KAC-RICE FORMULAS AND THE NUMBER OF SOLUTIONS OF PARAMETRIZED SYSTEMS OF POLYNOMIAL EQUATIONS

ELISENDA FELIU[1], AMIRHOSEIN SADEGHIMANESH[2]

ABSTRACT. Kac-Rice formulas express the expected number of elements a fiber of a random field has in terms of a multivariate integral. We consider here parametrized systems of polynomial equations that are linear in enough parameters, and provide a Kac-Rice formula for the expected number of solutions of the system when the parameters follow continuous distributions. Combined with Monte Carlo integration, we apply the formula to partition the parameter region according to the number of solutions or find a region in parameter space where the system has the maximal number of solutions. The motivation stems from the study of steady states of chemical reaction networks and gives new tools for the open problem of identifying the parameter region where the network has at least two positive steady states. We illustrate with numerous examples that our approach successfully handles a larger number of parameters than exact methods.

Keywords: Kac-Rice formula, polynomial system, parameter region, Monte Carlo integration, multistationarity

INTRODUCTION

Systems of parametrized polynomial equations arise naturally in applications, and in particular in relation to steady states of polynomial ordinary differential equations (ODEs). We address here the problem of describing the function mapping a parameter vector to the number of solutions of the system specialized to the parameter vector. That is, given a parametrized system of n polynomial equations in n variables

$$(1) \qquad f_\kappa(t) = 0, \qquad t \in A, \quad \kappa \in B,$$

with $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$, we want to partition the parameter space B into regions where the number of solutions to the system in A is $0, 1, 2, \ldots, +\infty$.

The motivation stems from the study of the steady states of (bio)chemical reaction networks, where $A = \mathbb{R}^n_{>0}$ (the positive orthant), and B typically is $\mathbb{R}^m_{>0}$. In this setting, it is in particular of interest to understand for what parameter values the system describing the steady states of the network has at least two positive solutions (see Subsection 1.1). This property is termed *multistationarity*, and implies that the network, corresponding to a chemical system or mechanism in the cell, for instance, can potentially rest in two different states under exactly the same conditions. This property has received substantial attention in the context of systems and synthetic biology, for the implications in cell decision making [4, 8, 44]. In this context, only stable steady states are relevant, which leads to the concept of *bistability*, referring to the existence of two stable positive steady states. This usually implies that the network has at least three positive steady states, two of which are stable and one is unstable. There exist numerous approaches to determine whether multistationarity exists for some choice of parameter values, e.g. [10, 15, 17, 20, 21, 24, 25, 29, 33, 37, 42]. However, finding the parameter regions where the system displays multistationarity is a much harder question. Only very recently, approaches to partially

understand the region of multistationarity have been proposed, e.g. [7, 9, 12]. Furthermore, most methods to identify multistationarity return a parameter value in $\mathbb{R}^m_{>0}$ for which multistationarity occurs, but cannot be adapted to determine whether multistationarity also occurs for parameters in a given subset of $\mathbb{R}^m_{>0}$. This has the consequence that often, returned parameter values do not fall into biochemically relevant regions.

Theoretically, in order to partition the parameter space according to the number of solutions of the system, or to identify the parameter region of multistationarity, one might employ quantifier elimination or Cylindrical Algebraic Decomposition (CAD) [6]. However, these approaches have a high computational cost and become unfeasible already for small systems with three or four parameters and three or four variables.

In this work we explore the use of Kac-Rice formulas to study the parameter space. Kac introduced in [30] a formula to compute the expected number of real roots of a univariate polynomial with random coefficients. At about the same time, Rice introduced a similar formula for the number of crossings of ergodic stationary processes [38, 39]. These formulas became known as Kac-Rice formulas, and were extended in several directions, e.g. [3, 16, 47], see also [35]. This let to a metaformula for the expected number of elements of a fiber of a random field on a manifold under some conditions [1]. In general, a Kac-Rice formula expresses the expected number by means of a multivariate integral, and has found applications in many areas such as regression [45], the theory of random matrices [2], number theory [19] or enumerative geometry [5], to name a few.

In the first part of this work, we derive a Kac-Rice formula suited to polynomial systems with "suficient" linearity in the parameters (Theorem 1.1), which accommodates systems arising from reaction networks. The formula expresses, in terms of a multivariate integral, the expected number of solutions of system (1), when the parameters $\kappa_1, \ldots, \kappa_m$ are independent random variables with continuous distributions. We provide a direct proof of the formula that combines usual arguments in the derivation of Kac-Rice formulas [3]. We proceed to discuss how the Kac-Rice integral can be computed by employing numerical integration, mainly uisng Monte Carlo methods.

In the second part of this work, we provide numerous examples to illustrate how the Kac-Rice formula can be used in practice to address the following problems:

(i) Provide a grid partition of the parameter region B according to the number of solutions of system (1).

(ii) Find a parameter point or region for which the system has the maximal number of solutions, or at least M solutions (for some number M).

These questions are addressed by endowing the parameters with the uniform distribution in a box (product of intervals). Then the Kac-Rice formula gives the average number of solutions the system has when the parameters belong to the box. By making the boxes small, we can approximately partition the parameter region according to the number of solutions. We show that our approach can handle systems with over 10 parameters, where exact methods fail due to computational power. We also illustrate how parallelisation of our computations enables the study of complex systems.

The organization of the paper is as follows. In Section 1 the motivational setting of reaction networks is introduced, the statement of the Kac-Rice theorem is given, and we discuss Monte Carlo integration. Section 2 devises the strategy to use the Kac-Rice formula to study parameter regions, and illustrates it with numerous examples. Finally, Section 3 contains the proof of Theorem 1.1. Computational files can be accessed in the Github repository [43].

Notation. $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ refer to the non-negative and positive real numbers respectively. A box $B \subseteq \mathbb{R}^n$ is a Cartesian product $\prod_{i=1}^n B_i$ of (possibly unbounded) intervals of the real line. The intervals are allowed to be (half-)open or closed. For a set $B \subseteq \mathbb{R}^n$, we let $\chi_B(y)$ denote the indicator function being 1 if $y \in B$ and 0 otherwise.

1. Expected number of solutions using Kac-Rice formulas

1.1. Motivation: Reaction networks and multistationarity.

In this section we introduce the polynomial system of interest in the study of steady states of reaction networks. A *reaction network* on a set $\mathcal{S} = \{X_1, \ldots, X_n\}$ (*species set*) is a collection of *reactions* between linear combinations of species:

$$(2) \qquad \sum_{i=1}^{n} a_{ij} X_i \to \sum_{i=1}^{n} b_{ij} X_i, \qquad j = 1, \ldots, r,$$
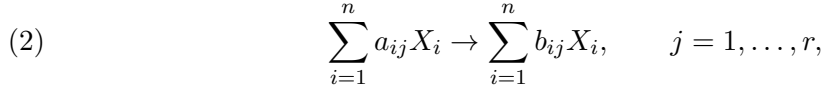
with $a_{ij}, b_{ij} \in \mathbb{Z}_{\geq 0}$. Let $x_i(t)$ denote the concentration of X_i at time t and $x(t) = (x_1(t), \ldots, x_n(t))$. Under the so-called *mass-action assumption* [22, 27], the evolution of the concentrations of the species in time is modeled by means of a polynomial system of autonomous ODEs in $\mathbb{R}^n_{\geq 0}$ of the form:

$$(3) \quad \frac{dx(t)}{dt} = F_k(x(t)), \qquad \text{where } F_{k,i}(x) = \sum_{j=1}^{r} (b_{ij} - a_{ij}) k_j \, x_1^{a_{1j}} \cdots x_n^{a_{nj}}, \quad i = 1, \ldots, n.$$

Here $k_j > 0$ are called *reaction rate constants*, and $0^0 = 1$ by convention. Typically, k_j are considered labels of the reactions, and by default their subindex indicates the order of the set of reactions. By letting $N \in \mathbb{Z}^{n \times r}$ be the matrix with entries $b_{ij} - a_{ij}$ for $i = 1, \ldots, n$, $j = 1, \ldots, r$, any vector ω in the left kernel of N gives rise to a linear first integral, as $\omega \cdot \frac{dx(t)}{dt} = 0$. Hence there are invariant linear subspaces with equations

$$W x = T, \qquad T \in \mathbb{R}^d,$$

for any matrix $W \in \mathbb{R}^{d \times n}$ whose rows form a basis of $\ker(N^t)$. These equations are called *conservation laws*, and T a vector of *total amounts*.

The *steady states* of the ODE system (3) in the invariant linear subspace with total amount T are the non-negative solutions to the system $F_k(x) = 0$, $Wx - T = 0$. As the conservation laws describe linear relations among the entries of F_k, d entries of F_k are redundant (linearly dependent of the rest) and can be removed. Let $\widetilde{F}_k(x)$ be a function with $n - d$ entries obtained in this way. Then the system of interest is square with n variables and n equations:

$$(4) \qquad \widetilde{F}_k(x) = 0, \qquad W x - T = 0.$$

The network is said to be *multistationary* if there exist $k = (k_1, \ldots, k_r) \in \mathbb{R}^r_{>0}$ and $T \in \mathbb{R}^d$ such that system (4) admits at least two positive solutions. Our ultimate goal is to understand how the number of positive solutions to (4) depends on $k \in \mathbb{R}^r_{>0}$ and $T \in \mathbb{R}^d$. This implies understanding the following map:

$$(5) \qquad \begin{aligned} \mathbb{R}^r_{>0} \times \mathbb{R}^d &\to \mathbb{N} \cup \{+\infty\} \\ \kappa := (k, T) &\mapsto \#\{x \in \mathbb{R}^n_{>0} \mid x \text{ is a solution to (4)}\}. \end{aligned}$$

The image of this map partitions the parameter space $\mathbb{R}^r_{>0} \times \mathbb{R}^d$.

1.2. The Kac-Rice formula.

We give here a Kac-Rice formula on the expected number of solutions of a polynomial system, which applies to polynomial systems with "sufficient" linearity in the parameters. This will later be applied to understand the map (5). In the following, measurability is with respect to the Borel σ-algebra on $\mathbb{R}^n$, and integrals are considered with respect to the Lebesgue measure on $\mathbb{R}^n$.

We consider functions of n polynomials in n variables and m parameters

$$f_\kappa(t) = (f_{\kappa,1}(t), \ldots, f_{\kappa,n}(t)), \qquad t = (t_1, \ldots, t_n) \in \mathbb{R}^n, \quad \kappa = (\kappa_1, \ldots, \kappa_m) \in \mathbb{R}^m,$$

with $m \geq n$, and such that the coefficients of the entries of $f_\kappa(t)$ are polynomials in κ. In the motivating scenario from Subsection 1.1, the polynomial map is given by the left-hand side of (4) and $\kappa = (k, T)$ such that $m = r + d$.

We assume that the parameters $\kappa_1, \ldots, \kappa_m$ are independent random variables with continuous distribution and density ρ_i in an interval B_i, for $i = 1, \ldots, m$. If t is such that the image of $f_\kappa(t)$ has non-zero measure in $\mathbb{R}^n$, then the values of $f_\kappa(t)$ for varying κ define a random variable taking values in $\mathbb{R}^n$ with a continuous distribution induced by the densities ρ_i. As $f_\kappa(t)$ is polynomial in κ, this image has non-zero measure if and only if it is Zariski dense in $\mathbb{R}^n$, or equivalently, the Jacobian of the polynomial map $f_\kappa(t) \colon \mathbb{R}^m \to \mathbb{R}^n$ with variable κ has maximal rank n. In particular, the image of $f_\kappa(t)$ in κ neither is constant nor lies in a proper algebraic variety of $\mathbb{R}^n$.

For a subset $A \subseteq \mathbb{R}^n$, consider the discrete random variable with state space $\mathbb{Z}_{\geq 0} \cup \{+\infty\}$ that counts the number of zeroes of f_κ in A, and let $\mathbb{E}\big(\#(f_\kappa^{-1}(0) \cap A)\big)$ be its expected value. We let $J_{f_\kappa}(t) = \big(\frac{\partial f_{\kappa,i}(t)}{\partial t_j}\big)_{i,j} \in \mathbb{R}^{n \times n}$ be the Jacobian matrix of $f_\kappa(t)$ with respect to t.

The following theorem gives a Kac-Rice formula for $\mathbb{E}\big(\#(f_\kappa^{-1}(0) \cap A)\big)$ for polynomial functions of a certain form, in line with the Kac-Rice metatheorem from [1, Ch 11]. The proof is given in Section 3.

Theorem 1.1 (Kac-Rice formula). *Let $A = A_1 \times \cdots \times A_n \subseteq \mathbb{R}^n$ be a box. Let $f_\kappa \colon A \to \mathbb{R}^n$ be a polynomial map whose coefficients are polynomials in $\kappa = (\kappa_1, \ldots, \kappa_m)$ with $m \geq n$. Assume that each parameter κ_i follows a continuous distribution with support on an interval B_i and density ρ_i, such that $\kappa_1, \ldots, \kappa_m$ are independently distributed. Assume ρ_i is a continuous function except maybe for a finite number of points of B_i, for $i = 1, \ldots, n$.*

Define $\widetilde{B} = B_{n+1} \times \cdots \times B_m$ and let $\bar{\kappa} = (\kappa_{n+1}, \ldots, \kappa_m)$. For each $i = 1, \ldots, n$, assume that there exist polynomials $h_i(\bar{\kappa}, t)$ and $q_i(\bar{\kappa}, t)$ in $\bar{\kappa}, t$, such that

$$(6) \qquad\qquad f_{\kappa,i}(t) = h_i(\bar{\kappa}, t)\kappa_i + q_i(\bar{\kappa}, t).$$

For $(\bar{\kappa}, t) \in \widetilde{B} \times A$ define

$$g_{\bar{\kappa},i}(t) := \frac{-q_i(\bar{\kappa}, t)}{h_i(\bar{\kappa}, t)}, \quad i = 1, \ldots, n, \qquad g_{\bar{\kappa}}(t) := (g_{\bar{\kappa},1}(t), \ldots, g_{\bar{\kappa},n}(t)),$$

and

$$\bar{\rho}(\bar{\kappa}, t) := \left(\prod_{i=1}^n \rho_i(g_{\bar{\kappa},i}(t))\right)\left(\prod_{i=n+1}^m \rho_i(\kappa_i)\right) \quad \text{if } \bar{\kappa} \in \widetilde{B}, \qquad \text{and} \quad \bar{\rho}(\bar{\kappa}, t) := 0 \quad \text{otherwise.}$$

Assume that

(i) $h_i(\bar{\kappa}, t)$ does not vanish in $\widetilde{B} \times A$ for $i = 1, \ldots, n$.

