



Research article

Parameter regions that give rise to $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states in the n -site phosphorylation system

Magalí Giaroli^{1,†}, Rick Rischter^{2,†}, Mercedes P. Millán¹ and Alicia Dickenstein^{1,*}

¹ Dto. de Matemática, FCEN, Universidad de Buenos Aires, and IMAS (UBA-CONICET), Ciudad Universitaria, Pab. I, C1428EGA Buenos Aires, Argentina

² Instituto de Matemática e Computação, IMC, Universidade Federal de Itajubá (UNIFEI), Av. BPS 1303, Bairro Pinheirinho, 37500-903, Itajubá, Minas Gerais, Brazil

† The authors contributed equally to this work.

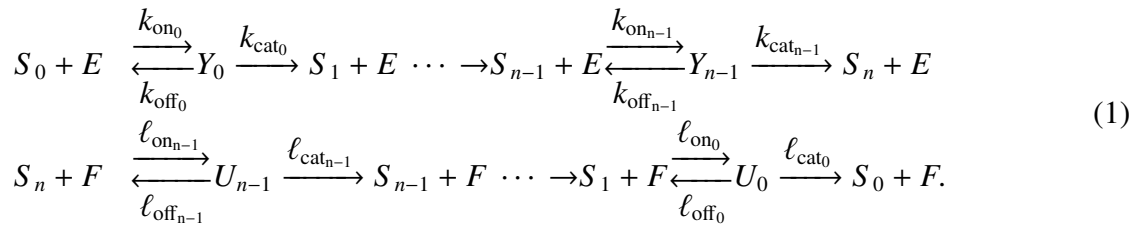
* **Correspondence:** Email: alidick@dm.uba.ar; Tel: +541152857590.

Abstract: The distributive sequential n -site phosphorylation/dephosphorylation system is an important building block in networks of chemical reactions arising in molecular biology, which has been intensively studied. In the nice paper of Wang and Sontag (2008) it is shown that for certain choices of the reaction rate constants and total conservation constants, the system can have $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states (that is, $n + 1$ positive steady states for n even and n positive steady states for n odd). In this paper we give open parameter regions in the space of reaction rate constants and total conservation constants that ensure these number of positive steady states, while assuming in the modeling that roughly only $\frac{1}{4}$ of the intermediates occur in the reaction mechanism. This result is based on the general framework developed by Bihan, Dickenstein, and Giaroli (2018), which can be applied to other networks. We also describe how to implement these tools to search for multistationarity regions in a computer algebra system and present some computer aided results.

Keywords: multistationarity; distributive sequential n -site phosphorylation/dephosphorylation system; steady states; regions of multistationarity; intermediate species

1. Introduction

Multisite phosphorylation cycles are common in cell regulation systems [1]. The important distributive sequential n -site phosphorylation network is given by the following reaction scheme:



Arrows correspond to chemical reactions. The labels in the arrows are positive numbers, known as *reaction rate constants*. This reaction scheme represents one substrate S_0 that can sequentially acquire up to n phosphate groups (giving rise to the phosphoforms S_1, \dots, S_n) via the action of the kinase E , and which can be sequentially released via the action of the phosphatase F , in both cases via an intermediate species (denoted by $Y_0, Y_1, \dots, Y_{n-1}, U_0, U_1, \dots, U_{n-1}$) formed by the interaction of the substrate and the enzyme.

The kinetics of this network is deduced by applying the law of mass action to the labeled digraph of reactions (1), which yields an autonomous polynomial system of ordinary differential equations depending on the positive reaction rate constants. This system describes the evolution in time of the concentrations of the different chemical species. We denote with lower case letters the concentrations of the $3n + 3$ species. The derivatives of the concentrations of the substrates satisfy (see e.g. [2] for the general definition):

$$\begin{aligned}
 \frac{ds_0}{dt} &= -k_{\text{on}_0}s_0e + k_{\text{off}_0}y_0 + \ell_{\text{cat}_0}u_0, \\
 \frac{ds_i}{dt} &= k_{\text{cat}_{i-1}}y_{i-1} - k_{\text{on}_i}s_i e + k_{\text{off}_i}y_i + \ell_{\text{cat}_i}u_i - \ell_{\text{on}_{i-1}}s_i f + \ell_{\text{off}_{i-1}}u_{i-1}, \quad i = 1, \dots, n-1, \\
 \frac{ds_n}{dt} &= k_{\text{cat}_{n-1}}y_{n-1} - \ell_{\text{on}_{n-1}}s_n f + \ell_{\text{off}_{n-1}}u_{n-1},
 \end{aligned}$$

the derivatives of the concentrations of the intermediate species satisfy:

$$\begin{aligned}
 \frac{dy_i}{dt} &= k_{\text{on}_i}s_i e - (k_{\text{off}_i} + k_{\text{cat}_i})y_i, \quad i = 0, \dots, n-1, \\
 \frac{du_i}{dt} &= \ell_{\text{on}_i}s_{i+1} f - (\ell_{\text{off}_i} + \ell_{\text{cat}_i})u_i, \quad i = 0, \dots, n-1,
 \end{aligned}$$

and $\frac{de}{dt} = -\frac{dy_0}{dt} - \dots - \frac{dy_{n-1}}{dt}$, $\frac{df}{dt} = -\frac{du_0}{dt} - \dots - \frac{du_n}{dt}$, which give two linear conservation relations. Indeed, there are three linearly independent conservation laws:

$$\sum_{i=0}^n s_i + \sum_{i=0}^{n-1} y_i + \sum_{i=0}^{n-1} u_i = S_{\text{tot}}, \quad e + \sum_{i=0}^{n-1} y_i = E_{\text{tot}}, \quad f + \sum_{i=0}^{n-1} u_i = F_{\text{tot}}, \tag{2}$$

where the *total conservation constants* $S_{\text{tot}}, E_{\text{tot}}, F_{\text{tot}}$ are positive for any trajectory of the system intersecting the positive orthant.

A steady state of the system is a common zero of the polynomials in the right-hand side of the differential equations. There are $6n + 3$ parameters: the reaction rate constants and the total conservation constants (which correspond to points in the positive orthant $\mathbb{R}_{>0}^{6n+3}$). It is known that for any $n \geq 2$, the sequential n -site system shows *multistationarity*. This means that there exists a choice

of parameters for which there are at least two positive steady states satisfying moreover the linear conservation relations (2) for the same S_{tot} , E_{tot} and F_{tot} . In this case, it is said that the steady states are *stoichiometrically compatible*, or that they lie in the same *stoichiometric compatibility class*. This is an important notion, because the occurrence of multistationarity allows for different cell responses with the same total conservation constants (that is, the same total amounts of substrate and enzymes).

The possible number of positive steady states of the n -site phosphorylation system (for fixed total conservation constants) has been studied in several articles. In the case $n = 2$, it is a well known fact that the number of nondegenerate positive steady states is one or three [3,4]. The existence of bistability is proved in [5]. In [6] and [7], the authors give conditions on the reaction rate constants to guarantee the existence of three positive steady states based on tools from degree theory, but this approach does not describe in general the total conservation constants for which there is multistationarity. In [7] Conradi and Mincheva give a sufficient condition on the reaction rate constants to guarantee multistationarity for suitable choices of total conservation constants. These total amounts are precisely given in implicit form. Specifically, after writing S_{tot} , E_{tot} and F_{tot} in terms of the concentrations at steady state of E , F and S_0 , as in the present paper, the values of total amounts that give rise to multistationarity are derived from those for which an explicit polynomial becomes negative, which allows to construct as many witnesses as wished. When $n \geq 3$, there could be more than three positive steady states, and our methods allows us to give lower bounds bigger than three for $n > 3$.

For an arbitrary number n of phosphorylation sites, it was shown in [4] that the system has at most $2n - 1$ positive steady states. In the same article, the authors showed that there exist reaction rate constants and total conservation constants such that the network has n (resp. $n + 1$) positive steady states for n odd (resp. even); that is, there are $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states for any value of n , where $\lfloor \cdot \rfloor$ denotes integer part.

In [8] (see also [9]) the authors showed parameter values such that for $n = 3$ the system has five positive steady states, and for $n = 4$ the system has seven steady states, obtaining the upper bound given in [4]. In the recent article [10] the authors show that if the n -site phosphorylation system is multistationary for a choice of rate constants and total conservation constants $(S_{tot}, E_{tot}, F_{tot})$ then for any positive real number c there are rate constants for which the system is multistationary when the total conservation constants are scaled by c . Concerning the stability, in [11] it is shown evidence that the system can have $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states with $\lfloor \frac{n}{2} \rfloor + 1$ of them being stable. Recently, a proof of this unlimited multistability was presented in [12], where the authors find a choice of parameters that gives the result for a smaller system, and then extend this result using techniques from singular perturbation theory.

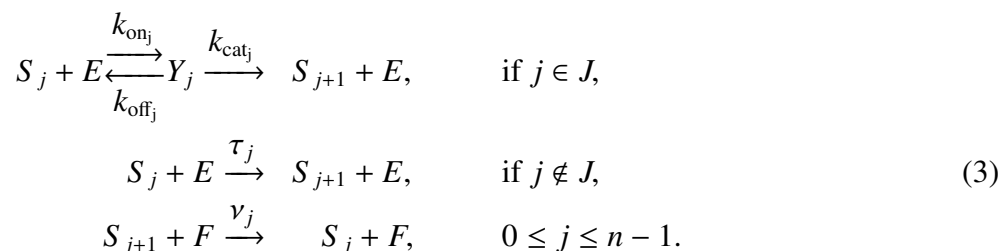
In the previous paper [13], open parameter regions on all the parameters are given for the occurrence of multistationarity for the n -site sequential phosphorylation system, but no more than three positive steady states are ensured. These conditions are based on a general framework to obtain multistationary regions jointly in the reaction rate constants and the total conservation constants. The description of these open sets is explicit on a subset of the reaction rate constants and the total conservation constants, while some other reaction rate constants can vary in an open set whose form is described but which is not completely quantified. Our approach in this article uses the systematic technique in [13], which we briefly summarize in Section 3.

The removal of intermediates was studied in [14]. More specifically, the emergence of multistationarity of the n -site phosphorylation system with less intermediates was studied in [15]. It is

known that the n -site phosphorylation network without any intermediates complexes has only one steady state for any choice of parameters. In [15], the authors showed which are the minimal sets of intermediates that give rise to a multistationarity system, but they do not give information about how many positive steady states can occur, and also, they do not describe the parameter regions for which these subnetworks are multistationary.

In this paper, we work with subnetworks of the sequential n -site phosphorylation system that only have intermediates in the E component (that is, in the connected component of the network where the kinase E reacts), see Definition 1. In case of the full mechanism on the E component or if we only assume that there are intermediate species that are formed between the phosphorylated substrates with *parity equal to n* (that is, roughly only $\frac{1}{4}$ of the intermediates of the n sequential phosphorylation cycle), we obtain precise conditions on all the parameters to ensure as many positive steady states as possible. Indeed, we show in Proposition 7 that the maximum number of complex solutions to the steady state equations intersected with the linear conservation relations is always $n + 1$, the maximum number of real roots is also $n + 1$, that could be all positive when n is even, while only n of them can be positive when n is odd, so the maximum number of positive steady states equals $2\lfloor \frac{n}{2} \rfloor + 1$ for any n . In Theorem 2 and Corollary 5 we describe open regions in all the parameters so that the associated phosphorylation/dephosphorylation system admits these number of positive steady states. This follows from Theorem 4. As we pointed out before, we give precise new conditions on some of the parameters, involving reaction rate constants and total conservation constants, while some other rate constants vary on open sets defined by inequalities that are not exactly quantified. In order to state these results, we need to introduce some notations.

