PREVALENCE OF MULTISTATIONARITY AND ABSOLUTE CONCENTRATION ROBUSTNESS IN REACTION NETWORKS

BADAL JOSHI, NIDHI KAIHNSA, TUNG D. NGUYEN, AND ANNE SHIU

ABSTRACT. For reaction networks arising in systems biology, the capacity for two or more steady states, that is, multistationarity, is an important property that underlies biochemical switches. Another property receiving much attention recently is absolute concentration robustness (ACR), which means that some species concentration is the same at all positive steady states. In this work, we investigate the prevalence of each property while paying close attention to when the properties occur together. Specifically, we consider a stochastic block framework for generating random networks, and prove edge-probability thresholds at which – with high probability – multistationarity appears and ACR becomes rare. We also show that the small window in which both properties occur only appears in networks with many species. Taken together, our results confirm that, in random reversible networks, ACR and multistationarity together, or even ACR on its own, is highly atypical. Our proofs rely on two prior results, one pertaining to the prevalence of networks with deficiency zero, and the other "lifting" multistationarity from small networks to larger ones.

Keywords: Multistationarity, absolute concentration robustness, random graph, stochastic block model, threshold function, reaction network.

1. INTRODUCTION

In biochemical reaction networks, multistationarity is often a desirable phenomenon, as it is associated with biochemical switches, cellular signaling, and decision-making [32]. A network is multistationary when there are two or more compatible positive steady states; "compatible" means that the steady states have the same conserved quantities such as total mass. Which reaction networks are multistationary? This question has a long history, and many results have been established (see the survey [25]).

Another significant property exhibited by some biochemical reaction networks is absolute concentration robustness (ACR), which refers to when a steady-state species concentration is maintained even when initial conditions are changed. The concept of ACR was first popularized by Shinar and Feinberg in 2010 [29] and has since attracted much interest both from the mathematical standpoint [2, 23, 28, 22] and in applications [13, 27].

While each of these two properties has been studied in isolation, their relationship is not well understood. Nonetheless, multistationarity and ACR can be viewed as opposite behaviors, as multiple steady states cannot be in general position if ACR is present. Indeed, known examples of networks having ACR (for instance, those in [29]) typically are non-multistationary.

Accordingly, the driving motivation of this article is to explore the relationship between multistationarity and ACR and to investigate the prevalence of networks with either property. However, it is generally challenging to assess multistationarity and ACR [25, 28]. Therefore, following the approach of Anderson and Nguyen [6, 7], we instead investigate the *prevalence* of these properties in

randomly generated reaction networks. Specifically, we prove asymptotic results on the probability that such a network has either property, as the number of species goes to infinity.

A summary of our results appears in Table 1, which pertains to when reaction networks are randomly generated by a certain *stochastic block model* in which the expected numbers of reactions of each "type" are roughly of the same magnitude. Here, the type refers to which forms of complexes – $0, X_i, 2X_i, X_i + X_j$ – appear in the reaction (notice that we restrict our attention to *at-most-bimolecular* networks, which encompass most reaction networks arising in biochemistry). We prove edge-probability thresholds for the resulting reaction networks to be nondegenerately multistationary or to preclude ACR (Theorem 4.9), and a restatement of this result in terms of the expected number of edges (that is, reactions) is in Table 1.

Expected number of reactions	Multistationary?	ACR?
Asymptotically greater than 1, but less than $n^{2/3}$	No	Yes
Asymptotically greater than n, but less than $\frac{2}{17}n(\log(n) - c(n))$, for some $c(n) \to \infty$	Yes	Yes
Greater than $n(\log(n) + c(n))$, for some $c(n) \to \infty$	Yes	No

TABLE 1. For random reaction networks, with n species, generated by a certain stochastic block model, this table lists ranges for the expected number of reactions, and whether – with high probability – such a network is multistationary or has ACR. For further details, see Theorem 4.9 and Remark 4.12.

We see from Table 1 that the window for having both multistationarity and ACR is relatively small: when the expected number of edges is asymptotically between n and $\frac{2}{17}n\log(n)$. In fact, in this window, which only appears when there are thousands of species (Remark 4.15), the ACR species either appears by itself without interacting with other species (in essence, the network decouples) or it appears as a catalyst (Remark 4.23). We therefore expect that reaction networks with both multistationarity and ACR, in which the ACR species interacts nontrivially with other species, are rare and may require special structures or constructions.

Our proofs rely crucially on two prior results. The first, due to Anderson and Nguyen, pertains to the prevalence of certain networks that are known to preclude multistationarity, namely, networks with deficiency zero (specifically, the result asserts that "sparse" reaction networks are likely to have deficiency zero) [6, 7]. The second result concerns "lifting" multistationarity from small networks to larger ones [8, 10, 24], which we use to show the high probability of multistationarity and the absence of ACR in "dense" reaction networks.

This article is structured as follows: Section 2 introduces reaction networks, multistationarity, and ACR. Section 3 contains our results on the prevalence of multistationarity and ACR in random reaction networks via edge-probability thresholds. In Section 4, we introduce the type-homogeneous stochastic block model, which we use to generate random reaction networks, and compute explicitly the thresholds for multistationarity and ACR. We end with a discussion in Section 5.

2. BACKGROUND

Here, we recall the basic setup and definitions involving reaction networks (Section 2.1), the dynamical systems they generate (Section 2.2), and absolute concentration robustness (Section 2.3). (A more detailed exposition can be found in [18].) We also discuss how multistationarity and ACR are affected when adding new reactions to an existing network (Section 2.4).

2.1. **Reaction networks.** A *reaction network* G is a directed graph in which the vertices are non-negative linear combinations of *species* $X_1, X_2, \ldots, X_n$. As is standard in reaction network theory,

we refer to each vertex as a *complex*, and we denote the i-th complex by $y_i = y_{i1}X_1 + y_{i2}X_2 + \cdots + y_{in}X_n$ or by $y_i = (y_{i1}, y_{i2}, \ldots, y_{in})$ (where $y_{ij} \in \mathbb{Z}_{\geq 0}$).

Edges of G represent the possible changes in the abundances of the species, and are referred to as *reactions*. It is standard to represent a reaction (y_i, y_i') by the notation $y_i \to y_i'$. In such a reaction, y_i is the *reactant complex*, and y_i' is the *product complex*. A species X_k is a *catalyst-only species* of a reaction (y_i, y_i') if the stoichiometric coefficient of X_k is the same in the product and reactant (that is, $y_{ik} = y_{ik}'$). Finally, a reaction network G' is a *subnetwork* of a network G if the sets of species, complexes, and reactions of G' are subsets of the respective sets of G.

In examples, it is often convenient to write species as $A, B, C, \ldots$ rather than $X_1, X_2, X_3, \ldots$. Additionally, we typically depict a reaction network by its set of reactions, in which case the sets of species and complexes are implied.

Example 2.1. The reaction network $\{A + B \to 2B, \ B \to A\}$ has 2 species (namely, A and B), 4 complexes ($A + B$, $2B$, B, A), and 2 reactions. $\triangle$

A reaction network is *reversible* if every edge in the graph is bidirected.

Example 2.2. The reaction network $\{A \leftrightarrows B + C, \ 0 \leftrightarrows A, \ 0 \leftrightarrows B, \ C \leftrightarrows 2C\}$ is reversible. $\triangle$

This article focuses on *at-most-bimolecular* reaction networks (or, for short, *bimolecular*), which means that every complex y_i of the network satisfies $y_{i1} + y_{i2} + \cdots + y_{in} \leq 2$ (where n is the number of species). Equivalently, each complex has the form 0, X_i, $X_i + X_j$, or $2X_i$ (where X_i and X_j are species). The reaction networks in Examples 2.1–2.2 are bimolecular.

2.2. **Dynamical system arising from a network.** Under the assumption of mass-action kinetics, each reaction network G defines a parametrized family of systems of ordinary differential equations (ODEs), as follows. Let r denote the number of reactions of G. We write the i-th reaction as $y_i \to y_i'$ and assign a positive *rate constant* $\kappa_i \in \mathbb{R}_{>0}$ to the corresponding reaction.

The *mass-action system*, denoted by (G, κ), where $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_r)$, is the dynamical system arising from the following ODEs:

$$(1) \qquad \frac{dx}{dt} \;=\; \sum_{i=1}^{r} \kappa_i x^{y_i}(y_i' - y_i) \;=:\; f_\kappa(x) \ ,$$

where $x_i(t)$ denotes the concentration of the i-th species at time t and $x^{y_i} := \prod_{j=1}^{n} x_j^{y_{ij}}$. The right-hand side of the ODEs (1) consists of polynomials $f_{\kappa,i}$, for $i = 1, 2, \ldots, n$ (where n is the number of species). For simplicity, we often write f_i instead of $f_{\kappa,i}$.

