations for signomial optimization. *SIAM J. Optim.*, 26:1147–1173, 2014.
[7] A. Dickenstein. Algebraic geometry tools in systems biology. *Not. Am. Math. Soc.*, 67:1, 12 2020.
[8] M. Dressler, S. Iliman, and T. de Wolff. An approach to constrained polynomial optimization via nonnegative circuit polynomials and geometric programming. *Journal of Symbolic Computation*, 91:149–172, 2019. MEGA 2017, Effective Methods in Algebraic Geometry, Nice (France), June 12-16, 2017.
[9] R. J. Duffin and E. L. Peterson. Geometric programming with signomials. *J. Optimiz. Theory. App.*, 11(1):3–35, 1973.
[10] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürück. The kinetic space of multistationarity in dual phosphorylation. *J. Dyn. Differ. Equ.*, 34:825–852, 2022.
[11] J. Forsgård and T. de Wolff. The algebraic boundary of the sonc cone. *arXiv*, (1905.04776), 2019.
[12] P. A. Helminck and Y. Ren. Generic root counts and flatness in tropical geometry. *arXiv*, (2206.07838), 2022.
[13] M. Henk, J. Richter-Gebert, and G. M. Ziegler. Basic properties of convex polytopes. In *Handbook of Discrete and Computational Geometry, 3rd Ed.*, 2017.
[14] J. Heuer and T. de Wolff. The duality of sonc: Advances in circuit-based certificates. *arXiv*, (2204.03918), 2022.
[15] D. R. Hilbert. Über die darstellung definiter formen als summe von formenquadraten. *Mathematische Annalen*, 32:342–350, 1888.
[16] S. Iliman and T. de Wolff. Amoebas, nonnegative polynomials and sums of squares supported on circuits. *Research in the Mathematical Sciences*, 3, 03 2016.
[17] P. Jell, C. Scheiderer, and J. Yu. Real Tropicalization and Analytification of Semialgebraic Sets. *Int. Math. Res. Notices*, 2022(2):928–958, 05 2020.
[18] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Springer, 01 2013.
[19] M. Laurent. *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, pages 157–270. Springer New York, New York, NY, 2009.
[20] D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2015.
[21] G. Mikhalkin. Enumerative tropical algebraic geometry in $\mathbb{R}^2$. *J. Am. Math. Soc.*, 18:313–378, 2003.
[22] G. Mikhalkin. Decomposition into pairs-of-pants for complex algebraic hypersurfaces. *Topology*, 43(5):1035–1065, 2004.
[23] M. Nisse. Maximally sparse polynomials have solid amoebas. *arXiv*, (0704.2216), 2008.

[24] Q.I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials*. London Mathematical Society monographs. Clarendon Press, 2002.
[25] B. Reznick. Forms derived from the arithmetic-geometric inequality. *Mathematische Annalen*, 283:431–464, 1989.
[26] R. T. Rockafellar. *Convex analysis*. Princeton University Press, 1972.
[27] I. Sahidul and A. M. Wasim. *Fuzzy Geometric Programming Techniques and Applications*. Springer, 2019.
[28] L. A. Steen and Seebach J. A. *Counterexamples in Topology*. Springer-Verlag New York, 2 edition, 1978.
[29] C. W. Wampler, A. Morgan, and A. J. Sommese. Complete solution of the nine-point path synthesis problem for four-bar linkages. *J. Mech. Des.*, 114:153–159, 1992.
[30] G. M. Ziegler. *Lectures on Polytopes*. Springer, 2007.

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: mlt@math.ku.dk

VI

Viro's patchworking and the signed reduced A-discriminant

Weixun Deng
Department of Mathematics
Texas A&M University

J. Maurice Rojas
Department of Mathematics
Texas A&M University

Máté L. Telek
Department of Mathematical Sciences
University of Copenhagen

Publication details

Submitted (2024)
Available on arXiv: https://doi.org/10.48550/arXiv.2403.08497

VIRO'S PATCHWORKING AND THE SIGNED REDUCED A-DISCRIMINANT

WEIXUN DENG, J. MAURICE ROJAS, AND MÁTÉ L. TELEK

Abstract. Computing the isotopy type of a hypersurface, defined as the positive real zero set of a multivariate polynomial, is a challenging problem in real algebraic geometry. We focus on the case where the defining polynomial has combinatorially restricted exponent vectors and fixed coefficient signs, enabling faster computation of the isotopy type. In particular, Viro's patchworking provides a polyhedral complex that has the same isotopy type as the hypersurface, for certain choices of the coefficients. So we present properties of the signed support, focussing mainly on the case of n-variate (n+3)-nomials, that ensure all possible isotopy types can be obtained via patchworking. To prove this, we study the signed reduced A-discriminant and show that it has a simple structure if the signed support satisfies some combinatorial conditions.

Keywords: exponential sum, signed support, Viro's patchworking, isotopy type

1. Introduction

Tropical geometry bridges the worlds of algebraic and polyhedral geometry. The key idea is to transform algebraic varieties into polyhedral objects, turning their algebraic structure functorially into a combinatorial structure. This approach has been fruitful in the case of varieties over algebraically closed fields [17, 22]. Recently, the tropicalization of semialgebraic sets over the real numbers (or more generally over real closed fields) has received increasing attention [4, 5, 18, 29, 31].

In the early 1980s, Oleg Viro showed that it is possible to associate a polyhedral complex with a parametrized polynomial in such a way that there exists *some* choice of coefficients such that the polyhedral complex and the positive real zero set of the polynomial with that choice of coefficients have the same isotopy type [33]. This result is known as *Viro's patchworking* in the literature and was one of the first examples of tropical geometry.

Classifying the isotopy type of real hypersurfaces is a challenging question in real algebraic geometry, tracing its origins to Hilbert's 16th problem [34]. In his original formulation, Hilbert asked for a classification of isotopy types of plane real algebraic curves. Based on his patchworking method, Viro provided such a classification for curves of degree 7 [32].

Using Viro's method, under certain conditions, one can determine the possible isotopy types for hypersurfaces defined by the positive roots of polynomials of the form $f(x) = \sum_{a\in\mathcal{A}} c_a x^a$, with $\mathcal{A} \subseteq \mathbb{Z}^n$ a finite set and c_a real and nonzero for all $a \in \mathcal{A}$. We call $\mathcal{A}$ the *support* of f, let $\varepsilon = (\varepsilon_a \mid a \in \mathcal{A}) \in \{\pm 1\}^{\mathcal{A}}$ denote the vector of signs of the coefficients c_a, and call the pair $(\mathcal{A}, \varepsilon)$ the *signed support* of f. To know whether one can build *all* possible isotopy types for a given $(\mathcal{A}, \varepsilon)$ via Viro's patchworking one must first understand how the isotopy type of a hypersurface changes while varying the coefficient vector c. We will abuse notation slightly by calling ε_a the *sign* of the exponent vector a when the underlying polynomial and support are clear. Hence we will sometimes speak of *positive* or *negative* exponents in this sense.

It is known that using discriminant varieties, one can decompose the coefficient space into a disjoint union of open connected sets — called *chambers* — such that in each chamber the topological type of any corresponding hypersurface is constant [3, 14]. More specifically, such a chamber decomposition is given by connected components of the complement of the union of the signed A-discriminants $\nabla_{A_F,\varepsilon_F}$ for the faces F of $\mathrm{Conv}(\mathcal{A})$. In [26], a reduced version of the signed A-discriminant, denoted $\Gamma_\varepsilon(A, B)$, was introduced. The complement of the union of the signed reduced A-discriminants has the same property as its non-reduced counterpart, but

has the advantage of reducing the number of parameters. We will review these constructions in Section 2.3.

When considering chambers (reduced or non-reduced), there will sometimes be chambers where the isotopy type of the hypersurfaces can *not* be obtained by Viro's patchworking (see, e.g., [11, Example 2.9] and [3, Theorem 7.8]). These chambers are called *inner* chambers, and reduced inner chambers are usually bounded sets.

Our main goal is to give conditions on the signed support $(\mathcal{A}, \varepsilon)$ enabling Viro's patchworking to find all isotopy types, i.e., conditions that obstruct the existence of inner chambers. For instance, we show that the signed A-discriminant is empty if and only if the exponent vectors in $\mathcal{A}$ with positive signs can be separated from those with negative signs by an affine hyperplane (Theorem 3.5). In that case, all hypersurfaces with signed support $(\mathcal{A}, \varepsilon)$ have the same isotopy type, and the isotopy type can be obtained by Viro's patchworking. Moreover, for a given set of exponent vectors $\mathcal{A}$, we give an upper bound on the number of sign distributions $\varepsilon \in \{\pm 1\}^{\mathcal{A}}$ such that the signed support $(\mathcal{A}, \varepsilon)$ does not have a separating hyperplane (Proposition 3.4).

As we will see, it will be convenient (and natural) to assume $\mathrm{Conv}(\mathcal{A})$ has dimension n and cardinality $n+k+1$, and then consider k as a measure of the complexity of the resulting family of isotopy types. The special cases $k \in \{0, 1\}$ are addressed in [2, 8], so we will contribute to the case $k=2$: We show that the complement of the signed reduced A-discriminant has exactly two connected components if $\mathcal{A}$ contains exactly one negative exponent vector (Theorem 4.11) or if the positive and the negative exponent vectors of $\mathcal{A}$ are separated by an n-simplex in a certain way (Theorem 4.12). Under the additional assumption that the signed A-discriminants associated to proper faces of $\mathrm{Conv}(\mathcal{A})$ are empty, one can use Viro's patchworking to find all possible isotopy types (Corollary 4.13).

Signed supports with a separating hyperplane or a separating simplex have previously been studied in [9]. In that work, the authors used these conditions to show that the set of points where the polynomial takes negative values has, at most, one connected component.

For $n = 2$ and $\mathcal{A}$ consisting 5 points, we show that if the negative and positive exponent vectors are separated by two pairs of affine lines (Theorem 4.10), then the complement of $\Gamma_\varepsilon(A, B)$ has at most two connected components, which are unbounded. In Section 5, we study further bivariate 5-nomials and show that for a bounded chamber to exist in the complement of $\Gamma_\varepsilon(A, B)$, the set of exponent vectors must satisfy very restrictive inequalities, see Theorem 5.5 and Remark 5.6.

We will work in the more general context of exponential sums on $\mathbb{R}^n$, instead of polynomials on $\mathbb{R}^n_{>0}$, and this will in fact simplify some of our arguments. Note that any real polynomial can be transformed into a real exponential sum while preserving topological properties of the corresponding zero sets: Any polynomial $f\colon \mathbb{R}^n_{>0} \to \mathbb{R}$ give rise to an exponential sum $\mathbb{R}^n \to \mathbb{R}$, $(x_1, \dots x_n) \mapsto f(e^{x_1}, \dots, e^{x_n})$. Since the map $\mathrm{Exp}\colon \mathbb{R}^n \to \mathbb{R}^n_{>0}$, $(x_1, \dots x_n) \mapsto (e^{x_1}, \dots, e^{x_n})$ is a homeomorphism, two subsets of $\mathbb{R}^n_{>0}$ have the same isotopy type if and only if their images under Exp have the same isotopy type.

Notation. For two vectors $v, w \in \mathbb{R}^n$, $v \cdot w$ denotes the Euclidean scalar product, and $v * w$ denotes the coordinate-wise product of v and w. The transpose of a matrix M will be denoted by $M^\top$. We denote the interior of a set $X \subseteq \mathbb{R}^n$ by $\mathrm{int}(X)$. If $X \subseteq \mathbb{R}^n$ is a polyhedron, $\mathrm{relint}(X)$ denotes the relative interior of X. By $\#S$ we denote the cardinality of a finite set S. For a differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}^m$, the Jacobian matrix at $x \in \mathbb{R}^n$ is denoted by $J_f(x)$.

2. Preliminaries

2.1. **Signed support of an exponential sum.** Let $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$ be a finite set. We think about the elements of $\mathcal{A}$ as the exponent vectors of an *exponential sum*:

$$f_c\colon \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto f_c(x) = \sum_{i=1}^{n+k+1} c_i e^{\alpha_i \cdot x}, \tag{1}$$

where $c \in (\mathbb{R} \setminus \{0\})^{n+k+1}$. We call $\varepsilon = \mathrm{sign}(c) \in \{\pm 1\}^{n+k+1}$ a *sign distribution*, and $(\mathcal{A}, \varepsilon)$ a *signed support*. For a fixed order of the exponent vectors $\alpha_1, \dots, \alpha_{n+k+1}$, there is an isomorphism

between the vector spaces $\mathbb{R}^{\mathcal{A}} \cong \mathbb{R}^{n+k+1}$. We might use these two notations interchangebly. For a fixed sign distribution $\varepsilon \in \{\pm 1\}^{n+k+1}$, we write

$$\mathbb{R}^{\mathcal{A}}_{\varepsilon} = \{c \in \mathbb{R}^{\mathcal{A}} \mid \operatorname{sign}(c) = \varepsilon\}.$$

for the orthant in $\mathbb{R}^{\mathcal{A}}$ containing the coefficients matching the signs given by ε.

We split the support set $\mathcal{A}$ into the sets of *positive* and *negative exponent vectors*, that is, we define

$$\mathcal{A}_+ := \{\alpha_i \in \mathcal{A} \mid \varepsilon_i = 1\}, \qquad \mathcal{A}_- := \{\alpha_i \in \mathcal{A} \mid \varepsilon_i = -1\}.$$

We call $\mathcal{A}$ *full-dimensional* if $\operatorname{Conv}(\mathcal{A})$ has dimension n. For a set $S \subseteq \mathbb{R}^n$, we denote the restriction of f_c to S by

$$f_{c|S} \colon \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto f_c(x) = \sum_{\alpha_i \in \mathcal{A} \cap S} c_i e^{\alpha_i \cdot x}.$$

Furthermore, we set $\mathcal{A}_S := \mathcal{A} \cap S$, and define ε_S to be the sign distribution containing the signs corresponding to the elements in $\mathcal{A}_S$.

The real zero set of f_c is denoted by

$$Z(f_c) := \{x \in \mathbb{R}^n \mid f_c(x) = 0\}.$$

We are interested in the possible *isotopy types* of $Z(f_c)$ when $c \in \mathbb{R}^{n+k+1}$ varies over all coefficients such that $\operatorname{sign}(c) = \varepsilon$. Two subsets $Z_0, Z_1 \subseteq \mathbb{R}^n$ are *isotopic* (ambient in $\mathbb{R}^n$) if there exists a continuous map $H \colon [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$, called an *isotopy*, such that

- $H(t, \cdot)$ is a homeomorphism for all $t \in [0,1]$,
- $H(0, \cdot)$ is the identity on $\mathbb{R}^n$,
- $H(1, Z_0) = Z_1$.

Being isotopic gives an equivalence relation on subsets of $\mathbb{R}^n$ [20, Chapter 10.1], which allows us to talk about their isotopy types.

2.2. **Viro's patchworking.** Viro's patchworking method provides possible isotopy types of $Z(f_c)$, $c \in \mathbb{R}^{\mathcal{A}}_{\varepsilon}$ for a fixed signed support $(\mathcal{A}, \varepsilon)$. In this section, we recall this method. We follow the notation used in [7].

Let $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n+k+1}\} \subseteq \mathbb{Z}^n$ be a finite set and $h \in \mathbb{R}^{n+k+1}$. We consider the *lifted* points

$$\mathcal{A}^h := \big\{(\alpha_i, h_i) \in \mathbb{R}^{n+1} \mid \alpha_i \in \mathcal{A}\big\}.$$

A face $F \subseteq \operatorname{Conv}(\mathcal{A}^h)$ is called an *upper face* if there exists a vector $(v_F, 1) \in \mathbb{R}^{n+1}$ such that

$$F = \big\{x \in \operatorname{Conv}(\mathcal{A}^h) \mid (v_F, 1) \cdot x \geq (v_F, 1) \cdot y \text{ for all } y \in \operatorname{Conv}(\mathcal{A}^h)\big\}.$$

For a generic choice of $h \in \mathbb{R}^{n+k+1}$, each upper face $F \subseteq \operatorname{Conv}(\mathcal{A}^h)$ contains exactly $n+1$ points of $\mathcal{A}^h$. The projection of upper faces of $\operatorname{Conv}(\mathcal{A}^h)$ onto $\mathbb{R}^n$ gives a polyhedral subdivision $\mathcal{P}$ of $\operatorname{Conv}(\mathcal{A})$. If h is generic, each polyhedron in $\mathcal{P}$ is a simplex.