(ii) For $\bar{\kappa}$ outside a Zariski closed set (relative to $\widetilde{B}$) of measure zero $\widetilde{P} \subseteq \widetilde{B}$, the numerator of $\det(J_{g_{\bar{\kappa}}}(t))$ is a non-zero polynomial in t; equivalently $\det(J_{f_\kappa}(t))_{|(\kappa_1, \ldots, \kappa_n) = g_{\bar{\kappa}}(t)} \neq 0$ as a rational function in t, c.f. (8).

Then for all $t \in A$, the image of $f_\kappa(t)$ has positive measure in $\mathbb{R}^n$ and

$$(7) \qquad \mathbb{E}(\#(f_\kappa^{-1}(0) \cap A)) = \int_A \int_{\widetilde{B}} \big|\det\big(J_{g_{\bar{\kappa}}}(t)\big)\big| \, \bar{\rho}(\bar{\kappa}, t) \, d\kappa_{n+1} \ldots d\kappa_m \, dt.$$

Equality (7) is called the *Kac-Rice formula*, and the integral on the right-hand side of the equality is called the *Kac-Rice integral*.

The proof of Theorem 1.1 is given in Section 3. We typically consider uniform or normal distributions on the parameters, hence the density functions are continuous outside a finite

number of points. Note that Theorem 1.1(ii) implies that $q_i(\bar{\kappa}, t) \neq 0$ as a polynomial in $\bar{\kappa}, t$ for $i = 1, \ldots, n$. An easy computation shows that for $\kappa \in B_1 \times \cdots \times B_m$, we have

$$(8) \qquad \det\left(J_{g_{\bar{\kappa}}}(t)\right) = \frac{(-1)^n}{\prod_{i=1}^n h_i(\bar{\kappa}, t)} \det(J_{f_\kappa}(t))_{|(\kappa_1, \ldots, \kappa_n) = g_{\bar{\kappa}}(t)}.$$

Using this, one can show that (7) agrees with the usual expression of Kac-Rice formulas,

$$\int_A \mathbb{E}\left(|\det(J_{f_\kappa}(t))| \mid f_\kappa(t) = 0\right) p_t(0) dt,$$

where $p_t(0)$ is the density of the random variable f_κ at 0. However, derivation of the formula in the form (7) is more straightforward and avoids considering random variables conditioned on a measure zero set.

Back to the motivating scenario, the next theorem guarantees that the Kac-Rice formula can be applying to study systems arising from reaction networks as in (4).

Theorem 1.2. *System* (4) *is equivalent to a system that admits a decomposition of the form* (6) *satisfying assumption (i) of Theorem 1.1 for $A \subseteq \mathbb{R}_{>0}^n$ and any choice of intervals $B_i \subseteq \mathbb{R}$ for $i = 1, \ldots, m$.*

Proof. Let $\kappa = (k, T)$. The i-th equation in $Wx - T = 0$ (indexed $n - d + i$ in (4)) decomposes as in (6) with parameter T_i and $h_{n-d+i}(\kappa, t) = 1$.

Consider now the entries of $\widetilde{F}_k$ and let $\widetilde{F}_k = \widetilde{N} \operatorname{diag}(k) x^Y$, where $Y = (a_{ij}) \in \mathbb{R}^{n \times r}$ is the matrix of coefficients of the reactants, c.f. (2). By construction, $\widetilde{N}$ is any choice of $n - d$ linearly independent rows of N. Hence there exist column indices $i_1, \ldots, i_{n-d}$ such that the submatrix $\widetilde{N}'$ of $\widetilde{N}$ given by these columns has full rank $n - d$. For simplicity assume $i_j = j$, and write $\widetilde{N} = (\widetilde{N}' | \widetilde{N}'')$. Consider the function

$$G_k(x) = (\widetilde{N}')^{-1} \widetilde{F}_k(x) = \left(\operatorname{Id}_{n-d} \mid (\widetilde{N}')^{-1} \widetilde{N}'' \right) \operatorname{diag}(k) x^Y.$$

The solutions to $\widetilde{F}_k(x) = 0$ and to $G_k(x) = 0$ agree. Furthermore, $G_k(x)$ admits a decomposition as in (6) with $\bar{\kappa} = (k_1, \ldots, k_{n-d})$, $h_i(\bar{\kappa}, x) = x_1^{a_{1i}} \cdots x_n^{a_{ni}}$ and $q_i(\bar{\kappa}, x)$ the i-th row of $(\widetilde{N}')^{-1} \widetilde{N}'' \operatorname{diag}(\bar{\kappa}) x^{\bar{Y}}$, with $\bar{Y}$ consisting of the last $r - (n - d)$ columns of Y. Clearly, $h_i(\bar{\kappa}, x)$ does not vanish in $\mathbb{R}_{>0}^n$. $\qquad \square$

We illustrate Theorem 1.1 with a couple of simple examples, before we turn to computing the Kac-Rice integral.

Example 1.3. Perhaps the simplest non-trivial example to consider is the linear polynomial $f_\kappa(t) = \kappa_2 t - \kappa_1$, which has one positive root if $\kappa_1 \kappa_2 > 0$. Assume κ_1, κ_2 follow a uniform distribution in $[0, 1]$ and that t is positive (that is, $A = \mathbb{R}_{>0}$). We apply Theorem 1.1, with $g_{\kappa_2, 1}(t) = \kappa_2 t$. Assumption (ii) holds with $\widetilde{P} = \{\kappa \in [0, 1]^2 \mid \kappa_2 = 0\}$. We obtain

$$\mathbb{E}\left(\#(f_\kappa^{-1}(0) \cap \mathbb{R}_{>0})\right) = \int_0^{+\infty} \int_0^1 \kappa_2 \, \chi_{[0,1]}(\kappa_2 t) d\kappa_2 \, dt = \int_0^{+\infty} \int_0^{\min(1, \frac{1}{t})} \kappa_2 d\kappa_2 \, dt$$

$$= \int_0^{+\infty} \frac{\min(1, \frac{1}{t})^2}{2} dt = \int_0^1 \frac{1}{2} dt + \int_1^{+\infty} \frac{1}{2t^2} dt = 1,$$

in agreement with our expectation.

Example 1.4. Consider this simple system of polynomial equations $f_\kappa(t) = 0$

$$\kappa_1 - \kappa_3 t_1 = 0, \qquad \kappa_2 - \kappa_3 t_1 t_2 = 0.$$

We use Theorem 1.1 to determine the average number of solutions $(t_1, t_2) \in [0, 1]^2$ in the parameter box $[0, 1]^3$. With the notation of Theorem 1.1, we have $\kappa = (\kappa_3)$, $g_{\kappa_3, 1}(t) = \kappa_3 t_1$ and $g_{\kappa_3, 2}(t) = \kappa_3 t_1 t_2$, with $h_1 = h_2 = 1$. We consider each $\kappa_1, \kappa_2, \kappa_3$ uniformly distributed

in $[0,1]$. We have $|\det(J_{g_{\tilde{\kappa}}}(t))| = \kappa_3^2 t_1$. Hence assumption Theorem 1.1(ii) holds with $\widetilde{P} = \{\kappa \in [0,1]^3 \mid \kappa_3 = 0\}$. This leads to the following:

$$\mathbb{E}\big(\#(f_\kappa^{-1}(0) \cap [0,1]^2)\big) = \int_0^1 \int_0^1 \int_0^1 \kappa_3^2 t_1 \rho_1(\kappa_3 t_1)\rho_2(\kappa_3 t_1 t_2)\rho_3(\kappa_3)d\kappa_3 dt_1 dt_2$$

$$= \int_0^1 \int_0^1 \int_0^1 \kappa_3^2 t_1\, d\kappa_3 dt_1 dt_2 = \tfrac{1}{6},$$

where we have used $\kappa_3 t_1 < 1$ and $\kappa_3 t_1 t_2 < 1$.

Note that the solution to the system is $t_1 = \frac{\kappa_1}{\kappa_3}, t_2 = \frac{\kappa_2}{\kappa_1}$. This solution belongs to $[0,1]$ if and only if $\kappa_2 < \kappa_1 < \kappa_3$. The volume of this region within the cube $[0,1]^3$ is precisely $\frac{1}{6}$, in accordance with the result given by the Kac-Rice integral.

If the second equation is replaced with $\kappa_2 t_2 - \kappa_3 t_1 t_2 = 0$, then $h_2(\kappa_3, t) = t_2$ vanishes in A and hence Theorem 1.1 does not apply. However, after factoring this equation as $t_2(\kappa_2 - \kappa_3 t_1)$, the set of solutions of the original system is the union of the solution sets of two systems, arising from each factor, and for each of these systems Theorem 1.1 applies.

1.3. **Monte Carlo integration for the Kac-Rice formula.** Although in some cases, such as in Examples 1.3 and 1.4, the exact value of the Kac-Rice integral can be found, this is typically not the case and one needs to rely on numerical integration. To this end, we use Monte Carlo integration with importance sampling.

Monte Carlo integration. We give the main ingredients of Monte Carlo integration relevant to this work (see [36] for details). We consider an integral on a region $M \subseteq \mathbb{R}^n$ of the form

$$(9) \qquad I = \int_M f(x_1, \ldots, x_n)dx_1 \ldots dx_n.$$

Given any probability distribution P with non-zero probability density function $p(x) = p(x_1, \ldots, x_n)$ on M, it holds

$$I = \int_M \tfrac{f(x)}{p(x)}\, p(x)\, dx = \mathbb{E}\left(\tfrac{f(x)}{p(x)}\right).$$

That is, the integral is expressed as the expected value of the function $Q(x) := \frac{f(x)}{p(x)}$ with respect to the chosen probability distribution. By the Law of Large Numbers, for large $N \in \mathbb{N}$, the integral I can be approximated by the average value of Q evaluated at randomly sampled points $x^{(1)}, \ldots, x^{(N)}$ from P, that is, by

$$(10) \qquad \widehat{I}_N = \tfrac{1}{N} \sum_{i=1}^N Q(x^{(i)}).$$

Furthermore, an estimate of the standard error of the approximation is

$$(11) \qquad \hat{e}_N = \sqrt{\frac{\sum_{i=1}^N \left(Q(x^{(i)}) - \widehat{I}_N\right)^2}{N(N-1)}} = \sqrt{\frac{\frac{1}{N}\sum_{i=1}^N Q(x^{(i)})^2 - \widehat{I}_N^2}{N-1}},$$

where the second equality is well known (an easy to derive, see [36]).

We apply the approximation in (10) to the Kac-Rice integral I in (7) of Theorem 1.1. To this end, we need to choose a probability distribution on the domain $M = A \times \widetilde{B}$. First we note that the integration over A can sometimes be considered as a bounded integral. This happens for example if there exists a bounded subset A' of A containing $f_\kappa^{-1}(0)$ for all $\kappa \in B_1 \times \cdots \times B_m$. In this case we sample t using the uniform distribution on A'. If the

integral is unbounded, then apply the following decomposition of an unbounded integral to a sum of bounded integrals

$$
(12) \quad
\begin{aligned}
\int_0^{+\infty} g(x)dx &= \int_0^1 g(x)dx + \int_0^1 \tfrac{1}{x^2} g(\tfrac{1}{x})dx, \\
\int_{-\infty}^{+\infty} g(x)dx &= \int_0^1 \tfrac{1}{x^2} g(-\tfrac{1}{x})dx + \int_0^1 g(-x)dx \int_0^1 g(x)dx + \int_0^1 \tfrac{1}{x^2} g(\tfrac{1}{x})dx,
\end{aligned}
$$

and use the uniform distribution on $[0,1]$ for each variable t_i. Let $\mu(t)$ denote the density of the chosen distribution for t.

For the integral over $\widetilde{B}$ in the parameters $\kappa_{n+1}, \ldots, \kappa_m$, we simply use the original density function $\rho_{n+1} \times \cdots \times \rho_m$. This choice makes the expression of the corresponding sums in (10) simpler, thereby increasing the computational speed. Specifically, with these choices, the function $Q = f/p$ used in (10) becomes

$$
Q(t, \bar{\kappa}) = \big| \det \big(J_{g_{\bar\kappa}}(t)\big) \big| \frac{\bar\rho(\bar\kappa, t)}{\rho_{n+1}(\kappa_{n+1}) \cdot \ldots \cdot \rho_m(\kappa_m)\mu(t)} = \big| \det \big(J_{g_{\bar\kappa}}(t)\big) \big| \frac{\prod_{i=1}^n \rho_i(g_{\bar\kappa,i}(t))}{\mu(t)}.
$$

If A is split into subregions, then there is one such expression for each region, with a corresponding density function $\mu(t)$.

Monte-Carlo in practice. To approximate I, we sample n variables and $m - n$ parameters from the given distributions, and obtain points $t_1^{(i)}, \ldots, t_n^{(i)}, \kappa_{n+1}^{(i)}, \ldots, \kappa_m^{(i)}$ for $i = 1, \ldots, N$. We then compute $\widehat{I}_N$ and the standard error $\hat{e}_N$. We increase N and sample new points until

$$
(13) \quad\quad\quad\quad\quad\quad \frac{\hat{e}_N}{\widehat{I}_N} < 10^{-2}.
$$

We report $\widehat{I}_N$ with two digits of significance. Some considerations on the minimal sample size are given below.