The *stoichiometric subspace* of G, which we denote by S, is the linear subspace of $\mathbb{R}^n$ spanned by all reaction vectors $y_i' - y_i$ (for $i = 1, 2, \ldots, r$). When $\dim(S) = n$, we say that G is *full dimensional*. Observe that the vector field of the mass-action ODEs (1) lies in S (more precisely, the vector of ODE right-hand sides is always in S). Hence, a forward-time solution $\{x(t) \mid t \geq 0\}$ of (1), with initial condition $x(0) \in \mathbb{R}_{>0}^n$, remains in the following *stoichiometric compatibility class* [18]:

$$P_{x(0)} \;:=\; (x(0) + S) \cap \mathbb{R}_{\geq 0}^n \ .$$

Example 2.1 (continued). The network $\{A + B \overset{\kappa_1}{\to} 2B, \ B \overset{\kappa_2}{\to} A\}$ generates the following mass-action ODEs (1):

$$\frac{dx_1}{dt} \;=\; -\kappa_1 x_1 x_2 + \kappa_2 x_2$$
$$\frac{dx_2}{dt} \;=\; \kappa_1 x_1 x_2 - \kappa_2 x_2.$$

Moreover, it has a one-dimensional stoichiometric subspace (spanned by the vector $(1, -1)^\top$). $\triangle$

A *steady state* of a mass-action system is a nonnegative concentration vector $x^* \in \mathbb{R}^n_{\geq 0}$ at which the right-hand side of the ODEs (1) vanishes: $f_\kappa(x^*) = 0$. Our primary interest in this work is in *positive* steady states $x^* \in \mathbb{R}^n_{>0}$. Also, a steady state x^* is *nondegenerate* if $\text{Im}(df_\kappa(x^*)|_S) = S$, where $df_\kappa(x^*)$ is the Jacobian matrix of f_κ evaluated at x^*, and S is the stoichiometric subspace.

We consider multistationarity at two levels: systems and networks. A mass-action system (G, κ) is *multistationary* (respectively, *nondegenerately multistationary*) if there exists some stoichiometric compatibility class having more than one positive steady state (respectively, nondegenerate positive steady state). A reaction network G is *multistationary* if there exist positive rate constants κ such that (G, κ) is multistationary. Similarly, a network G can be *nondegenerately multistationary*.

Example 2.3 (Multistationary system). Consider the mass-action system generated by:

$$(2) \qquad \left\{ A \overset{1}{\underset{1}{\rightleftharpoons}} B + C, \quad 0 \overset{1}{\underset{6}{\rightleftharpoons}} A, \quad 0 \overset{1}{\underset{27}{\rightleftharpoons}} B, \quad C \overset{1}{\underset{8}{\rightleftharpoons}} 2C \right\}.$$
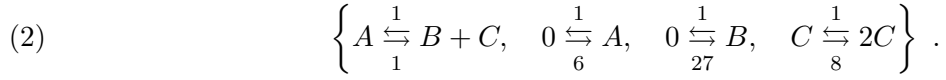
It is straightforward to check (or compute) that there are exactly three positive steady states, namely, $(13, 20, 1), (18, 15, 2), (21, 12, 3)$, and all three are nondegenerate. $\triangle$

2.3. Deficiency and absolute concentration robustness.

The *deficiency* of a reaction network G is $\delta = v - \ell - \dim(S)$, where v is the number of vertices (or complexes) of G, ℓ is the number of connected components of G, and S is the stoichiometric subspace. This invariant is central to many classical results pertaining to mass-action systems (1) [3, 4, 5, 17, 20, 21], including a structural criterion for absolute concentration robustness (ACR) [29], the topic we turn to next.

ACR, like multistationarity, is analyzed at the level of systems and also networks.

Definition 2.4 (ACR). Let X_i be a species of a reaction network G with r reactions.
  (1) For a fixed vector of positive rate constants $\kappa \in \mathbb{R}^r_{>0}$, the mass-action system (G, κ) has *absolute concentration robustness* (ACR) in X_i if (G, κ) has a positive steady state and in every positive steady state $x \in \mathbb{R}^n_{>0}$ of the system, the value of x_i (the concentration of X_i) is the same. This value of x_i is the *ACR-value* of X_i.
  (2) The reaction network G has *unconditional ACR* in species X_i if the mass-action system (G, κ) has ACR in X_i for all $\kappa \in \mathbb{R}^r_{>0}$.

When G has unconditional ACR in X_i, the property of ACR in X_i holds across all rate constants, but the ACR-value can (and typically does) change with rate constants, as in the next example.
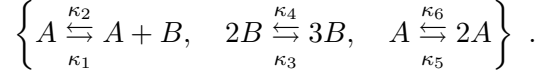
Example 2.1 (continued). We return to the following network: $\{A + B \xrightarrow{\kappa_1} 2B, \ B \xrightarrow{\kappa_2} A\}$. This network is a classical example of a network with ACR [29]. Indeed, at all positive steady states, the concentration of species A is κ_2/κ_1, and hence the network has unconditional ACR in A.
$\triangle$

Remark 2.5 (ACR and reversible networks). In Definition 2.4, ACR requires the existence of a positive steady state. This requirement is not included in some definitions of ACR in the literature. However, in this work, we focus on the reversible networks, which guarantees the existence of positive steady states (this result is due to Deng *et al.* [14] and Boros [12]). Hence, our results are valid with or without the requirement of positive steady states.

In the literature, reaction networks with ACR are typically not multistationary. Nonetheless, a network can both have ACR (in some species) and be multistationary. A simple example can be constructed by joining two networks, with disjoint species sets, where one network has ACR and the other is multistationary. A less trivial example can be generated by having the ACR species

participate as an enzyme – more precisely, as a catalyst-only species – in the multistationary network. We illustrate this in the following example.

Example 2.6 (A network with multistationarity and ACR)**.** Consider the following network, in which A is a catalyst-only species in the first two reactions:

$$\left\{ A \underset{\kappa_1}{\overset{\kappa_2}{\rightleftharpoons}} A + B, \quad 2B \underset{\kappa_3}{\overset{\kappa_4}{\rightleftharpoons}} 3B, \quad A \underset{\kappa_5}{\overset{\kappa_6}{\rightleftharpoons}} 2A \right\} .$$

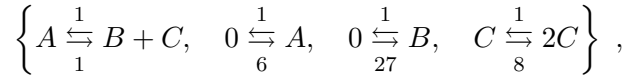This network, which we call G, generates the following mass-action ODEs (1):

$$\begin{aligned} \frac{dx_1}{dt} &= \kappa_5 x_1 - \kappa_6 x_1^2 \\ \frac{dx_2}{dt} &= \kappa_1 x_1 - \kappa_2 x_1 x_2 + \kappa_3 x_2^2 - \kappa_4 x_2^3. \end{aligned}$$

One can check directly that G has unconditional ACR in species A with ACR-value κ_5/κ_6. Moreover, for reaction rates $(\kappa_1, \kappa_2, \ldots, \kappa_6) = (\frac{1}{512}, \frac{1}{16}, 1, 1, 2, 1)$, we obtain exactly 3 positive steady states, with the following approximate values: $(2, 0.050987)$, $(2, 0.0890928)$, and $(2, 0.85992)$.   △

It is not straightforward to find non-trivial examples of reaction networks with multistationarity and unconditional ACR where the network cannot be decomposed into individual pieces, each with only one of the two properties. Nonetheless, such networks do exist and deserve an independent in-depth study. We will report on several families of such networks and their operating principles in future work.

In this work, we are interested in asymptotic results (as the size of the network grows) on the prevalence of multistationarity and ACR. An important tool we use for proving thresholds for these properties (or the lack thereof) is the network in the following example.

Example 2.3 **(continued).** Consider again the following mass-action system:

$$\left\{ A \underset{1}{\overset{1}{\rightleftharpoons}} B + C, \quad 0 \underset{6}{\overset{1}{\rightleftharpoons}} A, \quad 0 \underset{27}{\overset{1}{\rightleftharpoons}} B, \quad C \underset{8}{\overset{1}{\rightleftharpoons}} 2C \right\} ,$$

which we saw has three positive steady states: $(13, 20, 1), (18, 15, 2), (21, 12, 3)$. By inspection, this system has no ACR (in any species) and hence the network does not have unconditional ACR.   △

We end this subsection by recalling what is known about multistationarity and ACR for networks with deficiency 0. In the following result, part (1) follows from the deficiency-zero theorem [20] and part (2) is immediate from a recent result of Joshi and Craciun [23, Theorem 6.1].