The *tropical hypersurface* associated to $\mathcal{A}$ and $h \in \mathbb{R}^{\mathcal{A}}$ is defined as

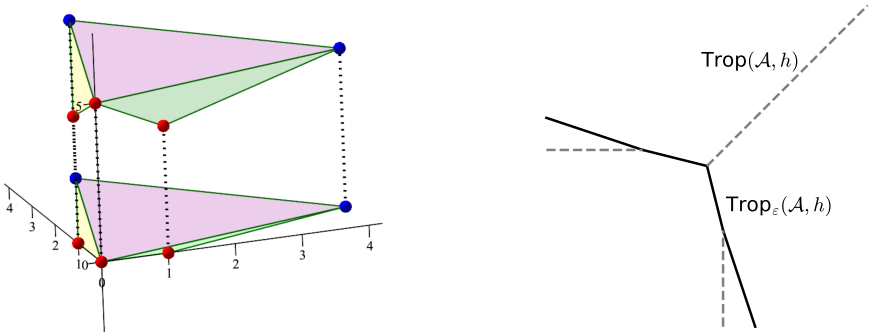
$$\operatorname{Trop}(\mathcal{A}, h) := \big\{v \in \mathbb{R}^n \mid \max_{i=1,\ldots,n+k+1} (v \cdot \alpha_i + h_i) \text{ is attained at least twice.}\big\}.$$

It is dual to the $(n-1)$-skeleton of the subdivision $\mathcal{P}$ induced by h. For a sign distribution $\varepsilon \in \{\pm 1\}^{\mathcal{A}}$, we define the *signed tropical hypersurface*

$$\operatorname{Trop}_{\varepsilon}(\mathcal{A}, h) := \Big\{v \in \mathbb{R}^n \mid \max_{i=1,\ldots,n+k+1} (v \cdot \alpha_i + h_i) \text{ is attained for some } \alpha \in \mathcal{A}_+ \text{ and } \alpha' \in \mathcal{A}_-\Big\}.$$

Example 2.1. Consider the following set of exponent vectors

$$\mathcal{A} = \Big\{\alpha_1 = \begin{bmatrix}0\\0\end{bmatrix}, \quad \alpha_2 = \begin{bmatrix}1\\0\end{bmatrix}, \quad \alpha_3 = \begin{bmatrix}0\\1\end{bmatrix}, \quad \alpha_4 = \begin{bmatrix}4\\1\end{bmatrix}, \quad \alpha_5 = \begin{bmatrix}1\\4\end{bmatrix}\Big\}, \tag{2}$$



(a) (b)

FIGURE 1. (a) Upper convex hull of the lifted points from Example 2.1 and the induced polyhedral subdivision. (b) Signed tropical hypersurface associated to the points in (a).

the sign distribution $\varepsilon = (1, 1, 1, -1, -1)$ and $h = (5, 4, 4, 5, 5)$. The upper faces of the convex hull of the lifted points

$$\mathcal{A}^h = \left\{ \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 5 \\ 1 \\ 4 \end{bmatrix} \right\},$$

are shown in Figure 1(a). The induced subdivision of $\mathrm{Conv}(\mathcal{A})$ contains 3 triangles, 7 edges, and 5 vertices. The tropical hypersurface $\mathrm{Trop}(\mathcal{A}, h)$, which is dual to the 1 skeleton of the subdivision, contains 3 vertices and 7 edges. We depicted $\mathrm{Trop}(\mathcal{A}, h)$ and the signed tropical hypersurface $\mathrm{Trop}_\varepsilon(\mathcal{A}, h)$ in Figure 1(b).

Viro's patchworking is usually stated for polynomials and their zero sets in the positive orthant $\mathbb{R}^n_{>0}$, however using the coordinate-wise exponential map $\mathrm{Exp}\colon \mathbb{R}^n \to \mathbb{R}^n_{>0}$ it possible to translate Viro's theorem to exponential sums.

Theorem 2.2. *[33][14, Ch.11 Theorem 5.6][17, Theorem 2.19] Let $(\mathcal{A}, \varepsilon)$ be a signed support such that $\mathcal{A} \subseteq \mathbb{Z}^n$ and let $h \in \mathbb{R}^{\mathcal{A}}$ be generic. For $t \in \mathbb{R}$, consider the exponential sum*

$$g_t\colon \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \sum_{i=1}^{n+k+1} \varepsilon_i e^{h_i t} e^{\alpha_i \cdot x}.$$

Then the signed tropical hypersurface $\mathrm{Trop}_\varepsilon(\mathcal{A}, h)$ is isotopic to $Z(g_t)$ for $t \gg 0$ sufficiently large.

2.3. **Signed A-discriminant.** The goal of this subsection is to recall the notion of the *A-discriminant* from [3, 14, 26]. Let f_c be an exponential sum as in (1). A point $x \in \mathbb{R}^n$ is a *singular zero* of f_c if and only if

$$f_c(x) = \frac{\partial f_c(x)}{\partial x_1} = \cdots = \frac{\partial f_c(x)}{\partial x_n} = 0. \tag{3}$$

We denote the set of singular zeros of f_c by $\mathrm{Sing}(f_c)$. For a fixed signed support $(\mathcal{A}, \varepsilon)$, we define the *signed A-discriminant* as

$$\nabla_{\mathcal{A},\varepsilon} := \left\{ c \in \mathbb{R}^{\mathcal{A}}_\varepsilon \mid \mathrm{Sing}(f_c) \neq \emptyset \right\}.$$

Thus, $\nabla_{\mathcal{A},\varepsilon}$ contains all coefficients $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ such that the exponential sum f_c has a singular zero in $\mathbb{R}^n$.

For the sake of completeness, we recall that the signed A-discriminant does not change under affine transformations of the exponent vectors.

Proposition 2.3. *Let f_c be an exponential sum with support $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$. For an invertible matrix $M \in \mathbb{R}^{n\times n}$ and $v \in \mathbb{R}^n$ consider the exponential sum*

$$g_c \colon \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto g_c(x) = \sum_{i=1}^{n+k+1} c_i e^{(M\alpha_i + v)\cdot x}.$$

Then we have:

(i) *If $\det(M) > 0$, then the hypersurfaces $Z(f_c)$ and $Z(g_c)$ are isotopic.*
(ii) $\operatorname{Sing}(f_c) = M^\top \operatorname{Sing}(g_c)$.
(iii) *For all $x \in \operatorname{Sing}(g_c)$ the Hessian matrices $\operatorname{Hess}_{f_c}(M^\top x)$ and $\operatorname{Hess}_{g_c}(x)$ have the same number of positive, negative and zero eigenvalues.*

Proof. Note that $g_c(x) = e^{v\cdot x} f_c(M^\top x)$ for all $x \in \mathbb{R}^n$. Since $e^{v\cdot x} \neq 0$, we have

$$\begin{aligned} Z(g_c) = Z(f_c(M^\top x)) &= \{x \in \mathbb{R}^n \mid \sum_{i=1}^{n+k+1} c_i e^{\alpha_i \cdot (M^\top x)} = 0\} \\ &= \{(M^\top)^{-1} y \in \mathbb{R}^n \mid \sum_{i=1}^{n+k+1} c_i e^{\alpha_i \cdot y} = 0\} = (M^\top)^{-1} Z(f_c) \end{aligned} \tag{4}$$

Since the group of invertible real $n\times n$ matrices with positive determinant is path-connected (see, e.g. [36, Theorem 3.68]), there exists a continuous path from the identity matrix to $(M^\top)^{-1}$, which induces an isotopy. This shows (i).

Applying the product and the chain rule from calculus, we have

$$J_{g_c}(x) = v^\top f_c(M^\top x) + e^{v\cdot x} J_{f_c}(M^\top x) M^\top.$$

Using (4) and that $M^\top$ is invertible, for $x \in Z(g_c)$ it follows that $J_{g_c}(x) = 0$ if and only if $J_{f_c}(M^\top x) = 0$, which implies (ii).

For the rest of the proof, we assume that $x \in \operatorname{Sing}(g_c)$. From [28, Corollary 1] it follows that

$$\operatorname{Hess}_{g_c}(x) = e^{v\cdot x} M \operatorname{Hess}_{f_c}(M^\top x) M^\top.$$

Thus, $\operatorname{Hess}_{g_c}(x)$ and $\operatorname{Hess}_{f_c}(M^\top x)$ have the same number of positive, negative and zero eigenvalues by Sylvester's law of inertia [25, Chapter 7].

□

Corollary 2.4. *Let $(\mathcal{A}, \varepsilon)$ be a signed support, $M \in \mathbb{R}^{n\times n}$ an invertible matrix and $v \in \mathbb{R}^n$. For $M\mathcal{A} + v = \{M\alpha + v \mid \alpha \in \mathcal{A}\}$, we have*

$$\nabla_{\mathcal{A},\varepsilon} = \nabla_{M\mathcal{A}+v,\varepsilon}.$$

Proof. The statement follows directly from Proposition 2.3(ii). □

Remark 2.5. Using Proposition 2.3, one might transform any full-dimensional support $\mathcal{A} \subseteq \mathbb{R}^n$ to a support containing the standard basis vectors $e_1, \dots, e_n \in \mathbb{R}^n$ and the zero vector without changing the isotopy types of the corresponding hypersurfaces.

To be more precise, from $\dim \operatorname{Conv}(\mathcal{A}) = n$ it follows that $\mathcal{A}$ contains $n+1$ affinely independent vectors $\alpha_1, \dots, \alpha_{n+1}$. Thus, there exists an invertible matrix $M \in \mathbb{R}^{n\times n}$ such that $M(\alpha_1 - \alpha_{n+1}) = e_1, \dots, M(\alpha_n - \alpha_{n+1}) = e_n$. If $\det(M) < 0$, we change the order of $\alpha_1, \dots, \alpha_n$ such that the corresponding matrix M has positive determinant. For $v := -M\alpha_{n+1}$, the affine linear map $L \colon \mathbb{R}^n \to \mathbb{R}^n$, $\alpha \mapsto M\alpha + v$ satisfies $\{0, e_1, \dots, e_n\} \subseteq L(\mathcal{A})$.

For each face $F \subseteq \operatorname{Conv}(\mathcal{A})$, we define $\nabla_{\mathcal{A}_F,\varepsilon_F}$ in a similar way, and set

$$\tilde{\nabla}_{\mathcal{A}_F,\varepsilon_F} := \left\{(c_{\alpha_i})_{i=1,\dots,n+k+1} \in \mathbb{R}_\varepsilon^{n+k+1} \mid (c_{\alpha_i})_{\alpha_i \in \mathcal{A}_F} \in \nabla_{\mathcal{A}_F,\varepsilon_F}\right\}.$$

In [3], the authors proved the following statement regarding the topology of hypersurfaces corresponding to different connected components of the complement of the union of the signed A-discriminants $\tilde{\nabla}_{\mathcal{A}_F,\varepsilon_F}$.

Proposition 2.6. *[3, Proposition 2.9] Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support such that $\mathcal{A}\subseteq\mathbb{Z}^n$. If c and c' are in the same connected component of*

$$\mathbb{R}^{\mathcal{A}}_{\varepsilon}\setminus\Big(\bigcup_{F\subseteq\mathrm{Conv}(\mathcal{A})\ a\ face}\tilde{\nabla}_{\mathcal{A}_F,\varepsilon_F}\Big),$$

then the zero sets $Z(f_c)$ and $Z(f_{c'})$ are homeomorphic.

Finding the defining equalities of the signed A-discriminant is challenging, but an explicit parametrization is much simpler to find. First, let $\mathrm{diag}(c)$ denote the $\#\mathcal{A}\times\#\mathcal{A}$ diagonal matrix with (a,a)-entry c_a, and let us rewrite the equalities in (3) as:

$$\hat{A}\,\mathrm{diag}(c)\Big(e^{\alpha_i\cdot x}\Big)_{i=1,\dots,n+k+1}=\hat{A}_{\varepsilon}\,\mathrm{diag}(|c|)\Big(e^{\alpha_i\cdot x}\Big)_{i=1,\dots,n+k+1}=0, \tag{5}$$

where the matrix $\hat{A}$ is given by

$$\hat{A}=\begin{bmatrix}1&\dots&1\\ \alpha_1&\dots&\alpha_{n+k+1}\end{bmatrix}\in\mathbb{R}^{(n+1)\times(n+k+1)}, \tag{6}$$

$\varepsilon=\mathrm{sign}(c)\in\{\pm1\}^{n+k+1}$ and $\hat{A}_{\varepsilon}=\hat{A}\,\mathrm{diag}(\varepsilon)$. We refer to the equation system (5) as the *critical system* of $(\mathcal{A},\varepsilon)$. Note that the assumption $\dim\mathrm{Conv}(\mathcal{A})=n$ is equivalent to $\mathrm{rk}(\hat{A})=n+1$ [38]. If $\mathrm{rk}(\hat{A})=n+1$, the kernel of $\hat{A}$ has dimension k, which is usually called the *codimension* of $\mathcal{A}$.

If $x\in\mathbb{R}^n$ is a singular zero of f_c, then $\mathrm{diag}(c)\big(e^{\alpha_i\cdot x}\big)_{i=1,\dots,n+k+1}\in\ker(\hat{A})$. Choose a basis of $\ker(\hat{A})$ and write these vectors as columns of a matrix $B\in\mathbb{R}^{(n+k+1)\times k}$. Such a choice of B is called a *Gale dual matrix* of $\hat{A}$. With slight abuse of notation, we might call B a Gale dual matrix of $\mathcal{A}$. Since $\mathrm{im}(B)=\ker(\hat{A})$ for each $x\in\mathrm{Sing}(f_c)$ there exists $\lambda\in\mathbb{R}^k$ such that

$$B\lambda=\mathrm{diag}(c)\big(e^{\alpha_i\cdot x}\big)_{i=1,\dots,n+k+1}. \tag{7}$$

Since $e^{\alpha_i\cdot x}$ is positive for all $x\in\mathbb{R}^n$ and all $\alpha_1,\dots,\alpha_{n+k+1}$, the signs of the vector in (7) are given by $\mathrm{sign}(c)=\varepsilon$. Therefore, it is enough to look at the set

$$\mathcal{C}_{B,\varepsilon}:=\Big\{\lambda\in\mathbb{R}^k\mid\mathrm{sign}(B\lambda)=\varepsilon\Big\}.$$

We define the *signed Horn-Kapranov Uniformization* map as

$$\psi\colon\mathcal{C}_{B,\varepsilon}\times\mathbb{R}^n\to\mathbb{R}^{n+k+1}_{\varepsilon},\quad(\lambda,x)\mapsto B\lambda*\big(e^{\alpha_i\cdot(-x)}\big)_{i=1,\dots,n+k+1}, \tag{8}$$

where $*$ denotes the point-wise multiplication of two vectors.

Proposition 2.7. *Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support with Gale dual matrix B. For the signed Horn-Kapranov Uniformization map* (8)*, it holds that*

$$\mathrm{im}(\psi)=\nabla_{\mathcal{A},\varepsilon}.$$

Proof. If $\psi(\lambda,x)=c$, then

$$\hat{A}\,\mathrm{diag}(c)\big(e^{\alpha_i\cdot x}\big)_{i=1,\dots,n+k+1}=\hat{A}\Big(B\lambda*\big(e^{\alpha_i\cdot -x}\big)_{i=1,\dots,n+k+1}*\big(e^{\alpha_i\cdot x}\big)_{i=1,\dots,n+k+1}\Big)=\hat{A}B\lambda=0,$$

which implies that $x\in\mathrm{Sing}(f_c)$ and therefore $c\in\nabla_{\mathcal{A},\varepsilon}$.

On the contrary, if $c\in\nabla_{\mathcal{A},\varepsilon}$, then there exists a point $x\in\mathrm{Sing}(f_c)$. From (7) follows that there exists $\lambda\in\mathcal{C}_{B,\varepsilon}$ such that $\psi(\lambda,x)=c$. □

The signed A-discriminant lives in an ambient space of dimension $n+k+1$. Following [26], we reduce the dimension of the ambient space to k by quotienting out some homogeneities without losing essential information as follows. We define the *signed reduced A-discriminant* $\Gamma_{\varepsilon}(A,B)$ [26, Definition 2.5] to be

$$\Gamma_{\varepsilon}(A,B):=B^{\top}\,\mathrm{Log}|\nabla_{\mathcal{A},\varepsilon}|,$$

where Log is the coordinate-wise natural logarithm map and $|\cdot|$ denotes the coordinate-wise absolute value map. In [26], bounded (resp. unbounded) connected components of $\mathbb{R}^k\setminus\Gamma_{\varepsilon}(A,B)$ have been called *signed reduced inner* (resp. *outer*) *chambers*.

