A Case Study on the Parametric Occurrence of Multiple Steady States

Bradford, R, Davenport, JH, England, M, Errami, H, Gerdt, V, Grigoriev, D, Hoyt, C, Kosta, M, Radulescu, O, Sturm, T & Weber, A

A Case Study on the Parametric Occurrence of Multiple Steady States

Russell Bradford
University of Bath, U.K.
R.J.Bradford@bath.ac.uk

James H. Davenport
University of Bath, U.K.
J.H.Davenport@bath.ac.uk

Matthew England
Coventry University, U.K.
Matthew.England@coventry.ac.uk

Hassan Errami
University of Bonn, Germany
errami@cs.uni-bonn.de

Vladimir Gerdt
JINR, Dubna, Russia
gerdt@jinr.ru

Dima Grigoriev
CNRS & University of Lille, France
dmitry.grigoryev@math.univ-lille1.fr

Charles Hoyt
b-it, Bonn, Germany
cthoyt@gmail.com

Marek Košta
Slovak Academy of Sciences, Slovakia
marek.kosta@savba.sk

Ovidiu Radulescu
DIMNP UMR CNRS/UM 5235,
University of Montpellier, France
ovidiu.radulescu@umontpellier.fr

Thomas Sturm
UL, CNRS, Inria & LORIA, France
MPII & Saarland University, Germany
thomas.sturm@loria.fr

Andreas Weber
University of Bonn, Germany
weber@cs.uni-bonn.de

ABSTRACT

We consider the problem of determining multiple steady states for positive real values in models of biological networks. Investigating the potential for these in models of the mitogen-activated protein kinases (MAPK) network has consumed considerable effort using special insights into the structure of corresponding models. Here we apply combinations of symbolic computation methods for mixed equality/inequality systems, specifically virtual substitution, lazy real triangularization and cylindrical algebraic decomposition. We determine multistationarity of an 11-dimensional MAPK network when numeric values are known for all but potentially one parameter. More precisely, our considered model has 11 equations in 11 variables and 19 parameters, 3 of which are of interest for symbolic treatment, and furthermore positivity conditions on all variables and parameters.

CCS CONCEPTS

• **Applied computing → Biological networks**; • **Computing methodologies → Equation and inequality solving algorithms**;

1 INTRODUCTION

The occurrence of multiple steady states (multistationarity) has important consequences on the capacity of signaling pathways to process biological signals, even in its elementary form of two stable steady states (bistability). Bistable switches can act as memory circuits storing the information needed for later stages of processing [28]. The response of bistable signaling pathways shows hysteresis, namely dynamic and static lags between input and output. Because of hysteresis one can have, at the same time, a sharp binary response and protection against chatter noise.

Bistability of signaling usually occurs as a result of activation of upstream signaling proteins by downstream components [3]. A different mechanism for producing bistability in signaling pathways was proposed by Markevich et al. [20], where bistability can be caused by multiple phosphorylation/dephosphorylation cycles that share enzymes. A simple, two-step phosphorylation/dephosphorylation cycle is capable of ultrasensitivity, a form of all or nothing response with no hysteresis (Goldbeter–Koshland mechanism). In multiple phosphorylation/dephosphorylation cycles, enzyme sharing provides competitive interactions and positive feedback that ultimately leads to bistability.

Algorithmically the task is to find the positive real solutions of a parameterized system of polynomial or rational systems, since the dynamics of the network is given by polynomial systems (arising from mass action kinetics) or rational functions (arising in signaling

networks when some intermediates of the reaction mechanisms are reduced). Due to the high computational complexity of this task [14] considerable work has been done to use specific properties of networks and to investigate the potential of multistationarity of a biological network out of the network structure. This only determines whether or not there exist rate constants allowing multiple steady states, instead of coming up with a semi-algebraic description of the range of parameters yielding this property. These approaches can be traced back to the origins of Feinberg's *Chemical Reaction Network Theory* (CRNT) whose main result is that networks of deficiency 0 have a unique positive steady state for all rate constants [9, 13]. We refer to [7, 17, 22] for the use of CRNT and other graph theoretic methods to determine potential existence of multiple positive steady states, with [18] giving a survey.

Given a bistable mechanism it is also important to compute the bistability domains in parameter space: the parameter values for which there is more than one stable steady state. The size of bistability domains gives the spread of the hysteresis and quantifies the robustness of the switches. The work of Wang and Xia [26] is relevant here: they used symbolic tools, including cylindrical algebraic decomposition as we do, to determine the number of steady states and their stability for several systems. They reported results up to a 5-dimensional system using specified parameter values, but their method is extensible to parametric questions. Higher-dimensional systems were studied using sign conditions on the coefficients of the characteristic polynomial of the Jacobian. In some cases these guarantee uniqueness of the steady state [8].

In this paper we use an 11-dimensional model of a mitogen-activated protein kinases (MAPK) cascade [20] as a case study to investigate properties of the system using algorithmic methods towards the goal of semi-algebraic descriptions of parameter regions for which multiple positive steady states exist.