Lemma 2.7. *If G is a reaction network that has deficiency 0, then:*

(1) G is not multistationary, and

(2) if G contains an inflow or outflow reaction (that is, a reaction of the form $0 \rightarrow X_i$ or $0 \leftarrow X_i$, for some species X_i), then G has unconditional ACR (in some species).

2.4. **Monotonicity of multistationarity and non-ACR with respect to adding reactions.** This subsection pertains to how multistationarity and ACR are affected as we add new reactions to a reaction network. The following proposition essentially follows from recent results on "lifting" multistationarity from smaller networks to larger ones [8, 10, 24].

Lemma 2.8 (Lifting multistationarity or non-ACR)**.** *Let G be a full-dimensional network, and let G' be a network obtained by adding to G a reaction that involves no new species (new complexes are allowed).*

(i) *If G is nondegenerately multistationary, then so is G'.*

(ii) *If there exists a vector of positive rate constants κ^* such that (G, κ^*) is nondegenerately multistationary and also does not have ACR (in any species), then the network G' does not have unconditional ACR in any species.*

Proof. Part (i) follows directly from [24, Theorem 3.1].

For part (ii), suppose that there exists κ^* such that (G, κ^*) does not have ACR (in any species) and also has nondegenerate, positive steady states $q_1, q_2, \ldots, q_m$, where $m \geq 2$. We denote each steady state by $q_i = (q_{i,1}, \ldots, q_{i,n}) \in \mathbb{R}^n_{\geq 0}$ (for $i = 1, \ldots, m$), where n is the number of species of G.

Let ϵ denote the rate constant of the reaction added to G to obtain G'. From the proof of [24, Theorem 3.1], there exists $\epsilon_0 > 0$ such that if $0 < \epsilon < \epsilon_0$, then $(G', (\kappa; \epsilon))$ has nondegenerate, positive steady states $q_1(\epsilon), q_2(\epsilon), \ldots, q_m(\epsilon)$ such that $\lim_{\epsilon \to 0^+} q_i(\epsilon) = q_i$ (for all i).

Next, for arbitrary $\ell \in \{1, 2, \ldots, n\}$, consider the species X_ℓ. As (G, κ^*) does not have ACR in X_ℓ, there exist steady states q_i and q_j at which the corresponding concentrations of X_ℓ differ (that is, $q_{i,\ell} \neq q_{j,\ell}$). So, as $\lim_{\epsilon \to 0^+} q_i(\epsilon) = q_i$ and $\lim_{\epsilon \to 0^+} q_j(\epsilon) = q_j$, there exists $\epsilon_\ell > 0$ (with $\epsilon_\ell < \epsilon_0$) such that if $0 < \epsilon < \epsilon_\ell$, then $|q_{i,\ell}(\epsilon) - q_{j,\ell}(\epsilon)| > 0$ and hence $(G', (\kappa^*; \epsilon))$ does not have ACR in X_ℓ.

Finally, we pick ϵ such that $0 < \epsilon < \min_\ell \epsilon_\ell$. By construction, the system $(G', (\kappa^*; \epsilon))$ does not have ACR in any species. Hence, G' does not have unconditional ACR. □

Remark 2.9. Lemma 2.8 implies that, given full dimensionality, nondegenerate multistationarity is a *monotone increasing* property (with respect to adding new edges/reactions).

3. Multistationarity and ACR in random reaction networks

In this section, we follow the approach in [6, 7] in which reaction networks are generated using a random-graph framework (Section 3.1). In Section 3.2, we prove the existence of thresholds for the presence or absence of nondegenerate multistationarity and unconditional ACR. These thresholds are with respect to increasing *graph density*, that is, the fraction of reactions present – among all possible reactions.

In this section, we use the following standard notation. For sequences of positive numbers $\{a_n\}$ and $\{b_n\}$, we write $a_n \ll b_n$ (or $b_n \gg a_n$) if

$$\lim_{n \to \infty} \frac{a_n}{b_n} = 0 \ ;$$

and we write $a_n \sim b_n$ if

$$\lim_{n \to \infty} \frac{a_n}{b_n} = c \ ,$$

for some positive constant c. Also, a sequence of events $\{A_n\}$ occurs *with high probability (w.h.p.)* if $\lim_{n \to \infty} \mathbb{P}(A_n) = 1$.

3.1. **Random reaction networks.** Consider the class of bimolecular reaction networks on n species $X_1, X_2, \ldots, X_n$. The set of all possible complexes is then

$$(3) \qquad V_n \ = \ \{0\} \ \cup \ \{X_i \mid 1 \leq i \leq n\} \ \cup \ \{2X_i \mid 1 \leq i \leq n\} \ \cup \ \{X_i + X_j \mid 1 \leq i < j \leq n\}.$$

The cardinality of V_n is therefore given by

$$|V_n| \ = \ 1 + n + n + \binom{n}{2} \ = \ \frac{n^2 + 3n + 2}{2} \ .$$

Definition 3.1 (Edge probabilities). Let n be a positive integer.

(1) Consider two distinct vertices $u, v \in V_n$. An *edge probability function* for the unordered pair $e = (u, v)$ is a non-decreasing function, $\phi_e(p_n)$, in a single parameter $p_n \in [0, 1]$.

(2) A *choice of edge probabilities* is a collection of edge probability functions, $\phi_e(p_n)$, one for each unordered pair $e = (u, v)$ of vertices in V_n.

Definition 3.2 (Random graph $G(V_n, p_n)$)**.** Fix a positive integer n, some $p_n \in [0, 1]$, and a choice of edge probabilities $\{\phi_e(p_n)\}$. We generate random (undirected) graphs, which we denote by $G(V_n, p_n)$, as follows:

- the vertex set is V_n, and
- the probability that there is an edge between two vertices $u, v \in V_n$ is given by the corresponding edge probability function (where $e = (u, v)$):

$$\mathbb{P}(e \text{ is an edge of } G(V_n, p_n)) = \phi_e(p_n).$$

Definition 3.3 (Random network G_n)**.** Each random graph $G(V_n, p_n)$ (generated by some choice of edge probabilities) defines a random reaction network, which we denote by G_n, consisting of reversible reactions, as follows:

- The set of species of G_n is $\{X_1, X_2, \ldots, X_n\}$.
- The (reversible) reactions of G_n correspond to the edges of $G(V_n, p_n)$.

Recall from Section 2.2 that G_n is full-dimensional if its dimension is n.
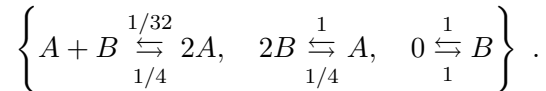
Remark 3.4. It is possible that some species of G_n appears in no complexes, especially when $G(V_n, p_n)$ is sparse (e.g., the network shown later in Figure 1). Such G_n are not full-dimensional.

3.2. **Thresholds for multistationarity and ACR.** In this subsection, we show that for the random reaction networks defined in the prior subsection, there exist thresholds for the presence or absence of nondegenerate multistationarity and unconditional ACR (Theorem 3.7). Subsequently, we discuss the challenges of computing such thresholds, and then describe a strategy for proving upper bounds on the thresholds (see Corollary 3.14).

The following definition is useful in the proof of Theorem 3.7 and also later in Proposition 3.13.

Definition 3.5 (S_n^*)**.** For $n \in \mathbb{Z}_{\geq 1}$, let S_n^* denote the set of all full-dimensional bimolecular networks G with exactly n species for which there exists a vector of rate constants κ such that (G, κ) is nondegenerately multistationary and also does not have ACR in any species.

Remark 3.6 (S_n^* is nonempty for $n \geq 2$)**.** The set S_1^* is empty [26], but for all $n \geq 2$, S_n^* is nonempty. This is shown for $n \geq 3$ in Proposition 3.13, and S_2^* contains the following network:

$$\left\{ A + B \underset{1/4}{\overset{1/32}{\leftrightarrows}} 2A, \quad 2B \underset{1/4}{\overset{1}{\leftrightarrows}} A, \quad 0 \underset{1}{\overset{1}{\leftrightarrows}} B \right\}.$$

Indeed, the indicated rate constants generate a system with 3 nondegenerate positive steady states – with approximate values $(0.419694, 1.11107)$, $(2.65005, 2.3128)$, and $(216.681, 27.5757)$ – and so this system is nondegenerately multistationarity and also does not have ACR.