Definition 1. For any natural number n , we write $I_n = \{0, \dots, n - 1\}$. Given $n \geq 1$, and a subset $J \subset I_n$, we denote by G_J the network whose only intermediate complexes are Y_j for $j \in J$, and none of the U_i . It is given by the following reactions:



where the labels of the arrows are positive numbers. We will also denote by G_J the associated system of differential equations with mass-action kinetics.

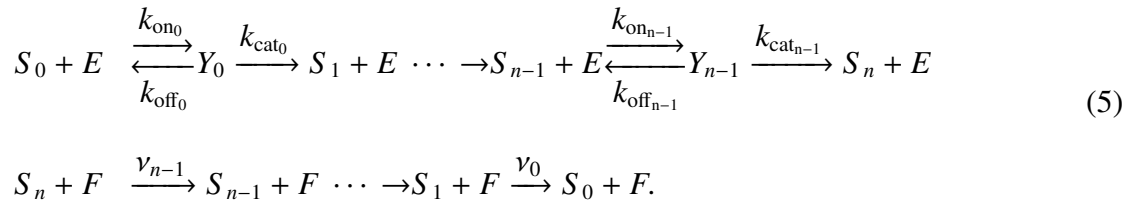
For all these systems G_J , there are always three linearly independent conservation laws for any value of n :

$$\sum_{i=0}^n s_i + \sum_{j \in J} y_j = S_{\text{tot}}, \quad e + \sum_{j \in J} y_j = E_{\text{tot}}, \quad f = F_{\text{tot}}, \quad (4)$$

where the total conservation constants $S_{\text{tot}}, E_{\text{tot}}, F_{\text{tot}}$ are positive for any trajectory of the system of differential equations which intersects the positive orthant. Note that the concentration of the phosphatase F is constant, equal to F_{tot} .

To get lower bounds on the number of positive steady states with fixed positive total conservation constants, we first consider the network G_{I_n} , that is, when all the intermediates in the E component

appear. It has the following associated digraph:



Theorem 2. Let $n \geq 1$. With the previous notation, consider the network G_{I_n} in (5), and suppose that the reaction rate constants k_{cat_i} and ν_i , $i = 0, \dots, n-1$, satisfy the inequality

$$\max_{i \text{ even}} \left\{ \frac{k_{\text{cat}_i}}{\nu_i} \right\} < \min_{i \text{ odd}} \left\{ \frac{k_{\text{cat}_i}}{\nu_i} \right\}.$$

For any positive values S_{tot} , E_{tot} and F_{tot} of the total conservation constants with

$$S_{\text{tot}} > E_{\text{tot}},$$

verifying the inequalities:

$$\max_{i \text{ even}} \left\{ \frac{k_{\text{cat}_i}}{\nu_i} \right\} < \left(\frac{S_{\text{tot}}}{E_{\text{tot}}} - 1 \right) F_{\text{tot}} < \min_{i \text{ odd}} \left\{ \frac{k_{\text{cat}_i}}{\nu_i} \right\}, \tag{6}$$

there exist positive constants B_1, \dots, B_n such that for any choice of positive constants $\lambda_0, \dots, \lambda_{n-1}$ satisfying

$$\frac{\lambda_j}{\lambda_{j-1}} < B_j \text{ for } j = 1, \dots, n-1, \quad \frac{1}{\lambda_{n-1}} < B_n, \tag{7}$$

rescaling of the given parameters k_{on_j} by $\lambda_j k_{\text{on}_j}$, for each $j = 0, \dots, n-1$, gives rise to a system with exactly $2\lfloor \frac{n}{2} \rfloor + 1$ nondegenerate positive steady states.

Remark 3. We will also show in the proof of Theorem 2, that for any reaction rate constants and total conservation constants satisfying (6), there exist $t_0 > 0$ such that for any value of $t \in (0, t_0)$, the system G_{I_n} has exactly $2\lfloor \frac{n}{2} \rfloor + 1$ nondegenerate positive steady states after modifying the constants k_{on_j} by $t^{j-n} k_{\text{on}_j}$ for each $j = 0, \dots, n-1$.

We now consider subnetworks G_J , with $J \subset J_n$ where

$$J_n := \{i \in I_n : (-1)^{i+n} = 1\}, \text{ for } n \geq 1, \tag{8}$$

that is, subsets J with indexes that have the same parity as n .

Theorem 4. Let $n \geq 1$, and consider a subset $J \subset J_n$. Let G_J be its associated system as in (3). Assume moreover that

$$S_{\text{tot}} > E_{\text{tot}}.$$

A multistationarity region in the space of all parameters for which the system G_J admits at least $1 + 2|J|$ positive steady states can be described as follows. Given any positive value of F_{tot} , choose any positive real numbers k_{cat_j}, ν_j , with $j \in J$ satisfying

$$\max_{j \in J} \left\{ \frac{k_{\text{cat}_j}}{\nu_j} \right\} < \left(\frac{S_{\text{tot}}}{E_{\text{tot}}} - 1 \right) F_{\text{tot}}. \tag{9}$$

Then, there exist positive constants B_1, \dots, B_n such that for any choice of positive constants $\lambda_0, \dots, \lambda_{n-1}$ satisfying

$$\frac{\lambda_j}{\lambda_{j-1}} < B_j \text{ for } j = 1, \dots, n-1, \quad \frac{1}{\lambda_{n-1}} < B_n, \quad (10)$$

rescaling of the given parameters k_{on_j} by $\lambda_j k_{\text{on}_j}$, for $j \in J$ and τ_j by $\lambda_j \tau_j$ if $j \notin J$ gives rise to a system with at least $1 + 2|J|$ positive steady states.

The following immediate Corollary of Theorem 4 implies that we can obtain a region in parameters space with $\lfloor \frac{n}{2} \rfloor$ intermediates, where the associated system has $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states.

Corollary 5. *Let $n \geq 1$, and consider the network G_{J_n} as in (3), with J_n as in (8). Assume moreover that*

$$S_{\text{tot}} > E_{\text{tot}}.$$

Then, there is a multistationarity region in the space of all parameters for which the network G_{J_n} admits $2\lfloor \frac{n}{2} \rfloor + 1$ steady states (with fixed total conservation constants corresponding to the coordinates of a vector of parameters in this region), described in the statement of Theorem 4.

Similar versions of all our results work if one assumes that some or all of the reverse unbinding reactions do not occur, so that the rate constants k_{off_i} are set to zero (see Remarks 9 and 15). We thank one of the referees for pointing this out.

We explain in Section 5 the computational approach to find new regions of multistationarity. This is derived from the framework in [13] that we used to get the previous results. We find new regions of multistationarity for the cases $n = 2, 3, 4$ and 5 , after some manual organization of the automatic computations. Our computer aided results are summarized in Propositions 17, 18, 19, 20, 21 and 22.

The paper is organized as follows. In Section 2 we prove Proposition 7 and we show how to lift regions of multistationarity from the reduced system G_J to the complete sequential n -site phosphorylation system. We also show in Lemma 8 that even with a single intermediate Y_0 it is possible to make a choice of all parameters such that the system has $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states. This result has been independently found by Elisenda Feliu (personal communication). This says that while Corollary 5 is optimal, the regions obtained for any subset J with indexes of the same parity of n in Theorem 4 properly contained in J_n , only ensure $2|J| + 1$ positive steady states. However, note that we are able to describe open regions in parameter space and Lemma 8 only allows us to get choices of parameters.

In Section 3 we briefly recall the framework in [13], which is the basis of our approach. In Section 4 we give the proofs of Theorems 2 and 4. Finally, as we mentioned above, we explain in Section 5 how to implement the computational approach to find new regions of multistationarity. We end with a discussion where we further compare our detailed results with previous results in the literature.

2. Upper bounds and extension of multistationarity

In this section we collect three results that complete our approach to describe multistationarity regions giving lower bounds for the number of positive steady states with fixed linear conservation relations for the systems G_J in Definition 1 for any $J \subset I_n$. We first show a positive parametrization of the concentrations at steady state of the species of the systems G_J , which allows us to translate the

equations defining the steady states in all the concentrations plus the linear conservation relations into a system of two equations in two variables. In Proposition 7, we prove that any mass-action system G_J , has at most $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states. Lemma 8 shows that having a single intermediate is enough to get that number of positive steady states, for particular choices of the reaction rate constants. However, this computation by reduction to the univariate case is not systematic as the general approach from [13] that we use to describe multistationarity regions in Theorems 2 and 4, which can be applied to study other quite different mechanisms. It is known that if a system G_J has m nondegenerate positive steady states for a subset $J \subset I_n$, then it is possible to find parameters for the whole n -site phosphorylation system that also give at least m positive steady states (see [14]). In Theorem 10, we give precise conditions on the rate constants to lift the regions of multistationarity for the reduced networks to regions of multistationarity with $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states (with fixed total conservation constants) of the complete n -sequential phosphorylation cycle.

2.1. Parametrizing the steady states

The following lemma gives a positive parametrization of the concentration of the species at steady state for the systems G_J , in terms of the concentrations of the unphosphorylated substrate S_0 and the kinase E . It is a direct application of the general procedure presented in Theorem 3.5 in [16], and generalizes the parametrization given in Section 4 of [13].

Lemma 6. *Given $n \geq 1$ and a subset $J \subset I_n$, consider the system G_J as in Definition 1. Denote for each $j \in J$*

$$K_j = \frac{k_{\text{on}j}}{k_{\text{off}j} + k_{\text{cat}j}}, \quad \tau_j = k_{\text{cat}j} K_j, \quad (11)$$

set $T_{-1} = 1$, and for any $i = 0, \dots, n-1$:

$$T_i = \prod_{j=0}^i \frac{\tau_j}{\nu_j}. \quad (12)$$

Then, the parametrization of the concentrations of the species at a steady state in terms of s_0 and e is equal to:

$$s_i = \frac{T_{i-1}}{(F_{\text{tot}})^i} s_0 e^i, \quad i = 1, \dots, n, \quad y_j = \frac{K_j T_{j-1}}{(F_{\text{tot}})^j} s_0 e^{j+1}, \quad j \in J, \quad (13)$$

The inverses K_j^{-1} of the constants K_j in (11) in the statement of Lemma 6 are usually called *Michaelis-Menten constants*.