This method easily allows parallelization. Specifically, the second expression for the standard error in (14) allows for an iterative computation of $\hat{e}_N$ without storing all sampled points, using the cumulative values of $\sum_{i=1}^N Q(x^{(i)})^2$ and $\sum_{i=1}^N Q(x^{(i)})$.

As indicated in [36, §2.3], the computation of $\hat{e}_N$ using (11) might lead to an imprecise value when $\hat{e}_N$ is very small. A way to bypass this problem is to consider $J_1 = Q(x^{(1)})$ and $S_1 := 0$, and iteratively construct the following functions for every new sampled point $x^{(i)}$, $i \geq 2$:

$$
(14) \quad\quad \delta_i = Q(x^{(i)}) - J_{i-1}, \quad\quad J_i = J_{i-1} + \tfrac{1}{i}\delta_i, \quad\quad S_i = S_{i-1} + \tfrac{i-1}{i}\delta_i^2.
$$

An easy computation shows that $\widehat{I}_N = J_N$ and $\hat{e}_N = \sqrt{\frac{S_N}{N(N-1)}}$ (see [36, §2.3]).

Note that division by $N(N-1)$ for N large may also cause numerical errors. Hence $\hat{e}_N$ is computed by first dividing S_N by N, and then by $N-1$.

When the sample size is too small, then $\widehat{I}_N$ might be an imprecise approximation of the integral I, even if the standard error is small. This happens when the integrand $f(x)$ in (9) is nearly zero on $M \setminus M'$, and the density $p(x)$ of the chosen probability distribution is small on M'. Then (10) and (11) are both close to 0 if the sample size is too small to cover M' properly (as $f(x)$ will be close to zero for most sampled points) [36, Ch 9].

In practice, it may be difficult to choose the "best" probability function. In this work, we adopt the following thumb rule for the minimum sample size. We compute $\widehat{I}_N$ and $\hat{e}_N$ for $N = 10^n$, starting with $n = 1$. We increase n until $\widehat{I}_N$ belongs to a reasonable interval. For example, if I is the Kac-Rice integral of a polynomial system that *we know* has between 1 and 3 solutions in A, then we expect $\widehat{I}_N \in [1,3]$. After this initial check on minimum sample size, we consider the termination condition (13).

Antithetic Monte Carlo. When the probability density function is symmetric, then one might use antithetic Monte Carlo [36, §8.2]. Specifically, for our setting, consider n independent uniform or (truncated) normal distributions on intervals $[a_i, b_i]$, such that the means of the truncated normal distributions are the centers of the respective intervals. Let $c = (\frac{a_1+b_1}{2}, \cdots, \frac{a_n+b_n}{2})$ be the center of the product of intervals and ρ the probability density function. Then $\rho(x) = \rho(2c - x)$.

Antithetic Monte Carlo consists in sampling $\frac{N}{2}$ points (for N even) and evaluating the function of interest in each sampled point and its reflection. If computing the reflection of a point is faster than sampling a point, then antithetic Monte Carlo is faster than simple Monte Carlo. Besides, the standard error of antithetic Monte Carlo is between 0 and $\sqrt{2}\hat{e}$, if $\hat{e}$ is the standard error of simple Monte Carlo [36, §8.2].

Implementation. In our computations, we considered simple and antithetic Monte Carlo implemented manually on different platforms[1]: `Maple`, `Python` (with and without the package `Numba`), `C++` and `Julia`. Additionally, we considered the `CUBA` package [28], as already implemented in all these platforms[2], which has four advanced numerical integration techniques: Vegas (Monte Carlo integration with importance sampling), Suave (Monte Carlo integration with globally adaptive subdivision and importance sampling), Divonne (Monte Carlo integration with stratified sampling and numerical optimization), and Cuhre (deterministic integration with globally adaptive subdivision).

Despite the advanced techniques implemented in `CUBA`, the computation of the Kac-Rice integral in `CUBA` often gives inaccurate answers in examples, while a manual implementation works well. We attribute the problem to the fact that we cannot choose the distribution for sampling in `CUBA`. This observation additionally supports the appropriateness of our choice of distribution.

The integrals reported in Section 2 have been computed using our own implementation[3] with `Julia`. In Subsection 2.2, we compare the speed of computation of the Kac-Rice integral for a specific example using `Maple`, `Python` (with and without `Numba`), `C++` and `Julia`. The analysis showed that `Numba` was the fastest option, competing closely with `Julia`. However, manual parallelization using `Julia` and the package `Distributed` is easier than with `Numba` using the module `multiprocessing`. Therefore we favoured `Julia` over `Numba`. For testing parallelization, we have used a server consisting of 64 cpus, AMD Opteron(tm) Processor 6380.

The code for the computations in Section 2 can be found in a GitHub repository archived by Zenodo [43]. A separate folder contains the relevant files for each subsection.

2. PARAMETER REGIONS USING KAC-RICE FORMULAS

As stated in the introduction and motivated in Subsection 1.1, our main goal is to understand the parameter region in relation to the number of positive solutions of a polynomial system. For a parametrized polynomial system $f_\kappa(t) = 0$, we focus on determining the expected number of positive solutions when the parameters κ belong to a bounded box B. To this end, we compute the Kac-Rice formula after endowing all parameters with uniform distributions. That is, let $B = B_1 \times \cdots \times B_m$ with B_i bounded intervals, consider $\kappa_i \sim U(B_i)$ for $i = 1, \ldots, m$, and let

$$\hat{r}(B) = \mathbb{E}(f_\kappa^{-1}(0) \cap \mathbb{R}_{>0}^n).$$

Then $\hat{r}(B)$ is the average number of positive solutions of the system $f_\kappa(t) = 0$ for $\kappa \in B$.

---

[1]Versions: `Maple 2020`, `Python 3.7.4`, `C++11`, `Numba 0.48.0` and `Julia 1.4.2`.

[2]List of platforms providing the CUBA package: `http://www.feynarts.de/cuba/`.

[3]Computations performed per default on Windows 10, Intel(R) Core(TM) i7-2670QM CPU @ 2.20GHz 2.20 GHz, x64-based processor, 6.00GB (RAM)

Let $M_{\max}$ and $M_{\min}$ be the maximal and minimal number of positive solutions the system $f_\kappa(t) = 0$ generically admits (that is, in some open set of $\mathbb{R}^m$). If $\hat{r}(B) = M_{\min}$ resp. $M_{\max}$, then for almost all parameter values in B, the system has $M_{\min}$, resp. $M_{\max}$ solutions. If $M_{\min} < \hat{r}(B) < M_{\max}$, then all we can assert is that B contains parameter values where the system has more than $M_{\min}$ solutions. In general, if $\hat{r}(B) > M$ for some M, then B contains parameter points κ where $f_\kappa(t) = 0$ has more than M positive solutions.

We aim at dividing the parameter region into areas where

(15)   (i) $\hat{r}(B) = M_{\max}$,   (ii) $\hat{r}(B) = M_{\min}$,   and   (iii) $M_{\min} < \hat{r}(B) < M_{\max}$,

or in the setting of reaction networks, into areas where

(16)   (i) $\hat{r}(B) = M_{\max}$,   (ii) $\hat{r}(B) \leq 1$,   and   (iii) $1 < \hat{r}(B) < M_{\max}$.

In this scenario, cases (i) and (iii) include the region of multistationarity if $M_{\max} > 1$.

With this in mind, we use $\hat{r}(B)$ (if it is well defined and can be computed) to address the following two problems.

Problem I: Coarse description of parameter regions of multistationarity. Let $\delta_1, \ldots, \delta_m > 0$ be the desired precision for each parameter. Consider a grid partition of some box $B = B_1 \times \cdots \times B_m$ in small sub-boxes $C_1, \ldots, C_\ell$ of side length at most δ_i for the i-th variable. We approximate the classification of the parameter points according to the number of solutions to $f_\kappa(t) = 0$ by computing $\hat{r}(C_i)$ for $i = 1, \ldots, \ell$ and classifying it into cases (i)-(iii) as in (15) or (16). In the setting of reaction networks, using (16) we obtain a coarse approximation of the real parameter region of multistationarity, as well as the region where multistationarity does not occur.

In order to optimize the speed of computation, we use a **bisect strategy**. If $\hat{r}(B) \neq M_{\min}, M_{\max}$, then we bisect B along one direction, and obtain two sub-boxes C_1, C_2. We compute $\hat{r}(C_1)$ and $\hat{r}(C_2)$. If C_1 belongs to cases (i) or (ii) of (15) or (16), then we have classified this box and move onto C_2. Otherwise, if some side of the box is larger than δ_i, we repeat with $C = C_1$. We perform the same procedure with C_2.

We start by considering the maximal number of steps for each parameter value κ_i, as given by the precision δ_i:

(17) $$L_i := \text{ceiling}\left(\log_2 \left(\tfrac{\text{length}(B_i)}{\delta_i} \right) \right), \qquad i = 1, \ldots, m.$$

At the j-th step, the direction of bisection is the axis along the parameter κ_i for which $i = j \pmod{m}$. If the direction of κ_i has already been bisected L_i times, then this direction is skipped.

In this way, larger boxes already belonging to (i) or (ii) are not subdivided and hence the computational time is reduced dramatically. This approach considers smaller boxes containing the boundary separating regions where the number of solutions to $f_\kappa(t) = 0$ changes.

With a grid description of the parameter region of multistationarity, one can derive a semialgebraic set defined by a single polynomial, which contains the multistationarity region, and with the minimal volume (see [41]). Additionally, the boundary of the region of multistationarity can be approximated by the hypersurface given by the polynomial in the superlevel set representation.

Problem II: Parameter point with multistationarity. We aim at finding a parameter point or sub-box for which the system has $M_{\max}$ solutions in a given bounded box C of interest, or conclude that no such parameter choice exists. To this end, we apply the bisect strategy, but keeping at each step the sub-box with largest $\hat{r}$, and stopping when $\hat{r}(C) = M_{\max}$ (approximately) or the maximal number of divisions has been reached for all parameters.

If the precision is small enough, this strategy is guaranteed to work if the system only has two possible number of solutions for generic parameter values. If that is not the case, then we might not identify a box with $M_{\max}$ solutions. For example, if the system generically admits one, three or five positive solutions, and at one step the two boxes C_1 and C_2 to consider are such that C_2 belongs to the region with three solutions, while C_1 intersects the regions with one and with five but such that $\hat{r}(C_2) > \hat{r}(C_1)$, we will miss the region with five solutions. To bypass this problem, we should search the parameter region as in **Problem I**, and keep both boxes unless $\hat{r}$ equals $M_{\min}$.

This approach can also be used to numerically determine the maximal number of positive solutions the system admits in a box C, and to search for parameter points for which the system has a given number of solutions M. If $M \neq M_{\max}, M_{\min}$, then $\hat{r}(C) \simeq M$ does not guarantee that all parameters in the box give rise to M solutions, so one needs to pick a point and verify the number of solutions by solving the system. Alternatively, in [41, Lemma 5.4] it is shown that by considering a distribution on κ different from the uniform, one can check whether $\hat{r}(C) \simeq M$ implies that all parameters in the box give rise to M solutions.

Theoretically, these two problems can be addressed using CAD [6, 13, 32]. However, this method is impractical as it is double exponential in the total number of variables and parameters, and depends also on the number and degree of the polynomials [18].

In what follows we provide several examples (mainly arising from reaction networks) to illustrate how to address the two problems described above by computing $\hat{r}(C)$ using the Kac-Rice formula. We start in Subsection 2.1 with an illustrative reaction network with eight parameters where the number of positive steady states is generically one or three, and the system (4) can be reduced to one polynomial equation. For illustration purposes, we start by fixing the value of six parameters, finding the parameter regions of interest, and comparing them to CAD. Afterwards, we show that **Problems I and II** can also be solved with eight free parameters.
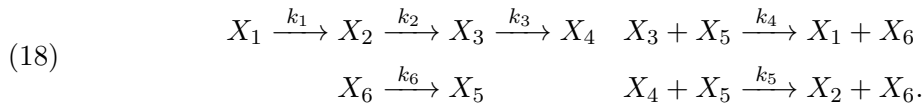
We proceed with another reaction network in Subsection 2.2 with five free parameters. We find a parameter point with multistationarity and compare the performance of simple and antithetic Monte Carlo in different platforms.

We next study a polynomial in one variable and two parameters that admits five positive roots (Subsection 2.3). We study the partition of the parameter space according to the number of positive roots of the polynomial as given by the Kac-Rice formula and compare the result to CAD.

Finally, we study two relevant reaction networks in Subsection 2.4 and 2.5, with a higher number of parameters and variables.

The computation of parameter regions using CAD has been done using the package `RootFinding[Parametric]` of `Maple 2020` [26].