Theorem 3.7 (Thresholds for full-dimensionality, multistationarity, and non-ACR)**.** *Consider the setup for generating random reaction networks G_n, described in Section 3.1, for some choice of edge probabilities. Then there exist threshold functions ("thresholds", for short) $0 < r_0(n) \leq r_1(n) \leq r_2(n) \ll 1$, such that for any $\{p_n\}_{n \geq 1}$:*

- *(0) If $p_n \gg r_0(n)$, then G_n is full-dimensional w.h.p.*
- *(1) If $p_n \gg r_1(n)$, then G_n is full-dimensional and nondegenerately multistationary w.h.p.*
- *(2) If $p_n \gg r_2(n)$, then G_n is full-dimensional, is nondegenerately multistationary, and does **not** have unconditional ACR (in any species) w.h.p.*

Proof. Being a full-dimensional network is a monotone increasing property with respect to adding reactions (with no new species). This fact, combined with a well-known result from the theory of threshold functions [11], proves part (0).

From Remark 2.9, the property (for full-dimensional networks) of being nondegenerately multistationary is monotonically increasing with respect to adding reactions (with no new species). Exploiting again the theory of threshold functions [11], we obtain part (1).

Finally, Lemma 2.8(ii) and the theory of threshold functions together imply that there exists a threshold function $r_2(n)$ for G_n to contain a subnetwork $H \in S_n^*$. This implies part (2).   □

Theorem 3.7 implies that when a random network is sufficiently dense, it is multistationary w.h.p. (after a threshold $r_1(n)$) and also lacks unconditional ACR (after a threshold $r_2(n)$). However, computing these thresholds is generally difficult, because it is challenging to determine whether a large reaction network is multistationary and whether it precludes ACR. In fact, while there are sufficient conditions for ACR, such as the Shinar-Feinberg criterion [29], there are no easy-to-check necessary conditions for ACR (for general networks) [28, Section 2].

Nevertheless, there is a fruitful strategy for establishing upper bounds on the thresholds $r_1(n)$ and $r_2(n)$, which we describe in detail in the remainder of this subsection. The underlying idea comes from the fact (stated earlier in Lemma 2.8) that multistationarity can sometimes be *lifted* from a small subnetwork to the whole network. Therefore, in lieu of determining when multistationarity of the entire network emerges (as edge probabilities increase), we instead investigate when a small multistationary subnetwork emerges. The choice of edge probabilities dictates which such subnetworks emerge first. For the edge probabilities we consider in the next section, we focus on a particular multistationary subnetwork, as follows.

Definition 3.8 (Sets $S_{M,n}$ of multistationary motifs)**.** For $n \in \mathbb{Z}_{\geq 1}$, let $S_{M,n}$ denote the set of all networks of the following form:

$$(4) \qquad \{X_i \leftrightarrows X_j + X_k, \quad 0 \leftrightarrows X_i, \quad 0 \leftrightarrows X_j, \quad X_k \leftrightarrows 2X_k\} ,$$

where i, j, k are distinct indices with $1 \leq i, j, k \leq n$. Each network (4) is a *multistationary motif*.

Remark 3.9. Recall from Example 2.3 and Example 2.3 (continued) that each multistationary motif (4) is full-dimensional (3-dimensional) and nondegenerately multistationary, and does not have unconditional ACR (in any species).

Our next aim is to show (in Proposition 3.13 below) how to join a multistationary motif to another network (which we call a "lifting component") so that the resulting network again is full-dimensional and nondegenerately multistationary, and lacks unconditional ACR.

Definition 3.10 (Sets $S_{L,k}$ of lifting components)**.** For $k \in \mathbb{Z}_{\geq 1}$, let $S_{L,k}$ be the set of all reversible reaction networks for which the associated graph is a tree on k vertices (that is, complexes), where each of the vertices has the form X_i (for $i \in \mathbb{Z}_{\geq 1}$). Every network in $S_{L,k}$ is a *lifting component*.

Example 3.11. An example of a network (lifting component) in $S_{L,3}$ is $\{X_2 \leftrightarrows X_3 \leftrightarrows X_4\}$.   △

Next, we describe a set of networks obtained by joining a multistationary motif (4), which has dimension 3, to a lifting component of dimension $n - 3$, so that the result is full-dimensional.

Definition 3.12 (Sets $S_{J,n}$ of joined multistationary motifs and lifting components)**.** For $n \geq 3$, let $S_{J,n}$ be the set of all reaction networks whose reactions can be written as the union of a network $G_M \in S_{M,n}$ and a network $G_L \in S_{L,n-2}$, such that G_M and G_L have exactly one species in common.

An example of such a joined network is shown later (see Figure 3 and Example 4.14).

Proposition 3.13 (Properties of $S_{J,n}$)**.** *For $n \geq 3$, the following hold:*

   (1) $S_{J,n} \subseteq S_n^*$. Consequently, every network $H \in S_{J,n}$ is full-dimensional and nondegenerately multistationary, and does not have unconditional ACR (in any species).

(2) *If some $H \in S_{J,n}$ is a subnetwork of a network G with n species, then G is full-dimensional and nondegenerately multistationary, and does not have unconditional ACR (in any species).*

Proof. We first prove part (1). Let $H \in S_{J,n}$. By definition, there exist $G_M \in S_{M,n}$ and $G_L \in S_{L,n-2}$, with exactly one species in common, such that the reactions of H are a union of those in G_M and G_L. Relabel the species of H, if needed, so that G_M is the following network:

$$\{X_1 \leftrightarrows X_2 + X_3, \quad 0 \leftrightarrows X_1, \quad 0 \leftrightarrows X_2, \quad X_3 \leftrightarrows 2X_3\} \ ,$$
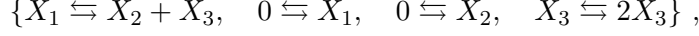
and also that the species of G_L are $X_\ell, X_4, X_5, \ldots, X_n$, for some $\ell \in \{1, 2, 3\}$.

It is straightforward to check that H is full-dimensional (that is, has dimension n). Thus, to show that $H \in S_n^*$, it suffices to show that there exists a vector of rate constants κ such that (H, κ) is nondegenerately multistationary and does not have ACR in any species. Accordingly, we define κ as follows. First, we choose the rate constants for reactions in G_M as in (2) in Example 2.3, so that G_M has three nondegenerate, positive steady states: $(13, 20, 1), (18, 15, 2), (21, 12, 3)$. Next, fix all rate constants for reactions in G_L to be 1. Using the fact that G_L is a spanning tree, a simple computation shows that the positive steady states are $(x_\ell, x_4, x_5, \ldots, x_n) = (c, c, c, \ldots, c)$, where c is any positive real number, and these steady states are all nondegenerate.

We consider first the case when the common species is $X_\ell = X_3$. We claim that, in this case, the following are nondegenerate steady states of (H, κ):

$$(5) \qquad\qquad (13, 20, 1, 1, \ldots, 1), \quad (18, 15, 2, 2, \ldots, 2), \quad (21, 12, 3, 3, \ldots, 3) \ .$$

To see this, let $\dot{x}_i = f_i$ for $i = 1, 2, 3$, and $\dot{x}_i = g_i$ for $i = 3, 4, \ldots, n$, denote the ODEs for G_M and G_L, respectively, with rate constants as defined above. Hence, $(13, 20, 1), (18, 15, 2), (21, 12, 3)$ satisfy $f_1 = f_2 = f_3 = 0$, and $(1, 1 \ldots, 1), (2, 2, \ldots, 2), (3, 3, \ldots, 3)$ satisfy $g_3 = g_4 = \cdots = g_n = 0$. Next, the ODEs of (H, κ) are as follows:

$$\begin{aligned} \dot{x}_i &= f_i & \text{for } i = 1, 2 \\ \dot{x}_3 &= f_3 + g_3 \\ \dot{x}_j &= g_j & \text{for } j = 4, \ldots, n. \end{aligned}$$

Hence, the vectors in (5) indeed are steady states of (H, κ).