2 THE MAPK NETWORK

The model of the MAPK cascade we are investigating can be found in the Biomodels database [19].[1] We have renamed the species names to $x_1, \ldots, x_{11}$ and the rate constants to $k_1, \ldots, k_{16}$ to facilitate reading:

$$\dot{x}_1 = k_2 x_6 + k_{15} x_{11} - k_1 x_1 x_4 - k_{16} x_1 x_5$$
$$\dot{x}_2 = k_3 x_6 + k_5 x_7 + k_{10} x_9 + k_{13} x_{10} - x_2 x_5 (k_{11} + k_{12}) - k_4 x_2 x_4$$
$$\dot{x}_3 = k_6 x_7 + k_8 x_8 - k_7 x_3 x_5$$
$$\dot{x}_4 = x_6 (k_2 + k_3) + x_7 (k_5 + k_6) - k_1 x_1 x_4 - k_4 x_2 x_4$$
$$\dot{x}_5 = k_8 x_8 + k_{10} x_9 + k_{13} x_{10} + k_{15} x_{11} -$$
$$\qquad x_2 x_5 (k_{11} + k_{12}) - k_7 x_3 x_5 - k_{16} x_1 x_5$$
$$\dot{x}_6 = k_1 x_1 x_4 - x_6 (k_2 + k_3)$$
$$\dot{x}_7 = k_4 x_2 x_4 - x_7 (k_5 + k_6)$$
$$\dot{x}_8 = k_7 x_3 x_5 - x_8 (k_8 + k_9)$$
$$\dot{x}_9 = k_9 x_8 - k_{10} x_9 + k_{11} x_2 x_5$$
$$\dot{x}_{10} = k_{12} x_2 x_5 - x_{10} (k_{13} + k_{14})$$
$$\dot{x}_{11} = k_{14} x_{10} - k_{15} x_{11} + k_{16} x_1 x_5. \tag{1}$$

The Biomodels database also gives us meaningful values for the rate constants:

$$\begin{aligned}
&k_1 = 0.02, &&k_2 = 1, &&k_3 = 0.01, &&k_4 = 0.032, \\
&k_5 = 1, &&k_6 = 15, &&k_7 = 0.045, &&k_8 = 1, \\
&k_9 = 0.092, &&k_{10} = 1, &&k_{11} = 0.01, &&k_{12} = 0.01, \\
&k_{13} = 1, &&k_{14} = 0.5, &&k_{15} = 0.086, &&k_{16} = 0.0011.
\end{aligned} \tag{2}$$

Some of these values are measured and some are well-educated guesses. For the purpose of our study we assume they are suitable.

We add three linear conservation constraints introducing three further constant parameters k_{17}, k_{18}, k_{19}:

$$x_5 + x_8 + x_9 + x_{10} + x_{11} = k_{17}$$
$$x_4 + x_6 + x_7 = k_{18}$$
$$x_1 + x_2 + x_3 + x_6 + x_7 + x_8 + x_9 + x_{10} + x_{11} = k_{19}. \tag{3}$$

Computations to produce these in MathWorks SimBiology use the left-null space of the stoichiometric matrix under positivity conditions, see for example [23].

Meaningful values for k_{17}, k_{18}, k_{19} are harder to obtain than the constants in (2). The following are some realistic values estimated by ourselves on the basis of our understanding of the biological model:

$$k_{17} = 100, \qquad k_{18} = 50, \qquad k_{19} \in \{200, 500\}. \tag{4}$$

The long-term goal of our research is to treat all three of these together parametrically, although in the present work we focus on situations with one free-parameter.

The steady state problem for the MAPK cascade can now be formulated as a real algebraic problem. That is, we replace the left hand sides of all equations in (1) with 0. This together with the equations in (3) yields an algebraic system with polynomials in

$$F \subset \mathbb{Z}[k_1, \ldots, k_{19}][x_1, \ldots, x_{11}].$$

All entities in our model are strictly positive, which yields in addition a system

$$P = \{k_1, \ldots, k_{19}, x_1, \ldots, x_{11}\} \subset \mathbb{Z}[k_1, \ldots, k_{19}][x_1, \ldots, x_{11}]$$

establishing a side condition on the solutions of F that all variables x_i and parameters k_i of P be positive. In terms of first-order logic our specification of F and P yields a quantifier-free Tarski formula

$$\varphi = \bigwedge_{f \in F} f = 0 \wedge \bigwedge_{v \in P} v > 0. \tag{5}$$

The estimations for the rate constants in (2) formally establish a substitution rule $\sigma = [0.02/k_1, \ldots, 0.0011/k_{16}]$, which can be applied to F, P, or φ in postfix notation.

2.1 Symbolic Determination of Occurrences of Multiple Steady States

In this section we are going to analyze the system for multiple positive steady states. As we will not include a priori information about the stability of the fixed points, we do not only have to consider (at least) two stable fixed points but also unstable fixed points, i.e., we investigate the existence of at least three different roots $\mathbf{x} \in]0, \infty[^{11}$ of F for given choices $\mathbf{k} \in]0, \infty[^{19}$ of parameters.

We present two investigations: one using the Redlog package in Reduce and the other using the Regular Chains Library in Maple.

Both will make use of Cylindrical Algebraic Decomposition (CAD) [1] to solve the problem. The worst-case time complexity of CAD is doubly exponential.[2] Our approaches admit, in principle, arbitrary numbers of indeterminates. However, for the sake of realistic computation times we must restrict ourselves to one free parameter. Even then, the number of variables present is too large for contemporary CAD implementations. We make progress by combining CAD with additional symbolic methods. Our first approach uses virtual substitution techniques and the second real triangularization. In both cases we have combined the corresponding methods by hand, but automation is clearly possible.

2.1.1 Real Quantifier Elimination in Redlog.
Real Quantifier Elimination (QE) can directly handle the parametric existence of steady states, taking as input $\exists x_1 \ldots \exists x_{11} \varphi$, possibly with substitutions for some parameters. However, we are not only interested in the existence but also in the number of solutions. We are going to combine Virtual Substitution (VS) [27] with CAD. The former smoothly eliminates the majority of the quantifiers while the latter allows us to count numbers of solutions via decomposition of the remaining low-dimensional spaces. That combination of methods requires the solution of several QE runs with each problem and some combinatorial arguments. Throughout this subsection we are using Redlog [10].