2.1. Illustrative example: two component system.

The following reactions define a reaction network representing a simplified model of a two-component system with hybrid histidine kinase as considered in [31]:

$$(18) \quad \begin{aligned} X_1 \xrightarrow{k_1} X_2 \xrightarrow{k_2} X_3 \xrightarrow{k_3} X_4 \quad X_3 + X_5 \xrightarrow{k_4} X_1 + X_6 \\ X_6 \xrightarrow{k_6} X_5 \qquad\qquad X_4 + X_5 \xrightarrow{k_5} X_2 + X_6. \end{aligned}$$

The system of parametrized polynomial equations (4) is

$$\begin{aligned} k_4 x_3 x_5 - k_1 x_1 &= 0, & k_5 x_4 x_5 + k_1 x_1 - k_2 x_2 &= 0, \\ -k_4 x_3 x_5 + k_2 x_2 - k_3 x_3 &= 0, & -k_4 x_3 x_5 - k_5 x_4 x_5 + k_6 x_6 &= 0, \\ x_1 + x_2 + x_3 + x_4 - T_1 &= 0, & x_5 + x_6 - T_2 &= 0. \end{aligned}$$
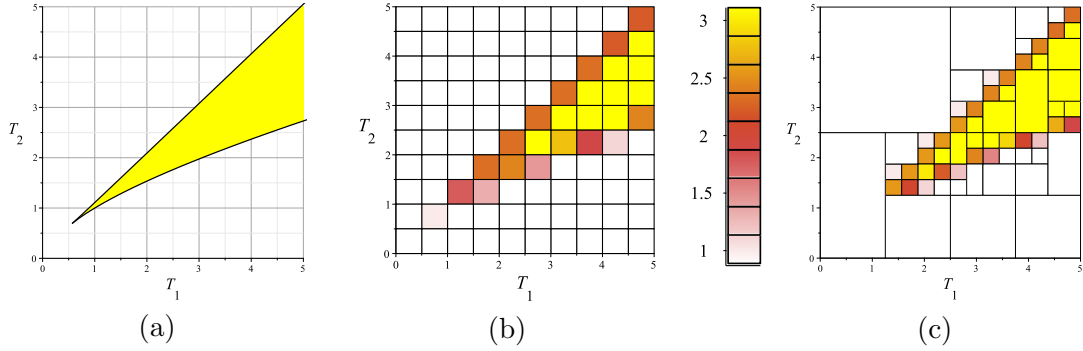
FIGURE 1. Parameter regions in T_1, T_2 according to the number of positive steady states for network (18) with k as in (20). (a) Obtained using CAD. The network has three positive steady states in the yellow region and one in the white region (and two on the boundary between the two regions). (b-c) Approximation of the parameter region using the Kac-Rice formula and numerical integration on sub-boxes. Yellow corresponds to three positive steady states and white to one, see bar diagram. For (b) the sub-boxes are constructed by grid partitioning, while for (c) the sub-boxes are constructed by the bisect strategy with termination condition of 4 divisions per parameter.

It is shown in [31] that the positive solutions to this system are in one-to-one correspondence with the positive solutions to the following univariate polynomial of degree three in $t = x_5$:

$$
\begin{aligned}
(19) \quad f_\kappa(t) = {}&(k_1 + k_2)k_4k_5k_6t^3 + (T_1k_1k_2k_4k_5 - T_2(k_1 + k_2)k_4k_5k_6 + k_1(k_2 + k_3)k_5k_6)t^2 \\
&+ (T_1k_1k_2k_3k_5 - T_2k_1k_5k_6(k_2 + k_3) + k_1k_2k_3k_6)t - T_2k_1k_2k_3k_6.
\end{aligned}
$$

Thus, in this example, the goal is to study the number of positive roots of a degree three polynomial, as function of the eight parameters $k_1, \ldots, k_6 > 0$ and T_1, T_2. As shown in [31], there exist parameter values for which (19) has three positive roots, and it always has at least one. Observe that we necessarily have $T_1, T_2 > 0$ for positive solutions to exist. CAD is computationally prohibitive with 8 parameters $k_1, \ldots, k_6, T_1, T_2$ on a standard computer. As we will see below, the Kac-Rice formula combined with Monte Carlo integration can cope with this situation.

Identifying the region of multistationarity. For *illustrative purposes*, we first fix the reaction rate constants k and understand the region defined by the parameters (T_1, T_2) according to the number of positive roots of the polynomial. In [9] it is shown that there exists a choice of $(T_1, T_2) \in \mathbb{R}^2_{>0}$ for which the network is multistationary if and only if $k_1 < k_3$. So we fix the following reaction rate constants (from [31, Fig. 2C]):

$$
(20) \qquad (k_1, \ldots, k_6) = (0.7329, 100, 73.29, 50, 100, 5).
$$

Evaluating the univariate polynomial (19) at (20) gives a polynomial $f_{T_1, T_2}(t)$ of degree 3 in t, whose coefficients depend on the two parameters T_1 and T_2:

$$
\begin{aligned}
(21) \quad f_{T_1, T_2}(t) = {}&(2518322.5)t^3 + \big((366450)T_1 - (2518322.5)T_2 + 63502.1205\big)t^2 \\
&+ \big((537142.41)T_1 - (63502.1205)T_2 + 26857.1205\big)t - (26857.1205)T_2.
\end{aligned}
$$

The analysis of this polynomial is addressable using CAD, which provides an explicit description of the region where (21) has three positive solutions. For $T_1, T_2 \in (0, 5)$, the region is depicted in Figure 1(a).

We consider now the same problem using the Kac-Rice formula and Monte Carlo integration. In the notation of Theorem 1.1, by letting $\bar{\kappa} = T_2$ and $n = 1$, we have

$$h(t) = 366450t^2 + 537142.41t,$$

$$q(T_2, t) = -(2518322.5t^2 + 63502.1205t + 26857.1205)(T_2 - t).$$

Then, for any bounded box $B = [a_7, b_7] \times [a_8, b_8]$, we have

$$(22) \qquad \hat{r}(B) = \int_0^{+\infty} \int_{a_8}^{b_8} \frac{|J_{T_2}(t)|}{(b_8 - a_8)(b_7 - a_7)} \chi_{[a_7, b_7]}\left(\frac{-q(T_2, t)}{h(t)}\right) dT_2 \, dt,$$

where $J_{T_2}(t) = \frac{\partial}{\partial t}\left(\frac{-q(T_2, t)}{h(t)}\right)$. As $t = x_5$ and $0 < x_5, x_6$, $x_5 + x_6 = T_2$, any positive root of (21) for parameter values in B satisfies $t < b_8$. Hence we choose $\mu(t)$ (c.f. Subsection 1.3) to be the density of the uniform distribution on $(0, b_8)$. The Kac-Rice integral is then approximated by the following sum for randomly sampled points $t^{(i)}, T_2^{(i)}$ for $i = 1, \dots, N$ and N large:

$$(23) \qquad \frac{b_8}{(b_7 - a_7)} \sum_{i=1}^{N} \left|J_{T_2^{(i)}}(t^{(i)})\right| \chi_{[a_7, b_7]}\left(\frac{-q(T_2^{(i)}, t^{(i)})}{h(t^{(i)})}\right).$$

We consider the box $B = [0, 5] \times [0, 5]$, subdivide it into 100 sub-boxes (of side length 0.5), and for each sub-box compute $\hat{r}(B)$ using (23). It took 46 seconds and 100 integrals were computed. We depict the output in Figure 1(b), where we color each sub-box with a graduation of yellow, orange and white: yellow means the expected number is three, and white means it is one.

Clearly, Figure 1(b) approximates Figure 1(a), which displays the exact region. In Figure 1(b) the sub-boxes that cross the thick line separating the yellow and white regions in Figure 1(a) have an orange-like color, because the sub-box contains parameters with both one and three positive steady states. By making the size of the sub-boxes smaller, we would get more accurate approximations of Figure 1(a).

Figure 1(c) has been found using the bisect strategy. In order to have a precision of at least 0.5, (17) gives that 4 bisections are (at most) required for each parameter. The process took 52 seconds, computed 111 integrals and returned 56 sub-boxes.

This example illustrates how the Kac-Rice formula can be used to approximate the parameter region. The advantage is that the numerical integrals we need to compute require, in principle, less computer power than performing CAD.

Finding a multistationary point in a box. We consider now the problem of finding a parameter value where (19) has three positive roots. We follow the approach outlined for **Problem II** at the beginning of this section.

Again for illustrative purposes, fix the reaction rate constants in (20) and consider the box $B = [1, 3] \times [2, 4]$ in the parameter space for T_1 and T_2. Computing the Kac-Rice integral (22) we find $\hat{r}(B) \simeq 1.29$, and hence there are parameter values in B yielding more than one positive steady state. We proceed to iteratively bisect B and compute the Kac-Rice integral for the two resulting sub-boxes, until we obtain a sub-box B' with $\hat{r}(B') \simeq 3$.

Figure 2(a) shows the result of an implementation of this process, and Figure 2(b) depicts the sub-boxes considered in the process and highlights the found sub-box. Figure 2(b) shows the real region with three positive steady states from Figure 1(a) in the background, such that one can visually interpret the expected number of solutions given in Figure 2(a), and verify that the final sub-box is entirely inside of the multistationarity region.

Step	Sub-box B	$\hat{r}(B)$	Chosen sub-box
0	$[1,3] \times [2,4]$	$\simeq 1.29$	✓
1	$[1,2] \times [2,4]$	$\simeq 1.00$	
	$[2,3] \times [2,4]$	$\simeq 1.58$	✓
2	$[2,3] \times [2,3]$	$\simeq 2.16$	✓
	$[2,3] \times [3,4]$	$\simeq 1.00$	
3	$[2,2.5] \times [2,3]$	$\simeq 1.68$	
	$[2.5,3] \times [2,3]$	$\simeq 2.65$	✓
4	$[2.5,3] \times [2,2.5]$	$\simeq 3.00$	✓
	$[2.5,3] \times [2.5,3]$	$\simeq 2.30$	

(a)
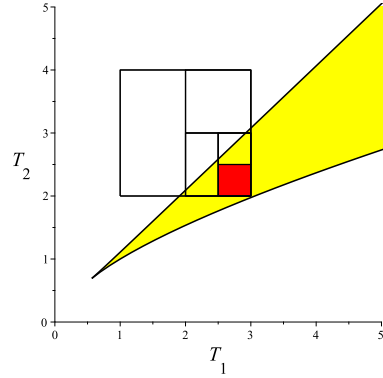
(b)

FIGURE 2. **Problem II** for network (18) with k as in (20). At each step, $\hat{r}$ is computed for the considered sub-boxes, and the one with the largest $\hat{r}$ is bisected, until $\hat{r}$ is 3 with two decimal digits of precision. At step 0, the test verifies that $\hat{r} > 1$, otherwise the box does not intersect the region of multistationarity. Here the T_1-axis is chosen in odd steps, and the T_2-axis in even steps. (a) Table description of the considered sub-boxes and their $\hat{r}$. (b) Visual depiction of the sub-boxes in (a). The final sub-box is colored in red and is entirely inside the multistationary region (colored in yellow). The sub-boxes with $\hat{r} = 1$ are outside the yellow region, and the sub-boxes with $1 < \hat{r} < 3$, have intersection with both the white and yellow regions.

With 8 parameters. In the previous analysis of network (18), we kept only 2 parameters free to be able to visually illustrate our approach, as well as to compare with CAD. We show here that we can find a parameter point/box where the network has three positive steady states, also when all 8 parameters are free.

We consider the following box for the parameter vector $(k_1, k_2, k_3, k_4, k_5, k_6, T_1, T_2)$:

$$B = (0,1) \times (0,200) \times (0,100) \times (0,100) \times (0,200) \times (0,10) \times (0,5) \times (0,5).$$

Computing the Kac-Rice integral we find $\hat{r}(B) \simeq 1.2$ with two digits of significant. Therefore it has intersection with the multistationarity region. We apply the algorithm for **Problem II**, which, after 22 iterations in less than 58 seconds, returns the following sub-box:

$$C = (0.125, 0.25) \times (125, 150) \times (75, 87.5) \times (12.5, 25) \times (175, 200)$$
$$\times (2.5, 3.75) \times (3.75, 5) \times (3.75, 5).$$

For almost all parameter values in this box, the network has three positive steady states, because $\hat{r}(C) = 2.97$ with standard error $\hat{e} = 0.009$.

We address also **Problem I** with the bisect strategy to obtain a coarse description of the parameter region of multistationarity inside the following box B,

$$(0.125, 0.375) \times (100, 125) \times (75, 100) \times (12.5, 37.5) \times (150, 200) \times (1.25, 3.75) \times (0,5) \times (0,5).$$

As input precision, we considered $\delta = (\delta_1, \ldots, \delta_8)$ (the upper bound of the minimal length of the intervals for $\kappa_1, \ldots, \kappa_8$) as follows:

$$\delta = (0.125, 12.5, 12.5, 12.5, 25, 1.25, 1.25, 1.25).$$

The algorithm terminates after computing 1787 integrals in 2164 seconds. The obtained partition consists of 894 sub-boxes, 204 of which are out of the multistationary region, 334 are inside the multistationary region, and the remaining 356 have intersection with both the region of multistationarity and monostationarity.