To show that the steady states (5) are nondegenerate, we must show that the Jacobian matrix of (H, κ), when evaluated at each of these steady states, is nonsingular (recall that H is full-dimensional). Since G_L has mass conservation among the species $X_3, X_4, \ldots, X_n$, we have $g_3 + g_4 + \cdots + g_n = 0$. Adding rows $4, 5, \ldots, n$ to row 3 of the Jacobian matrix yields a triangular block matrix $\begin{bmatrix} A & 0 \\ B & C \end{bmatrix}$, where A is the Jacobian matrix of G_M and C is obtained from the Jacobian matrix of G_L by setting x_3 to 0. As both A and C are nonsingular, when evaluated at any of the positive steady states of its corresponding system, we conclude that the Jacobian matrix of (H, κ), when evaluated at any of the steady states (5), is nonsingular, as desired.

By inspection of the steady states (5), we see that there is no ACR in any species. As for the remaining cases, when X_ℓ (the common species of G_M and G_L) is X_1 or X_2, the argument is very similar to what is shown above and so the result holds in those cases, too.

Finally, part (2) follows directly from part (1) and Lemma 2.8. $\qquad\square$

Corollary 3.14 (Bound on threshold r_2)**.** *Consider the setup for generating random reaction networks G_n, described in Section 3.1, for some choice of edge probabilities. Let $r_2(n)$ be a threshold as defined in Theorem 3.7; and let $r_2'(n)$ be a threshold for G_n to contain, as a subnetwork, some $H \in S_{J,n}$ (more precisely, $r_2'(n)$ is a function such that when $p_n \gg r_2'(n)$, w.h.p. G_n has a subnetwork H*

such that $H \in S_{J,n}$). If $p_n \gg r_2'(n)$, then G_n is full-dimensional, is nondegenerately multistationary, and does not have unconditional ACR (in any species) w.h.p. Consequently, $\limsup \frac{r_2(n)}{r_2'(n)} < \infty$.

Proof. This result follows directly from Proposition 3.13 and Theorem 3.7. □

Remark 3.15. In the next section, we show that for a certain choice of edge probabilities, the thresholds for full-dimensionality, multistationarity, and containing (as a subnetwork) a multistationary motif in $S_{M,n}$ coincide. That is, in this scenario, $r_0(n) = r_1(n) = r_1'(n)$, where r_1' denotes the threshold for G_n to contain a multistationary motif. (In fact, in this setting, once we pass the threshold for full-dimensionality, a multistationary motif emerges and furthermore its multistationarity can be lifted; see Theorem 4.9.) Intuitively, the reason a motif in $S_{M,n}$ is among the first small multistationary subnetworks to emerge is that it is fairly "generic": it contains only a pair of reversible reactions and some flow reactions involving those species. Meanwhile, other small multistationary networks, such as the following (from [24]):

$$\{A \leftrightarrows A + B \leftrightarrows 2A, \qquad 0 \leftrightarrows A, \qquad 0 \leftrightarrows B\},$$

may contain specific pathways such as $A \leftrightarrows A + B \leftrightarrows 2A$, where a species (here, A) must appear in all three complexes, and therefore such networks are expected to emerge at higher thresholds.

4. Multistationarity and ACR in type-homogeneous stochastic block model

The prior section considered general random graph models without specifying the edge probabilities. In this subsection, we introduce a specific choice of edge probabilities (Section 4.1) and then compute the resulting thresholds from Theorem 3.7 (see Theorem 4.9 in Section 4.2).

4.1. **A stochastic block model.** One possible choice of edge probabilities comes from the Erdős-Rényi random graph model; here, the edge probabilities are uniform (that is, every edge is equally likely). In this framework, reactions of the form $X_i + X_j \leftrightarrows X_h + X_k$ are overwhelmingly the most prominent [7]. However, this situation is unlikely to occur in biochemistry. Indeed, in applications, one expects to see various types of complexes and reactions, such as inflow and outflow reactions $0 \leftrightarrows X_i$ or association and disassociation reactions $X_i + X_j \leftrightarrows X_k$.

Therefore, we instead consider a model in which *reaction types* are equally represented. This model is a specific case of the *stochastic block models* [19] introduced in [6]. To define this model, we need the following partitions of sets of vertices and edges:

Definition 4.1 (C_i and $E_{i,j}$). Let $n \geq 1$. Consider the following partition of the set of vertices V_n, as in (3), into 3 subsets:

(1) $C_0 = \{0\}$,
(2) $C_1 = \{aX_i \mid 1 \leq i \leq n \ \text{and} \ a \in \{1, 2\}\}$,
(3) $C_2 = \{X_i + X_j \mid 1 \leq i, j \leq n \ \text{and} \ i \neq j\}$.

Let $E_{i,j}$ denote the set of (undirected) edges (u, v) with $u \in C_i$ and $v \in C_j$; in particular:

(1) $E_{0,1} = \{0 \leftrightarrows aX_i \mid 1 \leq i \leq n \ \text{and} \ a \in \{1, 2\}\}$,
(2) $E_{0,2} = \{0 \leftrightarrows X_i + X_j \mid 1 \leq i, j \leq n \ \text{and} \ i \neq j\}$,
(3) $E_{1,1} = \{aX_i \leftrightarrows bX_j \mid 1 \leq i, j \leq n, \ a, b \in \{1, 2\}, \ \text{and} \ (a, i) \neq (b, j)\}$,
(4) $E_{1,2} = \{aX_i \leftrightarrows X_j + X_k \mid 1 \leq i, j, k \leq n, \ a \in \{1, 2\}, \ \text{and} \ j \neq k\}$,
(5) $E_{2,2} = \{X_i + X_j \leftrightarrows X_k + X_h \mid 1 \leq i, j, k, h \leq n \ \text{and} \ i \neq j, k \neq h, \ \text{and} \ (i, j) \neq (k, h) \neq (j, i)\}$.

Two reactions in the same set $E_{i,j}$ have the same *type*.

Remark 4.2. The sets $E_{0,1}, E_{0,2}, E_{1,1}, E_{1,2}, E_{2,2}$ partition the set of all possible edges of a graph with vertex set V_n. Also, $|C_0| = 1$, $|C_1| \sim n$, and $|C_2| \sim n^2$. So, $|E_{i,j}| \sim n^{i+j}$ for $0 \leq i \leq j \leq 2$.

In what follows, we denote the minimum of two numbers a and b as follows:

$$a \wedge b \; := \; \min(a, b) \; .$$

Definition 4.3. The *type-homogeneous stochastic block model* generates random graphs $G(V_n, p_n)$ with vertex set V_n (as described in Definition 3.2) via the following choice of edge probabilities:

$$(6) \qquad\qquad \text{if } e \in E_{i,j}, \text{ then } \mathbb{P}(e \text{ is an edge of } G(V_n, p_n)) \; = \; n^{4-i-j} p_n \wedge 1 \; ,$$

for all $0 \le i \le j \le 2$.

Remark 4.4. In Definition 4.3, for vertices $u \in C_i$ and $v \in C_j$, the edge probability function (as in Definition 3.1) for the edge $e = (u, v)$ is given by $\phi_e(p_n) = n^{4-i-j} p_n \wedge 1$. This edge probability function is readily seen to be non-decreasing in p_n.

Remark 4.5 (Motivation for type-homogeneous models)**.** Recall from Definition 3.3 that each random graph $G(V_n, p_n)$ generates a random reaction network G_n. We choose the edge probabilities (6) to ensure that, in G_n, the expected numbers of edges of each type are of the same magnitude (namely, $\sim n^4 p_n$), whenever possible.

Example 4.6. When $p_n = \frac{1}{n^{3.5}}$, the expected number of reactions of each type in G_n is $\sim \sqrt{n}$. $\quad\triangle$

Example 4.7. When $p_n = \frac{1}{n^{2.9}}$, G_n contains all reactions in $E_{0,1}$ (since $n^{4-0-1} p_n \wedge 1 = 1$), and the expected number of reactions for each of the remaining types is $\sim n^{1.1}$. $\quad\triangle$

Remark 4.8. When $p_n \gg \frac{1}{n^3}$ (for instance, $p_n = \frac{1}{n^{2.9}}$ in Example 4.7), G_n contains all reactions in $E_{0,1}$, including all inflows/outflows $0 \leftrightarrows X_i$, and so G_n is full-dimensional.

In the next subsection, we see that the choice of p_n in Example 4.6 generates networks G_n that are not multistationary w.h.p., while the choice of p_n in Example 4.7 generates networks that are multistationary and lack ACR w.h.p. (see Theorem 4.9).