Theorem 2.8. *[26, Theorem 3.8.] Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support with Gale dual matrix B and let $c, c' \in \mathbb{R}^{\mathcal{A}}_{\varepsilon}$. If $B^\top \operatorname{Log}|c|$ and $B^\top \operatorname{Log}|c'|$ are in the same connected component of*

$$\mathbb{R}^k \setminus \Big(\bigcup_{F \subseteq \operatorname{Conv}(\mathcal{A}) \text{ a face}} B^\top \operatorname{Log}|\tilde{\nabla}_{\mathcal{A}_F, \varepsilon_F}| \Big),$$

then the zero sets $Z(f_c)$ and $Z(f_{c'})$ are ambiently isotopic in $\mathbb{R}^n$.

The signed reduced A-discriminant admits a parametrization as well.

Proposition 2.9. *Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support with Gale dual matrix B. The image of the map*

$$\xi_{B,\varepsilon} \colon \mathcal{C}_{B,\varepsilon} \to \mathbb{R}^k, \quad \lambda \mapsto B^\top \operatorname{Log}|B\lambda| \tag{9}$$

is the signed reduced A-discriminant $\Gamma_\varepsilon(A, B)$.

Proof. Let $\operatorname{pr}_k \colon \mathcal{C}_{B,\varepsilon} \times \mathbb{R}^n \to \mathcal{C}_{B,\varepsilon}$ be the natural coordinate projection and ψ be the signed Horn-Kapranov Uniformization (8). Furthermore, denote $\tilde{A}$ the matrix obtained from $\hat{A}$ in (6) by removing its top row, that is, the columns of $\tilde{A}$ are given by the vectors in $\mathcal{A}$. For every $(\lambda, x) \in \mathcal{C}_{B,\varepsilon} \times \mathbb{R}^n$, we have

$$\psi(\lambda, x) = B^\top \operatorname{Log}|B\lambda * (e^{\alpha_i \cdot (-x)})_{i=1,\dots,n+k+1}| = B^\top \operatorname{Log}|B\lambda| - B^\top \tilde{A}^\top x = \xi_{B,\varepsilon}(\lambda),$$

where the last equality holds, since $\tilde{A}B = 0$. Thus, the following diagram commutes

$$\begin{array}{ccccc} \mathcal{C}_{B,\varepsilon} \times \mathbb{R}^n & \xrightarrow{\psi} & \mathbb{R}^{n+k+1}_{\varepsilon} & \xrightarrow{\operatorname{Log}|\cdot|} & \mathbb{R}^{n+k+1} \\ \downarrow{\scriptstyle \operatorname{pr}_k} & & & & \downarrow{\scriptstyle B^\top} \\ \mathcal{C}_{B,\varepsilon} & & \xrightarrow{\qquad\qquad \xi_{B,\varepsilon} \qquad\qquad} & & \mathbb{R}^k \end{array} \tag{10}$$

which gives that

$$\operatorname{im}(\xi_{B,\varepsilon}) = \operatorname{im}(\xi_{B,\varepsilon} \circ \operatorname{pr}_k) = \operatorname{im}(B^\top \circ \operatorname{Log}|\cdot| \circ \psi) = B^\top \operatorname{Log}|\nabla_{\mathcal{A},\varepsilon}| = \Gamma_\varepsilon(A, B),$$

where the first equality holds since pr_k is surjective, and the second-to-last equality follows from Proposition 2.7. □

Since the first row of the matrix $\hat{A}$ is given by the all one vector $\mathbb{1} \in \mathbb{R}^{n+k+1}$, we have $B^\top \mathbb{1} = 0$, which implies that the map $\xi_{B,\varepsilon}$ is homogeneous, i.e. for all $a \in \mathbb{R}$:

$$\xi_{B,\varepsilon}(a\lambda) = B^\top \operatorname{Log}|B(a\lambda)| = \log(|a|) B^\top \mathbb{1} + B^\top \operatorname{Log}|B\lambda| = B^\top \operatorname{Log}|B\lambda| = \xi_{B,\varepsilon}(\lambda).$$

Thus, one could projectivize the domain $\mathcal{C}_{B,\varepsilon} \subseteq \mathbb{R}^k$ of $\xi_{B,\varepsilon}$.

Remark 2.10. Modifying a Gale dual matrix using elementary column operations gives another choice of Gale dual matrix. Thus, one can assume without any restriction that the last row of the Gale dual matrix B has the form $B_{n+k+1} = (0, \dots, 0, -1)$. Such a choice of B fixes the sign $\varepsilon_{n+k+1} = -1$. Since $Z(f_c) = Z(f_{-c})$, we can always fix one of the signs of the coefficients.

Since $\xi_{B,\varepsilon}$ is homogeneous, one can replace $\mathbb{R}^k$ (assuming $B_{n+k+1} = (0, \dots, 0, -1)$) by the upper half of the $(k-1)-$sphere

$$C_{k-1} = \{\lambda \in \mathbb{R}^k \mid \|\lambda\| = 1,\ \lambda_k > 0\},$$

or by the $(k-1)-$dimensional affine subspace

$$\{\lambda \in \mathbb{R}^k \mid \lambda_k = 1\}.$$

In Section 4, we will prefer this latter choice and work with the map

$$\bar{\xi}_{B,\varepsilon} \colon \{\mu \in \mathbb{R}^{k-1} \mid \operatorname{sign}(B \begin{bmatrix} \mu \\ 1 \end{bmatrix}) = \varepsilon\} \to \mathbb{R}^k, \quad \lambda \mapsto B^\top \operatorname{Log}|B \begin{bmatrix} \mu \\ 1 \end{bmatrix}|. \tag{11}$$

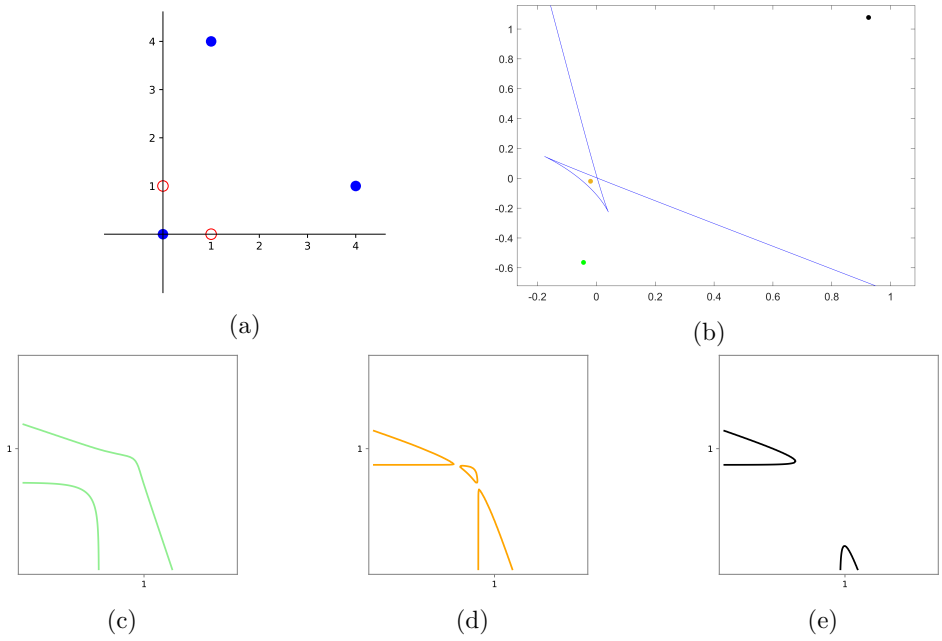


FIGURE 2. (a) Signed support from Example 2.11. The positive and negative exponent vectors are depicted by red circles and blue dots respectively. (b) Signed reduced A-discriminant of the signed support in (a). (c)(d)(e) Hypersurfaces $Z(f_c)$ corresponding to different connected components of the complement of the signed reduced A-discriminant.

Example 2.11. To give an illustration of the signed reduced A-discriminant, we recall [11, Example 2.9]. Consider the same set of exponent vectors as in Example 2.1

$$\mathcal{A} = \left\{\alpha_1 = \begin{bmatrix}0\\0\end{bmatrix}, \quad \alpha_2 = \begin{bmatrix}1\\0\end{bmatrix}, \quad \alpha_3 = \begin{bmatrix}0\\1\end{bmatrix}, \quad \alpha_4 = \begin{bmatrix}4\\1\end{bmatrix}, \quad \alpha_5 = \begin{bmatrix}1\\4\end{bmatrix}\right\}. \tag{12}$$

Unlike in Example 2.1, here we consider the sign distribution $\varepsilon = (-1, 1, 1, -1, -1)$. We depicted the signed support $(\mathcal{A}, \varepsilon)$ in Figure 2(a). Since $\mathcal{A}$ has codimension 2, the signed reduced A-discriminant $\Gamma_\varepsilon(A, B)$ is in the plane $\mathbb{R}^2$. For an illustration we refer to Figure 2(b). The complement of $\Gamma_\varepsilon(A, B)$ has 3 connected components. For the coefficient vectors $c = (-1, 6, 3, -1, -1)$,$(-1, 1, 1, -1, -1)$,$(-1, 0.5, 1, -1, -1)$, their projection $B^\top \operatorname{Log}|c|$ lies in different connected components of $\mathbb{R}^2 \backslash \Gamma_\varepsilon(A, B)$. The corresponding hypersurfaces $Z(f_c)$ are shown in Figure 2(c),(d),(e) respectively.

One remarkable property of this particular signed support is that the isotopy type of the hypersurface in Figure 2(d), corresponding to the coefficient $c = (-1, 1, 1 - 1, -1)$, cannot be obtained by Viro's patchworking (cf. Theorem 2.2). All possible signed tropical hypersurfaces $\operatorname{Trop}_\varepsilon(\mathcal{A}, h)$, with $h \in \mathbb{R}^{\mathcal{A}}$ generic, consist of 2 unbounded connected components, but the hypersurface $Z(f_c)$, $c = (-1, 1, 1 - 1, -1)$ has 3 connected components, one bounded and two unbounded.

2.4. **Some useful results from topology.** In the proof of Proposition 4.6, we need some classical results from topology. Let us also introduce these briefly here (see, for example, Chapter I.11 & Chapter IV.19 in [6]).

Lemma 2.12. *Suppose that X and Y are locally compact, Hausdorff spaces and that $f\colon X \to Y$ is continuous. Let X^+ be the one-point compactification space of X. Then f is proper (i.e., the preimage of any compact subset is compact) $\iff$ f extends to a continuous map $f^+\colon X^+ \to Y^+$ by setting $f^+(\infty_X) = \infty_Y$.*

Lemma 2.13. *(Jordan-Brouwer Separation Theorem) If $f\colon \mathbf{S}^{n-1} \to \mathbf{S}^n$ (where $\mathbf{S}^n$ denotes n-sphere) is an injective continuous map, then $\mathbf{S}^n \setminus f(\mathbf{S}^{n-1})$ consists of exactly two connected components. Moreover, $f(\mathbf{S}^{n-1})$ is the topological boundary of each of these components.*

Corollary 2.14. *If $f\colon \mathbb{R}^n \to \mathbb{R}^{n+1}$ is injective, continuous and proper, then $\mathbb{R}^{n+1} \setminus f(\mathbb{R}^n)$ consists of exactly two unbounded connected components.*

Proof. Note that the one point compactification of $\mathbb{R}^n$ is $\mathbf{S}^n$. By Lemma 2.12, f can be extended to $f^+\colon \mathbf{S}^n \to \mathbf{S}^{n+1}$ with $f^+(\infty) = \infty$. Then f^+ is also injective. By Lemma 2.13, $\mathbf{S}^{n+1} \setminus f^+(\mathbf{S}^n)$ consists of two connected components and the point ∞ is in the boundary. Since the point ∞ is in the boundary of each component, we can always find a sequence $\{x_l\} \subseteq \mathbb{R}^{n+1} = \mathbf{S}^{n+1} \setminus \{\infty\}$ in each of the components such that $x_l \to \infty$. Therefore, these two components are unbounded in $\mathbb{R}^{n+1}$. □

Finally, we need the following version of the mean value theorem:

Lemma 2.15. *(Cauchy's Mean Value Theorem) If the functions $f, g\colon [a,b] \to \mathbb{R}$ are both continuous and differentiable on the open interval (a,b), then there exists some $c \in (a,b)$, such that*

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

3. Signed supports without singular zeros

In this section, we give a necessary and sufficient condition on the signed support $(\mathcal{A}, \varepsilon)$ such that the signed reduced A-discriminant $\nabla_{\mathcal{A},\varepsilon}$ is empty. Building on this result and Theorem 2.8, we give conditions on $(\mathcal{A}, \varepsilon)$ such that for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$ the hypersurfaces $Z(f_c)$ have the same isotopy type (Theorem 3.5). First, we start with a simple observation.

Proposition 3.1. *Let $\mathcal{A} = \{\alpha_1, \dots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$ be a set of exponent vectors and $\varepsilon \in \{\pm 1\}^{n+k+1}$ be a fixed sign distribution. Let $\hat{A}$ the matrix defined in* (6).

(a) *If $\ker(\hat{A}) \cap \mathbb{R}_\varepsilon^{n+k+1} = \ker(\hat{A}_\varepsilon) \cap \mathbb{R}_{>0}^{n+k+1} = \emptyset$, then for all $c \in \mathbb{R}_\varepsilon^{n+k+1}$ the critical system* (5) *does not have any solution $x \in \mathbb{R}^n$.*

(b) *If $\ker(\hat{A}_\varepsilon) \cap \mathbb{R}_{>0}^{n+k+1} \neq \emptyset$, then there exists $c \in \mathbb{R}_\varepsilon^{n+k+1}$ such that the critical system has a solution $x \in \mathbb{R}^n$.*

Proof. Part (a) follows directly, since for any $c \in \mathbb{R}_\varepsilon^{n+k+1}$ and any solution $x \in \mathbb{R}^n$ of (5), we have $\operatorname{diag}(c)\big(e^{\alpha_i \cdot x}\big)_{i=1,\dots,n+k+1} \in \ker(\hat{A}) \cap \mathbb{R}_\varepsilon^{n+k+1}$.

If $v \in \ker(\hat{A}_\varepsilon) \cap \mathbb{R}_{>0}^{n+k+1}$, then for $c = \operatorname{diag}(\varepsilon)v$, the point $x = (0, \dots, 0)$ is a solution of (5). □

In the following, we interpret the conditions in Proposition 3.1 in terms of the geometry of the support $\mathcal{A}$ and the sign distribution ε. An *affine hyperplane* is a set of the form

$$\mathcal{H}_{v,a} := \{\mu \in \mathbb{R}^n \mid v \cdot \mu = a\},$$

for some $v \in \mathbb{R}^n \setminus \{0\}$ and $a \in \mathbb{R}$. Each affine hyperplane defines two half-spaces

$$\mathcal{H}_{v,a}^+ := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \geq a\}, \quad \mathcal{H}_{v,a}^- := \{\mu \in \mathbb{R}^n \mid v \cdot \mu \leq a\}.$$

Following [9], we call $\mathcal{H}_{v,a}$ a *separating hyperplane* of $(\mathcal{A}, \varepsilon)$ if

$$\mathcal{A}_+ \subseteq \mathcal{H}_{v,a}^+, \quad \text{and} \quad \mathcal{A}_- \subseteq \mathcal{H}_{v,a}^- . \tag{13}$$

A separating hyperplane $\mathcal{H}_{v,a}$ is called *non-trivial*, if at least one of the open half-spaces $\operatorname{int}(\mathcal{H}_{v,a}^+)$, $\operatorname{int}(\mathcal{H}_{v,a}^-)$ contains a point of $\mathcal{A}$. A non-trivial separating hyperplane is called *very strict* if $\mathcal{H}_{v,a}$ does not contain any point in $\mathcal{A}$. For an illustration of separating hyperplanes, we refer to Figure 3.

Proposition 3.2. *A signed support $(\mathcal{A}, \varepsilon)$ has a non-trivial separating hyperplane if and only if $\ker(\hat{A}_\varepsilon) \cap \mathbb{R}_{>0}^{n+k+1} = \emptyset$, where $\hat{A} \in \mathbb{R}^{(n+1)\times(n+k+1)}$ denotes the matrix from* (6).