Parameter-Free Computations. We consider

$$\varphi_{500} = \varphi\sigma[100/k_{17}, 50/k_{18}, 500/k_{19}]$$

where all parameters have been substituted with rational numbers. The closed formula $\bar{\varphi}_{500} = \exists x_1 \ldots \exists x_{11} \varphi_{500}$ states the existence of a suitable real solution. In a first step, we solve for $i \in \{1, \ldots, 11\}$ the following eleven QE problems using VS:

$$\varphi_{500}^{(i)} = \text{VS}(\exists x_1 \ldots \exists x_{i-1} \exists x_{i+1} \ldots \exists x_{11} \varphi_{500}).$$

Each $\varphi_{500}^{(i)}$ is a univariate quantifier-free formula describing all possible real choices for x_i for which there exist real choices for all other variables such that φ_{500} holds. CAD can easily decompose the corresponding one-dimensional spaces. It happens that for each x_i there are exactly three zero-dimensional cells $a_i, b_i, c_i \in \mathbb{R}$ where $\varphi_{500}^{(i)}$ holds. We extract all a_i, b_i, and c_i as *real algebraic numbers*, i.e., univariate defining polynomials with integer coefficients plus isolating intervals. By combinatorial arguments it is not hard to see that the following holds for the set S_{500} of real solutions of φ_{500}:

$$3 \le |S_{500}| \quad \text{and} \quad S_{500} \subseteq \prod_{i=1}^{11} \{a_i, b_i, c_i\}.$$

Notice that at this point we have proven multistationarity for $k_{19} = 500$. We can furthermore compute S_{500} by plugging the 3^{11} candidates from the Cartesian product into φ_{500}. A straightforward approach requires arithmetic with real algebraic numbers followed by the determination of the signs of the results, which is quite inefficient in practice. We use instead a heuristic approach combining refinements of the isolating intervals of the real algebraic numbers with interval arithmetic. This excludes $3^{11} - 3$ of

the candidate solutions. The three remaining candidates require no further checking since we already know that $|S_{500}| \ge 3$. The overall CPU time is 71.3 seconds for 11 runs of VS plus 11 runs of CAD, followed by 16 hours for checking candidates.[3] Our checking procedure is a file-based prototype starting a Reduce process for every single of the 3^{11} candidates; there is considerable room for optimization.

For $k_{19} = 200$ instead of 500 all eleven univariate CAD computations yield unique solutions which can be straightforwardly combined to one unique solution for the corresponding φ_{200}. The overall CPU time here is 66.4 seconds for 11 runs of VS plus 11 runs of CAD. Machine float approximations of all our solutions are given in Table 1.

Parametric Analysis for k_{19}. We consider

$$\varphi_{k_{19}} = \varphi\sigma[100/k_{17}, 50/k_{18}]$$

leaving k_{19} as a parameter. Again, we solve for $i \in \{1, \ldots, 11\}$ eleven QE problems using VS:

$$\varphi_{k_{19}}^{(i)} = \text{VS}(\exists x_1 \ldots \exists x_{i-1} \exists x_{i+1} \ldots \exists x_{11} \varphi_{k_{19}}).$$

This time each $\varphi_{k_{19}}^{(i)}$ is a bivariate quantifier-free formula in k_{19} and the corresponding x_i. This time we construct a two-dimensional CAD for each $\varphi_{k_{19}}^{(i)}$. The projection order is important: we first project x_i, then the CAD base phase decomposes the k_{19}-axis, followed by an extension phase that decomposes the x_i-space over the k_{19}-cells obtained in the base phase. This is feasible with one limitation: we do not extend over zero-dimensional k_{19}-cells. In other words, we accept finitely many blind spots in parameter space, which we can explicitly read off from the CAD so that in the end we know exactly what we are missing.

Figure 1 shows our CAD tree for $\varphi_{k_{19}}^{(2)}$. The first layer next to the root shows the decomposition of the k_{19}-axis. The five zero-dimensional (rectangular) cells are the previously mentioned blind spots, among which the smallest one with negative value of k_{19} is not relevant. Those zero-dimensional cells also establish the limits of the full dimensional (oval) cells in between. The cylinders over those one-dimensional k_{19}-cells each contain either one or three zero-dimensional x_2-cells where $\varphi_{k_{19}}^{(2)}$ holds. We have deleted from the tree all x_2-cells where $\varphi_{k_{19}}^{(2)}$ does not hold. We make two observations, important for a qualitative analysis of our system:

(i) For all positive choices of k_{19}—extending to infinity—there is at least one positive solution for x_2.

(ii) There is a break point around $k_{19} = 409.253$ where the system changes from unique solutions to exactly three solutions.

Recall that for all floating point numbers given here as approximations we in fact know exact real algebraic numbers. For instance, the exact break point is the only real zero in the interval $(409, 410)$ of an irreducible defining polynomial

$$\sum_{i=0}^{10} c_i k_{19}^i \text{ with integer coefficients } c_i \text{ as in Table 2.} \tag{6}$$

---

[2]Traditionally, doubly exponential in the number of variables. However recent progress on CAD in the presence of equational constraints [11], such as (1) with 0 for left-hand side, allows us to conclude it is actually doubly-exponential the number of variables minus the number of equational constraints at different levels of the projection [12].

[3]All QE-related computations have been carried out on a 2.4 GHz Intel Core i7 with 3 GB RAM or cores on a compute server with similar speed and memory limitations.