4.2. **Thresholds for multistationarity and ACR.** For the type-homogeneous stochastic block model, the thresholds for nondegenerate multistationarity and (no) ACR are stated in the following theorem (which is proven later in Section 4.3).

Theorem 4.9 (Type-homogeneous stochastic block model)**.** *Consider the setup for generating random reaction networks G_n, described in Section 3.1, for the edge probabilities given by (6). Then, for any $\{p_n\}_{n \geq 1}$:*

(i) *(Sparse regime) If $\dfrac{1}{n^4} \ll p_n \ll \dfrac{1}{n^{10/3}}$, then w.h.p. G_n has deficiency zero, is not multistationary, and has unconditional ACR (in some species).*

(ii) *(Dense regime, window of co-existence) If there exists a function $c(n)$ such that $c(n) \to \infty$ and $\dfrac{1}{n^3} \ll p_n \leq \dfrac{\frac{2}{17} \log(n) - c(n)}{n^3}$, then w.h.p. G_n is nondegenerately multistationary and has unconditional ACR in some species.*

(iii) *(Dense regime) If there exists a function $c(n)$ such that $c(n) \to \infty$ and $p_n \geq \dfrac{\log(n - 2) + c(n)}{n^2 (n - 2)}$, then w.h.p. G_n is nondegenerately multistationary and does **not** have unconditional ACR (in any species).*

Remark 4.10. When $p_n \ll \frac{1}{n^4}$, the expected number of reactions in G_n is $\ll 1$. Accordingly, we do not consider this interval in Theorem 4.9.

Remark 4.11. Theorem 4.9 immediately yields the following thresholds $r_0(n), r_1(n), r_2(n)$ (as defined in Theorem 3.7) for networks generated by the type-homogeneous stochastic block model:

$$r_0(n) \; = \; r_1(n) \; = \; \frac{1}{n^3} \quad \text{and} \quad r_2(n) \; = \; \frac{\log(n)}{n^3} \;.$$

Remark 4.12 (Thresholds via number of reactions)**.** Theorem 4.9 can be rephrased in terms of the expected number of edges (reactions) instead of edge-probability thresholds, as follows. For random reaction networks G_n (with n species) generated by the type-homogeneous stochastic block model, the following are implied directly by Theorem 4.9:

(i) If the expected number of reactions of each type is $\gg 1$ and $\ll n^{2/3}$, then G_n has deficiency zero and thus is not multistationary w.h.p.

(ii) If the expected number of reactions of each type is $\gg n$ but less than $\frac{2}{17}n(\log(n) - c(n))$ for some $c(n) \to \infty$, then G_n is nondegenerately multistationariy and has unconditional ACR in some species w.h.p.

(iii) If the expected number of reactions of each type is greater than $n(\log(n) + c(n))$ for some $c(n) \to \infty$, then G_n is nondegenerately multistationary and does **not** have unconditional ACR (in any species) w.h.p.
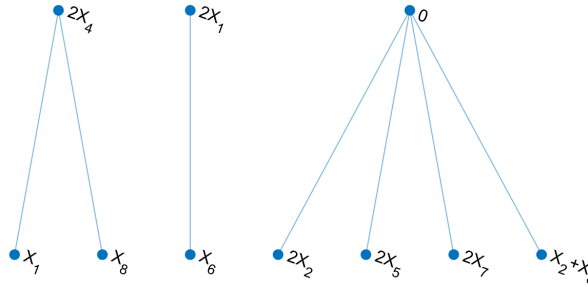

FIGURE 1.   A realization of a random reaction network generated by the type-homogeneous stochastic block model in the sparse regime ($n = 8$ species and $p_n = \frac{0.5}{n^{3.5}}$). Edges represent reversible reactions.

Example 4.13 (Sparse regime)**.** Figure 1 shows a realization of a random reaction network G_n generated by the type-homogeneous stochastic block model with $n = 8$ and $p_n = \frac{0.5}{n^{3.5}}$ (which is in the sparse regime). The following properties of G_n are as expected from Theorem 4.9: The deficiency is $\delta = 10 - 3 - 7 = 0$ and so G_n is not multistationary, and it is easy to check that G_n has unconditional ACR in all species except X_3 (which does not appear in any complex).   △

Example 4.14 (Dense regime)**.** Figure 2 shows a realization of a random reaction network G_n generated by the type-homogeneous stochastic block model with $n = 8$ and $p_n = \frac{2.5}{n^3}$ (which is in the dense regime). Figure 3 depicts a subnetwork of G_n that is a union of a multistationary motif and a lifting component with one species in common. We can now use Proposition 3.13 and Lemma 2.8 to assert that this subnetwork is multistationary and lacks ACR, and then "lift" these properties to G_n. Indeed, this approach underlies our proof of Theorem 4.9 in the dense regime.   △

Remark 4.15 (Window of co-existence)**.** If the number of species satisfies $n < e^{17/2} \approx 4914.8$, then the small window between $\frac{1}{n^3}$ and $\frac{\frac{2}{17}\log(n) - c(n)}{n^3}$, in Theorem 4.9(ii), does not exist. Therefore, in the type-homogeneous stochastic block model, it is unlikely to observe a random network with both multistationarity and ACR, unless it has many species.
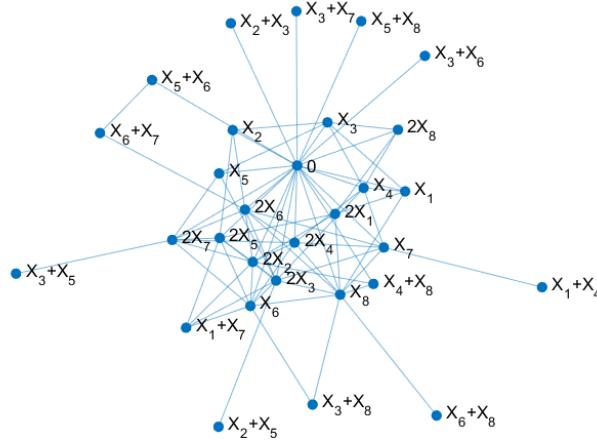
FIGURE 2. A realization of a random reaction network generated by the type-homogeneous stochastic block model in the dense regime ($n = 8$ and $p_n = \frac{2.5}{n^3}$).
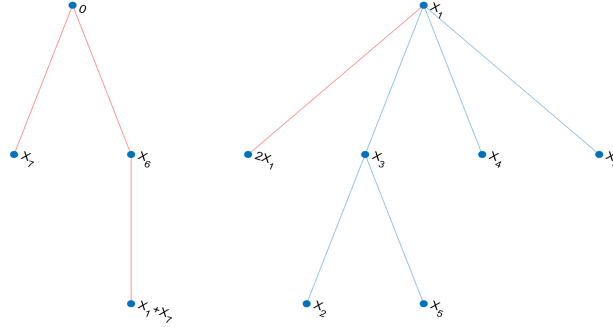

FIGURE 3. A subnetwork of the network in Figure 2, which is a union of a multistationary motif M (red edges) and a lifting component L (blue edges). Notice that M and L share exactly one species, namely, X_1.

4.3. **Proof of Theorem 4.9.** Theorem 4.9 follows directly from Propositions 4.17 and 4.20–4.22 below. This subsection is devoted to proving these propositions, which requires the following lemma.

Lemma 4.16. *For all $n \geq 1$ and $0 \leq x \leq 1$, the following inequality holds:*

$$(1 - x)^n \ \leq \ e^{-nx} \ .$$

Proof. If $x = 1$, the inequality holds. For $0 \leq x < 1$, the result follows directly from the inequality $\log(1 - x) \leq -x$ (which is easy to check) and the fact that the log function is increasing. $\qquad\square$

4.3.1. *Sparse regime.* Our result for the sparse regime is Proposition 4.17 below. Its proof uses Lemma 2.7 and recent results on the prevalence of deficiency-zero networks [6] .