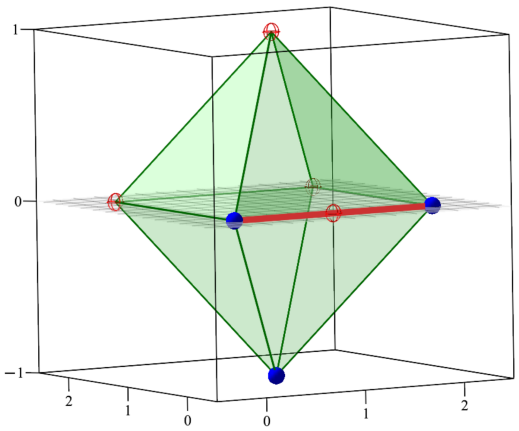


FIGURE 3. The hyperplane $\mathcal{H}_{v,0}$ with $v = (0,0,1)$ is a non-trivial separating hyperplane of $\mathcal{A}_+ = \{(1,0,0)^\top, (2,2,0)^\top, (0,2,0)^\top, (1,1,1)^\top\}$ (depicted as red circles) and $\mathcal{A}_- = \{(0,0,0)^\top, (2,0,0)^\top, (1,1,-1)^\top\}$ (blue dots). For the face $F = \operatorname{Conv}\left((0,0,0)^\top, (0,0,2)^\top\right)$ (marked by thick line segment), the restricted signed support $\mathcal{A}_{F,+} = \{(1,0,0)^\top\}$, $\mathcal{A}_{F,-} = \{(0,0,0)^\top, (2,0,0)^\top\}$ does not have any non-trivial separating hyperplane.

Proof. By Stiemke's Theorem [30] (which is a version of Farkas' Lemma, see also [38, Section 6.2]), exactly one of the following holds. Either there exists $w \in \mathbb{R}^{n+1}$ such that

$$\hat{A}_\varepsilon^\top w \geq 0, \tag{14}$$

and at least one of the inequalities is strict, or there exists $u \in \mathbb{R}^{n+k+1}_{>0}$ such that

$$\hat{A}_\varepsilon u = 0, \tag{15}$$

but not both. Condition (15) is equivalent to $\ker(\hat{A}_\varepsilon) \cap \mathbb{R}^{n+k+1}_{>0} \neq \emptyset$. Note that one can rewrite (14) as

$$\varepsilon_i((w_2, \ldots, w_{n+1}) \cdot \alpha_i) \geq \varepsilon_i(-w_1),$$

for all $\alpha_i \in \mathcal{A}$. Thus, if such a w exists, then $\mathcal{H}_{v,a}$ with $v = (w_2, \ldots, w_{n+1})$, $a = -w_1$ is a non-trivial separating hyperplane of $(\mathcal{A}, \varepsilon)$. On the other hand, if $\mathcal{H}_{v,a}$ is a non-trivial hyperplane of $(\mathcal{A}, \varepsilon)$, then $w = (-a, v)$ satisfies (14). □

Theorem 3.3. *Let $(\mathcal{A}, \varepsilon)$ be a signed support with Gale dual matrix B. Then the following are equivalent:*

(i) $\nabla_{\mathcal{A},\varepsilon} = \emptyset$
(ii) $\Gamma_\varepsilon(A, B) = \emptyset$
(iii) *$(\mathcal{A}, \varepsilon)$ has a non-trivial separating hyperplane.*

Proof. The equivalence between (i) and (ii) follows directly from the definition of the signed reduced A-discriminant, since $\Gamma_\varepsilon(A, B) = B^\top \operatorname{Log}|\nabla_{\mathcal{A},\varepsilon}|$.

From Proposition 3.1 it follows that $\nabla_{\mathcal{A},\varepsilon} = \emptyset$ if and only if $\ker(\hat{A}_\varepsilon) \cap \mathbb{R}^{n+k+1}_{>0} = \emptyset$, which is equivalent to the existence of a non-trivial separating hyperplane of $(\mathcal{A}, \varepsilon)$ by Propositon 3.2. This shows that (i) and (ii) are equivalent. □

For fixed set of exponent vectors $\mathcal{A} \subseteq \mathbb{R}^n$, using the correspondence between hyperplane arrangements and zonotopes, one derives a bound on the number of sign distributions for which $(\mathcal{A}, \varepsilon)$ does not have a non-trivial separating hyperplane.

Proposition 3.4. *Let $\mathcal{A} = \{\alpha_1, \ldots, \alpha_{n+k+1}\} \subseteq \mathbb{R}^n$ be a finite set such that $\dim \operatorname{Conv}(\mathcal{A}) = n$. The number of sign distributions $\varepsilon \in \{\pm 1\}^{n+k+1}$ for which $(\mathcal{A}, \varepsilon)$ does not have a non-trivial*

separating hyperplane is bounded above by:

$$2\sum_{i=0}^{k-1}\binom{n+k}{i}.$$

Proof. Let $\hat{A}$ be the matrix defined in (6). By Proposition 3.2, the signed support $(\mathcal{A},\varepsilon)$ does not have a non-trivial separating hyperplane if and only if $\ker(\hat{A})\cap\mathbb{R}_\varepsilon^{n+k+1}\neq\emptyset$. So all we have to do is to count how many orthants $\mathbb{R}_\varepsilon^{n+k+1}$ $\ker(\hat{A})$ intersects.

The assumption $\dim \mathrm{Conv}(\mathcal{A}) = n$ implies that $\dim\ker(\hat{A}) = k$. Let $B\in\mathbb{R}^{(n+k+1)\times k}$ be Gale dual to $\hat{A}$ and denote by $B_1,\dots,B_{n+k+1}$ the rows of B. By [12, Lemma 0.16] (see also [38, Corollary 7.17]), the orthants $\mathbb{R}_\varepsilon^{n+k+1}$ that $\mathrm{im}(B)=\ker(\hat{A})$ intersects, correspond one-to-one to the vertices of the zonotope

$$[-B_1,B_1]+\cdots+[-B_{n+k+1},B_{n+k+1}]\subseteq\mathbb{R}^k.$$

By [13, Table 2.1] (see also [37]) such a zonotope can have at most

$$2\sum_{i=0}^{k-1}\binom{n+k}{i}$$

many vertices. □

We finish the section by characterizing signed supports for which all corresponding hypersurfaces have the same isotopy type.

Theorem 3.5. *Let $(\mathcal{A},\varepsilon)$ be a signed support. If for all faces $F\subseteq\mathrm{Conv}(\mathcal{A})$ the signed support $(\mathcal{A}_F,\varepsilon_F)$ has a non-trivial separating hyperplane, then for all $c\in\mathbb{R}_\varepsilon^{\mathcal{A}}$ the hypersurfaces $Z(f_c)$ have the same isotopy type.*

Proof. From Theorem 3.3 follows that the signed reduced A-discriminants associated to the faces $F\subseteq\mathrm{Conv}(\mathcal{A})$ are empty. Thus, all the hypersurfaces $Z(f_c)$, $c\in\mathbb{R}_\varepsilon^{\mathcal{A}}$ have the same isotopy type by Theorem 2.8. □

Corollary 3.6. *If a signed support $(\mathcal{A},\varepsilon)$ has a very strict separating hyperplane, then the hypersurfaces $Z(f_c)$ have the same isotopy type for all $c\in\mathbb{R}_\varepsilon^{\mathcal{A}}$.*

Proof. If $\mathcal{H}_{v,a}$ is a very strict separating hyperplane of $(\mathcal{A},\varepsilon)$, then it is also a very strict separating hyperplane of $(\mathcal{A}_F,\varepsilon_F)$ for all faces $F\subseteq\mathrm{Conv}(\mathcal{A})$. Now, the statement follows from Theorem 3.5. □

Example 3.7. The signed support $(\mathcal{A},\varepsilon)$ from Example 2.1 has a very strict separating hyperplane. Thus, by Corollary 3.6, all hypersurfaces $Z(f_c)$, $c\in\mathbb{R}_\varepsilon^{\mathcal{A}}$ have the same isotopy type. From Theorem 2.2 follows that this isotopy type agrees with the isotopy type of the signed tropical hypersurface $\mathrm{Trop}_\varepsilon(\mathcal{A},h)$ for every generic $h\in\mathbb{R}^{\mathcal{A}}$. We refer to Figure 1(b) for an illustration of $\mathrm{Trop}_\varepsilon(\mathcal{A},h)$ with $h=(5,4,4,5,5)$.

Remark 3.8. If a signed support $(\mathcal{A},\varepsilon)$ has a non-trivial separating hyperplane, it might happen that for one of the faces $F\subseteq\mathrm{Conv}(\mathcal{A})$ the restricted signed support $(\mathcal{A}_F,\varepsilon_F)$ does not have a non-trivial separating hyperplane. For such an example, we revisit the signed support from Figure 3. The face $F=\mathrm{Conv}((0,0,0)^\top,(2,0,0)^\top)$, contains two negative exponent vectors $\alpha_1=(0,0,0)^\top,\alpha_2=(2,0,0)^\top$ and one positive exponent vector $\alpha_3=(1,0,0)^\top$. Since α_3 lies in the relative interior of $\mathrm{Conv}(\alpha_1,\alpha_2)$, it follows that $\{\alpha_1,\alpha_2\}$ and $\{\alpha_3\}$ cannot be separated by an affine hyperplane.

4. A-discriminants with two signed reduced outer chambers

The goal of this section is to describe conditions on the signed support $(\mathcal{A},\varepsilon)$ that ensure $\mathbb{R}^k\setminus\Gamma_\varepsilon(A,B)$ has at most two connected components, which are unbounded. In Section 4.1, we focus on the case where $\mathcal{A}$ has exactly $n+3$ exponent vectors. We show that the complement of

$\Gamma_\varepsilon(A,B)$ has at most two chambers if the parametrization map $\bar{\xi}_{B,\varepsilon}$ has at most one critical point (Proposition 4.6). It is known that $\bar{\xi}_{B,\varepsilon}$ can have at most n critical points [27], however there did not exist any known example in the literature where this bound is attained. In Example 4.3, we describe a family of signed supports such that $\bar{\xi}_{B,\varepsilon}$ has n critical points for every $n \in \mathbb{N}$.

In Section 4.2, we investigate the relation between critical points of $\bar{\xi}_{B,\varepsilon}$ and degenerate singular points of $Z(f_c)$, and show that if $Z(f_c)$ has no degenerate singular point for all $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ and the codimension of $\mathcal{A}$ is 2, then $\mathbb{R}^k \setminus \Gamma_\varepsilon(A,B)$ has at most two connected components (Theorem 4.9). In Section 4.3, we give several conditions on the geometry of the signed support $(\mathcal{A},\varepsilon)$ precluding the existence of degenerate singular points in $Z(f_c)$, $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$.

During the whole section, we assume that the Gale dual matrix $B \in \mathbb{R}^{(n+k+1)\times k}$ is chosen in a way such that its last row has the form $(0,\dots,0,-1)$ (cf. Remark 2.10).

4.1. **Critical points of the signed reduced A-discriminant.** Let $\mathcal{A} = \{\alpha_1,\dots,\alpha_{n+k+1}\} \subseteq \mathbb{R}^n$ be a set of exponent vectors such that $\dim \mathrm{Conv}(\mathcal{A}) = n$ and fix a sign distribution $\varepsilon \in \{\pm 1\}^{n+k+1}$. Let $\hat{A} \in \mathbb{R}^{(n+1)\times(n+k+1)}$ be as given in (6) and choose a Gale dual matrix $B \in \mathbb{R}^{(n+k+1)\times k}$ with rows $B_1,\dots,B_{n+k+1}$ such that its last row has the form $B_{n+k+1} = (0,\dots,0,-1)$.

Let $\bar{\xi}_{B,\varepsilon}$ be the parametrization map of $\Gamma_\varepsilon(A,B)$ as defined in (11). Following [1, Section 1.2], we call a point $\mu \in \mathbb{R}^{k-1}$ a *critical point* of $\bar{\xi}_{B,\varepsilon}$ if $\bar{\xi}_{B,\varepsilon}(\mu)$ is well-defined and the Jacobian matrix $J_{\bar{\xi}_{B,\varepsilon}}(\mu)$ does not have full rank, that is, it has rank strictly less than $k-1$.

Lemma 4.1. *Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support with Gale dual matrix B, and let $\bar{\xi}_{B,\varepsilon}$ be as defined in* (11)*. Then for each $\mu \in \mathbb{R}^{k-1}$ where $\bar{\xi}_{B,\varepsilon}$ is defined, we have the following equality for the Jacobian matrix*

$$J_{\bar{\xi}_{B,\varepsilon}}(\mu) = B^\top \operatorname{diag}\left(\tfrac{1}{B\hat{\mu}}\right)\tilde{B}, \tag{16}$$

where $\tilde{B}$ denotes the matrix obtained from B by deleting its last column.

Proof. By definition (cf. (11)), the $j-$th coordinate of $\bar{\xi}_{B,\varepsilon}$ is given by

$$\big(\bar{\xi}_{B,\varepsilon}(\mu)\big)_j = \sum_{i=1}^{n+k+1} \log|B_i\cdot\hat{\mu}|B_{i,j},$$

where $\hat{\mu} = \begin{bmatrix}\mu\\1\end{bmatrix}$. Thus, the partial derivatives of $\bar{\xi}_{B,\varepsilon}$ have the form

$$\frac{\partial\big(\bar{\xi}_{B,\varepsilon}(\mu)\big)_j}{\partial\mu_\ell} = \sum_{i=1}^{n+k+1}\frac{B_{i,j}B_{i,\ell}}{B_i\cdot\hat{\mu}} \tag{17}$$

for all $j = 1,\dots,k$, and $\ell = 1,\dots,k-1$. Comparing (17) with the entries of the right-hand side of (16) the result follows. □

Using Lemma 4.1, an easy computation shows that

$$\hat{\mu}^\top J_{\bar{\xi}_{B,\varepsilon}}(\mu) = \left(B\hat{\mu}\right)^\top \operatorname{diag}\left(\tfrac{1}{B\hat{\mu}}\right)\tilde{B} = \mathbb{1}^\top\tilde{B} = 0. \tag{18}$$

Therefore, if $\bar{\xi}_{B,\varepsilon}$ is differentiable at μ, then a normal vector at $\bar{\xi}_{B,\varepsilon}(\mu)$ is given by $\hat{\mu} = \begin{bmatrix}\mu\\1\end{bmatrix}$. This statement was proven by Kapranov [21, Theorem 2.1].

In the remaining of the section, we focus on the case $k = 2$. Under this assumption, we have

$$J_{\bar{\xi}_{B,\varepsilon}}(\mu) = \begin{bmatrix} b_1^\top \operatorname{diag}\left(\frac{1}{B\hat{\mu}}\right)b_1 \\ b_2^\top \operatorname{diag}\left(\frac{1}{B\hat{\mu}}\right)b_1 \end{bmatrix},$$

where b_1, b_2 denote the first and second column of the Gale dual matrix B.

Lemma 4.2. *Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support of codimension 2 with Gale dual $B\in\mathbb{R}^{(n+3)\times 2}$, and let $\bar{\xi}_{B,\varepsilon}$ be as defined in (11). For $\mu\in\mathbb{R}\setminus\{0\}$ the following are equivalent.*

(i) *μ is a critical point of $\bar{\xi}_{B,\varepsilon}$.*

(ii) *$\operatorname{sign}(B\hat{\mu})=\varepsilon$ and μ is a zero of the univariate polynomial*

$$q_B(\mu) := \Big(\prod_{i=1}^{n+3}(B\hat{\mu})_i\Big) b_1^\top \operatorname{diag}\Big(\tfrac{1}{B\hat{\mu}}\Big) b_1. \tag{19}$$

(iii) *$\operatorname{sign}(B\hat{\mu})=\varepsilon$ and μ is a zero of the univariate polynomial*

$$\tilde{q}_B(\mu) := \Big(\prod_{i=1}^{n+3}(B\hat{\mu})_i\Big) b_2^\top \operatorname{diag}\Big(\tfrac{1}{B\hat{\mu}}\Big) b_1. \tag{20}$$

Moreover, q_B has degree at most n and $\bar{\xi}_{B,\varepsilon}$ has at most n critical points.

Proof. Note that $\bar{\xi}_{B,\varepsilon}(\mu)$ is defined only if $\operatorname{sign}(B\hat{\mu})=\varepsilon$. Furthermore, the factor $\prod_{i=1}^{n+3}(B\hat{\mu})_i$ clears the denominator of $b_1^\top \operatorname{diag}\Big(\frac{1}{B\hat{\mu}}\Big)b_1$ and of $b_2^\top \operatorname{diag}\Big(\frac{1}{B\hat{\mu}}\Big)b_1$. Moreover, $\prod_{i=1}^{n+3}(B\hat{\mu})_i\neq 0$ if $\operatorname{sign}(B\hat{\mu})=\varepsilon$. From (18) follows that

$$\mu b_1^\top \operatorname{diag}\Big(\tfrac{1}{B\hat{\mu}}\Big)b_1 = -b_2^\top \operatorname{diag}\Big(\tfrac{1}{B\hat{\mu}}\Big)b_1.$$

Thus, $J_{\bar{\xi}_{B,\varepsilon}}(\mu)=0$ if and only if $q_B(\mu)=0$, which is also equivalent to $\tilde{q}_B(\mu)=0$.