2.2. **Method and platform comparison: An example with 5 parameters.** Consider the following reaction network

$$(24) \qquad 2X_1 + X_2 \underset{k_4}{\overset{k_1}{\rightleftarrows}} 3X_1 \xrightarrow{k_2} X_1 + 2X_2 \xrightarrow{k_3} 3X_2 \ .$$
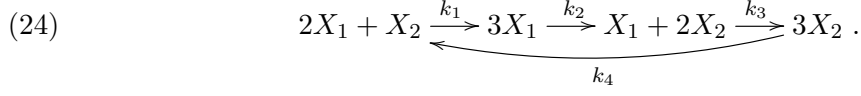
System (4) becomes a parametrized polynomial system in five parameters and two variables:

$$(25) \qquad \kappa_1 t_1^2 t_2 - 2\kappa_2 t_1^3 - \kappa_3 t_1 t_2^2 + 2\kappa_4 t_2^3 = 0, \qquad\qquad t_1 + t_2 - \kappa_5 = 0,$$

where $\kappa_i = k_i > 0$, $i = 1, \ldots, 4$ and $\kappa_5 > 0$. We find $\hat{r}(B)$ for $B = \prod_{i=1}^{5}(a_i, b_i)$ a box in the parameter space with $a_i \geq 0$. In order to apply Theorem 1.1, we choose κ_1 and κ_5 as the linear parameters, which gives $\bar{\kappa} = (\kappa_2, \kappa_3, \kappa_4)$,

$$g_{\bar{\kappa},1}(t) = \tfrac{1}{t_1^2 t_2}(2\kappa_2 t_1^3 + \kappa_3 t_1 t_2^2 - 2\kappa_4 t_2^3), \qquad\qquad g_{\bar{\kappa},2}(t) = t_1 + t_2,$$

and

$$\det(J_{g_{\bar{\kappa}}}(t)) = \tfrac{1}{t_1^3 t_2^2}(t_1 + t_2)(2\kappa_2 t_1^3 - \kappa_3 t_1 t_2^2 + 4\kappa_4 t_2^4).$$

The numerator of $\det(J_{g_{\bar{\kappa}}}(t))$ is not identically zero as long as $\bar{\kappa} \neq 0$. As the hypotheses of Theorem 1.1 hold, the expected number of positive solutions to the system for parameters in B is given by the Kac-Ric integral (7)

$$\hat{r}(B) = \int_0^{+\infty} \int_0^{+\infty} \int_{a_2}^{b_2} \int_{a_3}^{b_3} \int_{a_4}^{b_4} \big| \det\big(J_{g_{\bar{\kappa}}}(t)\big) \big| \chi_{[a_1,b_1]}(g_{\bar{\kappa},1}(t)) \cdot \chi_{[a_5,b_5]}(g_{\bar{\kappa},2}(t)) \cdot$$

$$\left(\prod_{i=1}^{5} \tfrac{1}{b_i - a_i}\right) d\kappa_2 \, d\kappa_3 \, d\kappa_4 \, dt_2 \, dt_1.$$

As $0 < t_1, t_2$ in A and $t_1 + t_2 = \kappa_5$, the values of t_1 and t_2 as solutions to the system are bounded above by b_5. In the computation of $\widehat{I}_N$ using Monte Carlo, we consider $\mu(t_1, t_2)$ to be the density of $U(0, b_5) \times U(0, b_5)$. Given sampled points $t_1^{(i)}, t_2^{(i)}, \kappa_2^{(i)}, \kappa_3^{(i)}, \kappa_4^{(i)}$ for $i = 1, \ldots, N$ and N large, the Kac-Rice integral is approximated by the following sum:

$$(26) \qquad \tfrac{b_5^2}{(b_1 - a_1)(b_5 - a_5)} \sum_{i=1}^{N} \big| \det\big(J_{g_{\bar{\kappa}(i)}}(t^{(i)})\big) \big| \chi_{[a_1,b_1]}\big(g_{\bar{\kappa}(i),1}(t^{(i)})\big) \chi_{[a_5,b_5]}\big(g_{\bar{\kappa}(i),2}(t^{(i)})\big).$$

To illustrate this, consider the bounded box

$$(27) \qquad B = (0, 100) \times (0, 2) \times (0, 200) \times (0, 100) \times (0, 2).$$

Table 1(a) summarises the computed approximation of $\hat{r}(B)$ using (26) with simple and antithetic Monte Carlo, as N is increased. This shows that the expected number of positive solutions to the system for parameters in B is around 1.4. Antithetic Monte Carlo is about 68% faster than simple Monte Carlo in this case with the same accuracy.

For comparison, we considered also truncated uniform distributions on the parameters $\kappa_i \sim \bar{N}_{(a_i, b_i)}(\mu_i, \sigma_i)$ with probability density function ρ_i. Then the Kac-Rice integral can be approximated with the following Monte Carlo sum

$$b_5^2 \sum_{i=1}^{N} \big| \det\big(J_{g_{\bar{\kappa}(i)}}(t^{(i)})\big) \big| \rho_1(g_{\bar{\kappa}(i),1}(t^{(i)})) \rho_5(g_{\bar{\kappa}(i),2}(t^{(i)})),$$

after sampling using $t_1 \sim U(0, b_5)$, $t_2 \sim U(0, b_5)$ and $\kappa_i \sim \bar{N}_{(a_i, b_i)}(\mu_i, \sigma_i)$ for $i = 2, 3, 4$. Results are shown in Table 1(b) for the box in (27), the mean μ_i the center of each interval, and $\sigma_i = 0.1$. Antithetic Monte Carlo is about 49% faster than simple Monte Carlo in this case with the same accuracy.

N	(a) Uniform distribution						(b) Truncated normal distribution					
	Simple Monte Carlo			Antithetic Monte Carlo			Simple Monte Carlo			Antithetic Monte Carlo		
	$\widehat{I}_N$	$\hat{e}_N$	Time	$\widehat{I}_N$	$\hat{e}_N$	Time	$\widehat{I}_N$	$\hat{e}_N$	Time	$\widehat{I}_N$	$\hat{e}_N$	Time
10	0.527	0.583	0.000023	0.199	0.172	0.000008	0.000	0.000	0.000016	0	0	0.000023
10^2	2.542	1.884	0.000021	1.100	0.415	0.000015	0.000	0.000	0.000060	0.000	0.000	0.000057
10^3	2.470	0.745	0.000165	0.942	0.147	0.000095	0.000	0.000	0.000543	0.000	0.000	0.000391
10^4	1.468	0.175	0.001570	1.662	0.333	0.000950	1.102	1.101	0.005587	0.002	0.002	0.004718
10^5	1.990	0.595	0.015469	1.392	0.055	0.008920	0.127	0.069	0.054673	2.027	0.817	0.038053
10^6	1.432	0.031	0.150625	1.449	0.034	0.090292	1.019	0.178	0.520236	1.021	0.171	0.372723
10^7	1.422	0.007	1.536203	1.419	0.009	0.939078	0.963	0.056	5.218110	0.965	0.057	3.681651
10^8	1.413	0.003	15.49415	1.415	0.003	10.04099	1.020	0.019	52.14723	1.034	0.019	37.95210
10^9	1.419	0.001	155.5443	1.418	0.001	92.33368	1.010	0.006	537.2604	0.989	0.006	371.1164

TABLE 1. Output for network (24) with computations done in `Julia`. The time is reported in seconds and rounded. In (b) the minimum considered sample size is 10^5 (see Subsection 1.3). The first cells satisfying the stop condition are highlighted. With two digits of significance, $\hat{r}$ is 1.4 and 1.0 in (a) and (b) respectively.

N	Time	
	Monte Carlo	Antithetic Monte Carlo
`Maple 2020` 10^5	177.116	96.411
`Python 3.7.4` 10^7	89.74713	78.52874
`C++11 Dev-Cpp 5.11` 10^8	142.395	94.6376
`Numba 0.48.0` 10^9	93.41889	59.82621
`Julia 1.4.2` 10^9	156.2837	88.81576
`Julia 1.4.2` parallelized with 2 workers 10^9	76.16551	47.35521
`Julia 1.4.2` parallelized with 32 workers 10^{10}	116.7884	34.55502

TABLE 2. Computation time (in seconds) for $\widehat{I}_N$ in (26) and $\hat{e}_N$ using (14). Computation is performed using both simple and antithetic Monte Carlo sampling and different platforms. We report the computation time of the largest sample size $N = 10^n$ taking less than 200 seconds. The antithetic Monte Carlo is faster on all platforms, and `Numba` is the fastest of the considered platforms. The parallel implementation in `Julia` using the package `Distributed` with n workers increases the speed of computations by a factor of $\max(n, d/2)$ where d is the number of cpus of the computer.

We have also used this example to compare the time it takes to compute $\widehat{I}_N$ and $\hat{e}_N$ using the algorithm (14) on different platforms. In Table 2 we report the largest value of the type $N = 10^n$ that can be computed under 200 seconds. Among the five considered platforms, `Numba` is the fastest.

Finding a point in multistationary region. This network admits between one and three positive steady states. We use the Kac-Rice integral and Monte Carlo integration (in `Julia`) to find a parameter point where the network has three positive steady states. To this end, we consider the following starting box:

$$B = (0, 100) \times (0, 2) \times (0, 200) \times (0, 100) \times (0, 2).$$

The algorithm outlined for **Problem II** concludes with the box

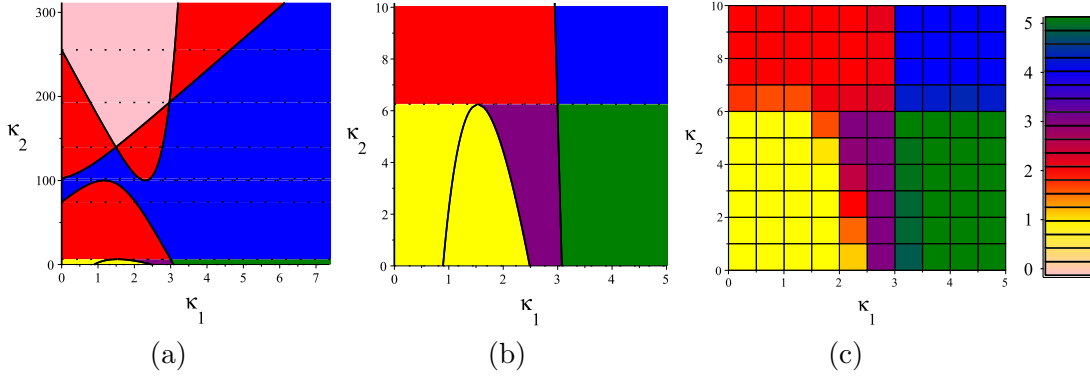$$C = (50, 75) \times (1, 1.5) \times (150, 200) \times (0, 50) \times (1, 2),$$

FIGURE 3. Partition of the parameter region of (28) according to the number of positive roots. The color indicates the number of positive roots as given by the bar code on the right. (a-b) Partition obtained using CAD for two different boxes. (c) Grid partition and computation of $\hat{r}$ with the Kac-Rice formula and Monte Carlo integration.

after 8 steps in less than 7.5 seconds. Therefore, for almost all parameters in C, the network has three positive steady states, and hence multistationarity.

2.3. **Finding** 5 **solutions.** We now analyse an example where the maximal number of solutions is five. Consider the following parametrized univariate polynomial of degree five in the variable t and parameters κ_1, κ_2:

$$
(28) \qquad \begin{aligned}
f_\kappa(t) = t^5 - (\kappa_1 + \tfrac{9}{2})t^4 + (\tfrac{9}{2}\kappa_1 + \tfrac{21}{4})t^3 + (-\tfrac{23}{4}\kappa_1 + \tfrac{3}{8})t^2 \\
+ (\tfrac{15}{8}\kappa_1 - \tfrac{23}{8})t + (\tfrac{1}{100}\kappa_2 - \tfrac{1}{16}).
\end{aligned}
$$

Using CAD, we know that $f_\kappa(t)$ generically has 0, 1, 2, 3, 4 or 5 positive roots for suitable choices of the parameter vector $(\kappa_1, \kappa_2) \in \mathbb{R}^2_{>0}$ (see Figure 3(a-b)). The polynomial $f_\kappa(t)$ is linear in κ_2 with coefficient $h(\kappa_1, t) = 100$, and the hypotheses of Theorem 1.1 hold for $A = \mathbb{R}_{>0}$. Hence, for a box $B = [a_1, b_1] \times [a_2, b_2]$ with $0 < a_i$, $\hat{r}(B)$ is given by:

$$
(29) \qquad \hat{r}(B) = \tfrac{1}{b_2 - a_2} \int_0^{+\infty} \int_{a_1}^{b_1} |g'_{\kappa_1}(t)| \chi_{[a_2, b_2]}(g_{\kappa_1}(t))\, d\kappa_1\, dt,
$$

where

$$
g_{\kappa_1}(t) = -100\left(t^5 + (-\kappa_1 - \tfrac{9}{2})t^4 + (\tfrac{9}{2}\kappa_1 + \tfrac{21}{4})t^3 + (-\tfrac{23}{4}\kappa_1 + \tfrac{3}{8})t^2 + (\tfrac{15}{8}\kappa_1 - \tfrac{23}{8})t - \tfrac{1}{16}\right).
$$

In order to approximately partition the box $B = [0, 5] \times [0, 10]$ according to the number of positive roots of $f_\kappa(t)$, we subdivide it into 100 sub-boxes and compute $\hat{r}$ using Monte Carlo integration. By (12), integral (29) can be written as the following bounded integral:

$$
\hat{r}(B) = \tfrac{1}{b_2 - a_2} \int_0^1 \int_{a_1}^{b_1} \left(|g'_{\kappa_1}(t)| \chi_{[a_2, b_2]}(g_{\kappa_1}(t)) + \tfrac{1}{t^2} |g'_{\kappa_1}(\tfrac{1}{t})| \chi_{[a_2, b_2]}(g_{\kappa_1}(\tfrac{1}{t})) \right) d\kappa_1\, dt.
$$

By choosing t and κ_1 to follow uniform distributions on $[0, 1]$ and $[a_1, b_1]$ respectively, $\hat{r}(B)$ is approximated by

$$
\tfrac{1}{b_2 - a_2} \sum_{i=1}^N \left(|g'_{\kappa_1^{(i)}}(t^{(i)})| \chi_{[a_2, b_2]}(g_{\kappa_1^{(i)}}(t^{(i)})) + \tfrac{1}{(t^{(i)})^2} |g'_{\kappa_1^{(i)}}(\tfrac{1}{t^{(i)}})| \chi_{[a_2, b_2]}(g_{\kappa_1^{(i)}}(\tfrac{1}{t^{(i)}})) \right)
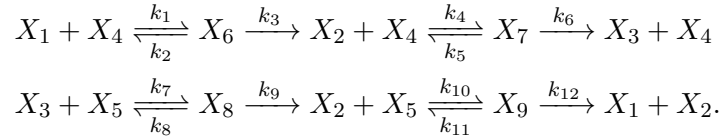$$

for sampled points $t^{(i)}$, $\kappa_1^{(i)}$ for $i = 1, \dots, N$ and N large.