Let $n \geq 1$ and a subset $J \subset I_n$. If we substitute the monomial parametrization of the concentration of the species at steady state (13) into the conservation laws, we obtain a system of two equations in two variables s_0 and e . We have:

$$s_0 + \sum_{j \in J} \left(\frac{T_j}{(F_{\text{tot}})^{j+1}} + \frac{K_j T_{j-1}}{(F_{\text{tot}})^j} \right) s_0 e^{j+1} + \sum_{j \notin J} \frac{T_j}{(F_{\text{tot}})^{j+1}} s_0 e^{j+1} - S_{\text{tot}} = 0, \quad (14)$$

$$e + \sum_{j \in J} \frac{K_j T_{j-1}}{(F_{\text{tot}})^j} s_0 e^{j+1} - E_{\text{tot}} = 0.$$

We can write system (14) in matricial form:

$$C \begin{pmatrix} s_0 & e & s_0 e & s_0 e^2 & \dots & s_0 e^n & 1 \end{pmatrix}^t = 0, \quad (15)$$

where $C \in \mathbb{R}^{2 \times (n+3)}$ is the matrix of coefficients:

$$C = \begin{pmatrix} 1 & 0 & \frac{T_0}{F_{tot}} + \beta_0 & \frac{T_1}{(F_{tot})^2} + \beta_1 & \dots & \frac{T_{n-1}}{(F_{tot})^n} + \beta_{n-1} & -S_{tot} \\ 0 & 1 & \beta_0 & \beta_1 & \dots & \beta_{n-1} & -E_{tot} \end{pmatrix}, \quad (16)$$

with

$$\beta_j = \frac{K_j T_{j-1}}{(F_{tot})^j} \text{ for } j \in J, \text{ and } \beta_j = 0 \text{ if } j \notin J. \quad (17)$$

Note that if we order the variables s_0, e , the support of the system (that is, the exponents of the monomials that occur) is the following set \mathcal{A} :

$$\mathcal{A} = \{(1, 0), (0, 1), (1, 1), (1, 2), \dots, (1, n), (0, 0)\} \subset \mathbb{Z}^2, \quad (18)$$

independently of the choice of $J \subset I_n$.

2.2. Upper bounds on the number of positive steady states

We first recall Kushnirenko Theorem, a fundamental result about sparse systems of polynomial equations, which gives a bound on the number of complex solutions with nonzero coordinates. Given a finite point configuration $\mathcal{A} \subset \mathbb{Z}^d$, denote by $\text{ch}(\mathcal{A})$ the convex hull of \mathcal{A} . We write vol to denote Euclidean volume, and set $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$.

Kushnirenko Theorem [17]: Given a finite point configuration $\mathcal{A} \subset \mathbb{Z}^d$, a sparse system of d Laurent polynomials in d variables with support \mathcal{A} has at most $d! \text{vol}(\text{ch}(\mathcal{A}))$ isolated solutions in $(\mathbb{C}^*)^d$ (and exactly this number if the polynomials have *generic* coefficients.)

We also recall the classical Descartes rule of signs.

Descartes rule of signs: Let $p(x) = c_0 + c_1 x + \dots + c_m x^m$ be a nonzero univariate real polynomial with r positive real roots counted with multiplicity. Denote by s the number of sign variations in the ordered sequence of the signs $\text{sign}(c_0), \dots, \text{sign}(c_m)$ of the coefficients, i.e., discard the 0's in this sequence and then count the number of times two consecutive signs differ. Then $r \leq s$ and r and s have the same parity, which is even if $c_0 c_m > 0$ and odd if $c_0 c_m < 0$.

We then have that $2\lfloor \frac{n}{2} \rfloor + 1$ is an upper bound for the number of positive real solutions of the system of equations defining the steady states of any system G_J in Definition 1:

Proposition 7. *For any choice of rate constants and total conservation constants, the dynamical system G_J associated with any subset $J \subset I_n$ has at most $2\lfloor \frac{n}{2} \rfloor + 1$ isolated positive steady states. In fact, the polynomial system of equations defining the steady states of G_J can have at most $n+1$ isolated solutions in $(\mathbb{C}^*)^d$.*

Proof. The number of positive steady states of the subnetwork G_J is the number of positive solutions of the sparse system (14) of two equations and two variables. The support of the system is (18) whose convex hull has Euclidean volume equal to $\frac{n+1}{2}$. By Kushnirenko Theorem, the number of isolated

solutions in $(\mathbb{C}^*)^2$ is at most $2! \frac{(n+1)}{2} = n + 1$. In particular, the number of isolated positive solutions is at most $n + 1$.

Moreover, when all positive solutions are nondegenerate, their number is necessarily odd by Corollary 2 in [6], which is based on the notion of Brouwer's degree. Indeed, in our case, we can bypass the condition of nondegeneracy because we can use Descartes rule of signs in one variable. In fact, from the first equation of system (14), we can write:

$$s_0 = \frac{S_{tot}}{p(e)}, \quad (19)$$

where $p(e)$ is the following polynomial of degree n on the variable e :

$$p(e) := 1 + \sum_{i=0}^{n-1} (\alpha_i + \beta_i) e^{i+1}, \quad (20)$$

with

$$\alpha_i = \frac{T_i}{(F_{tot})^{i+1}}, \quad i = 0, \dots, n-1, \quad (21)$$

and $\beta_i = \frac{K_j T_{j-1}}{(F_{tot})^j}$ if $j \in J$ or $\beta_j = 0$ if $j \notin J$ were defined in (17). Note that for any $e > 0$ it holds that $p(e) > 0$, and so $s_0 > 0$. If we replace (19) in the second equation of (14), we have:

$$e + \sum_{i=0}^{n-1} \beta_i \frac{S_{tot}}{p(e)} e^{i+1} - E_{tot} = 0. \quad (22)$$

The number of positive solutions of Eq (22) is the same if we multiply by $p(e)$:

$$q(e) := e p(e) + \sum_{i=0}^{n-1} \beta_i S_{tot} e^{i+1} - E_{tot} p(e) = 0. \quad (23)$$

This last polynomial q has degree $n + 1$, with leading coefficient equal to $\alpha_{n-1} + \beta_{n-1} > 0$ and constant coefficient equal to $-E_{tot} < 0$. The sign variation of the coefficients of q has the same parity as the sign variation of the leading coefficient and the constant coefficient, which is one. So, by Descartes rule of signs, as the sign variation is odd, the number of positive solutions is also odd. \square

2.3. One intermediate is enough in order to obtain $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states

As we mentioned in the introduction, the following result has been independently found by Elisenda Feliu (personal communication).

Lemma 8. *If $J = \{0\}$, then there exists parameter values such that the system G_J admits $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states.*

Proof. Assume n is even, then $n + 1$ is odd. As we did in the proof of Proposition 7, the positive solutions of the system (14) are in bijection with the positive solutions of the polynomial $q(e)$ in (23). Here $\beta_0 = K_0$ and $\beta_i = 0$ for $i \neq 0$. We will consider the polynomial $\tilde{q}(e) := \frac{q(e)}{E_{tot}}$, with constant coefficient equal to -1 .

Consider any polynomial of degree $n + 1$

$$c_{n+1}e^{n+1} + c_n e^n + \dots + c_1 e - 1, \quad (24)$$

with $n + 1$ distinct positive roots, and with constant term equal to -1 . Then, we have that $c_i(-1)^{i+1} > 0$, and in particular, $c_{n+1} > 0$.

Our goal is to find reaction rate constants and total conservation constants such that the polynomial (24) coincides with the polynomial $\tilde{q}(e)$. Comparing the coefficient of e^i , for $i = 1, \dots, n + 1$ in both polynomials, we need to have:

$$\begin{aligned} \frac{\alpha_{n-1}}{E_{tot}} &= c_{n+1}, \\ \frac{\alpha_{i-2}}{E_{tot}} - \alpha_{i-1} &= c_i, \text{ for } i = 3, \dots, n, \\ \frac{\alpha_0 + K_0}{E_{tot}} - \alpha_1 &= c_2, \\ \frac{1 + S_{tot}K_0}{E_{tot}} - (\alpha_0 + K_0) &= c_1. \end{aligned} \quad (25)$$

Keep in mind that the values of c_i are given. We may solve (25) in terms of E_{tot} and of the $c_i, i = 1, \dots, n + 1$:

$$\begin{aligned} \alpha_{n-1-k} &= E_{tot} \sum_{i=0}^k c_{n+1-i} (E_{tot})^{k-i}, \text{ for each } k = 0, 1, \dots, n - 2, \\ \alpha_0 + K_0 &= E_{tot} \sum_{i=0}^{n-1} c_{n+1-i} (E_{tot})^{n-1-i}, \\ 1 + S_{tot}K_0 &= E_{tot} \sum_{i=0}^n c_{n+1-i} (E_{tot})^{n-i}. \end{aligned} \quad (26)$$

Note that if we take any value for $E_{tot} > 0$, then the values of α_i for $i = 1, \dots, n - 1$, $\alpha_0 + K_0$ and $S_{tot}K_0$ are completely determined. So, we find an appropriate value of E_{tot} such that the previous values α_i , K_0 and S_{tot} are all positive. For that, we choose $K_0 = 1$, and we take E_{tot} large enough such that

$$\begin{aligned} \sum_{i=0}^k c_{n+1-i} (E_{tot})^{k-i} &> 0, \text{ for each } k \in \{0, 1, \dots, n - 2\} \text{ with } k \text{ odd,} \\ E_{tot} \sum_{i=0}^{n-1} c_{n+1-i} (E_{tot})^{n-1-i} &> 1, \quad E_{tot} \sum_{i=0}^{n-1} c_{n+1-i} (E_{tot})^{n-i} > 1. \end{aligned}$$

This is possible since $c_{n+1} > 0$ and that imposing the first condition just on k odd is enough to ensure that it holds for all $k \in I_{n-1}$ as well because of the signs of the c_i . With these conditions, and using Eqs (26), the values of α_i for each $i = 0, \dots, n - 1$ and S_{tot} are determined and are all positive.

Now, it remains to show that we can choose reaction rate constants such that the values of $\alpha_i, i = 0, \dots, n - 1$ are the given ones. Recall that $\alpha_i = \frac{T_i}{(F_{tot})^{i+1}}$, where $T_i = \prod_{j=0}^i \frac{\tau_j}{v_j}$ for $i = 0, \dots, n - 1$ and $T_{-1} = 1$, where $\tau_0 = k_{cat_0}K_0 = k_{cat_0}$ (we have chosen $K_0 = 1$). Take for example $F_{tot} = 1, k_{on_0} = 2$,

$k_{\text{off}_0} = 1$ and $k_{\text{cat}_0} = 1$ (to obtain $K_0 = 1$). Then, $\tau_0 = 1$, so we take $\nu_0 = \frac{1}{\alpha_0}$. As $\alpha_{i+1} = \alpha_i \frac{\tau_{i+1}}{\nu_{i+1}}$, for $i = 0, \dots, n-2$, we can choose any positive values of τ_{i+1}, ν_{i+1} such that $\frac{\tau_{i+1}}{\nu_{i+1}} = \frac{\alpha_{i+1}}{\alpha_i}$, and we are done.

When n is odd, with a similar argument, we can find reaction rate constants and total conservation constants such that the polynomial $\tilde{q}(e)$ gives a polynomial like (24) (but with n distinct positive roots and one negative root). \square

Remark 9. If we assume that $k_{\text{off}_0} = 0$ (so, if there is no reverse unbinding reaction), Lemma 8 is still true. In this case, it is enough to set $K_0 = \frac{k_{\text{on}_0}}{k_{\text{cat}_0}}$ and then choose $k_{\text{on}_0} = 1$ instead of $k_{\text{on}_0} = 2$ at the end of the proof.

2.4. Lifting regions of multistationarity

Multistationarity of any of the subsystems G_J can be extended to the full n -site phosphorylation system (for instance, by Theorem 4 in [14]). We give a precise statement of this result in Theorem 10.