The polynomial q_B has been studied previously in [27], where it has been shown that its degree is at most n. □

In [27, Theorem 3.10], the author constructed several matrices $B\in\mathbb{R}^{(n+3)\times 2}$ such that $q_B(\mu)$ has exactly n real roots. However, these roots correspond to different sign distributions ε. To show that the bound on the critical points of $\bar{\xi}_{B,\varepsilon}$ in Lemma 4.2 is attained, one needs to construct $B\in\mathbb{R}^{(n+3)\times 2}$ such that q_B (or $\tilde{q}_B$) has n real roots $\mu_1,\dots,\mu_n$ such that $\operatorname{sign}(B\hat{\mu}_1)=\cdots=\operatorname{sign}(B\hat{\mu}_n)=\varepsilon$ for some fixed $\varepsilon\in\{\pm 1\}^{\mathcal{A}}$. We provide such a construction in the following example.

Example 4.3. Let $n\in\mathbb{N}$ and let $\mu_1,\dots,\mu_n$ be distinct positive numbers different from 1. Consider the univariate polynomials $f(\mu)=(\mu-\mu_1)(\mu-\mu_2)\cdots(\mu-\mu_n)$, $g(\mu)=(\mu+\mu_1)(\mu+\mu_2)\cdots(\mu+\mu_n)(\mu+1)\in\mathbb{R}[\mu]$. Since $\deg(f)=n<\deg(g)=n+1$, the fraction $\frac{f(\mu)}{g(\mu)}$ admits a partial fraction decomposition:

$$\frac{(\mu-\mu_1)(\mu-\mu_2)\cdots(\mu-\mu_n)}{(\mu+\mu_1)(\mu+\mu_2)\cdots(\mu+\mu_n)(\mu+1)} = \frac{a_1}{\mu+\mu_1}+\frac{a_2}{\mu+\mu_2}+\cdots+\frac{a_n}{\mu+\mu_n}+\frac{a_{n+1}}{\mu+1}, \tag{21}$$

where $a_1,\dots,a_{n+1}\in\mathbb{R}$. The a_i's satisfy the following properties:

(1) $\dfrac{a_1}{\mu_1}+\dfrac{a_2}{\mu_2}+\dfrac{a_3}{\mu_3}+\cdots+\dfrac{a_n}{\mu_n}+a_{n+1}=(-1)^n$,

(2) $a_1+a_2+\cdots+a_{n+1}=1$.

(3) $a_i\neq 0$, $i=1,\dots,n+1$.

Property (1) follows by plugging in $\mu=0$, (2) follows by comparing the leading coefficients of numerators on both sides, (3) follows by comparing the degree of denominators on both sides.

We use the a_i's to build the matrix:

$$B=\begin{bmatrix} \frac{a_1}{\mu_1} & \frac{a_2}{\mu_2} & \frac{a_3}{\mu_3} & \cdots & \frac{a_n}{\mu_n} & a_{n+1} & (-1)^{n+1} & 0\\ a_1 & a_2 & a_3 & \cdots & a_n & a_{n+1} & 0 & -1\end{bmatrix}^\top.$$

Properties (1) and (2) imply that $\mathbb{1}^\top B=0$, thus it is possible to choose a matrix $\hat{A}\in\mathbb{R}^{(n+1)\times(n+3)}$ as in (6) such that B is its Gale dual. Denoting by b_1,b_2 the columns of B, we have:

$$b_2^\top \operatorname{diag}\Big(\tfrac{1}{B\hat{\mu}}\Big)b_1 = \frac{\frac{a_1^2}{\mu_1}}{\frac{a_1}{\mu_1}\mu+a_1}+\frac{\frac{a_2^2}{\mu_2}}{\frac{a_2}{\mu_2}\mu+a_2}+\frac{\frac{a_3^2}{\mu_3}}{\frac{a_3}{\mu_3}\mu+a_3}+\cdots+\frac{\frac{a_n^2}{\mu_n}}{\frac{a_n}{\mu_n}\mu+a_n}+\frac{a_{n+1}^2}{a_{n+1}\mu+a_{n+1}}.$$

The right-hand side of this equality agrees with the right-hand side of (21), as the a_i's are nonzero. Therefore, the zeros of $\tilde{q}_B(\mu) = b_2^\top \operatorname{diag}\left(\frac{1}{B\hat{\mu}}\right) b_1$ are $\mu_1, \ldots, \mu_n$. Since $\mu_1, \ldots, \mu_n$ are positive, it follows for all $i = 1, \ldots, n$ that

$$\operatorname{sign}(B\hat{\mu}_i) = (\operatorname{sign}(a_1), \operatorname{sign}(a_2), \ldots, \operatorname{sign}(a_{n+1}), (-1)^{n+1}, -1) =: \varepsilon.$$

We conclude that $\bar{\xi}_{B,\varepsilon}$ has n critical points using Lemma 4.2.

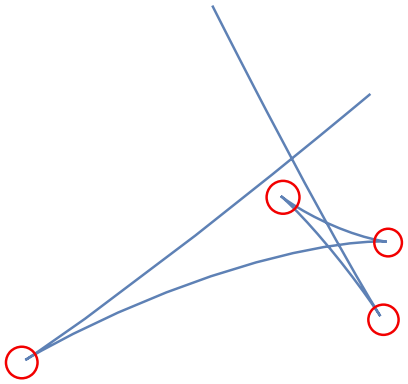
Let $n = 4$ and pick $\mu_1 = 5, \mu_2 = 6, \mu_3 = 7, \mu_4 = 8$. We have

$$\begin{aligned}&\frac{(\mu-5)(\mu-6)(\mu-7)(\mu-8)}{(\mu+5)(\mu+6)(\mu+7)(\mu+8)(\mu+1)}\\ =&\frac{-715}{\mu+5} + \frac{\frac{12012}{5}}{\mu+6} + \frac{-2730}{\mu+7} + \frac{1040}{\mu+8} + \frac{\frac{18}{5}}{\mu+1}\end{aligned}$$

and

$$B = \begin{bmatrix} -143 & \frac{2002}{5} & -390 & 130 & \frac{18}{5} & -1 & 0 \\ -715 & \frac{12012}{5} & -2730 & 1040 & \frac{18}{5} & 0 & -1 \end{bmatrix}^\top.$$

By the above, the map $\bar{\xi}_{B,\varepsilon}$ has 4 critical points for $\varepsilon = (-1, 1, -1, 1, 1, -1, -1)$. The signed reduced A-discriminant is drawn to the right and its critical points are highlighted by red circles.



In Example 2.11 (cf. Figure 2(b)), we saw that the complement of $\Gamma_\varepsilon(A, B)$ has three connected components if $\bar{\xi}_{B,\varepsilon}$ has 2 critical points. In the following, we show that if $\bar{\xi}_{B,\varepsilon}$ has at most one critical point, then $\mathbb{R}^k \setminus \Gamma_\varepsilon(A, B)$ cannot have more then two connected components. Before we get to this result, let us prove the following lemma for self-intersections of curves.

Lemma 4.4. *Let $\varphi\colon \mathbb{R} \to \mathbb{R}^2$ be a smooth map such that the Jacobian matrix $J_\varphi(\mu)$ has full rank for all $\mu \in \mathbb{R}$ except for at most one point. Let $S \subseteq \mathbb{R}^2$ be the curve parametrized by φ. If there exist two distinct points $a, b \in \mathbb{R}$ such that $\varphi(a) = \varphi(b)$, then there exist two distinct points $\mu_1, \mu_2 \in \mathbb{R}$ such that the tangent lines of S at $\varphi(\mu_1)$ and at $\varphi(\mu_2)$ are parallel.*

Proof. Denote φ_1, φ_2 the first and the second coordinate of φ. Suppose $a < b$, and assume there exists $t \in \mathbb{R}$ such that $\varphi_1'(t) = \varphi_2'(t) = 0$, that is, $J_\varphi(t)$ does not have full rank. We start by choosing $c \in \mathbb{R}$ such that $a < c < b$, $\varphi(c) \neq \varphi(a)$ and φ is smooth on both intervals (a, c) and (c, b). If $t \leq a$ or $t \geq b$, such c exists since the Jacobian matrix J_φ has full rank on (a, b). If $a < t < b$ and $\varphi(t) = \varphi(a)$, then the curve is smooth on the interval (a, t), and we pick a c as before. Finally, if $a < t < b$ and $\varphi(t) \neq \varphi(a)$, then we choose $c = t$. If φ does not have any singular point, then we pick c as in the case $t \leq a$ or $t \geq b$.

By Lemma 2.15, on the interval (a, c), there exists some $\mu_1 \in (a, c)$ such that

$$(\varphi_1(c) - \varphi_1(a))\varphi_2'(\mu_1) = (\varphi_2(c) - \varphi_2(a))\varphi_1'(\mu_1).$$

Similarly, on the interval (c, b), there exists some $\mu_2 \in (c, b)$ such that

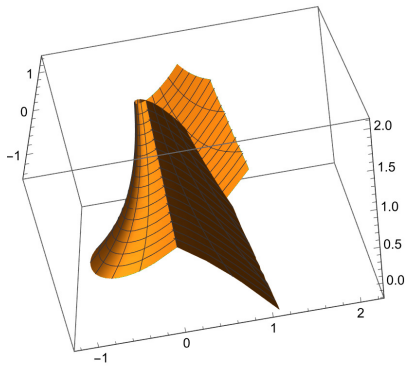
$$(\varphi_1(c) - \varphi_1(b))\varphi_2'(\mu_2) = (\varphi_2(c) - \varphi_2(b))\varphi_1'(\mu_2).$$

Thus, since $\varphi(a) = \varphi(b)$ and $\varphi(c) \neq \varphi(a)$, we have

$$\varphi_1'(\mu_1)\varphi_2'(\mu_2) = \varphi_1'(\mu_2)\varphi_2'(\mu_1)$$

and hence the tangent lines at μ_1, μ_2 are parallel. □

Remark 4.5. Lemma 4.4 is not true for hypersurfaces in $\mathbb{R}^n$ when $n \geq 3$. The surface given by $\varphi\colon (t,s) \mapsto (e^{-s}(t^2-1), e^{-s}t(t^2-1), s)$ is a counterexample for $n=3$. The map φ is not injective but there are no pairs of points with parallel tangent planes. The image of φ is shown on the right.



Now we are able to prove the following result bounding the number of connected components of $\mathbb{R}^k \setminus \Gamma_\varepsilon(A,B)$.

Proposition 4.6. *Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support of codimension 2 with Gale dual matrix $B \in \mathbb{R}^{(n+3)\times 2}$. If $\bar{\xi}_{B,\varepsilon}$ has at most one critical point, then the complement of the signed reduced A-discriminant $\Gamma_\varepsilon(A,B)$ has at most two connected components, which are unbounded.*

Proof. Recall that for $\mu \in \mathbb{R}$, we used the notation $\hat{\mu} = \begin{bmatrix}\mu\\1\end{bmatrix}$. If $(\mathcal{A},\varepsilon)$ has a non-trivial separating hyperplane, then $\Gamma_\varepsilon(A,B) = \emptyset$ by Theorem 3.3. Thus, $\mathbb{R}^k \setminus \Gamma_\varepsilon(A,B)$ has one connected component. If $(\mathcal{A},\varepsilon)$ does not have a non-trivial separating hyperplane, then from Proposition 3.2 follows that there exist $\mu_1,\mu_2 \in \mathbb{R}$ such that $\operatorname{sign}(B\hat{\mu}_1) = \operatorname{sign}(B\hat{\mu}_2) = \varepsilon$. By (18), $\hat{\mu}_1$ and $\hat{\mu}_2$ are normal vectors at $\bar{\xi}_{B,\varepsilon}(\mu_1)$ and at $\bar{\xi}_{B,\varepsilon}(\mu_2)$ respectively. If the tangent lines at $\bar{\xi}_{B,\varepsilon}(\mu_1)$ and at $\bar{\xi}_{B,\varepsilon}(\mu_2)$ are parallel, then $\hat{\mu}_1 = \lambda\hat{\mu}_2$ for some $\lambda \in \mathbb{R}\setminus\{0\}$, which implies that $\mu_1 = \mu_2$. This shows that there is no pair of points in $\Gamma_\varepsilon(A,B)$ with parallel tangent lines.

Lemma 4.4 implies that $\bar{\xi}_{B,\varepsilon}$ is injective. Also, $\bar{\xi}_{B,\varepsilon}$ maps an open interval of $\mathbb{R}$ to $\mathbb{R}^2$, and $\bar{\xi}_{B,\varepsilon}(\mu) \to \infty$ as μ approaches the endpoints of the interval. Therefore, $\bar{\xi}_{B,\varepsilon}$ is proper by Lemma 2.12, which implies that the complement of $\Gamma_\varepsilon(A,B)$ has exactly two unbounded connected components by Corollary 2.14. □

4.2. **Critical points and degenerate singularities.** Let f_c be an exponential sum as in (1). A singular point $x \in \operatorname{Sing}(f_c)$ is called *degenerate* if the Hessian matrix $\operatorname{Hess}_{f_c}(x)$ is not invertible. We have the following relationship between critical points of the signed reduced A-discriminant and degenerate singular points in the corresponding hypersurface.

Lemma 4.7. *Let $(\mathcal{A},\varepsilon)$ be a full-dimensional signed support with Gale dual matrix $B \in \mathbb{R}^{(n+k+1)\times k}$. If $\mu^* \in \mathbb{R}^{k-1}$ is a critical point of $\bar{\xi}_{B,\varepsilon}$, then for $c^* = B\begin{bmatrix}\mu^*\\1\end{bmatrix}$, the point $x^* = (0,\dots,0) \in \mathbb{R}^n$ is a degenerate singular point of f_{c^*}.*

Proof. Since $\operatorname{diag}(c^*)\left(e^{\alpha_i\cdot x^*}\right)_{i=1,\dots,n+k+1} = B\begin{bmatrix}\mu^*\\1\end{bmatrix} \in \ker(\hat{A})$, we have that x^* is a singular point of f_{c^*} (cf.(5)). Thus, we only have to show that it is a degenerate singular point. Let ψ denote the Horn-Kapranov Uniformization map (8). From [10, Theorem 3.4, Theorem 3.5], it follows that x^* is a degenerate singular point if

$$\operatorname{rk} J_\psi(\hat{\mu}^*, x^*) \leq n+k-1, \tag{22}$$

where $\hat{\mu}^* = \begin{bmatrix}\mu^*\\1\end{bmatrix}$.

We prove (22) in two steps. First we show that

$$\operatorname{rk} J_{\xi_{B,\varepsilon}}(\hat{\mu}^*) = \operatorname{rk} J_{\bar{\xi}_{B,\varepsilon}}(\mu^*). \tag{23}$$

To see this, a similar computation as in (16) shows

$$J_{\xi_{B,\varepsilon}}(\hat{\mu}^*) = B^\top \operatorname{diag}\left(\tfrac{1}{B\hat{\mu}^*}\right) B.$$

Thus, the first $k-1$ columns of $J_{\xi_{B,\varepsilon}}(\hat{\mu}^*)$ and $J_{\bar{\xi}_{B,\varepsilon}}(\mu^*)$ are the same. To show that the two matrices have the same rank, it is enough to show that the last column of $J_{\xi_{B,\varepsilon}}(\hat{\mu}^*)$ is contained in the linear space spanned by the columns of $J_{\bar{\xi}_{B,\varepsilon}}(\mu^*)$, which holds since

$$B^\top \operatorname{diag}\left(\tfrac{1}{B\hat{\mu}^*}\right) B\hat{\mu}^* = B^\top \mathbb{1} = 0,$$

where the last equality holds since $B^\top \hat{A}^\top = 0$ and the first column of $\hat{A}^\top$ equals $\mathbb{1}$. This shows (23).

In the second part of the proof, we show (22). Using that the diagram (10) commutes and the chain rule, we have

$$J_{B^\top \circ \operatorname{Log}|\cdot| \circ \psi}(\hat{\mu}^*, x^*) = J_{\xi_{B,\varepsilon} \circ \operatorname{pr}_k}(\hat{\mu}^*, x^*) = J_{\xi_{B,\varepsilon}}(\mu^*) J_{\operatorname{pr}_k}(\hat{\mu}^*, x^*).$$

Using (23) and that $\operatorname{rk} J_{\operatorname{pr}_k}(\hat{\mu}^*, x^*) = k$, it follows that

$$\operatorname{rk} J_{B^\top \circ \operatorname{Log}|\cdot| \circ \psi}(\hat{\mu}^*, x^*) = \operatorname{rk} J_{\xi_{B,\varepsilon}}(\hat{\mu}^*) = \operatorname{rk} J_{\bar{\xi}_{B,\varepsilon}}(\mu^*) \leq k-2,$$

where the last inequality holds since μ^* is a critical point of $\bar{\xi}_{B,\varepsilon}$.