Table 1: The unique solution $x^{(200)}$ for $k_{19} = 200$ and the three solutions $x_1^{(500)}, \ldots, x_3^{(500)}$ for $k_{19} = 500$. We have actually computed real algebraic numbers, which are pairs of univariate polynomials and isolated intervals. For convenience we are giving machine float approximations here, which can be made arbitrarily precise.

$x^{(200)} = (90.6512, 2.67311, 10.4996, 17.8545, 35.9695, 32.0501, 0.0954536, 15.5631, 2.39331, 0.641001, 45.4331)$

$x_1^{(500)} = (17.6392, 6.97675, 367.57, 36.6772, 5.50874, 12.811, 0.511775, 83.4416, 8.06095, 0.25622, 2.73253)$

$x_2^{(500)} = (122.034, 14.6721, 234.974, 14.5102, 7.16952, 35.064, 0.42579, 69.4223, 7.43877, 0.70128, 15.2681)$

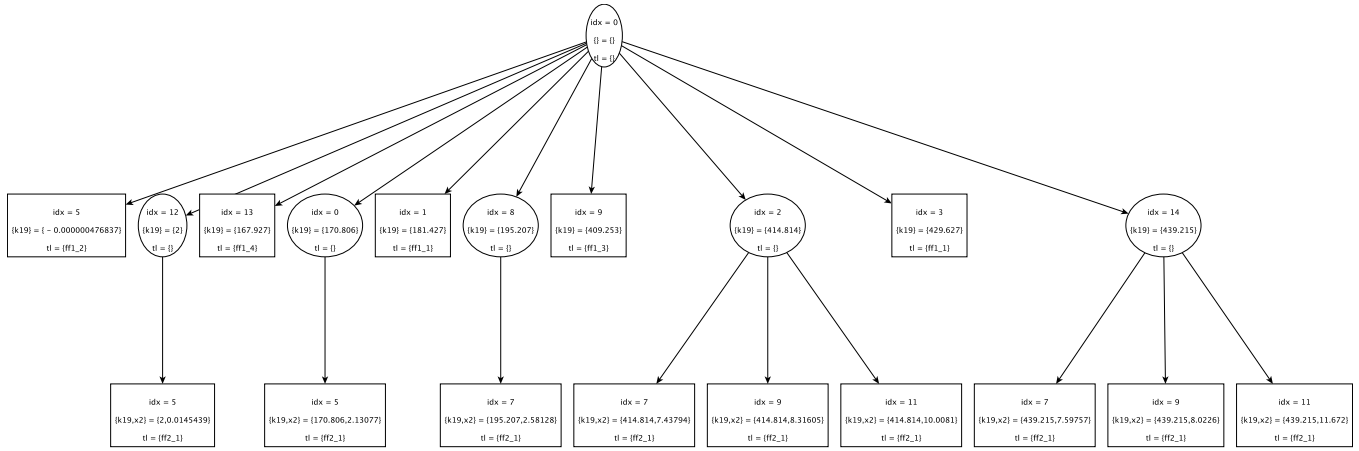$x_3^{(500)} = (323.761, 9.49621, 37.1013, 6.72938, 13.6295, 43.1428, 0.127807, 20.8381, 3.21139, 0.862856, 61.4581)$

Figure 1: The pruned CAD tree for x_2. Ellipses and rectangles are full-dimensional and zero-dimensional cells, respectively. We have removed cells where k_{19} is negative or where the input formula is false.

All other CAD trees are quite similar to the one just discussed. Even the break point from one to three solutions for x_i is identical for all $i \in \{1, \ldots, 11\}$ so that we can generalize our observations:

(i) For all positive choices of k_{19}—extending to infinity—there is at least one positive solution for $(x_1, \ldots, x_{11})$.

(ii) There is a break point β around $k_{19} = 409.253$ where the system changes its qualitative behavior. We have exactly given β as a real algebraic number in Equation (6). For $k_{19} < \beta$ there is exactly one positive solution for $(x_1, \ldots, x_{11})$. For $k_{19} > \beta$ there are at least 3 and at most 3^{11} positive solutions for $(x_1, \ldots, x_{11})$.

The overall computation time for our parametric analysis is 4.3 minutes. It is strongly dominated by 2.8 minutes for the computation of one particular CAD tree, for $\varphi_{k_{19}}^{(11)}$. It turns out that the suitable projection order with x_i eliminated first is computationally considerably harder than projecting the other way round. As a preprocessing step we apply CAD-based simplification of the $\varphi_{k_{19}}^{(i)}$ with the opposite, faster, projection order. Here we use Qepcad B, which performs better than Redlog at simple solution formula construction.

2.1.2 Triangular Decomposition methods with the Regular Chains Library. We now describe an alternative approach to the solution using regular chains methods. Regular chains are the triangular decompositions of systems of polynomial equations (triangular in terms of the variables in each polynomial). Highly efficient methods for working in complex space have been developed based on these; see [25] for a survey.

Recent work by Chen et al. [4] proposes adaptations of these tools to the real analogue: semi-algebraic systems. They describe two algorithms to decompose any real polynomial system into finitely many regular semi-algebraic systems. The first does so directly while the second, Lazy Real Triangularize (LRT) produces the highest dimension solution component and unevaluated function calls, which if all evaluated would combine to give the full solution. These algorithms are implemented in the Regular Chains Library[4] in Maple which we use throughout this subsection.