Consider the full n -site phosphorylation network (1), with a given vector of reaction rate constants $\kappa \in \mathbb{R}^{6n}$:

$$\kappa = (k_{\text{on}_0}, k_{\text{off}_0}, k_{\text{cat}_0}, \dots, k_{\text{on}_{n-1}}, k_{\text{off}_{n-1}}, k_{\text{cat}_{n-1}}, \ell_{\text{on}_0}, \ell_{\text{off}_0}, \ell_{\text{cat}_0}, \dots, \ell_{\text{on}_{n-1}}, \ell_{\text{off}_{n-1}}, \ell_{\text{cat}_{n-1}}).$$

We define the following rational functions of κ :

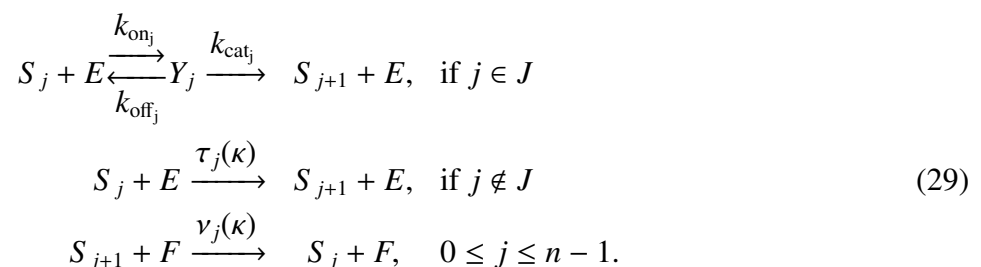
$$\tau_j(\kappa) = k_{\text{cat}_j} \mu_j(\kappa) \text{ if } j \notin J \text{ and } \nu_j(\kappa) = \ell_{\text{cat}_j} \eta_j(\kappa) \text{ for } j = 0, \dots, n-1, \quad (27)$$

where $\mu_j(\kappa)$ and $\eta_j(\kappa)$ are in turn the following rational functions:

$$\mu_j(\kappa) = \frac{k_{\text{on}_j}}{k_{\text{off}_j} + k_{\text{cat}_j}} \text{ if } j \notin J \text{ and } \eta_j(\kappa) = \frac{\ell_{\text{on}_j}}{\ell_{\text{off}_j} + \ell_{\text{cat}_j}} \text{ for } j = 0, \dots, n-1. \quad (28)$$

We denote by $\varphi: \mathbb{R}_{>0}^{6n} \rightarrow \mathbb{R}_{>0}^{2n+2|J|}$ the function that takes κ and gives a vector of (positive) reaction rate constants with the following order: first, the constants $k_{\text{on}_j}, k_{\text{off}_j}, k_{\text{cat}_j}, j \in J$, then $\tau(\kappa), j \notin J$, and then $\nu_j(\kappa), j = 0, \dots, n-1$.

Given a subset $J \subset I_n$ and a vector of reaction rate constants $\kappa \in \mathbb{R}_{>0}^{6n}$, we consider the subnetwork $G_J^{\varphi(\kappa)}$ as in Definition 1, with rate constants $\varphi(\kappa)$:



Applying a version of Theorem 4 in [14] that will appear in our forthcoming paper [18], we get the following lifting result.

Theorem 10. Consider the full n -site phosphorylation network (1) with fixed reaction rate constants κ^0 and the network $G_J^{\varphi(\kappa^0)}$, both with total conservation amounts $S_{tot}, E_{tot}, F_{tot} > 0$. Suppose that system $G_J^{\varphi(\kappa^0)}$ admits m nondegenerate positive steady states.

Then, there exists $\varepsilon_0 > 0$ such that for any choice of rate constants κ such that $\varphi(\kappa) = \varphi(\kappa^0)$ and

$$\max_{j \notin I} \mu_j(\kappa), \quad \max_{j \in I_n} \eta_j(\kappa) < \varepsilon_0, \quad (30)$$

the n -site sequential phosphorylation system admits m positive nondegenerate steady states in the stoichiometric compatibility class defined by S_{tot}, E_{tot} and $F_{tot} > 0$. Moreover, the set of rate constants κ verifying $\varphi(\kappa) = \varphi(\kappa^0)$ and (30) is nonempty.

Note that the function φ above is surjective, that is, any vector of reaction rate constants for the reduced network G_J can be obtained as $\varphi(\kappa)$, for some vector κ of reaction rate constants of the full n -site phosphorylation network. For instance, given $\tau_j \in \mathbb{R}_{>0}$, we can take $k_{onj} = 2\tau_j, k_{offj} = k_{catj} = 1$, and then $\tau_j(\kappa) = \tau_j$. Similarly, we can do this for the other reaction rate constants of G_J . Then, Corollary 10 allows us to obtain multistationary regions for the complete n -site phosphorylation system, combining the conditions on the parameters given in Theorem 2 and Theorem 4, with conditions (30) of Corollary 10. In particular, let $J_n \subset I_n$ as in (8). By lifting a multistationarity region for the system G_{J_n} in Corollary 5, we get a multistationarity region of parameters of the n -site phosphorylation cycle with $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states in the same stoichiometric compatibility class.

3. Positive solutions of sparse polynomial systems

As we saw in Subsection 2.1, the steady states of the systems G_J correspond to the positive solutions of the sparse polynomial system (14) in two variables. In this section, we briefly recall the general setting from [13, 19] to find lower bounds for sparse polynomial systems, that we will use in the proof of Theorems 2 and 4 in Section 4 and also in Section 5. For detailed examples of this approach, we refer the reader to Section 2 in [20].

Consider a finite point configuration $\mathcal{A} = \{a_1, \dots, a_n\} \subset \mathbb{Z}^d$, with $n \geq d + 2$. A sparse *polynomial system* of d Laurent polynomial equations with *support* \mathcal{A} is a system $f_1(x) = \dots = f_d(x) = 0$ in d variables $x = (x_1, \dots, x_d)$, with

$$f_i(x) = \sum_{j=1}^n c_{ij} x^{a_j} \in \mathbb{R}[x_1, \dots, x_d], \quad i = 1, \dots, d, \quad (31)$$

where the exponents belong to \mathcal{A} . We call $C = (c_{ij}) \in \mathbb{R}^{d \times n}$ the *coefficient matrix* of the system and we assume that no column of C is identically zero. Recall that a zero of (31) is nondegenerate when it is not a zero of the Jacobian determinant of $f = (f_1, \dots, f_d)$.

Our method to obtain a lower bound on the number of positive steady states, based on [13, 19], is to restrict our polynomial system (31) to subsystems which have a positive solution and then extend these solutions to the total system, via a deformation of the coefficients. The first step is thus to find conditions in the coefficient matrix C that guarantee a positive solution to each of the subsystems. The hypothesis of having subsystems supported on *simplices* of a *regular subdivision* of a finite point

configuration \mathcal{A} is the key to the existence of an open set in the space of coefficients where all these solutions can be extended. We recall below only the concepts that we need for our work. A comprehensive treatment of this subject can be found in [21]. Following Section 3 in [19], we define:

Definition 11. Given a matrix $M \in \mathbb{R}^{d \times (d+1)}$, we denote by $\text{minor}(M, i)$ the determinant of the square matrix obtained by removing the i -th column of M . The matrix M is called positively spanning if all the values $(-1)^i \text{minor}(M, i)$, for $i = 1, \dots, d + 1$, are nonzero and have the same sign.

Equivalently, a matrix M is positively spanning if all the coordinates of any nonzero vector in $\ker(M)$ are non-zero and have the same sign.

A d -simplex with vertices in a finite point configuration \mathcal{A} is a subset of $d + 1$ points of \mathcal{A} which is affinely independent. By Proposition 3.3 in [19], a system of d polynomial equations in d variables with a d -simplex as support, has one non-degenerate positive solution if and only if its $d \times (d + 1)$ matrix of coefficients is positively spanning. We further define, following [19]:

Definition 12. Let $C \in \mathbb{R}^{d \times n}$ and a finite point configuration $\mathcal{A} = \{a_1, \dots, a_n\} \subset \mathbb{Z}^d$, with $n \geq d + 2$. We say that a d -simplex $\Delta = \{a_{i_1}, \dots, a_{i_{d+1}}\}$ is positively decorated by C if the $d \times (d + 1)$ submatrix of C with columns $\{i_1, \dots, i_{d+1}\}$ is positively spanning.

Given a finite point configuration \mathcal{A} , take a *height function*

$$h : \mathcal{A} \rightarrow \mathbb{R}, \quad h = (h(a_1), \dots, h(a_n)).$$

Consider the *lower convex hull* \mathcal{L} of the n lifted points $(a_j, h(a_j)) \in \mathbb{R}^{d+1}$, $j = 1, \dots, n$ (see Figure 1). Project to \mathbb{R}^d the subsets of points in each of the faces of \mathcal{L} . These subsets define a *regular subdivision* of \mathcal{A} induced by h . When the height vector h is generic, the regular subdivision is a regular triangulation, in which all the subsets are simplices. For the specifics on this notions we refer the reader to Section 2.2 of [21].

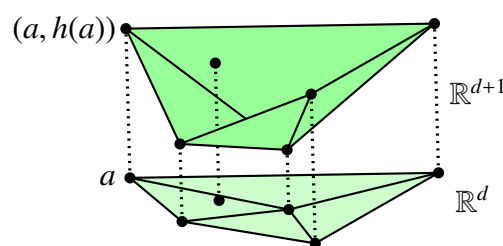


Figure 1. Regular triangulation.

It can be proved that the set of all height vectors inducing a regular subdivision of \mathcal{A} that contains certain d -simplices $\Delta_1, \dots, \Delta_p$ is defined by a finite number of linear inequalities, see Chapter 2 of [21]. Thus, this set is a finitely generated convex cone $C_{\Delta_1, \dots, \Delta_p}$ in \mathbb{R}^n with apex at the origin, which can be obtained as follows. Given a d -simplex Δ with vertices in \mathcal{A} , we consider height vectors $h \in \mathbb{R}^n$, where each coordinate h_j of h gives the value of the corresponding lifting function at the point a_j of \mathcal{A} .

Denote by $\varphi_{\Delta,h}$ the unique affine function that agrees with h on the points of Δ , that is, $\varphi_{\Delta,h}(a_j) = h_j$ for all $a_j \in \Delta$. We associate with Δ the following cone:

$$C_{\Delta} = \{h = (h_1, \dots, h_n) \in \mathbb{R}^n : \varphi_{\Delta,h}(a_j) < h_j \text{ for all } a_j \notin \Delta\}.$$

Note that each inequality $\varphi_{\Delta,h}(a_j) < h_j$ is linear in h and it can thus be written as $\langle m, h \rangle > 0$, for a certain vector m that depends on the points in Δ ($\langle \cdot, \cdot \rangle$ denotes the standard inner product in \mathbb{R}^n). As $C_{\Delta_1, \dots, \Delta_p} = \bigcap_{i=1}^p C_{\Delta_i}$, this gives a set of linear inequalities that define the cone $C_{\Delta_1, \dots, \Delta_p}$. We refer the reader to Example 2.5 in [20] for a complete computation. The set of all height vectors inducing a regular subdivision Γ of \mathcal{A} is a finitely generated convex cone in \mathbb{R}^n , which we call C_{Γ} . In particular, if the simplices $\Delta_1, \dots, \Delta_p$ completely determine the regular subdivision Γ , the cone $C_{\Delta_1, \dots, \Delta_p}$ is equal to the cone C_{Γ} .