Proposition 4.17 (Sparse regime)**.** *Consider random reaction networks G_n generated by edge probabilities given by (6). If $\frac{1}{n^4} \ll p_n \ll \frac{1}{n^{10/3}}$, then, w.h.p. G_n has deficiency zero, is not multistationary, and has unconditional ACR (in some species).*

Proof. Assume that $\frac{1}{n^4} \ll p_n \ll \frac{1}{n^{10/3}}$. It follows from [6, Theorem 5.1 and Example 10] that, w.h.p., the deficiency of G_n is 0. Thus, w.h.p., the deficiency-zero theorem (Lemma 2.7(1)) applies and so G_n is not multistationary. Additionally, by Lemma 2.7(2), to show that w.h.p. G_n has unconditional ACR in some species, it suffices to show that w.h.p. G_n contains an edge in $E_{0,1}$. The probability that any given edge in $E_{0,1}$ appears in G_n is $n^{4-0-1}p_n = n^3 p_n$, and there are $|E_{0,1}| = 2n$ such edges, so:

$$(7) \qquad \mathbb{P}(G_n \text{ contains an edge in } E_{0,1}) \;=\; 1 - \left(1 - n^3 p_n\right)^{2n} \;\geq\; 1 - e^{-2n^4 p_n} \; .$$

(The inequality in (7) is due to Lemma 4.16.) Finally, using (7) and the assumption $p_n \gg \frac{1}{n^4}$, we obtain that $\lim_{n\to\infty} \mathbb{P}(G_n \text{ contains an edge in } E_{0,1}) = 1$. This concludes the proof. $\qquad\square$

4.3.2. *Dense regime.* The proofs in this subsection make frequent use of the well-known second moment method (for example, see [1]). We summarize this approach in the following lemma.

Lemma 4.18. *Let $\{T_n\}$ be a sequence of non-negative random variables. If $\mathrm{Var}(T_n) \ll (\mathbb{E}T_n)^2$, then $\lim_{n\to\infty} \mathbb{P}(T_n > 0) = 1$.*

Proof. From the second moment method, we have

$$\mathbb{P}(T_n > 0) \;\geq\; 1 - \frac{\mathrm{Var}(T_n)}{(\mathbb{E}T_n)^2}.$$

Taking the limit (as $n \to \infty$) completes the proof. $\qquad\square$

Next, we show that networks generated in the dense regime contain multistationary motifs (4) w.h.p. (see Figures 2 and 3 for an example).

Lemma 4.19. *Consider random reaction networks G_n generated by edge probabilities given by (6). If $p_n \gg \frac{1}{n^3}$, then w.h.p. some multistationary motif in the set $S_{M,n}$ is a subnetwork of G_n.*

Proof. Assume $p_n \gg \frac{1}{n^3}$. Then G_n contains all inflows/outflows $0 \leftrightarrows X_i$ (recall Remark 4.8), so it suffices to show that w.h.p. G_n contains a subnetwork of the following form, for some choice of distinct indices i, j, k:

$$(8) \qquad\qquad \{X_k \leftrightarrows 2X_k, \; X_i \leftrightarrows X_j + X_k\} \; .$$

Consider the reactions in (8). First, $X_k \leftrightarrows 2X_k$ is in $E_{1,1}$, so its edge probability is $n^2 p_n \wedge 1$. Next, for a fixed k, with $1 \leq k \leq n$, there are $(n-1)(n-2)$ reactions of the form $X_i \leftrightarrows X_j + X_k$ with i, j, k distinct. Each such reaction belongs to $E_{1,2}$ and so its edge probability is $n p_n \wedge 1$.

We can reduce to considering only three cases: (1) when $p_n > \frac{1}{n}$ for all n, (2) when $\frac{1}{n^2} \leq p_n \leq \frac{1}{n}$ for all n, and (3) when $p_n < \frac{1}{n^2}$ for all n.

Case 1: $p_n > \frac{1}{n}$ for all n. In this case, $n^2 p_n \wedge 1 = 1 = n p_n \wedge 1$. So, for all $n \geq 3$, G_n contains a subnetwork of the form (8) (in fact, G_n contains all possible such subnetworks).

Case 2: $\frac{1}{n^2} \leq p_n \leq \frac{1}{n}$ for all n. In this case, $n^2 p_n \wedge 1 = 1$, so G_n contains all reactions of the form $X_k \leftrightarrows 2X_k$. Hence, we need only show that w.h.p. G_n contains at least one reaction of the form $X_i \leftrightarrows X_j + X_k$ with i, j, k distinct. The probability of this event, which we call E_n, is as follows:

$$\mathbb{P}(E_n) \;=\; 1 - (1 - n p_n)^{n(n-1)(n-2)} \;\geq\; 1 - e^{-n^2(n-1)(n-2)p_n} \;\to\; 1 \quad (\text{as } n \to \infty),$$

where the inequality is due to Lemma 4.16, and the limit comes from the fact that $p_n \geq \frac{1}{n^2}$.

Case 3: $p_n < \frac{1}{n^2}$ for all n. In this case, the edge probability for each reaction $X_k \leftrightarrows 2X_k$ (respectively, $X_i \leftrightarrows X_j + X_k$) is $n^2 p_n$ (respectively, $n p_n$).

For $1 \le k \le n$, let A_k denote the event that G_n contains a subnetwork of the form (8), where $i \neq j$ and $i, j \neq k$. (The notation $A_{k,n}$ would be better for A_k, but we prefer to avoid excessive subscripts.) It follows that the probability of A_k is:

$$
\text{(9)} \qquad \mathbb{P}(A_k) \;=\; n^2 p_n \left(1 - (1 - n p_n)^{(n-1)(n-2)} \right) \; .
$$

Define the random variable $T_n := \sum_{k=1}^{n} 1_{A_k}$. We wish to show that $\lim_{n \to \infty} \mathbb{P}(T_n > 0) = 1$. By Lemma 4.18, it is enough to prove $\mathrm{Var}(T_n) \ll (\mathbb{E} T_n)^2$. To this end, we first compute $\mathbb{E} T_n$, using (9):

$$
\text{(10)} \qquad \mathbb{E} T_n \;=\; \sum_{k=1}^{n} \mathbb{P}(A_k) \;=\; n^3 p_n \left(1 - (1 - n p_n)^{(n-1)(n-2)} \right) \; .
$$

Next, Lemma 4.16 yields the first inequality here:

$$
\text{(11)} \qquad (1 - n p_n)^{(n-1)(n-2)} \;\le\; e^{-n(n-1)(n-2) p_n} \;\ll\; 1 \; ,
$$

and the second inequality (limit) comes from the assumption that $p_n \gg \frac{1}{n^3}$. Hence, using (10), we obtain $\mathbb{E} T_n \sim n^3 p_n$.

To compute $\mathrm{Var}(T_n)$, we consider the event $A_h \cap A_k$, where $h \neq k$. It is straightforward to check that $A_h \cap A_k$ occurs if and only if G_n contains the reactions $X_h \leftrightarrows 2 X_h$ and $X_k \leftrightarrows 2 X_k$ and also one of the following:

(1) a reaction of the form $X_i \leftrightarrows X_h + X_k$ (for some $i \neq h, k$), or
(2) a reaction of the form $X_i \leftrightarrows X_j + X_k$ (for some $j \neq k, h$ and $i \neq j, k$) and a reaction of the form $X_l \leftrightarrows X_m + X_h$ (for some $m \neq h, k$ and $l \neq m, h$).

A direct computation now yields the following probability:

$$
\begin{aligned}
\text{(12)} \qquad \mathbb{P}(A_h \cap A_k) \;&=\; (n^2 p_n)^2 (1 - (1 - n p_n)^{n-2} + (1 - n p_n)^{n-2} (1 - (1 - n p_n)^{(n-2)^2})^2) \\
&=\; n^4 p_n^2 (1 - 2(1 - n p_n)^{(n-1)(n-2)} + (1 - n p_n)^{(n-2)(2n-3)}) \; .
\end{aligned}
$$

Now we use equations (9), (10), and (12) to compute $\mathrm{Var}(T_n)$, as follows:

$$
\begin{aligned}
\mathrm{Var}(T_n) \;&=\; \mathbb{E}(T_n^2) - (\mathbb{E} T_n)^2 \\
&=\; \sum_{k=1}^{n} \mathbb{P}(A_k) + \sum_{h \neq k} \mathbb{P}(A_h \cap A_k) - (\mathbb{E} T_n)^2 \\
&=\; n^3 p_n \left(1 - (1 - n p_n)^{(n-1)(n-2)} \right) + (n-1) n^5 p_n^2 \left(1 - 2(1 - n p_n)^{(n-1)(n-2)} + (1 - n p_n)^{(n-2)(2n-3)} \right) \\
\text{(13)} \qquad &\quad - n^6 p_n^2 \left(1 - (1 - n p_n)^{(n-1)(n-2)} \right)^2 \; .
\end{aligned}
$$

We claim that $\mathrm{Var}(T_n) \ll n^6 p_n^2$. Indeed, this follows in a straightforward way from (11) and (13), the limit $(1 - n p_n)^{(n-2)(2n-3)} \ll 1$ (which is closely related to (11)), and the assumption $p_n \gg \frac{1}{n^3}$. Finally, having shown $\mathrm{Var}(T_n) \ll n^6 p_n^2$ and $\mathbb{E} T_n \sim n^3 p_n$, we get, as desired, $\mathrm{Var}(T_n) \ll (\mathbb{E} T_n)^2$. $\qquad \square$

Lemma 4.19 allows us to establish the threshold for nondegenerate multistationarity, as follows:

Proposition 4.20 (Multistationarity in dense regime). *Consider random reaction networks G_n generated by edge probabilities in (6). If $p_n \gg \frac{1}{n^3}$, then G_n is nondegenerately multistationary w.h.p.*

Proof. Assume $p_n \gg \frac{1}{n^3}$. By Lemma 4.19, w.h.p., G_n contains (as a subnetwork) a multistationary motif $M \in S_{M,n}$ (which is 3-dimensional and nondegenerately multistationary, as noted in Remark 3.9). Relabeling species, if needed, we may assume that the species of M are X_1, X_2, X_3.