We apply LRT on the quantifier-free formula (5) evaluated with the parameter estimates for $k_1, \ldots, k_{18}$ given at the start of Section 2, so we have one free parameter as in the previous section. We need to choose a variable ordering: our analysis requires that k_{19} be the indeterminate considered alone; the remaining variables are placed in lexicographical order (the in-built heuristics to make the choice could suggest nothing better). The solutions must hence contain constraints in k_{19}, constraints in (x_1, k_{19}), in (x_2, x_1, k_{19}) and so on. We define the *main variable* of a constraint to be the highest one present in this ordering.

---

[4]www.regularchains.org

$c_{10} = 3515909345027402909368950332670171587360603139406930766501553712504$ 11

$c_9 = -2136990728521576742839975277746395583273033983170426080574800781989093156$

$c_8 = 2537485164122055477425960563505346943258210988396501580407711911095 8034090$

$c_7 = 1297249301830002270702763926780425925123599161802985288033000450856 4391594000$

$c_6 = -8468945963692802414226427249726123493448372439778349029355636316929687020660000$

$c_5 = 22310982703374064506703016631726643334214408338758486214236832656638 46533079600000$

$c_4 = -37626500890411225829031917319379205201489948552899492596588589551183187344 4245100000$

$c_3 = 39262101548790869407057994985320156500968958361396178908180026842806643766783104000000$

$c_2 = -24926239907430292349743540812702961063096034624515170577798775968424448287799337600000000$

$c_1 = 709788507358874734591769971861759784258732672467600232129406169246431718684780800000 00000$

$c_0 = -1062871192838985876948077114923898204990434138901495394834749613184670362810368000000000000$

$d_6 = 16838105723097694257603469$

$d_5 = -24078605201553273505077988 k_{19} + 772396796964497789614868 6580$

$d_4 = 81762026387357691270321 69 k_{19}^2 - 772341166546354447770149 9460 k_{19} + 1232154357941338876156606812900$

$d_3 = 1465408757440589841803452380 k_{19}^2 - 798169557586805582842481309800 k_{19} + 8315265524000276772955047 7640000$

$d_2 = 854625249012768461072516694 00 k_{19}^2 - 352664114014276568345720951400 00 k_{19} + 25568053548533183321974896 36000000$

$d_1 = 1631685649719702672282505500000 k_{19}^2 - 721989571100461862477342320000000 k_{19} + 28843755938318780823218400000000000$

$d_0 = -70131041394599108765205000000000000 k_{19}.$

LRT produces one solution component and 6 unevaluated function calls in less than 3 seconds. In the evaluated component: for each of $x_2, \ldots, x_{11}$ there is a single equation which had this as the main variable. Further, these are all linear in their main variable meaning they can be easily rearranged into the solution formulae in Table 3.

The constraints on (x_1, k_{19}) are that $x_1 > 0$ and that a polynomial equation of degree 6 be satisfied:

$$f(x_1, k_{19}) = \sum_{i=0}^{6} d_i x_1^i = 0 \qquad (7)$$

where the coefficients d_i are univariate polynomials in k_{19} of maximum degree 2 as given in Table 2.

Finally, the constraints on k_{19} are that it be positive; it not be a root of the polynomial in Equation (6); nor two other polynomials as described in Table 4. Thus this solution component is valid for all positive values of k_{19} excluding three points. As before, we could give these as exact algebraic numbers but for brevity give float approximations: 409.253, 16473.337, and 25084.536.

Three of the six unevaluated function calls define the solutions at these points, however evaluating these solutions is not possible in reasonable time. The other three define empty solution sets (evaluating to discover this is instantaneous). So, as with our previous approach, we proceed accepting a small number of blind spots.

The output of LRT has quickly given us the structure of the solution space valid at all but three isolated values of k_{19}. However, it does not identify where the number of real solutions change: although the break point identified earlier has been rediscovered

there is no information from which we can infer its significance; and there is no significance in our application of the other two isolated points.

To finish the analysis we need to decompose (x_1, k_{19})-space according to the real roots of $f(x_1, k_{19})$; and also x_1 and k_{19} since the constraints $x_1 > 0$ and $k_{19} > 0$ were specified separately in the output (the case for these variables only). CAD is ideally suited for this task. Using the Regular Chains algorithm [5] in Maple a CAD for $f(x_1, k_{19})$ divides the plane into 135 cells in a few seconds. This CAD decomposes the k_{19} axis into 11 cells, i.e. identifying five points which approximate to: -379.993, -87.776, 0, 409.253, and 25084.536.

On the cell for $k_{19} \in \,]0, 409.253[$, the cylinder above in the (x_1, k_{19}) plane is divided into 11 cells: three of which cover $x_1 > 0$ (two 2d sectors and a 1d section). This indicates that $f(x_1, k_{19})$ has a single positive real solution for such k_{19}. However, on the two cells for $k_{19} \in \,]409.253, 25084.536[$ and $k_{19} \in \,]25084.536, \infty[$ the cylinders above are divided into 15 cells; seven of which cover $x_1 > 0$. So in these cases $f(x_1, k_{19})$ has three positive real solutions for such k_{19}.