Given a finite point configuration $\mathcal{A} \subset \mathbb{Z}^d$, consider a system of sparse real polynomials $f_1 = f_2 = \dots = f_d = 0$ as in (31) with support \mathcal{A} . Let Γ be a regular subdivision of \mathcal{A} and h any height function that induces Γ . We define the following family of real polynomial systems parametrized by a positive real number $t > 0$:

$$\sum_{j=1}^n c_{ij} t^{h(a_j)} x^{a_j} = 0, \quad i = 1, \dots, d. \quad (32)$$

We also consider the following family of polynomial systems parametrized by $\gamma \in \mathbb{R}_{>0}^n$:

$$\sum_{j=1}^n c_{ij} \gamma_j x^{a_j} = 0, \quad i = 1, \dots, d. \quad (33)$$

Note that each polynomial system in the families (32) and (33) has again support \mathcal{A} .

Theorem 13 (Theorem 3.4 of [19] and Theorem 2.11 of [13]). *Consider $\Delta_1, \dots, \Delta_p$ d -simplices which occur in a regular subdivision Γ of a finite point configuration $\mathcal{A} \subset \mathbb{Z}^d$, and which are positively decorated by a matrix $C \in \mathbb{R}^{d \times n}$.*

1. *Let h be any height function that defines Γ . Then, there exists a positive real number t_0 such that for all $0 < t < t_0$, the number of nondegenerate positive solutions of (32) is at least p .*
2. *Let $m_1, \dots, m_{\ell} \in \mathbb{R}^n$ be vectors that define a presentation of the cone $C_{\Delta_1, \dots, \Delta_p}$:*

$$C_{\Delta_1, \dots, \Delta_p} = \{h \in \mathbb{R}^n : \langle m_j, h \rangle > 0, \quad j = 1, \dots, \ell\}.$$

Then, for any $\varepsilon \in (0, 1)^{\ell}$ there exists $t_0(\varepsilon) > 0$ such that for any γ in the open set

$$U = \bigcup_{\varepsilon \in (0, 1)^{\ell}} \{\gamma \in \mathbb{R}_{>0}^n ; \gamma^{m_j} < t_0(\varepsilon)^{\varepsilon_j}, \quad j = 1, \dots, \ell\},$$

system (33) has at least p nondegenerate positive solutions.

We remark that in the first item in Theorem 13 (Theorem 3.4 in [19]), we describe a piece of curve in the space of coefficients as we vary $t > 0$ where the associated system has at least p positive solutions, while in the second item in Theorem 13 (Theorem 2.11 in [13]), we describe a subset with nonempty interior in the space of coefficients where we can bound from below by p the number of positive solutions of the associated system.

4. Proofs of Theorems 2 and 4

We start this section with a lemma.

Lemma 14. Consider $\mathcal{A} = \{(1, 0), (0, 1), (1, 1), (1, 2), \dots, (1, n), (0, 0)\} \subset \mathbb{Z}^2$. The triangulation Γ of \mathcal{A} with the following 2-simplices:

$$\{(1, 0), (1, 1), (0, 0)\}, \{(1, 1), (1, 2), (0, 0)\}, \dots, \{(1, n-1), (1, n), (0, 0)\}, \{(1, n), (0, 1), (0, 0)\}$$

is regular (see Figure 2).

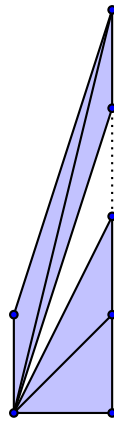


Figure 2. Triangulation Γ of \mathcal{A} .

Proof. We can take $h: \mathcal{A} \rightarrow \mathbb{R}$, with $h(0, 0) = 0$, $h(0, 1) = n$ and $h(1, j) = \frac{j(j-1)}{2}$, for $j = 0, \dots, n-1$. It is easy to check h defines a regular triangulation that is equal to Γ . □

The idea in the proofs of Theorem 2 and Theorem 4 is to detect positively decorated simplices in the regular triangulation Γ .

Proof of Theorem 2. By Proposition 7, the number of positive solutions of the system G_{I_n} is at most $2\lfloor \frac{n}{2} \rfloor + 1$. So, it is enough to prove that this number is also a lower bound.

The number of positive steady states of the system G_{I_n} is the number of positive solutions of the system (14). As we saw before, the support of this last system is

$$\mathcal{A} = \{(1, 0), (0, 1), (1, 1), (1, 2), \dots, (1, n), (0, 0)\} \subset \mathbb{Z}^2,$$

with coefficient matrix C (16). Note that if one multiplies a column of C by a positive number, then a simplex is positively decorated by C if and only if it is positively decorated by the new matrix. After multiplying the columns by convenient positive numbers, we obtain the following matrix from C :

$$C_{simple} = \begin{pmatrix} 1 & 0 & M_0 & \dots & M_{n-1} & -S_{tot} \\ 0 & 1 & 1 & \dots & 1 & -E_{tot} \end{pmatrix},$$

where $M_i = \frac{k_{cat_i}}{v_i F_{tot}} + 1$, for each $i = 0, \dots, n-1$. We will work with this new matrix C_{simple} .

We consider the regular triangulation Γ in Lemma 14. The simplex $\{(1, 0), (1, 1), (0, 0)\}$ of Γ is positively decorated by *Csimple* if and only if $E_{tot}M_0 - S_{tot} < 0$. The simplex $\{(1, j), (1, j + 1), (0, 0)\}$, for $j = 1, \dots, n - 1$, corresponds to the submatrix

$$Csimple_j = \begin{pmatrix} M_{j-1} & M_j & -S_{tot} \\ 1 & 1 & -E_{tot} \end{pmatrix},$$

and it is positively decorated by *Csimple* if and only if $E_{tot}M_{j-1} - S_{tot}$ and $E_{tot}M_j - S_{tot}$ have opposite signs. The last simplex $\{(0, 1), (1, n), (0, 0)\}$ is positively decorated by *Csimple* if and only if $E_{tot}M_{n-1} - S_{tot} > 0$.

Therefore we always have at least n positively decorated simplices using all simplices of Γ but the last one, just by imposing

$$(E_{tot}M_i - S_{tot})(-1)^i < 0, \quad \text{for } i = 0, \dots, n - 1. \quad (34)$$

We can include the last simplex if and only if n is even (because otherwise the inequalities are not compatible), and in this case we have at least $n + 1$ positively decorated simplices. We can obtain $2\lfloor \frac{n}{2} \rfloor + 1$ positively decorated simplices if the inequalities (34) are satisfied. These inequalities are equivalent to the inequalities (6) in the statement.

Assume (6) holds. Given any height function h inducing the triangulation Γ , by item (1) in Theorem 13 there exists t_0 in $\mathbb{R}_{>0}$ such that for all $0 < t < t_0$, the number of positive nondegenerate solutions of the deformed system as in (32) with support \mathcal{A} and coefficient matrix C_t , with $(C_t)_{ij} = t^{h(\alpha_j)}c_{ij}$ (with $\alpha_j \in \mathcal{A}$, $C = (c_{ij})$), is at least $2\lfloor \frac{n}{2} \rfloor + 1$. In particular if we choose h as in the proof of Lemma 14, there exists t_0 in $\mathbb{R}_{>0}$, such that for all $0 < t < t_0$, the system

$$\begin{aligned} s_0 + \sum_{j=0}^{n-1} \left(\frac{T_j}{(F_{tot})^{j+1}} + \frac{K_j T_{j-1}}{(F_{tot})^j} \right) t^{\frac{j(j+1)}{2}} s_0 e^{j+1} - S_{tot} &= 0, \\ t^n e + \sum_{j=0}^{n-1} \frac{K_j T_{j-1}}{(F_{tot})^j} t^{\frac{j(j+1)}{2}} s_0 e^{j+1} - E_{tot} &= 0, \end{aligned} \quad (35)$$

has at least $2\lfloor \frac{n}{2} \rfloor + 1$ positive solutions. If we change the variable $\bar{e} = t^n e$, we get the following system:

$$\begin{aligned} s_0 + \sum_{j=0}^{n-1} \left(\frac{T_j}{(F_{tot})^{j+1}} + \frac{K_j T_{j-1}}{(F_{tot})^j} \right) t^{(j+1)(\frac{j}{2}-n)} s_0 \bar{e}^{j+1} - S_{tot} &= 0, \\ \bar{e} + \sum_{j=0}^{n-1} \frac{K_j T_{j-1}}{(F_{tot})^j} t^{(j+1)(\frac{j}{2}-n)} s_0 \bar{e}^{j+1} - E_{tot} &= 0. \end{aligned} \quad (36)$$

It is straightforward to check that if we scale the constants K_j by

$$t^{j-n} K_j, \quad j = 0, \dots, n - 1, \quad (37)$$

while keeping fixed the values of the constants k_{cat_j} , ν_j for $j = 0, \dots, n - 1$ and the values of the total conservation constants E_{tot} , F_{tot} and S_{tot} (assuming that condition (6) holds), the intersection of the steady state variety and the linear conservation equations of the corresponding network is described by

system (36). It is easy to check that in order to get the scaling in (37), it is sufficient to rescale only the original constants k_{on_j} as follows: $t^{j-n}k_{\text{on}_j}$, for $j = 0, \dots, n-1$. Then, for these choices of constants, the system has at least $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states.

Now, we will show how to obtain the more general rescaling in the statement. The existence of the positive constants B_1, \dots, B_n follows from the inequalities that define the cone C_Γ of heights inducing the regular triangulation Γ and item (2) in Theorem 13. For instance, we can check that C_Γ is defined by n inequalities:

$$C_\Gamma = \{h = (h_1, \dots, h_{n+3}) \in \mathbb{R}^{n+3} : \langle m_j, h \rangle > 0, j = 1, \dots, n\},$$

where $m_1 = e_1 - 2e_3 + e_4$, $m_j = e_{j+1} - 2e_{j+2} + e_{j+3}$, for $j = 2, \dots, n-1$ and $m_n = e_2 + e_{n+1} - e_{n+2} - e_{n+3}$, where e_i denotes the i -th canonical vector of \mathbb{R}^{n+3} . Fix $\varepsilon \in (0, 1)^{n+3}$. As (6) holds, item (2) in Theorem 13 says that there exist positive numbers B_j for $j = 1, \dots, n$ (depending on ε), such that the system

$$\begin{aligned} \gamma_1 s_0 + \sum_{j=0}^{n-1} \left(\frac{T_j}{(F_{\text{tot}})^{j+1}} + \frac{K_j T_{j-1}}{(F_{\text{tot}})^j} \right) \gamma_{j+3} s_0 e^{j+1} - \gamma_{n+3} S_{\text{tot}} &= 0, \\ \gamma_2 e + \sum_{j=0}^{n-1} \frac{K_j T_{j-1}}{(F_{\text{tot}})^j} \gamma_{j+3} s_0 e^{j+1} - \gamma_{n+3} E_{\text{tot}} &= 0, \end{aligned} \quad (38)$$

has at least $2\lfloor \frac{n}{2} \rfloor + 1$ nondegenerate positive solutions, for any vector $\gamma \in \mathbb{R}^{n+3}$ satisfying $\gamma^{m_j} < B_j$, for all $j = 1, \dots, n$. In particular, this holds if we take $\gamma_1 = \gamma_2 = \gamma_{n+3} = 1$ and

$$\gamma_3^{-2} \gamma_4 < B_1, \quad \gamma_{j+1} \gamma_{j+2}^{-2} \gamma_{j+3} < B_j, \quad \text{for } j = 2, \dots, n-1, \quad \gamma_{n+1} \gamma_{n+2}^{-1} < B_n. \quad (39)$$

If we call $\lambda_0 = \gamma_3$ and $\lambda_j = \frac{\gamma_{j+3}}{\gamma_{j+2}}$ for $j = 1, \dots, n-1$, the inequalities in (39) are equivalent to the conditions (7). Then, if λ_j , $j = 0, \dots, n-1$, satisfy these inequalities, the rescaling of the given parameters k_{on_j} by $\lambda_j k_{\text{on}_j}$ for $j = 0, \dots, n-1$, gives rise to a system with exactly $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states. \square

The proof of Theorem 4 is similar to the previous one.