Figure 3(c) depicts each of these sub-boxes, colored corresponding to the approximated value of $\hat{r}$ for $N = 10^9$. Note that $\hat{r}(C) \simeq 2$ for the box $C = [2, 2.5] \times [2, 2.5]$. However this sub-box is not located inside or even have intersection with the (open) parameter region

where $f_\kappa(t)$ has two positive roots. It intersects only regions with one and three solutions, but the areas of the two intersections are almost equal. Only when $\hat{r}(C)$ is zero or five ($M_{\min}$ and $M_{\max}$ here), we can conclude that almost all parameter points in the box yield to zero or five positive roots. For example, the box $[3.5, 5] \times [0, 6]$ is entirely inside the parameter region with five positive roots.

For this computation, the standard error $\hat{e}_N$ increased with a_1, going from 0.002 to a maximal value of 0.02 for the box $[4.5, 5] \times [9, 10]$ independently of a_2.

2.4. **Dual phosphorylation.** We consider the following reaction network:

$$X_1 + X_4 \xrightleftharpoons[k_2]{k_1} X_6 \xrightarrow{k_3} X_2 + X_4 \xrightleftharpoons[k_5]{k_4} X_7 \xrightarrow{k_6} X_3 + X_4$$

$$X_3 + X_5 \xrightleftharpoons[k_8]{k_7} X_8 \xrightarrow{k_9} X_2 + X_5 \xrightleftharpoons[k_{11}]{k_{10}} X_9 \xrightarrow{k_{12}} X_1 + X_2.$$

This network models the distributive and sequential phosphorylation and dephosphorylation of a substrate that is phosphorylated at none, one or two sites (X_1, X_2 and X_3), catalysed by enzymes X_4, X_5. Generically, this network has either one or three positive steady states [46]. Even though substantial work has been done to understand the parameter region of multistationarity [7, 11, 12, 23, 34], a full description is still unknown.

This network has three conservation laws

$$x_1 + x_2 + x_3 + x_6 + x_7 + x_8 + x_9 = T_1, \quad x_4 + x_6 + x_7 = T_2, \quad x_5 + x_8 + x_9 = T_3.$$

System (4) can be simplified to a system of three polynomials in the three variables x_1, x_4, x_5 and 15 parameters, see [7]. Essentially, the equations $\widetilde{F}_k(x) = 0$ in (4) can be solved for x_1, x_4, x_5, and the output inserted into the three conservation laws. Hence the three polynomials are linear in T_1, T_2, T_3, respectively. With $\bar{\kappa} = (k_1, \ldots, k_{12})$, the hypotheses of Theorem 1.1 hold. The numerator of the rational function $\det(J_{g_{\bar{\kappa}}}(t))$ has total degree 18 (in the variables and parameters) and 165 terms. The denominator has total degree 10 and 9 terms.

By [11, Corollary 4.13], parameter choices where $T_1 < T_2$ and $T_1 < T_3$ do not yield multistationarity. Therefore for the following box $\hat{r}(B)$ is one:

$$B = (500, 1000) \times (25, 50) \times (25, 50) \times (5, 10) \times (5, 10) \times (5, 10) \times (5, 10) \times (1, 2)$$
$$\times (1, 2) \times (5, 10) \times (50, 100) \times (50, 100) \times (1, 2) \times (2, 4) \times (2, 4).$$

With Monte Carlo integration, it takes 3.2 seconds to approximate $\hat{r}(B)$ to one with two digits of significance using $N = 10^7$ (the minimum sample size is set to 10^3).

We consider this other box:

$$B = (0.5, 1.5) \times (509.5, 510.5) \times (1.5, 2.5) \times (1.5, 2.5) \times (0.5, 1.5) \times (0.5, 1.5)$$
$$\times (1.5, 2.5) \times (0.5, 1.5) \times (0.5, 1.5) \times (1.5, 2.5) \times (0.5, 1.5) \times (0.5, 1.5) \times (110, 150)$$
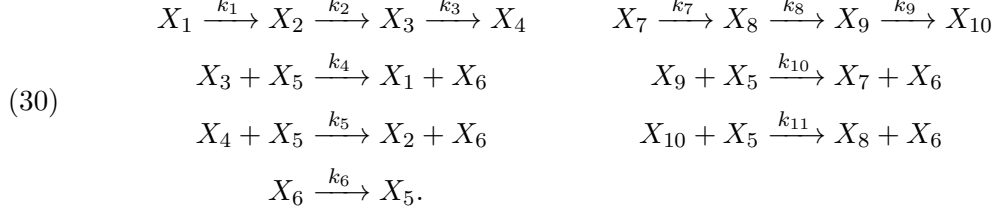$$\times (20, 30) \times (15, 25).$$

In this case, the minimum sample size for the computation of $\hat{r}(B)$ is $N = 10^7$. With antithetic Monte Carlo we obtain that $\hat{r}(B) = 1.45$ in 27564 seconds with standard error $\hat{e} = 0.015$ (1011.5 seconds using 32 workers). Therefore this box intersects the region of multistationarity.

In order to find a box inside the region of multistationarity, we have considered the bisect approach with the box B and parallelized with 32 workers. We obtain the box

$$(1.125, 1.25) \times (510.0, 510.25) \times (1.5, 1.75) \times (2.25, 2.5) \times (0.5, 0.75) \times (0.5, 0.75) \times (2.25, 2.5)$$
$$\times (0.5, 0.75) \times (0.75, 1.0) \times (2.0, 2.25) \times (0.75, 1.0) \times (1.25, 1.5) \times (117.5, 120.0)$$
$$\times (25.0, 27.5) \times (15.0, 20.0),$$

with $\hat{r}$ equal to 2.9 and $\hat{e} = 0.02$, with $N = 10^{12}$.

2.5. **Extended hybrid histidine-kinase network.** Finally, as a last example, we consider an extension of the hybrid histidine-kinase network studied in Subsection 2.1 as given in [31]:

$$
\begin{array}{lll}
X_1 \xrightarrow{k_1} X_2 \xrightarrow{k_2} X_3 \xrightarrow{k_3} X_4 & \qquad & X_7 \xrightarrow{k_7} X_8 \xrightarrow{k_8} X_9 \xrightarrow{k_9} X_{10} \\
X_3 + X_5 \xrightarrow{k_4} X_1 + X_6 & & X_9 + X_5 \xrightarrow{k_{10}} X_7 + X_6 \\
X_4 + X_5 \xrightarrow{k_5} X_2 + X_6 & & X_{10} + X_5 \xrightarrow{k_{11}} X_8 + X_6 \\
X_6 \xrightarrow{k_6} X_5.
\end{array}
$$
(30)

This network has three conservation laws

$$x_1 + x_2 + x_3 + x_4 = T_1, \qquad x_5 + x_6 = T_2, \qquad x_7 + x_8 + x_9 + x_{10} = T_3,$$

and hence involves a 14-dimensional parameter vector: $\kappa = (k_1, \ldots, k_{11}, T_1, T_2, T_3)$. By [31], network (30) admits between one and five positive steady states. The corresponding system (4) can be simplified to a univariate polynomial $f_\kappa(t)$ of degree five in $t = x_5$, which further is linear in each of κ_{12}, κ_{13} and κ_{14}. We choose to isolate $\kappa_{12} = T_1$, such that $\bar{\kappa} = (k_1, \ldots, k_{11}, T_2, T_3)$. The hypotheses of Theorem 1.1 hold.

We use Monte Carlo integration and the Kac Rice formula to address **Problem II** and find a multistationary point in the box

(31) $\quad B = (0, 2) \times (0, 100) \times (0, 50) \times (0, 10) \times (0, 100) \times (0, 50) \times (0, 5)$
$$\times (50, 100) \times (0, 100) \times (0, 100) \times (0, 100) \times (0, 100) \times (0, 50) \times (0, 100).$$

We have $\hat{r}(B) = 1.21$ with $\hat{e} = 0.010$. We set the termination condition of the search algorithm to be $\hat{r}(C) \geq 2.95$. After 32 bisections, computing 65 integrals in 3083 seconds, the algorithm terminates and returns the following sub-box:

$$C = (0.25, 0.5) \times (75, 87.5) \times (43.75, 50) \times (3.75, 5) \times (87.5, 100) \times (18.75, 25) \times (1.25, 2.5)$$
$$\times (62.5, 75.0) \times (75, 100) \times (25, 50) \times (75, 100) \times (75, 100) \times (25, 50) \times (50, 100).$$

As $\hat{r}(C) = 2.95$ with $\hat{e}_N = 0.008$, there must be parameter points yielding to three or five positive steady states. For 975 out of 1000 random parameter points in C, the polynomial has three positive roots (found numerically).

For the box B, the minimum sample size for Monte Carlo integration is 10. However, it is not always the case. For example, fix all parameters other than κ_{12} and κ_{14} as follows

(32) $\quad \kappa_1 = 0.1, \quad \kappa_2 = 120, \quad \kappa_3 = 17.95, \quad \kappa_4 = 0.1795, \quad \kappa_5 = 0.713, \quad \kappa_6 = 1,$
$\quad \kappa_7 = 0.002, \quad \kappa_8 = 500, \quad \kappa_9 = 160, \quad \kappa_{10} = 0.147, \quad \kappa_{11} = 4.15, \quad \kappa_{13} = 16.27.$

These values are taken from [31, Fig. 3B for $n = 2$], and, for the right choice of κ_{12}, κ_{14}, yield to five positive steady states. We have

$$g_{\kappa_{14}}(t) = 9377005242288437670000 t^5 + (18753935468835\kappa_{14} - 146395097463035713.8909)10^4 t^4$$
$$+ (204244474710835\kappa_{14} + 360657036215560.9291)10^6 t^3$$
$$+ (1332479842\kappa_{14} - 221423667689.22618)10^{11} t^2$$
$$+ (2860512\kappa_{14} - 206772.28792)10^{15} t - 112145856 \cdot 10^{14} \ /$$
$$\left(468.4598789 t^4 + 46854.40101 t^3 + 1087.040656 t^2 + 24572.832 t \right).$$

The coefficients of $g_{\kappa_{14}}(t)$ are of different scales ranging from 10^{23} to 10^3. We use Monte Carlo integration to approximate the average number of positive steady states when $(\kappa_{12}, \kappa_{14})$ belong to the following box:
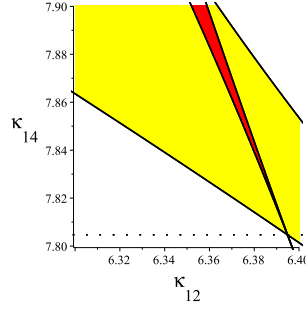
$$B = [6.3, 6.4] \times [7.8, 7.9].$$

FIGURE 4. In the red, yellow and white regions, network (30) with the choices in (32) has five, three and one positive steady states respectively, as obtained using CAD.

For any sample size from 10 to 10^{11}, we obtain $\widehat{I}_N = \hat{e}_N = 0$. It is clear from Figure 4 that $1 < \hat{r}(B)$, and hence different from zero. At this point it is unclear to us whether the problem arises because the minimum sample size is larger than 10^{11}, or due to numerical errors arising from the different orders of the coefficients of $g_{\kappa_{14}}$ and limited machine number sizes.

For comparison, instead of the values in (32), we let all the parameters other than κ_{12} and κ_{14} be equal to 1. Then for the box $B = [0,1] \times [0,1]$, we obtain $\widehat{I}_N = 1$ with $N = 10^5$ in 0.05 seconds and two digits of precision. Here $N = 10$ is already acceptable as sample size, and the function $g_{\kappa_{14}}(t)$ does not have coefficients of different order:

$$g_{\kappa_{14}}(t) = \frac{-4t^5 + (2\kappa_{14}+4)t^4 - 4\kappa_{14}t^3 + (3\kappa_{14}-4)t^2 + (\kappa_{14}-3)t+1}{2t^4 + 4t^3 + 3t^2 + t}.$$

For this example, the parameter region with five steady states is likely too small to be detectable using our approach. For instance, for the boxes $B = [0,100]^{14}$ or $[0.1,1]^{14}$, we obtain that $\hat{r}(B) = 1$ with 3 digits of significance, meaning the regions with three or five steady states are small. However, we illustrated with the box in (31) that parameter points yielding multistationarity can be easily found in a box of interest, even with 14 free parameters.

3. PROOF OF THEOREM 1.1

In this section we prove Theorem 1.1. The argument follows standard approaches to establish Kac-Rice formulas, see for example [3], Chapter 3 for the one-dimensional case, and Chapter 6 for the multivariate case.

Given a set $A \subseteq \mathbb{R}^n$, we let A°, $\bar{A}$ and ∂A denote the interior, closure and boundary set of A respectively, with respect to the Euclidean topology. Given a sequence of sets $\{S_r\}_{r \in \mathbb{N}}$ such that $S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots$ and $\cup_{r \in \mathbb{N}} S_r = S$, then we use the notation $\lim_{r \to +\infty} S_r = \cup_{r \in \mathbb{N}} S_r = S$.

Let π_1 and π_2 be the projections of $\mathbb{R}^m$ onto the first n components and the last $m - n$ components, respectively (such that $\bar{\kappa} = \pi_2(\kappa)$). For $t \in A$, it holds

(33) $\qquad f_\kappa(t) = 0 \quad$ if and only if $\quad \pi_1(\kappa) = g_{\pi_2(\kappa)}(t)$.

Before proving Theorem 1.1, we establish a series of lemmas.