At the end of this analysis we have rediscovered the break point where the system moves from a single positive real solution to three. We also have explicit solutions valid for all except three isolated k_{19} values. To obtain a solution select the k_{19} value of interest then identify the real roots of $f(x_1, k_{19})$ (we know in advance how many depending on the k_{19} value chosen); then for each x_1 solution substitute recursively into the equations of Table 3; starting from the bottom and including the new variable solution discovered from

$$x_{11} = -\frac{1}{60}x_2^2 + \frac{1}{600}(10k_{19} - 10x_1 - 37x_3 + 10x_4 - 2100)x_2 - \frac{9}{200}x_3^2 + \frac{1}{600}(-27x_1 + 27x_4 + 27k_{19} - 4650)x_3 - x_1 + x_4 + k_{19} - 50$$

$$x_{10} = \frac{1}{150}x_2(x_2 + x_3 - x_4 - k_{19} + x_1 + 150)$$

$$x_9 = \frac{1}{18200}(69x_3 + 182x_2)(x_2 + x_3 - x_4 - k_{19} + x_1 + 150)$$

$$x_8 = \frac{15}{364}(x_2 + x_3 - x_4 - k_{19} + x_1 + 150)x_3$$

$$x_7 = 50 - \frac{2}{101}x_4x_1 - x_4$$

$$x_6 = \frac{2}{101}x_4x_1$$

$$x_5 = x_2 + x_3 - x_4 - k_{19} + x_1 + 150$$

$$x_4 = \frac{2525000}{101x_2 + 1000x_1 + 50500}$$

$$x_3 = \frac{-101x_2^3 - (-101k_{19} + 1101x_1 + 65650)x_2^2 - (1000x_1^2 + (-1000k_{19} + 200500)x_1 - 50500k_{19} + 5050000)x_2 + 150000x_1)}{101x_2^2 + (1000x_1 + 50500)x_2}$$

$x_2 = \dfrac{n}{d}$ where

$\quad n = 30625833064790009548991419920x_1^5 + (-43795148662369306906962603840k_{19}$
$\qquad + 377499792254877318052736865046632000)x_1^4 + (148712110647782462053693235920k_{19}^2$
$\qquad - 1696333629369275091915491069067 2400k_{19} + 6815925407229297763234036009365120000)x_1^3$
$\qquad + (153832544822298322993053004920 0k_{19}^2 - 862702164104208291031357996000020000k_{19}$
$\qquad + 2792412190287203685788093362497480000 00)x_1^2 + (29370341694954648101085099000000k_{19}^2$
$\qquad - 1299581227980831352459216176000000 0k_{19} + 37059602821175232428867692137000000000 00)x_1$
$\qquad - 126235874510278395777736900000000 0000000k_{19}$

$\quad d = 232763663752113237974029404420089x_1^5 + (-33285361530104184557767163999022 8k_{19}$
$\qquad + 8864630321520507537630814702967 7220)x_1^4 + (11302476139945018694939062307 4789k_{19}^2$
$\qquad - 80843908028331498139954527761762 740k_{19} + 1168246506839176979663298692907277 6500)x_1^3$
$\qquad + (11455232309649034305597048791479 020k_{19}^2 - 5547251026060433566640620528023877000k_{19}$
$\qquad + 61914720758759700126802625440464 7600000)x_1^2 + (29024599706300155013019802645852 5000k_{19}^2$
$\qquad - 141348286758352762323489548674398 500000k_{19}$
$\qquad + 1454728852958138225258707154149460 0000000)x_1$
$\qquad - 1247498501818579946626756931775000 000000(k_{19} - 100)$

$x_1 \quad$ has at most 6 solutions for a given value of k_{19}, according to Equation (7).

$$k_{19} > 0 \ \wedge \ \text{polynomial in (6)} \neq 0 \ \wedge$$

$$231979894334195799994929k_{19}^2 - 8940740061545240945309 8800k_{19} - 482241930341916652549114 9190000 \neq 0 \ \wedge$$

$$5054655666224758676555478807865446379537904060599827 26185509k_{19}^4$$

$$- 12725780456964391893178560515183873684222178969868366 920505134120k_{19}^3$$

$$+ 1175510330915205241831243213231417517003073155628841 93657451445400k_{19}^2$$

$$- 281867359883676159811192082978541193600292804324596911878337972560000k_{19}$$

$$- 424343635702155874656684237015639321850510668927412079318793072000 00000 \neq 0$$

each substitution into the next. The solutions in Table 1 may be easily rediscovered this way.

Repeating the Process for Different Choices of the Lone Free Parameter/Fixed Parameter Values. We may repeat the approach described above for different choices of free parameter and different choices of fixed parameter values. For example:

- With k_{17} set to 95 instead of 100 we find that the break point between 1 and 3 real positive solutions moves to $k_{19} = 369.917$. With k_{17} set to 105 it moves to $k_{19} = 450.077$.
- Allowing k_{17} to be free and fixing $k_{19} = 200$ we find that there is only ever one positive real solution.
- Allowing k_{17} to be free and fixing $k_{19} = 500$ we find the number of positive real solutions moving from 1 to 3 to 1 breaking at $k_{17} = 85.988$ and $k_{17} = 110.869$.
- Similarly, allowing k_{18} to be free and fixing $k_{19} = 200$ we find there is only ever one positive real solution; but fixing $k_{19} = 500$ instead we find 3 real solutions between $k_{18} = 44.434$ and 58.329 and 1 otherwise.

The results above hint that there is a shape approximating a narrow paraboloid in (k_{17}, k_{18}, k_{19})-space within which bistability may occur; with bistability available for any k_{17} and k_{18} value but bounded from below in the k_{19} coordinate. We note that these additional experiments all produce results which, as with the one described in detail, are valid at all but a handful of isolated values of the free parameter.

2.2 Stability of the Fixed Points

We use the three linear conservation constraint equations (3) to eliminate x_1, x_7, and x_{11} from system (1) and symbolically compute the Jacobian $\tilde{J}$ of the obtained reduced system. We then numerically compute the eigenvalues of $\tilde{J}$ for the instances arising from the substitution of the different positive fixed points for the variables and the corresponding parameter values.