Proof of Theorem 4. Again, the number of positive steady states of our system is equal to the number of positive solutions of the system (14). Recall that the support of the system is

$$\mathcal{A} = \{(1, 0), (0, 1), (1, 1), (1, 2), \dots, (1, n), (0, 0)\} \subset \mathbb{Z}^2.$$

In this case, the coefficient matrix C (16) is equal, after multiplying the columns by convenient positive numbers, to the matrix

$$C_{\text{simple}} = \begin{pmatrix} 1 & 0 & M_0 & \dots & M_{n-1} & -S_{\text{tot}} \\ 0 & 1 & D_0 & \dots & D_{n-1} & -E_{\text{tot}} \end{pmatrix},$$

where $M_i = \frac{k_{\text{cat}_i}}{\nu_i F_{\text{tot}}} + 1$ and $D_i = 1$, for each $i \in J$, and $M_i = 1$ and $D_i = 0$, for each $i \notin J$.

We consider again the regular triangulation Γ in Lemma 14. Recall that $J \subset J_n$, see (8), and therefore each $j \in J$ has the same parity as n , in particular $0 \leq j \leq n-2$. For each $j \in J$, consider the simplices $\Delta_j = \{(1, j), (1, j+1), (0, 0)\}$ and $\Delta_{j+1} = \{(1, j+1), (1, j+2), (0, 0)\}$. Note that if $j \neq j'$ then $\{\Delta_j, \Delta_{j+1}\}$ and $\{\Delta_{j'}, \Delta_{j'+1}\}$ are disjoint since j, j' and n have the same parity.

The simplices are positively decorated by $Csimple$ (and then by C) if and only if the submatrices

$$Csimple_j = \begin{pmatrix} 1 & M_j & -S_{tot} \\ 0 & 1 & -E_{tot} \end{pmatrix}, \quad Csimple_{j+1} = \begin{pmatrix} M_j & 1 & -S_{tot} \\ 1 & 0 & -E_{tot} \end{pmatrix},$$

are positively spanning, and this happens if and only if $E_{tot}M_j - S_{tot} < 0$, where $M_j = \frac{k_{cat_j}}{v_j F_{tot}} + 1$, since $j \in J$. The simplex $\Delta_n = \{(0, 1), (1, n), (0, 0)\}$ is trivially positively decorated by $Csimple$. Then, by imposing the inequalities $E_{tot}M_j - S_{tot} < 0$ for $j \in J$, which are equivalent to the ones in the statement (9), we can obtain $2|J| + 1$ positively decorated simplices.

Assume (9) holds. Given any height function h inducing the triangulation Γ , by item (1) in Theorem 13 there exists t_0 in $\mathbb{R}_{>0}$, such that for all $0 < t < t_0$, the number of positive nondegenerate solutions of the deformed system with support \mathcal{A} and coefficient matrix C_t , with $(C_t)_{ij} = t^{h(\alpha_j)} c_{ij}$ (with $\alpha_j \in \mathcal{A}$, $C = (c_{ij})$) is at least $2|J| + 1$. In particular if we choose h as in the proof of Lemma 14, there exists t_0 in $\mathbb{R}_{>0}$, such that for all $0 < t < t_0$, the system

$$s_0 + \sum_{j \in J} \left(\frac{T_j}{(F_{tot})^{j+1}} + \frac{K_j T_{j-1}}{(F_{tot})^j} \right) t^{\frac{j(j+1)}{2}} s_0 e^{j+1} + \sum_{j \notin J} \frac{T_j}{(F_{tot})^{j+1}} t^{\frac{j(j+1)}{2}} s_0 e^{j+1} - S_{tot} = 0, \quad (40)$$

$$t^n e + \sum_{j \in J} \frac{K_j T_{j-1}}{(F_{tot})^j} t^{\frac{j(j+1)}{2}} s_0 e^{j+1} - E_{tot} = 0,$$

has at least $2|J| + 1$ positive solutions. If we change the variable $\bar{e} = t^n e$, we get the following system:

$$s_0 + \sum_{j \in J} \left(\frac{T_j}{(F_{tot})^{j+1}} + \frac{K_j T_{j-1}}{(F_{tot})^j} \right) t^{(j+1)(\frac{j}{2}-n)} s_0 \bar{e}^{j+1} + \sum_{j \notin J} \frac{T_j}{(F_{tot})^{j+1}} t^{(j+1)(\frac{j}{2}-n)} s_0 \bar{e}^{j+1} - S_{tot} = 0,$$

$$\bar{e} + \sum_{j \in J} \frac{K_j T_{j-1}}{(F_{tot})^j} t^{(j+1)(\frac{j}{2}-n)} s_0 \bar{e}^{j+1} - E_{tot} = 0. \quad (41)$$

Similarly as we did in the previous proof, if we scale the original parameters k_{on_j} , for $j \in J$, and τ_j if $j \notin J$ by

$$t^{j-n} k_{on_j} \text{ if } j \in J, \quad t^{j-n} \tau_j \text{ if } j \notin J, \quad (42)$$

respectively, and if we keep fixed the values of the remaining rate constants and the values of the total conservation constants E_{tot} , F_{tot} and S_{tot} , the intersection of the steady state variety and the linear conservation relations is described by system (41). Then, for these choices of constants the system G_J has at least $2|J| + 1$ positive steady states. The general rescaling that appears in the statement can be obtained in a similar way as we did in the proof of Theorem 2. \square

Remark 15. If some or all rate constants k_{off_i} are zero (so, if the corresponding reactions are not assumed to happen), Theorems 2 and 4 hold. The proofs are very similar, except that if $k_{off_i} = 0$, the constant K_i is equal to $K_i = \frac{k_{on_i}}{k_{cat_i}}$.

5. Computer aided results

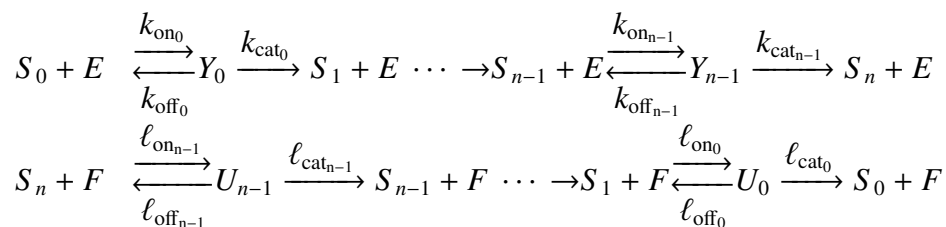
In this section we explore a computational approach to the multistationarity problem, more precisely we find new regions of multistationarity. The idea is to find good regular triangulations of

the point configuration corresponding to the exponents of the polynomial system given by the steady state equations and conservation laws, and deduce from them some regions of multistationarity. We first give the idea and then apply it for the n -site phosphorylation system for $n = 2, 3, 4$, and 5 , where we have successfully found several regions of multistationarity. This approach can be, in principle, applied to other systems if they satisfy certain hypotheses, see [13, Theorem 5.4], and are sufficiently small in order for the computations to be done in a reasonable amount of time.

The strategy is the following. Given a polynomial system with support \mathcal{A} and matrix of coefficients C , one first computes all possible regular triangulations of \mathcal{A} with the aid of a computer. The number of such triangulations can be very large depending on the size of \mathcal{A} , thus the next step is to discard in each such triangulation the simplices that obviously will not be positively decorated by C . With the reduced number of triangulations one can now search through all of them for the ones giving the biggest number of simultaneously positively decorated simplices. Each set of k simultaneously positively decorated simplices gives a candidate for a region of multistationarity with k positive steady states. If one finds m of such sets, then it is possible to have up to m such regions. Have in mind, however, that among these regions can be repetitions.

Next we apply this to the n -site phosphorylation system with all intermediates and explain more concretely this procedure in this case.

In Corollary 5 we obtained regions of multistationarity with $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states each using only $1/4$ of the intermediates, our objective now is to understand if it is possible to find more such regions with more intermediates. Consider the network G of the n -site phosphorylation system with all possible intermediates:



In Section 4 of [13], the concentration at steady state of all species are given in terms of the species s_0, e, f :

$$\begin{aligned}
 s_i &= T_{i-1} \frac{s_0 e^i}{f^i}, \quad i = 1, \dots, n, \\
 y_i &= K_i T_{i-1} \frac{s_0 e^{i+1}}{f^i}, \quad i = 0, \dots, n-1, \\
 u_i &= L_i T_i \frac{s_0 e^{i+1}}{f^i}, \quad i = 0, \dots, n-1,
 \end{aligned}$$

where $K_i = \frac{k_{\text{on}_i}}{k_{\text{off}_i} + k_{\text{cat}_i}}$, $L_i = \frac{\ell_{\text{on}_i}}{\ell_{\text{off}_i} + \ell_{\text{cat}_i}}$, $T_i = \prod_{j=0}^i \frac{\tau_j}{\nu_j}$ for each $i = 0, \dots, n-1$ (recall that K_i^{-1} and L_i^{-1} are usually called *Michaelis-Menten constants*) and $T_{-1} = 1$, where $\tau_i = k_{\text{cat}_i} K_i$ and $\nu_i = \ell_{\text{cat}_i} L_i$, for each $i = 0, \dots, n-1$.

This looks very similar to (13), where F_{tot} is replaced by f , which is now a variable, and so we need to work in dimension 3. The main difference is that the networks G_J considered in the previous sections have intermediates only in the E component and the network G we consider in this section has all intermediates.

Note from (2) that the support \mathcal{A} of this system, which has $2n + 4$ elements, ordering the variables

as s_0, e, f , is given by the columns of the following matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 & \dots & 1 & 1 & 1 & \dots & 1 & 0 \\ 0 & 1 & 0 & 1 & \dots & n & 1 & 2 & \dots & n & 0 \\ 0 & 0 & 1 & -1 & \dots & -n & 0 & -1 & \dots & 1-n & 0 \end{pmatrix},$$

and the corresponding matrix of coefficients for the system is

$$C = \begin{pmatrix} 1 & 0 & 0 & T_0 & \dots & T_{n-1} & K_0 + L_0 T_0 & K_1 T_0 + L_1 T_1 & \dots & K_{n-1} T_{n-2} + L_{n-1} T_{n-1} & -S_{tot} \\ 0 & 1 & 0 & 0 & \dots & 0 & K_0 & K_1 T_0 & \dots & K_{n-1} T_{n-2} & -E_{tot} \\ 0 & 0 & 1 & 0 & \dots & 0 & L_0 T_0 & L_1 T_1 & \dots & L_{n-1} T_{n-1} & -F_{tot} \end{pmatrix}.$$

Recall that if one multiplies a column of a matrix C by a positive number, then a simplex is positively decorated by C if and only if it is positively decorated by the modified matrix. So, in order to test whether a simplex with vertices in \mathcal{A} is positively decorated by C is enough to test if it is positively decorated by the following matrix

$$C_{simple} = \begin{pmatrix} 1 & 0 & 0 & 1 & \dots & 1 & 1 & 1 & \dots & 1 & -S_{tot} \\ 0 & 1 & 0 & 0 & \dots & 0 & N_0 & N_1 & \dots & N_{n-1} & -E_{tot} \\ 0 & 0 & 1 & 0 & \dots & 0 & 1-N_0 & 1-N_1 & \dots & 1-N_{n-1} & -F_{tot} \end{pmatrix},$$

where $0 < N_i = \frac{K_i T_{i-1}}{K_i T_{i-1} + L_i T_i} = \left(1 + \frac{k_{cat_i}}{l_{cat_i}}\right)^{-1} < 1$ for $i = 0, 1, \dots, n-1$. Here the matrix C_{simple} is obtained by dividing the fourth until the last column by its first entry.