Using again the chain rule

$$J_{B^\top \circ \operatorname{Log}|\cdot| \circ \psi}(\hat{\mu}^*, x^*) = B^\top J_{\operatorname{Log}|\cdot|}(\psi(\hat{\mu}^*, x^*)) J_\psi(\hat{\mu}^*, x^*).$$

Note that $B^\top$ has rank k and $J_{\operatorname{Log}|\cdot|}(\psi(\hat{\mu}^*, x^*))$ is a diagonal matrix with nonzero diagonal entries. Thus, $\operatorname{rk} B^\top J_{\operatorname{Log}|\cdot|}(\psi(\hat{\mu}^*, x^*)) = k$. From Sylvester's rank inequality follows that

$$\operatorname{rk} B^\top J_{\operatorname{Log}|\cdot|}(\psi(\hat{\mu}^*, x^*)) + \operatorname{rk} J_\psi(\hat{\mu}^*, x^*) - (n+k+1) \leq \operatorname{rk} J_{B^\top \circ \operatorname{Log}|\cdot| \circ \psi}(\hat{\mu}^*, x^*) \leq k-2,$$

which imples (22). □

Proposition 4.8. *Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support with Gale dual matrix B. If for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$, all singular points of $Z(f_c)$ are non-degenerate, then $\bar{\xi}_{B,\varepsilon}$ does not have any critical point.*

Proof. The statement is a direct consequence of Lemma 4.7. □

Theorem 4.9. *Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support of codimension 2 with Gale dual matrix B. If for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$, all singular points of $Z(f_c)$ are non-degenerate, then the complement of the signed reduced A-discriminant $\Gamma_\varepsilon(A, B)$ has at most two connected components, which are unbounded.*

Proof. The statement follows directly from Proposition 4.6 and Proposition 4.8. □

4.3. **Signed supports without degenerate singular points.** Now, we show that for certain signed supports $(\mathcal{A}, \varepsilon)$, the singular points of the hypersurfaces $Z(f_c)$ are non-degenerate singular for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$.

We call a pair of parallel affine hyperplanes $\mathcal{H}_{v,a}, \mathcal{H}_{v,b} \subseteq \mathbb{R}^n$ $(a \geq b)$ *enclosing hyperplanes of the positive exponents* $\mathcal{A}_+$ if

$$\mathcal{A}_+ \subseteq \mathcal{H}_{v,a}^- \cap \mathcal{H}_{v,b}^+ \quad \text{and} \quad \mathcal{A}_- \subseteq \mathbb{R}^n \setminus \left(\operatorname{int}(\mathcal{H}_{v,a}^-) \cap \operatorname{int}(\mathcal{H}_{v,b}^+)\right).$$

Enclosing hyperplanes $\mathcal{H}_{v,a}, \mathcal{H}_{v,b}$ are *strict enclosing hyperplanes* of $\mathcal{A}_+$ if additionally $\operatorname{int}(\mathcal{H}_{v,a}^+) \cap \mathcal{A}_- \neq \emptyset$ and $\operatorname{int}(\mathcal{H}_{v,b}^-) \cap \mathcal{A}_- \neq \emptyset$. We define *strict enclosing hyperplanes of the negative exponents* $\mathcal{A}_-$ in a similar way. For an illustration, we refer to Figure 4.

Our first statement concerns exponential sums in two variables.

Theorem 4.10. *Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support in $\mathbb{R}^2$ and assume that both $\mathcal{A}_+$ and $\mathcal{A}_-$ have a pair of strict enclosing hyperplanes. Then*

(i) *for every* $c \in \mathbb{R}^{\mathcal{A}}_{\varepsilon}$ *and* $x \in \mathrm{Sing}(f_c)$*, the Hessian matrix* $\mathrm{Hess}_{f_c}(x)$ *has a positive and a negative eigenvalue.*

(ii) *If* $\mathcal{A}$ *consists of* 5 *exponent vectors, then the complement of the signed reduced A-discriminant* $\Gamma_{\varepsilon}(A,B)$ *consists of at most two connected components.*

Proof. Let $\mathcal{H}_{v,a}, \mathcal{H}_{v,b}$ (resp. $\mathcal{H}_{w,a'}, \mathcal{H}_{w,b'}$) enclosing hyperplanes of $\mathcal{A}_+$ (resp. $\mathcal{A}_-$). Using an affine change of coordinates as in Proposition 2.3, we assume without loss of generality that $v = (1,0)^\top$, $w = (0,1)^\top$.

For $x^* \in \mathrm{Sing}(f_c)$ consider the univariate exponential sums:

$$f_{c,x^*,v} \colon \mathbb{R} \mapsto \mathbb{R}, \quad s \mapsto f_{c,x^*,v}(s) := \sum_{i=1}^{2+k+1} c_i e^{\alpha_i \cdot (x^* + sv)}$$

$$f_{c,x^*,w} \colon \mathbb{R} \mapsto \mathbb{R}, \quad s \mapsto f_{c,x^*,w}(s) := \sum_{i=1}^{2+k+1} c_i e^{\alpha_i \cdot (x^* + sw)}.$$

By construction it holds:

$$0 = f_c(x^*) = f_{c,x^*,v}(0) = f_{c,x^*,w}(0). \tag{24}$$

By denoting $\alpha_{1,i}, \alpha_{2,i}$ the first and the second coordinate of the vector α_i, it is easy to check that

$$\frac{\partial f_{c,x^*,v}}{\partial s}(s) = \sum_{i=1}^{2+k+1} c_i \alpha_{1,i} e^{\alpha_i \cdot (x^* + sv)}, \qquad \frac{\partial f_{c,x^*,w}}{\partial s}(s) = \sum_{i=1}^{2+k+1} c_i \alpha_{2,i} e^{\alpha_i \cdot (x^* + sw)},$$

It follows that

$$\frac{\partial f_{c,x^*,v}}{\partial s}(0) = \frac{\partial f_c}{\partial x_1}(x^*) = 0, \qquad \frac{\partial f_{c,x^*,w}}{\partial s}(0) = \frac{\partial f_c}{\partial x_2}(x^*) = 0, \tag{25}$$

since $x^* \in Z(f_c)$ is a singular point. Combining (24),(25), we have that 0 is a root of $f_{c,x^*,v}$ (resp. $f_{c,x^*,w}$) of multiplicity at least two.

The condition that $\mathcal{H}_{v,a}, \mathcal{H}_{v,b}$ (resp. $\mathcal{H}_{w,a'}, \mathcal{H}_{w,b'}$) are strict enclosing hyperplanes of $\mathcal{A}_+$ (resp. $\mathcal{A}_-$) implies that both exponential sums have at most two sign changes in their coefficient sequence. Since Descartes' rule of signs is valid for polynomials with real exponents [35], one can extend the result to exponential sums. Using Descartes' rule of signs, it follows that the multiplicity of 0 is exactly two for both $f_{c,x^*,v}$ and $f_{c,x^*,w}$. Furthermore,

$$f_{c,x^*,v}(s) < 0 \quad \text{and} \quad f_{c,x^*,w}(s) > 0 \qquad \text{for all } s \neq 0.$$

So 0 is a local maximum of $f_{c,x^*,v}$ and a local minimum of $f_{c,x^*,w}$. Therefore

$$\frac{\partial^2 f_c}{\partial x_1^2}(x) = \frac{\partial^2 f_{c,x^*,v}}{\partial s^2}(0) < 0, \qquad \frac{\partial^2 f_c}{\partial x_2^2}(x) = \frac{\partial^2 f_{c,x^*,w}}{\partial s^2}(0) > 0,$$

which implies that

$$\det(\mathrm{Hess}_{f_c}(x^*)) = \frac{\partial^2 f_c}{\partial x_1^2}(x^*) \frac{\partial^2 f_c}{\partial x_2^2}(x^*) - \Big(\frac{\partial^2 f_c}{\partial x_1 \partial x_2}(x^*)\Big)^2 < 0.$$

Thus, $\mathrm{Hess}_{f_c}(x^*)$ is invertible and must have a negative and a positive eigenvalue.

If $\mathcal{A}$ contains 5 exponent vectors, then the codimension of $\mathcal{A}$ is 2. Since all singular points of $Z(f_c)$ are non-degenerate for all $c \in \mathbb{R}^{\mathcal{A}}_{\varepsilon}$ by (i), part (ii) follows from Theorem 4.9. □

The next condition on the signed support that precludes the existence of degenerate singular points is valid for every number of variables n. Specifically, we require that there is only one exponent vector with negative sign.

Theorem 4.11. *Let* $(\mathcal{A}, \varepsilon)$ *be a full-dimensional signed support with Gale dual* B *such that* $\#\mathcal{A}_- = 1$*. Then we have*

(i) *for all* $c \in \mathbb{R}^{\mathcal{A}}_{\varepsilon}$ *and* $x \in \mathrm{Sing}(f_c)$ *the Hessian matrix* $\mathrm{Hess}_{f_c}(x)$ *has only positive eigenvalues.*

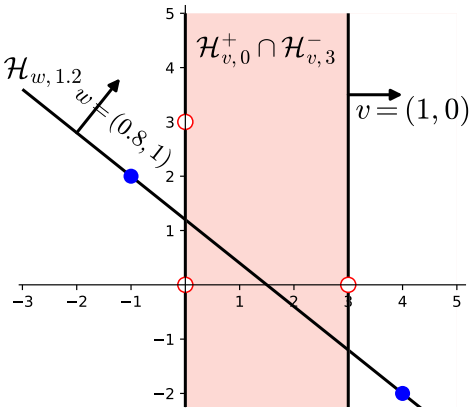


FIGURE 4. An illustration of strict enclosing hyperplanes. Consider the signed exponent vectors $\mathcal{A}_+ = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix} \right\}$ (depicted by red circles) and $\mathcal{A}_- = \left\{ \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -2 \end{bmatrix} \right\}$ (depicted by blue dots). The hyperplanes $\mathcal{H}_{v,3}, \mathcal{H}_{v,0}$ with $v = (1, 0)$ are strict enclosing hyperplanes of $\mathcal{A}_+$. The negative exponent vectors $\mathcal{A}_-$ also have a pair of strict enclosing hyperplanes given by $\mathcal{H}_{w,1.2}, \mathcal{H}_{w,1.2}$ with $w = (0.8, 1)$.

(ii) *If $(\mathcal{A}, \varepsilon)$ has codimension 2, then the complement of the signed reduced A-discriminant $\Gamma_\varepsilon(A, B)$ consists of at most two connected components.*

Proof. Write $\mathcal{A}_+ = \{\alpha_1, \ldots, \alpha_{n+k}\}$ and $\mathcal{A}_- = \{\alpha_{n+k+1}\}$. Using Proposition 2.3, we assume without loss of generality that $\alpha_{n+k+1} = 0$. Under this assumption, the Hessian of f_c at $x \in \mathrm{Sing}(f_c)$ is given by

$$\mathrm{Hess}_{f_c}(x) = \sum_{i=1}^{n+k} (e^{\alpha_i \cdot x} c_i) \alpha_i \cdot \alpha_i^\top = \tilde{A} \operatorname{diag}\left((e^{\alpha_i \cdot x} c_i)_{i=1,\ldots,n+k} \right) \tilde{A}^\top, \tag{26}$$

where

$$\tilde{A} = \begin{bmatrix} \alpha_1 & \ldots & \alpha_{n+k} \end{bmatrix} \in \mathbb{R}^{n \times (n+k)}.$$

Since the affine hull of $\alpha_1, \ldots, \alpha_{n+k+1}$ has dimension n and $\alpha_{n+k+1} = 0$, it follows that $\mathrm{rk}(\tilde{A}) = n$.

Since $e^{\alpha_i \cdot x} c_i$ is positive for $i = 1, \ldots, n+k$, their square root is a real number. This gives

$$\mathrm{Hess}_{f_c}(x) = \left(\tilde{A} \operatorname{diag}\left((\sqrt{e^{\alpha_i \cdot x} c_i})_{i=1,\ldots,n+k} \right) \right) \left(\tilde{A} \operatorname{diag}\left((\sqrt{e^{\alpha_i \cdot x} c_i})_{i=1,\ldots,n+k} \right) \right)^\top$$

and as $\tilde{A}$ has full rank

$$\mathrm{rk}\left(\tilde{A} \operatorname{diag}\left((e^{\alpha_i \cdot x} c_i)_{i=1,\ldots,n+k} \right) \tilde{A}^\top \right) = \mathrm{rk}(\tilde{A} \operatorname{diag}\left((\sqrt{e^{\alpha_i \cdot x} c_i})_{i=1,\ldots,n+k} \right)) = \mathrm{rk}\, \tilde{A} = n.$$

Thus, $\mathrm{Hess}_{f_c}(x)$ is positive semi-definite and of full rank, which implies that all of its eigenvalues are positive. This shows (i).

Since all singular points of $Z(f_c)$ are non-degenerate for all $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$, part (ii) follows from Theorem 4.9. □

Our final condition on the signed support, precluding the existence of degenerate singular points in the hypersurfaces $Z(f_c)$, requires the positive and negative exponent vectors to be separated by a simplex, as follows. We recall the definition of the negative vertex cone of a simplex from [9, Section 4]. For an n-simplex $P \subseteq \mathbb{R}^n$ with vertices $\mu_0, \ldots, \mu_n$, the *negative vertex cone* at vertex μ_k equals

$$P^{-,k} := \mu_k + \mathrm{Cone}\left(\mu_k - \mu_0, \ldots, \mu_k - \mu_n \right).$$

We write P^- for the union of $P^{-,0}, \ldots, P^{-n}$. We refer to Figure 5 for such a simplex and its negative vertex cones in the plane.

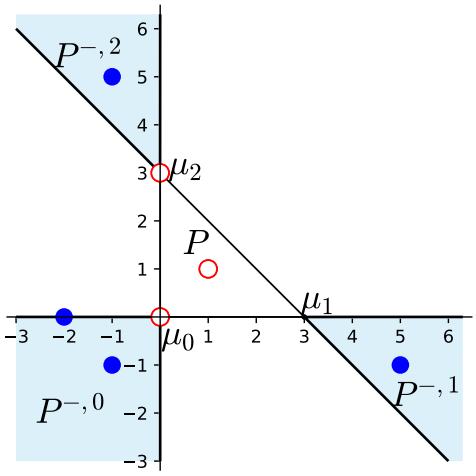


FIGURE 5. A simplex $P = \mathrm{Conv}((0,0),(3,0),(0,3))$ separating the signed exponent vectors $\mathcal{A}_+ = \left\{ \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}0\\3\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix} \right\}$ (marked by red circles) and $\mathcal{A}_- = \left\{ \begin{bmatrix}-1\\5\end{bmatrix}, \begin{bmatrix}5\\-1\end{bmatrix}, \begin{bmatrix}-1\\-1\end{bmatrix}, \begin{bmatrix}-2\\0\end{bmatrix} \right\}$ (marked by blue dots).

Theorem 4.12. *Let $P \subseteq \mathbb{R}^n$ be an n-simplex and let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support in $\mathbb{R}^n$ with Gale dual matrix B such that $\mathcal{A}_+ \subseteq P$, $\mathcal{A}_- \subseteq P^-$, and $\mathcal{A} \cap \mathrm{int}\left(P \cup P^-\right) \neq \emptyset$. Then*

(i) *for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$ and all singular points $x \in \mathrm{Sing}(f_c)$ the eigenvalues of $\mathrm{Hess}_{f_c}(x)$ are negative.*

(ii) *If $\mathcal{A}$ has codimension 2, then the complement of the signed reduced A-discriminant $\Gamma_\varepsilon(A, B)$ consists of at most two connected components.*

Proof. Note that the negative vertex cones are preserved under affine transformation of P, see e.g. [9, Lemma 4.5]. Thus by Proposition 2.3, we can assume without loss of generality that $P = \mathrm{Conv}(0, e_1, \ldots, e_n)$.