We have used the float approximations for the unique solution $x^{(200)}$ with $k_{19} = 200$ and the three solutions $x_1^{(500)}, \ldots, x_3^{(500)}$ for $k_{19} = 500$ in Table 1. For the single positive fixed point $x^{(200)}$ the Jacobian $\tilde{J}(x^{(200)})$ has eigenvalues with negative real part only and hence can be shown to be stable. For $k_{19} = 500$ one of the three positive fixed points $x_2^{(500)}$ can be shown to be unstable, as $\tilde{J}(x_2^{(500)})$ has one eigenvalue with positive real part; the other seven had negative real parts. In contrast $x_1^{(500)}$ and $x_3^{(500)}$ can be shown to be stable. Hence for $k_{19} = 500$ the system is indeed bistable.

A verification of the stability of the fixed points using exact real algebraic numbers by the well-known Routh–Hurwitz criterion is possible algorithmically [16], but seems to be out of range of current methods for this example. Notice that also in other studies on multistationarity of signalling pathways [7, 15] the question of stability has not been addressed.

2.3 Numerical Homotopy Methods

Finally, we compare our symbolic results with numerical ones obtained using the homotopy solver Bertini [2]. Bertini computes complex roots of polynomial systems using methods from numerical algebraic geometry [24].

For the parameter values as above and $k_{19} = 500$ we obtain six solutions, three of which are positive real solutions. For $k_{19} = 200$ we obtain a single positive solution. In both cases the relevant solutions coincide with the ones obtained with our symbolic analyses up to the used numeric precision.

However, for larger values of k_{19} Bertini produces incorrect results due to numerical instability. For instance, we falsely obtain exactly one positive real solution for $k_{19} = 6000$ and no positive real solution for $k_{19} = 10000$.

Figure 2 shows a Bertini-based grid sampling of parameter regions, varying k_{19} between 200 and 1000 and fixing one of k_{17} and k_{18} while varying the other among the default values (4). While this suffers from the discrete nature of sampling and potentially unreliable results as discussed, it is nevertheless useful for the generation of hypothesis about the nature of the parameter regions. Figure 2 seems to identify a region of bistability (in blue) within the parameter space, as hypothesised at the end of Section 2.1.2.

Bertini indicates holes in this region (the green dots within the blue). However, computation at these particular points reveals numerical errors: where an insufficiently high precision causes what is actually a positive real solution to appear to have a negative component. It seems there is scope for fruitful interplay between symbolic and numeric methods here; with numerics postulating hypotheses for the symbolic methods to check and refine.

3 CONCLUSIONS AND FUTURE WORK

We have shown that the determination of multistationarity of an 11-dimensional MAPK network can be achieved by combinations of currently available symbolic computation methods for mixed equality/inequality systems if, for all but potentially one parameter, numeric values are known. The aspiration of a semi-algebraic description of the ranges for all parameters in the conservation laws (3) yielding multistationarity will now be pursued, with the present results demonstrating that this aspiration may be within reach.

As there are many very relevant systems having dimensions between 10 and 20 it seems to be worth the effort to enhance and improve the present algorithmic methods, and in particular their combination, to solve such important application problems for symbolic computation.

Note Added in Press: Our attention has been drawn to the recent preprints [6, 21], which address some similar problems.
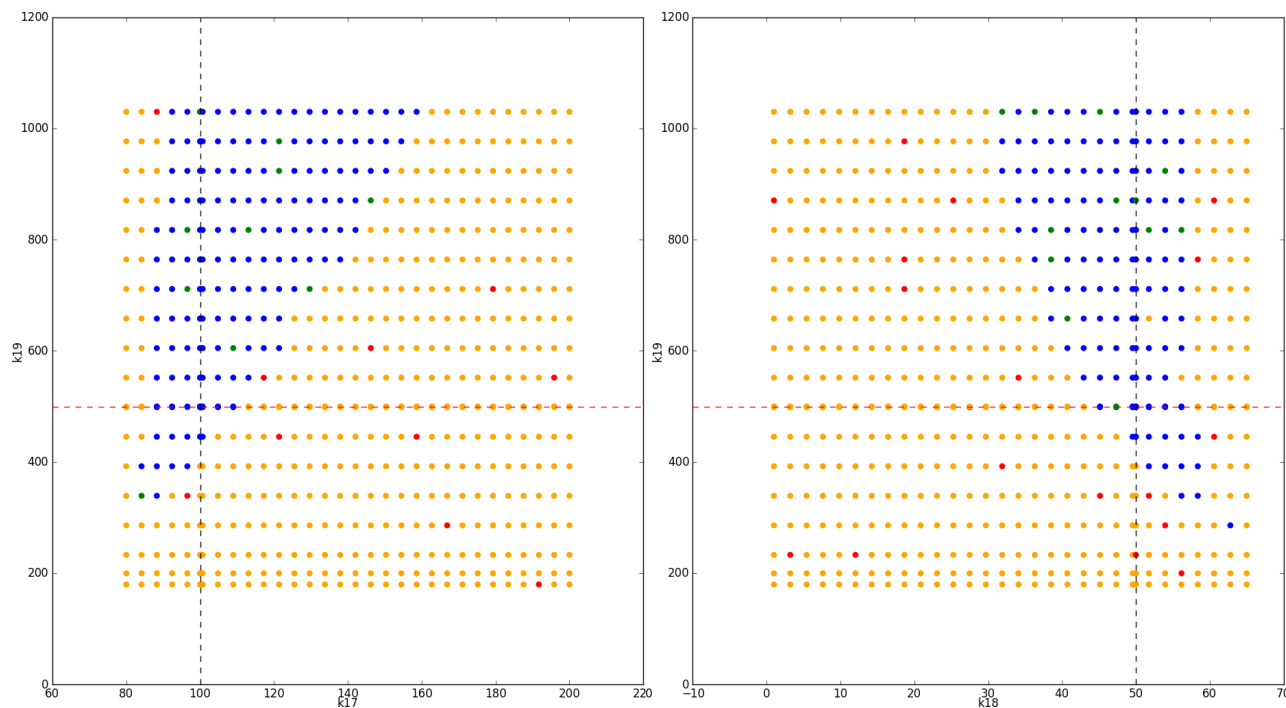
Figure 2: Grid sampling two-parameter regions using Bertini. We combine k_{19} with k_{17} (left) and with k_{18} (right). Colors indicate the computed numbers of positive real fixed points: blue 3, green 2, yellow 1, red 0. The dotted lines indicate values of the parameters as in Equation (4).