Now we compute all possible regular triangulations of \mathcal{A} and search through them looking for the ones with the maximal possible number of simplices simultaneously positively decorated by C_{simple} . Since the number of such triangulations grows very fast with n we approach it with the following strategy:

Algorithm 16. 1. Compute $L_1 := \{\text{all possible triangulations of } \mathcal{A}\}$.*

2. With L_1 compute L_2 by discarding all simplices which do not have the last vertex $(0, 0, 0)$. In fact we only need these simplices since a simplex not containing the last vertex cannot be positively decorated, because the corresponding coefficients of C_{simple} will be all positive.
3. Compute L_3 from L_2 by removing all simplices with the corresponding 3×4 submatrix of C_{simple} having a zero 3×3 minor. The reason for this is clear, such simplices will never be positively decorated by C_{simple} .
4. Compute L_4 from L_3 using the symmetry of C_{simple} . More precisely, change any index $4, 5, \dots, n+3$ on the simplex to 1 because on C_{simple} they yield the same column. Here we are using the easy-to-check fact that changing the order of indexes does not change the conditions for a simplex to be positively decorated.

5. Compute L_5 from L_4 removing all $T \in L_4$ such that there is another $T' \in L_4$ with $T \subset T'$.

*We are calling L_1, \dots, L_7 the sets defined in Algorithm 16. They are completely unrelated to the rational functions of the rate constants denoted with the same letters.

6. For each $T \in L_5$ and each set $S \subset T$ of simplices check if there is viable N_0, \dots, N_{n-1} such that all $\Delta \in S$ are positively decorated by *Csimple* at the same time. Call L_6 the list of such S 's.
7. If the maximum size of an element in L_6 is k , set $L_7 := \{T \in L_6 : \#T = k\}$. This k is the number of positive steady states and $m := \#L_7$ is the number of candidates for regions of multistationarity.

Step (1) can be done with the package TOPCOM inside SAGE [22], the other steps are quite simple to implement, for instance in MAPLE [23]. We show in Table 1 the number of elements in some of the lists and an approximation of the computation time for small values of n .

Table 1. Computation time and size of some lists while running Algorithm 16 .

n	$\#L_1$	$\#L_2$	$\#L_3$	$\#L_4$	$\#L_5$	$\#L_7$	k	regions of multistationarity	computation time
2	44	25	15	7	6	1	3	1	negligible
3	649	260	100	21	18	6	3	6	about 1 sec
4	9094	2728	682	62	53	5	5	4	about 2 min
5	122835	28044	4560	177	149	23	5	15	about 3 hours

The most computationally expensive part is to compute all regular triangulations, taking at least 90% of the time. These computations were done in a Linux virtual machine with 4 MB of RAM and with 4 cores of 3.2 GHz of processing. With a faster computer or more time one probably can do $n = 6$ or even $n = 7$ but probably not much more than this. For $n = 5$ just the file for the raw list L_1 of regular triangulations already has 10Mb.

An alternative path to Steps (6) and (7) is to set a number k and look for sets $T \in L_5$ and $S \subset T$ with $\#S \geq k$ such that there is viable N_0, \dots, N_{n-1} such that all $\Delta \in S$ are positively decorated by *Csimple* at the same time. We actually used this with $k = 2\lfloor \frac{n}{2} \rfloor + 1$. This other route depends upon a good guess one may previously have at how many positive steady states to expect.

After Step (7) one has to determine if there are any repetitions among the candidates for regions of multistationarity in L_7 and also if there are any superfluous candidates of regions, that is conditions C_1 and C_2 such that C_1 implies C_2 . In our case we did it by hand since the $\#L_7$ was quite small.

Once Step (7) is done, one has a list of inequalities for each element S of L_7 . These come from the conditions imposing that the simplices in S are positively decorated by *Csimple*. We are going to use these conditions to describe the regions of multistationarity. Because of the uniformity of *Csimple* the only kind of conditions that appear are

$$(I)_{i,j} \quad N_i - N_j > 0$$

$$(II)_i \quad S_{tot}N_i - E_{tot} > 0$$

$$(III)_i \quad E_{tot}N_i + F_{tot}N_i - E_{tot} > 0$$

$$(IV)_i \quad -S_{tot}N_i - F_{tot} + S_{tot} > 0$$

$$(V) \quad S_{tot} > E_{tot} + F_{tot},$$

or the opposite inequalities, and these translate from the N_i to the k_{cat_i}, ℓ_{cat_i} as follows

$$\begin{aligned}
 \text{(I)}'_{i,j} & \frac{k_{\text{cat}_j}}{\ell_{\text{cat}_j}} > \frac{k_{\text{cat}_i}}{\ell_{\text{cat}_i}} \\
 \text{(II)}'_i & \frac{S_{\text{tot}} - E_{\text{tot}}}{E_{\text{tot}}} > \frac{k_{\text{cat}_i}}{\ell_{\text{cat}_i}} \\
 \text{(III)}'_i & \frac{F_{\text{tot}}}{E_{\text{tot}}} > \frac{k_{\text{cat}_i}}{\ell_{\text{cat}_i}} \\
 \text{(IV)}'_i & \frac{F_{\text{tot}}}{S_{\text{tot}} - F_{\text{tot}}} < \frac{k_{\text{cat}_i}}{\ell_{\text{cat}_i}}.
 \end{aligned}$$

Note that

- conditions (III)_i and (V) together imply (II)_i;
- the opposite of condition (II)_i together with condition (V) imply the opposite of (III)_i;
- the opposite of condition (III)_i together with condition (V) imply (IV)_i;
- the opposite of condition (IV)_i together with condition (V) imply (III)_i;
- condition (III)_i and the opposite of (III)_j together imply (I)_{i,j}.

Using these properties it is easy to describe in a nice manner the regions of multistationarity and discard the repeated and superfluous ones. We sum up our findings on the following results which are proved in the same fashion as Theorems 2 and 4, once you have the regular triangulation obtained with the computer script. In the following propositions we describe the regions of multistationarity for $n = 2, 3, 4$ and 5. Note that we are not describing the precise rescalings; they are potentially different for different rate constants and similar to those in (7) and (10).

Proposition 17. *Let $n = 2$. Assume that $S_{\text{tot}} > E_{\text{tot}} + F_{\text{tot}}$. Then there is a choice of reaction rate constants for which the distributive sequential 2-site phosphorylation system admits 3 positive steady states. More explicitly, given rate constants and total concentrations such that*

$$\frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \quad (43)$$

after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 2-site phosphorylation system has 3 positive steady states.

Proposition 18. *Let $n = 3$. Assume that $S_{\text{tot}} > E_{\text{tot}} + F_{\text{tot}}$. Then, there is a choice of rate constants for which the distributive sequential 3-site phosphorylation system admits at least 3 positive steady states. More explicitly, if the rate constants and total concentrations are in one of the regions described below*

$$\begin{aligned}
 \text{(R}_{3.1}\text{)} & \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \\
 \text{(R}_{3.2}\text{)} & \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}}, \\
 \text{(R}_{3.3}\text{)} & \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}},
 \end{aligned}$$

then after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 3-site phosphorylation system has at least 3 positive steady states.

Proposition 19. Let $n = 3$. If the rate constants and total concentrations are in one of the regions described below

$$(R_{3.4}) \quad \max \left\{ \frac{F_{\text{tot}}}{E_{\text{tot}}}, \frac{F_{\text{tot}}}{S_{\text{tot}} - F_{\text{tot}}} \right\} < \min \left\{ \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}}, \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} \right\}, \quad S_{\text{tot}} > F_{\text{tot}},$$

$$(R_{3.5}) \quad \max \left\{ \frac{F_{\text{tot}}}{E_{\text{tot}}}, \frac{F_{\text{tot}}}{S_{\text{tot}} - F_{\text{tot}}} \right\} < \min \left\{ \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} \right\}, \quad S_{\text{tot}} > F_{\text{tot}},$$

$$(R_{3.6}) \quad \min \left\{ \frac{F_{\text{tot}}}{E_{\text{tot}}}, \frac{S_{\text{tot}} - E_{\text{tot}}}{E_{\text{tot}}} \right\} > \max \left\{ \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} \right\}, \quad S_{\text{tot}} > E_{\text{tot}},$$

then after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 3-site phosphorylation system has at least 3 positive steady states.

Proposition 20. Let $n = 4$. Assume that $S_{\text{tot}} > E_{\text{tot}} + F_{\text{tot}}$. Then, there is a choice of rate constants for which the distributive sequential 4-site phosphorylation system has at least 5 steady states. More explicitly, if the rate constants and total concentrations are in one of the regions described below

$$(R_{4.1}) \quad \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \min \left\{ \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \frac{k_{\text{cat}_3}}{\ell_{\text{cat}_3}} \right\},$$

$$(R_{4.2}) \quad \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \min \left\{ \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}}, \frac{k_{\text{cat}_3}}{\ell_{\text{cat}_3}} \right\},$$

$$(R_{4.3}) \quad \max \left\{ \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}}, \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} \right\} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_3}}{\ell_{\text{cat}_3}},$$

$$(R_{4.4}) \quad \max \left\{ \frac{k_{\text{cat}_0}}{\ell_{\text{cat}_0}}, \frac{k_{\text{cat}_2}}{\ell_{\text{cat}_2}} \right\} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \frac{k_{\text{cat}_1}}{\ell_{\text{cat}_1}},$$

then after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 4-site phosphorylation system has at least 5 steady states.

Proposition 21. Let $n = 5$. Assume that $S_{\text{tot}} > E_{\text{tot}} + F_{\text{tot}}$. Then, there is a choice of rate constants for which the distributive sequential 5-site phosphorylation system has at least 5 steady states. More explicitly, if the rate constants and total concentrations are in one of the 13 regions described below

$$(R_{5.(I,J)}) \quad \max_{i \in I} \left\{ \frac{k_{\text{cat}_i}}{\ell_{\text{cat}_i}} \right\} < \frac{F_{\text{tot}}}{E_{\text{tot}}} < \min_{j \in J} \left\{ \frac{k_{\text{cat}_j}}{\ell_{\text{cat}_j}} \right\},$$

with (I, J) in the following list (where we write e.g. 14 instead of $\{1, 4\}$):

$$(0, 14), (0, 24), (1, 24), (2, 13), (2, 14), (3, 14), (3, 024), (02, 3), (02, 4), (03, 1), (03, 2), (13, 2), (13, 4),$$

then after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 5-site phosphorylation system has at least 5 steady states.