Denote $\mathrm{Exp}\colon \mathbb{R}^n \to \mathbb{R}^n_{>0}$ and $\mathrm{Log}\colon \mathbb{R}^n_{>0} \to \mathbb{R}^n$ the coordinate-wise exponential and logarithm maps. From [24, Theorem 7], it follows that the Hessian of the function

$$f_c \circ \mathrm{Log}\colon \mathbb{R}^n_{>0} \to \mathbb{R}, \quad y \mapsto \sum_{i=1}^{n+k+1} c_i y^{\alpha_i}$$

is negative definite for all $y \in \mathbb{R}^n_{>0}$. If $x \in \mathbb{R}^n$ is a singular point of f_c, then from [28, Corollary 1] it follows that

$$\mathrm{Hess}_{f_c}(x) = \mathrm{Hess}_{f_c \circ \mathrm{Log} \circ \mathrm{Exp}}(x) = \left(J_{\mathrm{Exp}}(x)\right)^\top \mathrm{Hess}_{f_c \circ \mathrm{Log}}(\mathrm{Exp}(x)) J_{\mathrm{Exp}}(x).$$

Thus, all the eigenvalues of $\mathrm{Hess}_{f_c}(x)$ are negative by Sylvester's law of inertia [25, Chapter 7]. In particular, all singular points of $Z(f_c)$ are non-degenerate for all $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$. Part (ii) follows from Theorem 4.9. □

Using Theorem 4.11 and Theorem 4.12, we give conditions on the signed support $(\mathcal{A}, \varepsilon)$ such that all possible isotopy types of $Z(f_c)$, $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$ are given by some signed tropical hypersurface (cf. Theorem 2.2).

Corollary 4.13. *Let $(\mathcal{A}, \varepsilon)$ be a full-dimensional signed support of codimension 2 with Gale dual B such that either $\#\mathcal{A}_- = 1$ or $\mathcal{A}_+$ and $\mathcal{A}_-$ are separated by a simplex as in Theorem 4.12. If for each proper face $F \subsetneq \mathrm{Conv}(\mathcal{A})$ the restricted signed support $(\mathcal{A}_F, \varepsilon_F)$ has a non-trivial separating hyperplane, then for each smooth hypersurface $Z(f_c)$ with $c \in \mathbb{R}_\varepsilon^{\mathcal{A}}$ there exists $h \in \mathbb{R}^{\mathcal{A}}$ such that the signed tropical hypersurface $\mathrm{Trop}_\varepsilon(\mathcal{A}, h)$ and $Z(f_c)$ have the same isotopy type.*

Proof. In both cases, the complement of $\Gamma_\varepsilon(A,B)$ has at most two connected components by Theorem 4.11 or Theorem 4.12. Since $(\mathcal{A}_F, \varepsilon_F)$ has a non-trivial separating hyperplane for every proper face $F \subsetneq \mathrm{Conv}(\mathcal{A})$, we have $\nabla_{\mathcal{A}_F,\varepsilon_F} = \emptyset$ by Theorem 3.3. From Theorem 2.8 follows that the hypersurfaces $Z(f_c)$ with $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ have at most two different isotopy types.

First, we focus on the case $\#\mathcal{A}_- = 1$. Assume with out loss of generality that $\mathcal{A}_- = \{\alpha_{n+3}\}$. If α_{n+3} is contained in the boundary of $\mathrm{Conv}(\mathcal{A})$, then there exist a hyperplane $\mathcal{H}_{v,a} \subseteq \mathbb{R}^n$ such that $\alpha_{n+3} \in \mathcal{H}_{v,a}$ and $\mathcal{A} \subseteq \mathcal{H}^+_{v,a}$ (cf. [19, Corollary 2.5]). Thus $(\mathcal{A},\varepsilon)$ has a non-trivial separating hyperplane, which implies that all $Z(f_c)$ with $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ have the same isotopy type by Theorem 3.5.

If $\alpha_{n+3} \in \mathrm{int}(\mathrm{Conv}(\mathcal{A}))$, then choose a generic $h \in \mathbb{R}^{\mathcal{A}} \cong \mathbb{R}^{n+3}$ such that $h_{n+3} > h_i$ for $i = 1,\dots,n+2$. By construction, we have $\mathrm{Trop}_\varepsilon(\mathcal{A},h) \neq \emptyset$ and $\mathrm{Trop}_\varepsilon(\mathcal{A},-h) = \emptyset$. By Theorem 2.2, there exist $c_1, c_2 \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ such that $Z(f_{c_1})$ and $Z(f_{c_2})$ are isotopic to $\mathrm{Trop}_\varepsilon(\mathcal{A},h)$ and to $\mathrm{Trop}_\varepsilon(\mathcal{A},-h)$ respectively. Since the number of possible isotopy types is at most two, it follows that the possible isotopy types are given by $\mathrm{Trop}_\varepsilon(\mathcal{A},h)$ and $\mathrm{Trop}_\varepsilon(\mathcal{A},-h)$.

Next, we consider the case when $\mathcal{A}_+$ and $\mathcal{A}_-$ are separated by a simplex $P \subseteq \mathbb{R}^n$. If one of the negative simplex cones $P^{-,0},\dots,P^{-,n}$ does not contain any positive exponent vector, then $(\mathcal{A},\varepsilon)$ has a non-trivial separating hyperplane and all $Z(f_c)$ with $c \in \mathbb{R}^{\mathcal{A}}_\varepsilon$ have the same isotopy type by Theorem 3.5. If $P^{-,i} \cap \mathcal{A}_+ \neq \emptyset$ for each $i = 0,\dots,n$, then a similar argument as above shows that there exists two signed tropical hypersurfaces wich are not isotopic to each other. This concludes the proof. □

5. Bivariate 5-nomials

For bivariate 5-nomials, the signed reduced A-discriminant has at most 2 critical points by Lemma 4.2. If there is only one critical point, then $\mathbb{R}^k \setminus \Gamma_\varepsilon(A,B)$ has a simple structure, it has at most two connected components, which are unbounded (cf. Proposition 4.6). In this section, we give a complete description of the geometry of the signed support of a bivariate 5-nomial whose signed reduced A-discriminant has two critical points. In our experiments, if the signed reduced A-discriminant had two critical points, then its complement had a bounded chamber. We conjecture that this is always true, however we do not have a proof of this statement nor a counter example.

Conjecture 5.1. *Let $(\mathcal{A},\varepsilon)$ be the signed support of a bivariate 5-nomial and let*

$$\bar{\xi}_{B,\varepsilon}\colon \{\mu \in \mathbb{R} \mid \mathrm{sign}(B\begin{bmatrix}\mu\\1\end{bmatrix}) = \varepsilon\} \to \mathbb{R}^2$$

be the parametrization map of $\Gamma_\varepsilon(A,B)$ as defined in (11). If $\bar{\xi}_{B,\varepsilon}$ has two critical points, then the complement of $\Gamma_\varepsilon(A,B)$ has a bounded connected component.

Given a 2-simplex $P = \mathrm{Conv}(\mu_0,\mu_1,\mu_2)$, denote by $\mathcal{H}_{v_0,d_0}, \mathcal{H}_{v_1,d_1}, \mathcal{H}_{v_2,d_2}$ the supporting hyperplanes of the facets of P. We choose these hyperplanes such that

$$P = \mathcal{H}^+_{v_0,d_0} \cap \mathcal{H}^+_{v_1,d_1} \cap \mathcal{H}^+_{v_2,d_2}$$

and $\mu_i \notin \mathcal{H}_{v_i,d_i}$ for each $i = 0,1,2$. The complement of the union of the hyperplanes $\mathcal{H}_{v_0,d_0}, \mathcal{H}_{v_1,d_1}, \mathcal{H}_{v_2,d_2}$ has 7 chambers. One of these chambers is the simplex P. Three other chambers are the negative vertex cones $P^{-,0}, P^{-,1}, P^{-,2}$, as introduced in Section 4.3. For these chambers we have

$$P^{-,i} = \bigcap_{j=0,j\neq i}^{2} \mathcal{H}^-_{v_j,d_j} \cap \mathcal{H}^+_{v_i,d_i}, \quad \text{for } i = 0,1,2. \tag{27}$$

The three other chambers in the hyperplane arrangment can be written as

$$P^{+,i} = \bigcap_{j=0,j\neq i}^{2} \mathcal{H}^+_{v_j,d_j} \cap \mathcal{H}^-_{v_i,d_i}, \quad \text{for } i = 0,1,2. \tag{28}$$

For $i \neq j \in \{0,1,2\}$, we define the subset of $P^{+,i}$

$$P^{+,i,j} := \bigcap_{j=0, j\neq i}^{2} \mathcal{H}^{+}_{v_j,d_j} \cap \mathcal{H}^{-}_{v_i,d_i} \cap \mathcal{H}^{-}_{v_j,D_j}, \tag{29}$$

where $D_j := v_j \cdot \mu_j > d_j$. For an illustration of these chambers, we refer to Figure 6.

In the following lemmata, we focus on the special case when P is the standard 2-simplex $\Delta_2 = \mathrm{Conv}((0,0)^\top, (1,0)^\top, (0,1)^\top)$ and

$$\mathcal{A}_+ = \left\{ \begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix} \right\}, \quad \mathcal{A}_- = \left\{ \begin{bmatrix}x_1\\y_1\end{bmatrix}, \begin{bmatrix}x_2\\y_2\end{bmatrix} \right\}. \tag{30}$$

Afterward we extend the results to the general case in Theorem 5.5. Choose a Gale dual matrix corresponding to the exponent vectors in (30) as

$$B = \begin{bmatrix} 1-x_1-y_1 & 1-x_2-y_2 \\ x_1 & x_2 \\ y_1 & y_2 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{31}$$

With that choice of the Gale dual matrix, the polynomial from (19) is a quadratic polynomial $q_B(t) := at^2 + bt + c$, where

$$\begin{aligned} a &:= -x_1 y_1 \left(1 - x_1 - y_1\right), \\ b &:= -x_1^2 y_2^2 + 2x_1 x_2 y_1 y_2 - x_2^2 y_1^2 + x_1^2 y_2 + y_2^2 x_1 + x_2^2 y_1 + y_1^2 x_2 - x_1 y_2 - x_2 y_1, \\ c &:= -x_2 y_2 \left(1 - x_2 - y_2\right). \end{aligned} \tag{32}$$

A point $t \in \mathbb{R}$ is a critical point of $\bar{\xi}_{B,\varepsilon}$, $\varepsilon = (1,1,1,-1-1)$ if and only if $q_B(t) = 0$ and it satisfies the inequalities:

$$(1 - x_1 - y_1)t + 1 - x_2 - y_2 > 0, \quad x_1 t + x_2 > 0, \quad y_1 t + y_2 > 0, \quad t > 0. \tag{33}$$

Lemma 5.2. *Let $\mathcal{A}_+, \mathcal{A}_-$ be the set of exponent vectors as defined in* (30) *and let B the corresponding Gale dual matrix from* (31)*. If*

(i) *$\alpha_4 \in \Delta_2$ and $\alpha_5 \in \Delta_2^{+,i}$ for some $i \in \{0,1,2\}$, or*
(ii) *$\alpha_4 \in \Delta_2^{-,i}$ and $\alpha_5 \in \Delta_2^{+,i}$ for some $i \in \{0,1,2\}$,*

then $\bar{\xi}_{B,\varepsilon}$ has at most one critical point.

Proof. Let a, c denote the coefficients of q_B as in (32). Both in case (i) and (ii), we have $a \leq 0$, $c \geq 0$, which implies that q_B has at most one sign change in its coefficient sequence. From Descartes' rule of signs, it follows that q_B has at most 1 positive real root. By (33), every critical point of $\bar{\xi}_{B,\varepsilon}$ is a positive root of q_B. Therefore, $\bar{\xi}_{B,\varepsilon}$ has at most one critical point. □

Lemma 5.3. *Let $\mathcal{A}_+, \mathcal{A}_-$ be the set of exponent vectors as defined in* (30) *and let B the corresponding Gale dual matrix from* (31)*. If $\alpha_4 \in \mathrm{int}(\Delta_2)$ and $\alpha_5 \in \mathrm{int}(\Delta_2^{-,0})$, then $\bar{\xi}_{B,\varepsilon}$ has one critical point.*

Proof. The inequalities in (33) is equivalent to $M := \max\{\frac{-x_2}{x_1}, \frac{-y_2}{y_1}\} < t$. Note that $M > 0$, The number of critical points of $\bar{\xi}_{B,\varepsilon}$ is the same as the number of roots of q_B in the interval (M, ∞).

Let a, c denote the coefficients of q_B as in (32). Under the assumption of the lemma, we have $a < 0$ and $c < 0$. Thus, q_B has 0 or 2 sign changes in its coefficient sequence. By Descartes' rule of signs q_B has at most two positive roots. Moreover, if $q_B(M) > 0$, then q_B has exactly one root in the interval (M, ∞).

Evaluating q_B at $\frac{-x_2}{x_1}$ or at $\frac{-y_2}{y_1}$, depending which one is larger, we used the `Maple` function `IsEmpty` [23] and the `Mathematica` function `Reduce` [16], to verify that $q_B(M) > 0$. Thus, q_B has exactly one root in the interval (M, ∞), which concludes the proof. □

Lemma 5.4. *Let $\mathcal{A}_+, \mathcal{A}_-$ be the set of exponent vectors as defined in* (30)*, let B the corresponding Gale dual matrix from* (31) *and let a, b, c defined in* (32)*. Assume $\alpha_4 \in \operatorname{int}(\Delta_n^{+,1})$ and $\alpha_5 \in \operatorname{int}(\Delta_n^{+,2})$. The map $\bar{\xi}_{B,\varepsilon}$ has two critical points if and only if $\alpha_4 \in \operatorname{int}(P^{+,1,2})$ and $\alpha_5 \in \operatorname{int}(P^{+,2,1})$ and the coordinates of α_4 and α_5 satisfy the following inequalities:*

$$\begin{aligned} & b^2 - 4ac > 0 \\ & b^2 - 4ac < (2x_2y_1(1-x_1-y_1)+b)^2, \quad 0 < 2x_2y_1(1-x_1-y_1)+b \\ & b^2 - 4ac < (2x_1y_2(1-x_1-y_1)+b)^2, \quad 0 > 2x_1y_2(1-x_1-y_1)+b. \end{aligned} \tag{34}$$

Proof. If $\operatorname{relint}(\operatorname{Conv}(\{\alpha_4, \alpha_5\}) \cap \operatorname{relint}(\Delta_2) = \emptyset$, then $\mathcal{A}_+$ and $\mathcal{A}_-$ can be separated by an affine hyperplane [15, Section 2.2, Theorem 2] and therefore $\Gamma_\varepsilon(A, B) = \emptyset$ by Theorem 3.3. In particular, $\Gamma_\varepsilon(A, B)$ does not have any critical point.

If $\operatorname{relint}(P) \cap \operatorname{relint}(Q) \neq \emptyset$, then there exists a pair of strict enclosing hyperplanes of $\mathcal{A}_-$,which are parallel to the affine hull of α_4 and α_5. If additionally $\alpha_5 \in \operatorname{int}(\Delta_n^{+,2}) \setminus \operatorname{int}(\Delta_n^{+,2,1})$, then by perturbing the hyperplanes $\mathcal{H}_{e_1,0}, \mathcal{H}_{e_1,1}$ we get strict enclosing hyperplanes of $\mathcal{A}_+$. Thus, from Proposition 4.8 and Theorem 4.10 it follows that $\bar{\xi}_{B,\varepsilon}$ does not have any critical point.

A similar argument shows that $\bar{\xi}_{B,\varepsilon}$ does not have critical points if $\alpha_5 \in \operatorname{int}(\Delta_n^{+,1}) \setminus \operatorname{int}(\Delta_n^{+,1,2})$. This shows that to get a critical point of $\bar{\xi}_{B,\varepsilon}$ the negative exponent vectors must satisfy $\alpha_4 \in \operatorname{int}(\Delta_n^{+,1,2})$ and $\alpha_5 \in \operatorname{int}(\Delta_n^{+,2,1})$, which is equivalent to

$$\begin{aligned} & x_1 < 0, \quad 0 < y_1 < 1, \quad 1 - x_1 - y_1 > 0, \\ & 0 < x_2 < 1, \quad y_2 < 0, \quad 1 - x_2 - y_2 > 0. \end{aligned}$$

Assuming these inequalities are satisfied, the inequality $y_1 t + y_2 > 0$ in (33) implies $t > \frac{-y_2}{y_1} > 0$, and $t > 0$ implies $(1 - x_1 - y_1)t + 1 - x_2 - y_2 > 0$. Thus, the first and the last inequalities in (33) are redundant.

The roots of q_B are given by

$$t_1 = \tfrac{-b+\sqrt{b^2-4ac}}{2a}, \qquad t_2 = \tfrac{-b-\sqrt{b^2-4ac}}{2a}.$$

An easy computation shows that $t_1 \neq t_2$ and both satisfy the second and the third inequality in (33) if and only if

$$\begin{aligned} & b^2 - 4ac > 0 \\ & b^2 - 4ac < (2x_2y_1(1-x_1-y_1)+b)^2, \quad 0 < 2x_2y_1(1-x_1-y_1)+b \\ & b^2 - 4ac < (2x_1y_2(1-x_1-y_1)+b)^2, \quad 0 > 2x_1y_2(1-x_1-y_1)+b. \end{aligned}$$

This concludes the proof. □

Using Lemma 5.2, Lemma 5.3, and Lemma 5.4, we characterize the signed support of a bivariate 5-nomial such that signed reduced A-discriminant has two critical points.