Research Data Statement: Data supporting the research in this paper is available from http://doi.org/10.5281/zenodo.583280

REFERENCES

[1] D. S. Arnon, G. E. Collins, and S. McCallum. Cylindrical algebraic decomposition I: The basic algorithm. *SIAM J. Comput.*, 13(4):865–877, 1984.
[2] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. Bertini: Software for numerical algebraic geometry. doi:10.7274/R0H41PB5.
[3] U. S. Bhalla and R. Iyengar. Emergent properties of networks of biological signaling pathways. *Science*, 283(5400):381–387, 1999.
[4] C. Chen, J. Davenport, J. May, M. Moreno Maza, B. Xia, and R. Xiao. Triangular decomposition of semi-algebraic systems. *J. Symb. Comput.*, 49:3–26, 2013.
[5] C. Chen, M. Moreno Maza, B. Xia, and L. Yang. Computing cylindrical algebraic decomposition via triangular decomposition. In *Proceedings of the ISSAC 2009*, pages 95–102. ACM, 2009.
[6] C. Conradi, E. Feliu, M. Mincheva, and C. Wiuf. Identifying parameter regions for multistationarity. *Preprint: arXiv:1608.03993*, 2017.
[7] C. Conradi, D. Flockerzi, and J. Raisch. Multistationarity in the activation of a MAPK: parametrizing the relevant region in parameter space. *Math. Biosci.*, 211(1):105–31, 2008.
[8] C. Conradi and M. Mincheva. Catalytic constants enable the emergence of bistability in dual phosphorylation. *Journal of The Royal Society Interface*, 11(95), 2014.
[9] G. Craciun, A. Dickenstein, A. Shiu, and B. Sturmfels. Toric dynamical systems. *J. Symb. Comput.*, 44(11):1551–1565, 2009.
[10] A. Dolzmann and T. Sturm. Redlog: Computer algebra meets computer logic. *ACM SIGSAM Bulletin*, 31(2):2–9, 1997.
[11] M. England, R. Bradford, and J. Davenport. Improving the use of equational constraints in cylindrical algebraic decomposition. In *Proceedings of the ISSAC 2015*, pages 165–172. ACM, 2015.
[12] M. England and J. Davenport. The complexity of cylindrical algebraic decomposition with respect to polynomial degre. In *Proceedings of the CASC 2016*, volume 9890 of *LNCS*, pages 172–192. Springer, 2016.
[13] M. Feinberg. Stability of complex isothermal reactors–I. the deficiency zero and deficiency one theorems. *Chem. Eng. Sci.*, 42(10):2229–2268, 1987.

[14] D. Grigoriev and N. N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, 5:37–64, 1988.
[15] E. Gross, H. A. Harrington, Z. Rosen, and B. Sturmfels. Algebraic systems biology: A case study for the Wnt pathway. *Bull. Math. Biol.*, 78(1):21–51, 2016.
[16] H. Hong, R. Liska, and S. Steinberg. Testing stability by quantifier elimination. *J. Symb. Comput.*, 24(2):161–187, 1997.
[17] M. D. Johnston. A note on "MAPK networks and their capacity for multistationarity due to toric steady states". arXiv:1407.5651, 2014.
[18] 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.
[19] C. Li, M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. I. Stefan, J. L. Snoep, M. Hucka, N. Le Novère, and C. Laibe. BioModels database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, 4:92, 2010.
[20] 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(3):353–359, 2004.
[21] M. Pérez Millán and A. Dickenstein. The structure of MESSI biological systems. *Preprint: arXiv:1612.08763*, 2016.
[22] M. Pérez Millán and A. G. Turjanski. MAPK's networks and their capacity for multistationarity due to toric steady states. *Math. Biosci.*, 262:125–37, 2015.
[23] S. Schuster and T. Höfer. Determining all extreme semi-positive conservation relations in chemical reaction systems: a test criterion for conservativity. *J. Chem. Soc. Faraday T.*, 87(16):2561–2566, 1991.
[24] A. J. Sommese, J. Verschelde, and C. W. Wampler. Introduction to numerical algebraic geometry. In *Solving Polynomial Equations: Foundations, Algorithms, and Applications*, pages 301–337. Springer, 2005.
[25] D. Wang. *Elimination Methods*. Springer, 2000.
[26] D. Wang and B. Xia. Stability analysis of biological systems with real solution classification. In *Proceedings of the ISSAC 2005*, pages 354–361. ACM, 2005.
[27] V. Weispfenning. Quantifier elimination for real algebra—the quadratic case and beyond. *Appl. Algebr. Eng. Comm.*, 8(2):85–101, 1997.
[28] G. Weng, U. S. Bhalla, and R. Iyengar. Complexity in biological signaling systems. *Science*, 284(5411):92–6, 1999.