Proposition 22. *Let $n = 5$. If the rate constants and total concentrations are in one of the regions described below*

$$(R_{5.1}) \max \left\{ \frac{F_{tot}}{E_{tot}}, \frac{F_{tot}}{S_{tot} - F_{tot}} \right\} < \min \left\{ \frac{k_{cat_1}}{\ell_{cat_1}}, \frac{k_{cat_2}}{\ell_{cat_2}}, \frac{k_{cat_4}}{\ell_{cat_4}} \right\}, S_{tot} > F_{tot},$$

$$(R_{5.2}) \min \left\{ \frac{F_{tot}}{E_{tot}}, \frac{S_{tot} - E_{tot}}{E_{tot}} \right\} > \max \left\{ \frac{k_{cat_0}}{\ell_{cat_0}}, \frac{k_{cat_2}}{\ell_{cat_2}}, \frac{k_{cat_3}}{\ell_{cat_3}} \right\}, S_{tot} > E_{tot},$$

then after rescaling of the k_{on} 's and ℓ_{on} 's the distributive sequential 5-site phosphorylation system has at least 5 positive steady states.

Note that the conditions in this section describe *different* regions from the ones described by the inequalities in Theorem 2 and Theorem 4. For instance, in Propositions 17, 18, 20, 21 the inequalities between the reaction rate constants and total conservation constants do not involve the value of S_{tot} . In Propositions 19 and 22, the conditions on the total conservation constants are also different (e.g. on $\frac{F_{tot}}{E_{tot}}$ and $\frac{S_{tot}}{E_{tot}} - 1$ instead of the product $F_{tot}(\frac{S_{tot}}{E_{tot}} - 1)$). The inequalities in Theorem 2 and Theorem 4 hold for reactions rate constants of a reduced system G_J , but if we use Theorem 10 to extrapolate these conditions to the full n -site phosphorylation network, the regions are different as well.

6. Discussion

The problem of finding multistationarity regions for chemical reaction networks is in theory effectively computable but it is both hard to compute and to describe such regions in a useful fashion. It is well known that even for small systems, where multistationarity regions can be explicitly found, the expressions are very complex. So, partial approaches that give sufficient conditions on a subset of parameters are useful and relevant.

We developed in this paper both the theoretical setting based on [13, 19] and the algorithmic approach that follows from it, to describe multistationarity regions in the space of all parameters for subnetworks of the n -site sequential phosphorylation cycle, where there are up to $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states with fixed total conservation constants. We chose to assume that the subnetworks we consider have intermediate species only in the E component, but of course there is a symmetry in the network interchanging E with F , each S_i with S_{n-i} , the corresponding intermediates and rate constants, and completely similar results hold if we assume that there are only intermediates in the F component. These regions can be lifted to regions of multistationarity with at least these number of stoichiometrically compatible positive steady states of the full n -site sequential phosphorylation cycle. The main feature of our *polyhedral* approach is that it gives a systematic method applicable to other networks (for instance, to enzymatic cascades [20]), and it provides *open conditions* and not only *choices* of parameter values where *high* multistationarity occurs. Moreover, we show that it can be algorithmically implemented. However, as we have already remarked in the Introduction, the inequalities on some of the reaction rate constants are not completely explicit. Instead, they point out “directions” in which these parameters have to be scaled.

Theorem 1 in [4] states that for any positive values of S_{tot} , E_{tot} and F_{tot} , there exists $\varepsilon_0 > 0$ such that if $\frac{E_{tot}}{S_{tot}} < \varepsilon_0$, then there exists a choice of rate constants such that the system admits $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states. Also, in the recent paper [12], the authors prove the existence of parameters for which there are $2\lfloor \frac{n}{2} \rfloor + 1$ positive steady states and nearly half of them are stable and the other half

unstable, using a different degeneration of the full network. However, the maximum expected number of stoichiometrically compatible steady states for the n -site system is equal to $2n - 1$ (this is an upper bound by [4]), which has been verified for $n = 3$ and 4 in [8]. Probably, it is not possible to find a region in parameter space with this number of positive steady states using degeneration techniques.

In [7] and [6] it is proved that if $k_{\text{cat}_0}/\ell_{\text{cat}_0} < k_{\text{cat}_1}/\ell_{\text{cat}_1}$ then there exist constants $E_{\text{tot}}, F_{\text{tot}}, S_{\text{tot}}$ such that the distributive sequential 2-site phosphorylation presents multistationarity. In [24] the author gives new regions of multistationarity for $n = 2$, more precisely it is shown in Theorem 1 that given parameters $K_1, L_0, L_1, k_{\text{cat}_0}, k_{\text{cat}_1}, \ell_{\text{cat}_0}, \ell_{\text{cat}_1}$, for any K_0 small enough there exist constants $E_{\text{tot}}, F_{\text{tot}}, S_{\text{tot}}$ such that the distributive sequential 2-site phosphorylation presents multistationarity. How small K_0 has to be taken is explicitly given in terms of the other parameters by a rather involved equation (a similar result is obtained interchanging K_0 and L_1). Our Proposition 17 gives a similar result.

Proposition 17 is in agreement with [10, Corollary 4.11] which establishes that in order for the distributive sequential 2-site phosphorylation to present multistationarity the total concentration of the substrate needs to be larger than either the concentration of the kinase or the phosphatase ($S_{\text{tot}} > F_{\text{tot}}$ or $S_{\text{tot}} > E_{\text{tot}}$). In the regions of multistationarity we found for $n = 3, 4$ and 5 this is the case as well. Propositions 17, 18, 19, 20, 21 and 22 are of the same flavor as Theorems 2 and 4, in the sense that all of them give sufficient conditions on the rate constants and total conservation constants such that after rescaling of the k_{on} 's and ℓ_{on} 's, the distributive sequential n -site phosphorylation system is multistationary. Note that most of the conditions for multistationarity we found can be given a biochemical interpretation in terms of the different Michaelis-Menten maximal velocities $V_{\text{max}}^i(E) = k_{\text{cat}_i} E_{\text{tot}}, V_{\text{max}}^i(F) = \ell_{\text{cat}_i} F_{\text{tot}}$ for each independent phosphorylation/dephosphorylation step [25]. For instance, inequalities (43) in Proposition 17 can be written in classical biochemical terms as:

$$V_{\text{max}}^0(E) < V_{\text{max}}^0(F), \quad V_{\text{max}}^1(F) < V_{\text{max}}^1(E).$$

The computational approach described in Section 5 can be used for other networks of moderate size.

In conclusion, to the best of our knowledge other results (for instance [6], [7] and [24]) only give precise regions of multistationarity for the distributive sequential 2-site phosphorylation system. The present work gives new regions of multistationarity for the distributive sequential n -site phosphorylation system for any n , providing lower bounds in the number of positive steady states in such regions which are greater than three when $n > 3$. For small n we give several distinct regions of multistationarity computed via a Computer Algebra System basic implementation. Moreover, our approach is easily applicable to other networks, both analytically and computationally.

Acknowledgements

We are very thankful to Eugenia Ellis and Andrea Solotar for organizing the excellent project ‘‘Matemáticas en el Cono Sur’’, which led to this work. We also thank Eugenia Ellis for the warm hospitality in Montevideo, Uruguay, on December 3–7, 2018, where we devised our main results. We thank the referees and Jeremy Gunawardena for their useful comments that helped improve the manuscript. AD, MG and MPM were partially supported by UBACYT 20020170100048BA, CONICET PIP 11220150100473 and 11220150100483, and ANPCyT PICT 2016-0398, Argentina.

Conflict of interest

The authors declare no conflict of interest in this paper.

References

1. W. Lim, B. Meyer and T. Pawson, *Cellular signaling: principles and mechanisms*, Garland Science, 2014.
2. A. Dickenstein, Biochemical reaction networks: an invitation for algebraic geometers, MCA 2013, *Contemp. Math.*, **656** (2016), 65–83.
3. N. I. Markevich, J. B. Hoek and B. N. Kholodenko, Signaling switches and bistability arising from multisite phosphorylation in protein kinase cascades, *J. Cell Biol.*, **164** (2004), 353–359.
4. L. Wang and E. Sontag, On the number of steady states in a multiple futile cycle, *J. Math. Biol.*, **57** (2008), 29–52.
5. J. Hell and A. D. Rendall, A proof of bistability for the dual futile cycle, *Nonlinear Anal-Real.*, **24** (2015), 175–189.
6. C. Conradi, E. Feliu, M. Mincheva, et al., Identifying parameter regions for multistationarity, *PLOS Computat. Biol.*, **13** (2017), e1005751.
7. C. Conradi and M. Mincheva, Catalytic constants enable the emergence of bistability in dual phosphorylation, *J. R. Soc. Interface*, **11** (2014), 20140158.
8. D. Flockerzi, K. Holstein and C. Conradi, N-site phosphorylation systems with $2N-1$ steady states, *Bull. Math. Biol.*, **76** (2014), 1892–1916.
9. K. Holstein, D. Flockerzi and C. Conradi, Multistationarity in Sequential Distributed Multisite Phosphorylation Networks, *Bull. Math. Biol.*, **75** (2013), 2028–2058.
10. C. Conradi, A. Iosif and T. Kahle, Multistationarity in the space of total concentrations for systems that admit a monomial parametrization, *T. Bull. Math. Biol.*, 2019, 1–36.
11. M. Thomson and J. Gunawardena, Unlimited multistability in multisite phosphorylation systems, *Nature*, **460** (2009), 274.
12. E. Feliu, A. D. Rendall and C. Wiuf, A proof of unlimited multistability for phosphorylation cycles, preprint, arXiv:1904.02983.
13. F. Bihan, A. Dickenstein and M. Giaroli, Lower bounds for positive roots and regions of multistationarity in chemical reaction networks, *J. Algebra*, (2019), to appear.
14. E. Feliu and C. Wiuf, Simplifying biochemical models with intermediate species, *J. R. Soc. Interface*, **10** (2013), 20130484.
15. E. Feliu and A. Sadeghimanesh, The multistationarity structure of networks with intermediates and a binomial core network, *B. Math. Biol.*, **81** (2019), 2428–2462.
16. M. P. Millán and A. Dickenstein, The structure of MESSI biochemical networks, *SIAM J. Appl. Dyn. Syst.*, **17** (2018), 1650–1682.
17. A. G. Kušnirenko, Newton polyhedra and Bezout's theorem, *Funkcional. Anal. i Priložen*, **10** (1976), 82–83. English translation: *Funct. Anal. Appl.*, **10** (1977), 233–235.

18. A. Dickenstein, M. Giaroli, M. P. Millán, et al., Detecting the multistationarity structure in enzymatic networks, work in progress.
19. F. Bihan, F. Santos and P-J. Spaenlehauer, A polyhedral method for sparse systems with many positive solutions, *SIAM J. Appl. Algebra Geom.*, **2** (2018), 620–645.
20. M. Giaroli, F. Bihan and A. Dickenstein, Regions of multistationarity in cascades of Goldbeter-Koshland loops, *J. Math. Biol.*, **78** (2019), 1115–1145.
21. J. De Loera, J. Rambau and F. Santos, *Triangulations: Structures for Algorithms and Applications*, 2010, Series Algorithms and Computation in Mathematics, Springer, Berlin.
22. W. A. Stein, et al., *Sage Mathematics Software (Version 8.4)*, The Sage Development Team, 2018. Available from: <http://www.sagemath.org>.
23. Maple 18 (2014) Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario.
24. E. Feliu, On the reaction rate constants that enable multistationarity in the two-site phosphorylation cycle, preprint, [arXiv:1809.07275](https://arxiv.org/abs/1809.07275).
25. A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, 4th Ed., Wiley-Blackwell, 2012.



AIMS Press

©2019 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)