Lemma 3.1. *With the notation and assumptions of Theorem 1.1 the following holds:*

(i) *For every $\bar{\kappa} \in \widetilde{B}$ outside a Zariski closed set (relative to $\widetilde{B}$) of measure zero $\widetilde{P}'$, there exists a measure zero set $A(\bar{\kappa}) \subseteq A$ such that ρ_i is continuous at $g_{\bar{\kappa},i}(t)$ for all $t \notin A(\bar{\kappa})$ and for all $i = 1, \ldots, n$.*

(ii) *With $\widetilde{P} \subseteq \widetilde{B}$ as in Theorem 1.1(ii), for all $\bar{\kappa} \notin \widetilde{P}$ there exists a Zariski closed set (relative to $\pi_1(B)$) of measure zero $Q_{\bar{\kappa}} \subseteq \pi_1(B)$, such that if $u \in \pi_1(B) \setminus Q_{\bar{\kappa}}$, then the solution set to $g_{\bar{\kappa}}(t) = u$ in A consists of a finite number of simple points in the interior of A. In particular, $\det(J_{g_{\bar{\kappa}}}(t)) \neq 0$ at all solution points.*

Proof. To show (i), we use that ρ_i is a continuous function except maybe in a finite number of points $\xi_1, \ldots, \xi_\ell$. For a fixed $j \in \{1, \ldots, \ell\}$ and $\bar{\kappa} \in \widetilde{B}$, as $g_{\bar{\kappa}}$ is rational, the solutions to $g_{\bar{\kappa}}(t) = \xi_j$ in t form an algebraic variety given by n equations and n variables. As for almost all $\bar{\kappa}$, $g_{\bar{\kappa}}$ is not constant (by assumption Theorem 1.1(ii)), for almost all $\bar{\kappa}$ this algebraic variety has codimension at least 1 in $\mathbb{R}^n$. Hence (i) holds.

We turn now to (ii). By assumption Theorem 1.1(ii), for all $\bar{\kappa} \notin \widetilde{P}$, the polynomial $p_{\bar{\kappa}}(t)$ given by the numerator of $\det(J_{g_{\bar{\kappa}}}(t))$ is not identically zero. For a fixed $\bar{\kappa} \notin \widetilde{P}$, $p_{\bar{\kappa}}(t) = 0$ defines a real algebraic variety Y of codimension at least 1 in $\mathbb{R}^n$, and hence is not Zariski dense in A. Define $Q_{\bar{\kappa}}$ as the image of $g_{\bar{\kappa}}$ restricted to Y, which is not Zariski dense in $\pi_1(B)$. For any u outside $Q_{\bar{\kappa}}$, any solution t^* to $g_{\bar{\kappa}}(t) = u$ satisfies $p_{\bar{\kappa}}(t^*) \neq 0$, and hence is simple and isolated. In this case there is a finite number of solutions, as $g_{\bar{\kappa}}$ is rational.

As A is a box, the boundary of A can be decomposed into the union of subsets of coordinate hyperplanes. The restriction of $g_{\bar{\kappa}}$ to a coordinate hyperplane yields a rational function in (at most) $n - 1$ variables and n equations, and the image is not Zariski dense in $\pi_1(B)$. Now augment $Q_{\bar{\kappa}}$ to include the image of $g_{\bar{\kappa}}$ restricted to every coordinate hyperplane describing the boundary of A. Then the equation $g_{\bar{\kappa}}(t) = u$ has no boundary solutions if $u \notin Q_{\bar{\kappa}}$. Finally, redefine $Q_{\bar{\kappa}}$ to be its Zariski closure, which by construction is a real algebraic variety different from $\mathbb{R}^n$ and hence has measure zero. This concludes the proof of (ii). $\qquad\square$

Lemma 3.2. *Let $f_\kappa(t)$, B, and $A \subseteq \mathbb{R}^n$ as in Theorem 1.1. Assume that for any compact box $S \subseteq A$, the Kac-Rice formula (7) holds for the domain S. Then the Kac-Rice formula (7) holds for A.*

Proof. Write $A = I_1 \times \cdots \times I_n$. For $j = 1, \ldots, \ell$, let $a_j, b_j \in \mathbb{R} \cup \{\pm\infty\}$ be the lower and upper limits of I_j, and for $i \geq 1$ define $S_{i,j} := [c_j, d_j]$ with

$$
c_j = \begin{cases} a_j, & \text{if } a_j \in I_j, \\ a_j + \frac{1}{i}, & \text{if } a_j \in \mathbb{R}, \ a_j \notin I_j, \\ -i, & \text{if } a_j = -\infty, \end{cases} \qquad d_j = \begin{cases} b_j, & \text{if } b_j \in I_j, \\ b_j - \frac{1}{i}, & \text{if } b_j \in \mathbb{R}, \ b_j \notin I_j, \\ i, & \text{if } b_j = +\infty. \end{cases}
$$

Then $S_{i,j}$ is compact, $S_{i,j} \subseteq I_j$ for all i large enough, and $\lim_{i \to +\infty} S_{i,j} = I_j$. We define the compact set $S_i := S_{i,1} \times \cdots \times S_{i,n}$ such that $S_i \subseteq A$ for i large enough, and $\lim_{i \to +\infty} S_i = A$.

By hypothesis, formula (7) holds for the compact set S_i for i large enough. Let $\psi(t)$ be the function integrated on the right-hand side of (7) such that

$$
\mathbb{E}\big(\#\big(f_\kappa^{-1}(0) \cap S_i\big)\big) = \int_{S_i} \psi(t)dt \qquad \text{for all } i \in \mathbb{N}.
$$

By the monotone convergence theorem, we have

$$
\int_A \psi(t)dt = \lim_{i \to +\infty} \int_{S_i} \psi(t)dt \qquad \text{and} \qquad \mathbb{E}\big(\#\big(f_\kappa^{-1}(u) \cap A\big)\big) = \lim_{i \to +\infty} \mathbb{E}\big(\#\big(f_\kappa^{-1}(u) \cap S_i\big)\big),
$$

and hence

$$
\mathbb{E}\big(\#\big(f_\kappa^{-1}(u) \cap A\big)\big) = \int_A \psi(t)dt
$$

as desired. $\qquad\square$

Given $x \in \mathbb{R}^n$, $B(x, \delta)$ denotes the open ball centered at x with radius $\delta > 0$. Let V_δ denote the volume of such s ball.

Lemma 3.3. *Let $T \subseteq \mathbb{R}^n$ be an open set, $f \colon T \to \mathbb{R}$ a continuous function, and consider an increasing sequence of open sets $\{C_r\}_{r \in \mathbb{N}}$ such that $\bigcup_{r \in \mathbb{N}} C_r = T$.*
Then, for $y \notin \partial T$, it holds

$$\lim_{r \to +\infty} \int_{C_r} \frac{\chi_{B(x,1/r)}(y)}{V_{1/r}} f(x)\, dx = \begin{cases} f(y) & \text{if } y \in T, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. If $y \notin T$, then as $y \notin \partial T$, we have $y \in (\mathbb{R}^n \setminus T)^{\circ}$ and $y \notin B(x, \frac{1}{r})$ for all $x \in T$ and r large enough. Hence $\lim_{r \to +\infty} \int_{C_r} \frac{\chi_{B(x,1/r)}(y)}{V_{1/r}} f(x) dx = 0$.

Assume now $y \in T$. Observe that $y \in B(x, \frac{1}{r})$ if and only if $x \in B(y, \frac{1}{r})$. Let $m_0 > 0$ such that $y \in C_m$ for all $m \geq m_0$, and let $r_0 > 0$ such that $B(y, \frac{1}{r}) \subseteq C_{m_0}$ for all $r \geq r_0$. Then for $r > \max(m_0, r_0)$, it holds

$$(34) \qquad \int_{C_r} \frac{\chi_{B(x,1/r)}(y)}{V_{1/r}} f(x) dx = \int_{C_r} \frac{\chi_{B(y,1/r)}(x)}{V_{1/r}} f(x) dx = \int_{B(y,1/r)} \frac{f(x)}{V_{1/r}} dx.$$

Since f is continuous at y, for a fixed $\epsilon > 0$, there exists $\frac{1}{\max(m_0, r_0)} > \eta_{\epsilon} > 0$ such that for all $x \in B(y, \eta_{\epsilon})$ it holds $|f(x) - f(y)| < \epsilon$. Thus, for all $\epsilon > 0$, there exists η_{ϵ} such that for all n with $\frac{1}{r} < \eta_{\epsilon}$ it holds

$$\left| \left(\int_{B(y,1/r)} \frac{f(x)}{V_{1/r}} dx - f(y) \right) \right| < \epsilon.$$

This implies that the limit in r of (34) exists and equals $f(y)$. This concludes the proof. $\qquad \square$

Lemma 3.4. *Let $A \subseteq \mathbb{R}^n$ be a compact set and $f \colon \mathbb{R}^n \to \mathbb{R}^n$ a function with continuous first-order partial derivatives in A. For $u \in \mathbb{R}^n$, assume the solutions to $f(t) = u$ in A are isolated, satisfy $\det\big(J_f(t)\big) \neq 0$, and belong to A°. Then there exists $\mu > 0$ such that for all $\varepsilon < \mu$ and $v \in B(u, \mu)$ it holds*

$$\#\big(f^{-1}(v) \cap A\big) = \int_A \frac{\chi_{B(v,\varepsilon)}\big(f(t)\big)}{V_{\varepsilon}} |\det\big(J_f(t)\big)|\, dt.$$

Proof. For $w \in \mathbb{R}^n$, let S_w be the solution set of the equation $f(t) = w$ in A, that is $S_w = f^{-1}(w) \cap A$. In particular, S_u is finite as the solutions are isolated, and A is compact. Furthermore, as $S_u \subseteq A^{\circ}$ by assumption, there exists $\delta > 0$ such that for each $s \in S_u$, $B(s, \delta) \subseteq A^{\circ}$ and $B(s, \delta) \cap S_u = \{s\}$.

As the Jacobian of $f(t)$ does not vanish on the points in S_u, by the Inverse Mapping Theorem, δ can be chosen such that f is a diffeomorphism from each $B(s, \delta)$ to $f(B(s, \delta))$. Choose $\nu > 0$ such that $B(u, \nu) \subseteq \bigcap_{s \in S_u} f(B(s, \delta))$. It follows that for all $v \in B(u, \nu)$, the solutions to $f(t) = v$ also are isolated and the Jacobian of $f(t)$ does not vanish. By choosing ν smaller if necessary, we further guarantee that all solutions belong to A° as well. Hence $\#S_u = \#S_v$ and each set $B(s, \delta)$ contains one element of S_v.

Let $\mu = \frac{\nu}{3}$ and consider $\varepsilon < \mu$. For $v \in B(u, \mu)$, we have $B(v, \varepsilon) \subseteq B(u, \nu)$ and hence f is a diffeomorphism from each connected component of $f^{-1}(B(v, \varepsilon))$ to $B(v, \varepsilon)$. We denote these connected components by U_s, for $s \in S_v$ (which are Borel sets). A change of variables [40, Theorem 7.26] gives that

$$V_{\varepsilon} = \int_{B(v,\varepsilon)} \chi_{B(v,\varepsilon)}(x) dx = \int_{U_s} \chi_{B(v,\varepsilon)}\big(f(t)\big) |\det\big(J_f(t)\big)|\, dt.$$

Since $\chi_{B(v,\varepsilon)}\big(f(t)\big) = 0$ if $t \in A \setminus \cup_{s \in S_v} U_s$, and the union of U_s for $s \in S_v$ is disjoint, by summing over $s \in S_v$ we obtain

$$
\begin{aligned}
\#S_v &= \sum_{s \in S_v} \int_{B(v,\varepsilon)} \frac{\chi_{B(v,\varepsilon)}(x)}{V_\varepsilon} dx \\
&= \sum_{s \in S_v} \int_{U_s} \frac{\chi_{B(v,\varepsilon)}(f(t))}{V_\varepsilon} \,|\det\big(J_f(t)\big)|\, dt + \int_{A \setminus \cup_{s \in S_v} U_s} \frac{\chi_{B(v,\varepsilon)}(f(t))}{V_\varepsilon} \,|\det\big(J_f(t)\big)|\, dt \\
&= \int_A \frac{\chi_{B(v,\varepsilon)}(f(t))}{V_\varepsilon} \,|\det\big(J_f(t)\big)|\, dt.
\end{aligned}
$$

Hence, $\#\big(f^{-1}(v) \cap A\big)$ agrees with the integral above, for all $v \in B(u,\mu)$, and $\varepsilon < \mu$. This concludes the proof. $\qquad\square$

Proof of Theorem 1.1. We are now ready to proof Theorem 1.1. As the assumptions of Theorem 1.1 hold for a compact box $S \subseteq A$ if they hold for A, it is enough to prove Theorem 1.1 when A is compact by Lemma 3.2. Hence assume A is compact.

To show that the image of $f_\kappa(t)$ has positive measure, note that the first n columns of the Jacobian of $f_\kappa(t)$ with respect to κ form the diagonal matrix with entries $h_i(\bar\kappa, t)$, $i = 1, \ldots, n$, which by assumption do not vanish on $\widetilde{B} \times A$.

Using (33), the expected value of $\#(f_\kappa^{-1}(0) \cap A)$ is given by

$$
\begin{aligned}
\mathbb{E}(\#(f_\kappa^{-1}(0) \cap A)) &= \int_B \#(f_\kappa^{-1}(0) \cap A) \prod_{i=1}^m \rho_i(\kappa_i) d\kappa_1 \ldots d\kappa_m \\
&= \int_B \#\big(g_{\pi_2(\kappa)}^{-1}(\pi_1(\kappa)) \cap A\big) \prod_{i=1}^m \rho_i(\kappa_i) d\kappa_1 \ldots d\kappa_m =: (\star).
\end{aligned}
$$

By Lemma 3.1(ii), outside a measure zero set of B, the equation $f_\kappa(t) = 0$ has a finite number of solutions. As $f_\kappa(t)$ is polynomial, there is an upper bound M on the number of complex solutions depending only on the exponents of the monomials, and not on the coefficients (this follows for instance from Bernstein-Kushnirenko Theorem on the number of solutions in the torus $(\mathbb{C}^*)^n$ [14, Theorem 5.4]). Hence the integrand in $(\star)$, which is non-negative, is bounded above by M and the integral $(\star)$ is finite.