Theorem 5.5. *Let $(\mathcal{A}, \varepsilon)$ be the full-dimensional signed support of a bivariate 5-nomial with Gale dual matrix $B \in \mathbb{R}^{5\times 2}$. The map $\bar{\xi}_{B,\varepsilon}$ has two critical points only if $\#\mathcal{A}_+ = 3$, $\#\mathcal{A}_- = 2$ and $\dim \operatorname{Conv}(\mathcal{A}_+) = 2$, or $\#\mathcal{A}_+ = 2$, $\#\mathcal{A}_- = 3$ and $\dim \operatorname{Conv}(\mathcal{A}_-) = 2$.*

Assume that $\mathcal{A}_+ = \{\alpha_1, \alpha_2, \alpha_3\}$, $\mathcal{A}_- = \{\alpha_4, \alpha_5\}$ and $P = \operatorname{Conv}(\mathcal{A}_+)$ has dimension 2. Let $M \in \mathbb{R}^{2\times 2}$ be an invertible matrix such that $M(\alpha_2 - \alpha_1) = e_1$, $M(\alpha_3 - \alpha_1) = e_2$ and $v = -M\alpha_1$. Denote $(x_1, y_1)^\top = M\alpha_4 + v$, $(x_2, y_2)^\top = M\alpha_5 + v$ and a, b, c the expressions in x_1, y_1, x_2, y_2 from (32)*. The map $\bar{\xi}_{B,\varepsilon}$ has two critical points if and only if $\alpha_4 \in \operatorname{int}(P^{+,j,i})$ and $\alpha_5 \in \operatorname{int}(P^{+,i,j})$ for $i \neq j \in \{0, 1, 2\}$ and the coordinates of α_4 and α_5 satisfy the following inequalities:*

$$\begin{aligned} & b^2 - 4ac > 0 \\ & b^2 - 4ac < (2x_2y_1(1-x_1-y_1)+b)^2, \quad 0 < 2x_2y_1(1-x_1-y_1)+b \\ & b^2 - 4ac < (2x_1y_2(1-x_1-y_1)+b)^2, \quad 0 > 2x_1y_2(1-x_1-y_1)+b. \end{aligned} \tag{35}$$

Proof. If all the exponent vectors are positive (resp. negative), then $\Gamma_\varepsilon(A,B)=\emptyset$. If $\#\mathcal{A}_+=1$ or $\#\mathcal{A}_-=1$, then $\bar{\xi}_{B,\varepsilon}$ does not have any critical point by Theorem 4.11 and Proposition 4.8. Thus, in order to have a critical point of $\bar{\xi}_{B,\varepsilon}$ one needs $\#\mathcal{A}_+\geq 2$, $\#\mathcal{A}_-\geq 2$.

Write $P=\operatorname{Conv}(\mathcal{A}_+)$ and $Q=\operatorname{Conv}(\mathcal{A}_-)$ and assume $\dim P=\dim Q=1$. Since $\dim\operatorname{Conv}(P\cup Q)=\dim\operatorname{Conv}(\mathcal{A})=2$, the intersection $P\cap Q$ is either empty or a point. If $\operatorname{relint}(P)\cap\operatorname{relint}(Q)=\emptyset$, then P and Q can be separated by an affine hyperplane [15, Section 2.2, Theorem 2] and therefore $\Gamma_\varepsilon(A,B)=\emptyset$ by Theorem 3.3. If $\operatorname{relint}(P)\cap\operatorname{relint}(Q)\neq\emptyset$, then the affine hull of P (resp. Q) is a strict enclosing hyperplane of $\mathcal{A}_+$ (resp. $\mathcal{A}_-$). In that case, we use Theorem 4.10 and Proposition 4.8 to conclude that $\bar{\xi}_{B,\varepsilon}$ does not have any critical point. This shows the first part of the theorem.

In the rest of the proof, we assume $\mathcal{A}_+=\{\alpha_1,\alpha_2,\alpha_3\}$, $\mathcal{A}_-=\{\alpha_4,\alpha_5\}$ and $\dim P=2$. Choose the order of $\alpha_1,\alpha_2,\alpha_3$ so that $\det\begin{bmatrix}1&1&1\\ \alpha_1&\alpha_2&\alpha_3\end{bmatrix}>0$. Then there exists an invertible matrix $M\in\mathbb{R}^{2\times 2}$ with positive determinant such that $M(\alpha_2-\alpha_1)=e_1$ and $M(\alpha_3-\alpha_1)=e_2$. Let $v:=-M\alpha_1$. By construction, the affine linear map

$$L\colon\ \mathbb{R}^2\to\mathbb{R}^2,\quad a\mapsto Ma+v,$$

satisfies $L(\alpha_1)=0$, $L(\alpha_2)=e_1$, $L(\alpha_3)=e_2$ and $L(P)=\Delta_2$. By Proposition 2.3, we assume without loss of generality that

$$\alpha_1=\begin{bmatrix}0\\0\end{bmatrix},\alpha_2=\begin{bmatrix}1\\0\end{bmatrix},\alpha_3=\begin{bmatrix}0\\1\end{bmatrix},\alpha_4=\begin{bmatrix}x_1\\y_1\end{bmatrix},\alpha_5=\begin{bmatrix}x_2\\y_2\end{bmatrix},$$

and choose the Gale dual matrix B as in (31).

If α_4,α_5 are separated from Δ_2 by an affine hyperplane, then $\Gamma_\varepsilon(A,B)=\emptyset$ by Theorem 3.3. If $\alpha_4,\alpha_5\in\operatorname{int}(\Delta_2)$, then $\bar{\xi}_{B,\varepsilon}$ does not have any critical point by Theorem 4.12 and Proposition 4.8. If α_4 (resp. α_5) lies on a supporting hyperplane of a facet of P, then $a=0$ (resp. $c=0$). From this follows that q_B has at most two monomial terms, so q_B has at most one positive root. Thus, $\bar{\xi}_{B,\varepsilon}$ has at most one critical point.

In the following, we investigate the remaining cases:

(I) $\alpha_4\in\operatorname{int}(\Delta_2)$ and $\alpha_5\in\operatorname{int}(\Delta_2^{+,i})$ for some $i\in\{0,1,2\}$.
(II) $\alpha_4\in\operatorname{int}(\Delta_2^{-,i})$ and $\alpha_5\in\operatorname{int}(\Delta_2^{+,i})$ for some $i\in\{0,1,2\}$.
(III) $\alpha_4\in\operatorname{int}(\Delta_2)$ and $\alpha_5\in\operatorname{int}(\Delta_2^{-,i})$ for some $i\in\{0,1,2\}$.
(IV) $\alpha_4\in\operatorname{int}(\Delta_2^{+,i})$ and $\alpha_5\in\operatorname{int}(\Delta_2^{+,j})$ for some $i\neq j\in\{0,1,2\}$.

For (I) and (II), Lemma 5.2 implies that $\bar{\xi}_{B,\varepsilon}$ has at most one critical point. In case (III), by rotating the exponent vectors we assume without loss of generality that $\alpha_5\in\operatorname{int}(\Delta_2^{-,0})$. Now, it follows from Lemma 5.3 that $\bar{\xi}_{B,\varepsilon}$ has one critical point. Under the assumption in (IV), we rotate again the exponent vectors to achieve $\alpha_4\in\operatorname{int}(\Delta_2^{+,1})$ and $\alpha_5\in\operatorname{int}(\Delta_2^{+,2})$. This rotation does not change the number of critical points of $\bar{\xi}_{B,\varepsilon}$ by Corollary 2.4. We use Lemma 5.4 to conclude that $\bar{\xi}_{B,\varepsilon}$ has two critical points if and only if the inequalities in (35) are satisfied. □

Remark 5.6. Consider the signed support

$$\mathcal{A}_+=\left\{\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}1\\0\end{bmatrix},\begin{bmatrix}0\\1\end{bmatrix}\right\},\quad \mathcal{A}_-=\left\{\begin{bmatrix}x_1\\y_1\end{bmatrix},\begin{bmatrix}x_2\\y_2\end{bmatrix}\right\}.$$

Using the `Mathematica` function `Reduce` [16], we verified that for every fixed $(x_1,y_1)\in\operatorname{int}(\Delta_2^{+,1,2})$ and $0<x_2<1$, there always exists $y_2<0$ satisfying the inequalities in (35). With other words, for given $(x_1,y_1)\in\operatorname{int}(\Delta_2^{+,1,2})$ there exists a $(x_2,y_2)\in\operatorname{int}(\Delta_2^{+,2,1})$ such that $\bar{\xi}_{B,\varepsilon}$ has two critical points.

In Figure 6, we depicted the region of (x_2,y_2)'s satisfying the inequalities in (35) for $(x_1,y_1)=(-0.1,0.3)$.

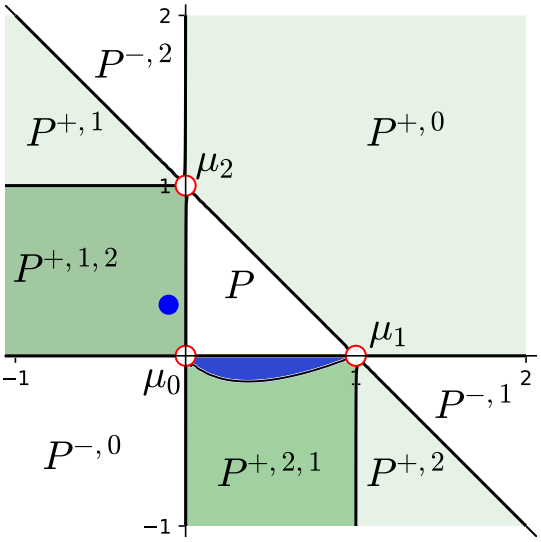


FIGURE 6. An illustration of the chambers as defined in (27),(28),(29) for $P = \mathrm{Conv}((0,0),(1,0),(0,1))$. The red circles denote positive exponent vectors $\mathcal{A}_+ = \{\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\}$, the blue dot denote a negative exponent vector $\begin{bmatrix}-0.1\\0.3\end{bmatrix}$. The blue region contains all negative exponent vectors $\begin{bmatrix}x_2\\y_2\end{bmatrix} \in P^{+,2,1}$ such that $\bar{\xi}_{B,\varepsilon}$ has two critical points.

Acknowledgements. WD and JMR were partially supported by NSF grant CCF-1900881. MLT thanks Elisenda Feliu for useful discussions and comments on the manuscript. MLT was supported by the European Union under the Grant Agreement no. 101044561, POSALG. Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or European Research Council (ERC). Neither the European Union nor ERC can be held responsible for them.

REFERENCES

[1] V.I. Arnold, S.M. Gusein-Zade, and A.N. Varchenko. *Singularities of Differentiable Maps, Volume 1: Classification of Critical Points, Caustics and Wave Fronts*. Modern Birkhäuser Classics. Birkhäuser Boston, 2012.

[2] F. Bihan, E. Croy, W. Deng, K. Phillipson, R. J. Rennie, and J. M. Rojas. Quickly computing isotopy type for exponential sums over circuits (extended abstract). *ACM Commun. Comput. Algebra*, 57:152–155, 2023.

[3] F. Bihan, T. Humbert, and S. Tavenas. New bounds for the number of connected components of fewnomial hypersurfaces. *arXiv*, 2208.04590, 2022.

[4] G. Blekherman, F. Rincón, R. Sinn, C. Vinzant, and J. Yu. Moments, sums of squares, and tropicalization. *arXiv*, 2203.06291, 2022.

[5] M. Brandenburg, G. Loho, and R. Sinn. Tropical positivity and determinantal varieties. *Algebr. Comb.*, 6(4):999–1040, 2023.

[6] G. E. Bredon. *Topology and Geometry*. Springer, 1993.

[7] A. A. Ergür and T. de Wolff. A polyhedral homotopy algorithm for real zeros. *Arnold Math. J.*, 9:305–338, 2022.

[8] A. A. Ergür, G Paouris, and J. M. Rojas. Feasibility of circuit polynomials without black swans. *submitted*, 2024.

[9] E. Feliu and M. L. Telek. On generalizing Descartes' rule of signs to hypersurfaces. *Adv. Math.*, 408(A), 2022.

[10] J. Forsgård. Defective dual varieties for real spectra. *J. Algebraic Comb.*, 49:49–67, 2019.

[11] J. Forsgård, M. Nisse, and J. M. Rojas. New subexponential fewnomial hypersurface bounds. *arXiv*, 1710.00481, 2017.
[12] K. Fukuda. Lecture notes on oriented matroids and geometric computation. 2004. RO-2004.0621, course of Doctoral school in Discrete System Optimization, EPFL 2004.
[13] K. Fukuda. *Lecture: Polyhedral Computation, Spring 2015*. 2015. An online version is available at https://people.inf.ethz.ch/fukudak/lect/pclect/notes2015/PolyComp2015.pdf.
[14] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. *Discriminants, Resultants, and Multidimensional Determinants*. Mathematics (Boston, Mass.). Birkhäuser, 1994.
[15] B. Grünbaum, V. Kaibel, V. Klee, and G. M. Ziegler. *Convex Polytopes*. Graduate Texts in Mathematics. Springer, 2003.
[16] Wolfram Research, Inc. Mathematica, Version 14.0. Champaign, IL, 2024.
[17] I. Itenberg, G. Mikhalkin, and E. I. Shustin. *Tropical Algebraic Geometry*. Springer Science, 2 edition, 2009.
[18] P. Jell, C. Scheiderer, and J. Yu. Real tropicalization and analytification of semialgebraic sets. *Int. Math. Res. Notices*, 2022(2):928–958, 2022.
[19] M. Joswig and T. Theobald. *Polyhedral and Algebraic Methods in Computational Geometry*. Springer, 2013.
[20] S. Kalajdzievski. *An Illustrated Introduction to Topology and Homotopy*. CRC Press, 2014.
[21] M. Kapranov. A characterization of A-discriminantal hypersurfaces in terms of the logarithmic Gauss map. *Math. Ann.*, 290:277–285, 1991.
[22] D. Maclagan and B. Sturmfels. *Introduction to Tropical Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2015.
[23] Maplesoft, a division of Waterloo Maple Inc. `Maple`. https://www.maplesoft.com, 2023.
[24] C. D. Maranas and C. A. Floudas. All solutions of nonlinear constrained systems of equations. *J. Global. Optim.*, 7:143–182, 1995.
[25] C. D. Meyer. *Matrix Analysis and Applied Linear Algebra (Solution)*. Society for Industrial and Applied Mathematics, 2004.
[26] J. M. Rojas and K. Rusek. A-discriminants for complex exponents and counting real isotopy types. *arXiv*, 1612.03458, 2017.
[27] K. Rusek. *A-Discriminant Varieties and Amoebae*. PhD thesis Texas A&M University, 2013.
[28] M. Skorski. Chain rules for Hessian and higher derivatives made easy by tensor calculus. *arXiv*, 1911.13292, 2019.
[29] D. Speyer and L. Williams. The tropical totally positive grassmannian. *J. Algebr. Comb.*, 22:189–210, 2005.
[30] E. Stiemke. Über positive Lösungen homogener linearer Gleichungen. *Math. Ann.*, 76:340–342, 1915.
[31] C. Vinzant. Real radical initial ideals. *J. Algebra*, 352(1):392–407, 2009.
[32] O. Y. Viro. Curves of degree 7, curves of degree 8 and the Ragsdale conjecture. *Dokl. Akad. Nauk SSSR*, 254:1305–1310, 1980.
[33] O. Y. Viro. *Constructing real algebraic varieties with prescribed topology*. Dissertation, LOMI, Leningrad, 1983. An english translation by the author is available at https://arxiv.org/abs/math/0611382.
[34] O. Y. Viro. From the sixteenth Hilbert problem to tropical geometry. *Japanese J. Math.*, 3:185–214, 2008.
[35] X. Wang. A simple proof of Descartes's rule of signs. *Am. Math. Mon.*, 111:525–526, 2004.
[36] F. W. Warner. *Foundations of Differentiable Manifolds and Lie Groups*. Springer Science, 1983.
[37] T. Zaslavsky. *Facing up to Arrangements: Face-Count Formulas for Partitions of Space by Hyperplanes*. American Mathematical Society: Memoirs of the American Mathematical Society. American Mathematical Society, 1975.
[38] G. M. Ziegler. *Lectures on Polytopes*. Springer, 1995.

Department of Mathematics, Texas A&M University, College Station, Texas, USA
Email address: deng15521037237@math.tamu.edu

Department of Mathematics, Texas A&M University, College Station, Texas, USA
Email address: rojas@math.tamu.edu

Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark
Email address: mlt@math.ku.dk