Let $\rho^{(1)}$, $\rho^{(2)}$ be the densities on $\pi_1(B)$ and $\pi_2(B)$, respectively. By Tonelli's theorem, the integral over B can be found iteratively over $\pi_2(B) = \widetilde{B}$ and $\pi_1(B)$ with variables y, z respectively:

$$
(\star) = \int_{\widetilde{B}} \left(\int_{\pi_1(B)} \#\big(g_y^{-1}(z) \cap A\big) \rho^{(1)}(z) dz \right) \rho^{(2)}(y) dy.
$$

By Lemma 3.1(i), for every y outside a measure zero set $\widetilde{P}'$, there exists a measure zero set $A(y)$ such that ρ_i is continuous at $g_i(y,t)$ for all $t \notin A(y)$ and for all $i = 1, \ldots, n$. Let $\widetilde{P}$ be the set in Lemma 3.1(ii), consider the (relative open) set $\widetilde{B}' := \widetilde{B} \setminus (\widetilde{P} \cup \widetilde{P}')$, and fix $y \in \widetilde{B}'$. We focus on the inner integral:

$$
(\star\star) := \int_{\pi_1(B)} \#\big(g_y^{-1}(z) \cap A\big) \rho^{(1)}(z) dz.
$$

As the denominator of g_y does not vanish on $\widetilde{B} \times A$, g_y has continuous first-order partial derivatives in $t \in A$ for all $y \in \widetilde{B}$. By Lemma 3.1(ii), there exists a measure zero set Q_y such that for $z \in \pi_1(B) \setminus Q_y$, the solutions to $g_y(t) = z$ are isolated, belong to A°, and the Jacobian does not vanish. Lemma 3.4 applies to g_y, with $u = z \in \pi_1(B) \setminus Q_y$. Hence, for

every such pair (z, y) and $\varepsilon > 0$ small enough, we have

$$(35) \qquad \#(g_y^{-1}(z) \cap A) = \int_A \frac{\chi_{B(z,\varepsilon)}(g_y(t))}{V_\varepsilon} |\det(J_{g_y}(t))| \, dt.$$

Let $B_1' := \pi_1(B) \setminus (Q_y \cup \partial\pi_1(B))$, which is an open set, and consider the integral $(\star\star)$ over B_1' instead. As $Q_y \cup \partial\pi_1(B)$ has measure zero, the value of the integral is the same. For every $r \in \mathbb{N}$, let $D_r \subseteq B_1'$ be the set of all points $z \in B_1'$ for which (35) holds for $\varepsilon \le \frac{1}{r}$, and let C_r be its interior. Clearly $C_1 \subseteq C_2 \subseteq \dots$ defines an increasing sequence of open sets. Their union is

$$\lim_{r \to +\infty} C_r = B_1'$$

by Lemma 3.4, as if $z \in B_1'$, there exists $r_0 > 0$ and $\delta > 0$ such that $B(z, \delta) \subseteq D_{r_0}$, hence $z \in C_{r_0}$. The above discussion, (35) and monotone convergence theorem gives that

$$(\star\star) = \lim_{r \to +\infty} \int_{C_r} \#(g_y^{-1}(z) \cap A) \, \rho^{(1)}(z) dz$$

$$= \lim_{r \to +\infty} \int_{C_r} \left(\int_A \frac{\chi_{B(z,1/r)}(g_y(t))}{V_{1/r}} |\det(J_{g_y}(t))| dt \right) \rho^{(1)}(z) dz.$$

We now claim that the following equalities, derived from interchanging limits, hold:

$$(36) \qquad (\star\star) = \lim_{r \to +\infty} \int_A \left(\int_{C_r} \frac{\chi_{B(z,1/r)}(g_y(t))}{V_{1/r}} \rho^{(1)}(z) dz \right) |\det(J_{g_y}(t))| \, dt$$

$$(37) \qquad = \int_A \lim_{r \to +\infty} \left(\int_{C_r} \frac{\chi_{B(z,1/r)}(g_y(t))}{V_{1/r}} \rho^{(1)}(z) dz \right) |\det(J_{g_y}(t))| dt$$

$$(38) \qquad = \int_A \rho^{(1)}(g_y(t)) |\det(J_{g_y}(t))| \, dt.$$

Let us show that (36)-(38) hold. Recall $y \in \widetilde{B}'$ is fixed. The interchange of limits in (36) follows again from Tonelli's theorem, as the integrand is a non-negative measurable function. For (37), we need to show that we can interchange the limit in r and the integral in A. To this end, we show that Lebesgue's dominated convergence theorem applies. Consider the sequence

$$\alpha_r := \int_{C_r} \frac{\chi_{B(z,1/r)}(g_y(t))}{V_{1/r}} \rho^{(1)}(z) dz.$$

Let $0 < M_1 < +\infty$ be an upper bound of $\rho^{(1)}$. We have

$$\alpha_r \le \int_{\pi_1(B)} \frac{\chi_{B(z,1/r)}(g_y(t))}{V_{1/r}} \rho^{(1)}(z) dz = \int_{\pi_1(B)} \frac{\chi_{B(g_y(t),1/r)}(z)}{V_{1/r}} \rho^{(1)}(z) dz$$

$$= \int_{B(g_y(t),1/r) \cap \pi_1(B)} \frac{1}{V_{1/r}} \rho^{(1)}(z) dz \le M_1.$$

As $|\det(J_{g_y}(t))|$ is integrable, the sequence $\alpha_r |\det(J_{g_y}(t))|$ is dominated by an integrable function. This gives (37), as long as the limit exists, but this follows from the proof of (38).

For (38), we apply Lemma 3.3 with $f = \rho^{(1)}$, point $g_y(t)$ with fixed t and set $T = B_1'$. Note that it is enough to prove that (38) holds for the integral over $A \setminus A(y)$ instead of over A, as $A(y)$ has measure zero. To this end, it is enough to verify that the hypotheses of Lemma 3.3 hold. First, the choices made above imply that for $t \notin A(y)$, $\rho^{(1)}$ is continuous at $g_y(t)$, and $g_y(t) \notin \partial B_1' \subseteq \partial\pi_1(B) \cup Q_y$. The last condition follows from Lemma 3.4.

Finally, using the expression found for $(\star\star)$, the definition of $\bar{\rho}$ in the statement, and that $\widetilde{P} \cup \widetilde{P}'$ has measure zero, we obtain

$$(\star) = \int_{\widetilde{B}} \left(\int_A |\det\left(J_{g_y}(t)\right)| \, \rho^{(1)}(g_y(t)) \, dt \right) \rho^{(2)}(y) dy.$$

All that is left is to justify that the integrals in A and $\widetilde{B}$ can be interchanged, but, again, this follows from Tonelli's theorem, as the integrand is non-negative and measurable. This concludes the proof of Theorem 1.1.

$\square$

References

[1] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer-Verlag New York, 1st edition, 2007.

[2] A. Auffinger, G. B. Arous, and Černý J. Random matrices and complexity of spin glasses. *Comm. Pure Appl. Math.*, 66(2):165–201, 2013.

[3] J.-M. Azaïs and M. Wschebor. *Level sets and extrema of random processes and fields*. Wiley, 1st edition, 2009.

[4] C. P. Bagowski, J. Besser, C. R. Frey, and J. E. Ferrell. The JNK cascade as a biochemical switch in mammalian cells: Ultrasensitive and all-or-none responses. *Curr. Biol.*, 13(4):315–320, 2003.

[5] S. Basu, A. Lerario, E. Lundberg, and C. Peterson. Random fields and the enumerative geometry of lines on real and complex hypersurfaces. *Math. Ann.*, 374(3):1773–1810, 2019.

[6] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in real algebraic geometry*. Springer-Verlag Berlin Heidelberg, 2nd edition, 2006.

[7] F. Bihan, A. Dickenstein, and M. Giaroli. Lower bounds for positive roots and regions of multistationarity in chemical reaction networks. *J. Algebra*, 542:367–411, 2018.

[8] V. Chickarmane, C. Troein, U. A. Nuber, H. M. Sauro, and C. Peterson. Transcriptional dynamics of the embryonic stem cell switch. *PLOS Comput. Biol.*, 2(9):e123, 2006.

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

[10] C. Conradi and D. Flockerzi. Switching in mass action networks based on linear inequalities. *SIAM J. Appl. Dyn. Syst.*, 11(1):110–134, 2012.

[11] C. Conradi, A. Iosif, and T. Kahle. Multistationarity in the space of total concentrations for systems that admit a monomial parametrization. *Bull. Math. Biol.*, 81(10):4174–4209, 2019.

[12] C. Conradi and M. Mincheva. Catalytic constants enable the emergence of bistability in dual phosphorylation. *J. R. S. Interface*, 11(95), 2014.

[13] S. Corvez and F. Rouillier. Using computer algebra tools to classify serial manipulators. In *4th International Workshop, ADG 2002, Hagenberg Castle, Austria, September 4-6, 2002, Revised Papers*, pages 31–43. Springer-Verlag Berlin Heidelberg, 2004.

[14] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer-Verlag New York, 2nd edition, 2005.

[15] P. Donnell, M. Banaji, A. Marginean, and C. Pantea. CoNtRol: an open source framework for the analysis of chemical reaction networks. *Bioinformatics*, 30(11):1633–1634, 2014.

[16] A. Edelman and E. Kostlan. How many zeros of a random polynomial are real? *Bull. Amer. Math. Soc.*, 32:1–37, 1995.

[17] P. Ellison, M. Feinberg, H. Ji, and D. Knight. Chemical reaction network toolbox, version 2.2. Available online at `http://www.crnt.osu.edu/CRNTWin`, 2012.

[18] M. England, R. Bradford, and J. H. Davenport. Improving the use of equational constraints in cylindrical algebraic decomposition. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, ISSAC '15, pages 165–172, New York, NY, USA, 2015. Association for Computing Machinery.

[19] E. Evans. The expected number of zeros of a random system of p-adic polynomials. *Electron. Commun. Probab.*, 11(29):278–290, 2006.

[20] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors–II. Multiple steady states for networks of deficiency one. *Chem. Eng. Sci.*, 43(1):1–25, 1988.

[21] M. Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Ration. Mech. Anal.*, 132(4):311–370, 1995.

[22] M. Feinberg. *Foundations of Chemical Reaction Network Theory.* Springer, Cham, 2019.

[23] E. Feliu, N. Kaihnsa, T. de Wolff, and O. Yürük. The kinetic space of multistationarity in dual phosphorylation. *J. Dyn. Differ. Equ.*, To appear, 2020.

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

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

[26] J. Gerhard, D. Jeffrey, and G. Moroz. A package for solving parametric polynomial systems. *ACM Commun. Comput. Algebra*, 43(3/4):61–72, 2010.

[27] J. Gunawardena. Chemical reaction network theory for in-silico biologists. Available online at `http://vcp.med.harvard.edu/papers/crnt.pdf`, 2003.

[28] T. Hahn. `CUBA`– a library for multidimensional numerical integration. *Comput. Phys. Commun.*, 168(2):78 – 95, 2005.

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

[30] M. Kac. On the average number of real roots of a random algebraic equation. *Bull. Amer. Math. Soc.*, 49(4):314–320, 1943.

[31] V. B. Kothamachu, E. Feliu, L. Cardelli, and O. S. Soyer. Unlimited multistability and boolean logic in microbial signalling. *J. R. Soc. Interface*, 12(108):20150234, 2015.

[32] D. Lazard and F. Rouillier. Solving parametric polynomial systems. *J. Symb. Comput.*, 42(6):636–667, 2007.

[33] S. Müller, E. Feliu, G. Regensburger, C. Conradi, A. Shiu, and A. Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16(1):69–97, 2016.

[34] K. M. Nam, B. M. Gyori, S. V. Amethyst, D. J. Bates, and J. Gunawardena. Robustness and parameter geography in post-translational modification systems. *PLOS Comput. Biol.*, 16(5):e1007573, 2020.

[35] L. Nicolaescu. On the Kac-Rice formula. Available online at `https://www.researchgate.net/publication/267039543_On_the_Kac-Rice_formula`, 2014.

[36] A. B. Owen. Monte Carlo theory, methods and examples. `http://statweb.stanford.edu/~owen/mc/`, 2013.

[37] M. Pérez Millán, A. Dickenstein, A. Shiu, and C. Conradi. Chemical reaction systems with toric steady states. *Bull. Math. Biol.*, 74(5):1027–1065, 2012.

[38] A. J. Rainal. Origin of Rice's formula. *IEEE T. Inform. Theory*, 34(6):1383–1387, 1988.

[39] S. O. Rice. Mathematical analysis of random noise. *Bell Syst. Tech. J.*, 23(3):282–332, 1944.

[40] W. Rudin. *Real and Complex analysis.* McGraw-Hill Inc., 3rd edition, 1987.

[41] A. H. Sadeghimanesh. Polynomial superlevel set representation of the multistationarity region of chemical reaction networks. *arXiv:2003.07764*, 2020.

[42] A. H. Sadeghimanesh and E. Feliu. The multistationarity structure of networks with intermediates and a binomial core network. *Bull. Math. Biol.*, 2019.

[43] A. H. Sadeghimanesh and E. Feliu. MCKR project, version 1.0.0. Available online at `https://doi.org/10.5281/zenodo.4026954`, 2020.

[44] T. Shiraishi, S. Matsuyama, and H. Kitano. Large-scale analysis of network bistability for human cancers. *PLOS Comput. Biol.*, 6(7):e1000851, 2010.

[45] J. E. Taylor, J. R. Loftus, and R. J. Tibshirani. Inference in adaptive regression via the Kac-Rice formula. *Ann. Statist.*, 44(2):743–770, 2016.

[46] L. Wang and E. D. Sontag. On the number of steady states in a multiple futile cycle. *J. Math. Biol.*, 57(1):29–52, 2008.

[47] N. D. Ylvisaker. The expected number of zeros of a stationary gaussian process. *Ann. Math. Statist.*, 36(3):1043–1046, 1965.