Next, $p_n \gg \frac{1}{n^3}$ implies that, w.h.p., G_n contains all inflow/outflow reactions $0 \leftrightarrows X_i$ (Remark 4.8), and in particular contains the $(n-3)$-dimensional subnetwork consisting of reactions $0 \leftrightarrows X_i$, for all $i = 4, 5, \ldots, n$, which we denote by G'. As M and G' have no species in common, the subnetwork of G formed by the union of their reactions, which we denote by N, is full-dimensional (n-dimensional). It is straightforward to check that N "inherits" nondegenerate multistationarity from M. (The proof is similar to that of Proposition 3.13(1).) Thus, by Lemma 2.8, G_n is nondegenerately multistationary w.h.p. □

Proposition 4.21 (ACR in dense regime). *Consider random reaction networks G_n generated by edge probabilities given by* (6). *If the following inequality holds:*

$$(14) \qquad p_n \;\geq\; \frac{\log(n-2) + c(n)}{n^2(n-2)} \;, \qquad \text{for some } c(n) \to \infty \;,$$

*then w.h.p. G_n does **not** have unconditional ACR (in any species).*

Proof. Assume that inequality (14) holds. By Proposition 3.13(2), it suffices to show that, w.h.p., there exists $H \in S_{J,n}$ (as in Definition 3.12) such that H is a subnetwork of G_n.

First, consider the case when $p_n \geq \frac{1}{n^2}$ for all n. Then, $p_n \gg \frac{1}{n^3}$, so by Lemma 4.19, G_n contains, as a subnetwork, some multistationary motif $M \in S_{M,n}$. Next, we show that G_n also contains all lifting components involving species $\{X_1, X_2, \ldots, X_n\}$. Indeed, reactions of the form $X_\ell \leftrightarrows X_m$ are in $E_{1,1}$ and hence have edge-probability $n^{4-1-1}p_n \wedge 1 = 1$ (since $p_n \geq \frac{1}{n^2}$), and so G_n contains all such reactions. Thus, as desired, w.h.p., G_n contains some $H \in S_{J,n}$ as a subnetwork.

To complete the proof, we need only consider the following case:

$$(15) \qquad \frac{\log(n-2) + c(n)}{n^2(n-2)} \;\leq\; p_n \;<\; \frac{1}{n^2} \;, \qquad \text{for some } c(n) \to \infty \;.$$

Recall from Remark 4.8 that, in this case, G_n contains all reactions of the form $0 \leftrightarrows X_\ell$. Thus, all vertices of the form X_ℓ appear in G_n. For positive integers $i \neq j$, let $G_n^{i,j}$ denote the subgraph of (the underlying graph of) G_n, induced by the following set of vertices of G_n: $\{X_\ell \mid \ell \in \{1, 2, \ldots, n\} \setminus \{i, j\}\}$. Next, for distinct $i, j, k \in \{1, 2, \ldots, n\}$, let $A_{k,i,j}$ denotes the event that (i) $G_n^{i,j}$ is connected and (ii) G_n contains the reactions $X_k \leftrightarrows 2X_k$ and $X_i \leftrightarrows X_j + X_k$.

We claim that, for n sufficiently large, the event $A_{k,i,j}$ implies that G_n contains some $H \in S_{J,n}$ as a subnetwork. To see this, first note that the inequality (15) implies that $p_n \gg \frac{1}{n^3}$ and so, for n large enough, G_n contains all flow reactions $0 \leftrightarrows X_\ell$ (Remark 4.8). So, condition (ii) guarantees a multistationary motif M, for n sufficiently large. Next, condition (i) and the fact that connected graphs have spanning trees yield a "complementary" lifting component L. By joining M and L, we obtain some $H \in S_{J,n}$ as a subnetwork of G_n, as claimed.

Let $T_n = \sum_{k,i,j} 1_{A_{k,i,j}}$ (where the sum is over distinct $i, j, k \in \{1, 2, \ldots, n\}$). To finish the proof, it suffices to show that $\lim_{n \to \infty} \mathbb{P}(T_n > 0) = 1$. By Lemma 4.18, we need only show $\text{Var}(T_n) \ll (\mathbb{E}T_n)^2$.

Again, we start by computing $\mathbb{E}T_n$. Each edge of $G_n^{i,j}$ belongs to $E_{1,1}$, and so its edge-probability (6) is $(n^2 p_n \wedge 1) = n^2 p_n \geq \frac{\log(n-2)+c(n)}{n-2}$ (here we use (15)). It is well known that $\frac{\log n}{n}$ is the edge-probability threshold for connectivity of random graphs with n vertices and uniform edge probabilities [16]. So, for any $i \neq j$ (with $1 \leq i, j \leq n$), we have:

$$(16) \qquad \lim_{n \to \infty} \mathbb{P}(G_n^{i,j} \text{ is connected}) \;=\; 1 \;.$$

We emphasize that the above probability does not depend on the choice of i, j. So, for convenience, we denote $d_n := \mathbb{P}(G_n^{i,j}$ is connected). Next, we compute the following probability using (6):

$$\mathbb{P}(A_{k,i,j}) = (n^2 p_n)(n p_n) d_n = n^3 p_n^2 d_n ,$$

which implies the following:

$$(17) \qquad \mathbb{E}T_n = \sum_{k,i,j} \mathbb{P}(A_{k,i,j}) = n(n-1)(n-2) n^3 p_n^2 d_n = n^4(n-1)(n-2) p_n^2 d_n .$$

Recall that $d_n \sim 1$ (from (16)), so we have $\mathbb{E}T_n \sim n^6 p_n^2$.

Next, we analyze $\mathrm{Var}(T_n)$ by first computing $\mathbb{P}(A_{k_1,i_1,j_1} \cap A_{k_2,i_2,j_2})$ where $(k_1, i_1, j_1) \neq (k_2, i_2, j_2)$. We have the following cases.

Case 1: $k_1 = k_2$. In this case, we have $(i_1, j_1) \neq (i_2, j_2)$, and the number of such pairs of events is $\leq n^5$. Each pair of events occurs precisely when G_n contains the (distinct) reactions $X_{k_1} \leftrightarrows 2X_{k_1}$, $X_{i_1} \leftrightarrows X_{j_1} + X_{k_1}$, $X_{i_2} \leftrightarrows X_{j_2} + X_{k_1}$, and both $G_n^{i_1,j_1}$ and $G_n^{i_2,j_2}$ are connected. Thus, for this case, we use (6) to compute:

$$\mathbb{P}(A_{k_1,i_1,j_1} \cap A_{k_2,i_2,j_2}) \leq n^2 p_n (n p_n)^2 d_n = n^4 p_n^3 d_n.$$

Case 2: $k_1 \neq k_2, i_1 = i_2, j_1 = k_2, j_2 = k_1$. The number of such pairs of events is $n(n-1)(n-2)$. Each pair of events occurs when G_n contains the (distinct) reactions $X_{k_1} \leftrightarrows 2X_{k_1}$, $X_{k_2} \leftrightarrows 2X_{k_2}$, $X_{i_1} \leftrightarrows X_{j_1} + X_{k_1}$, and both $G_n^{i_1,j_1}$ and $G_n^{i_2,j_2}$ are connected. Thus for this case

$$\mathbb{P}(A_{k_1,i_1,j_1} \cap A_{k_2,i_2,j_2}) \leq (n^2 p_n)^2 (n p_n) d_n = n^5 p_n^3 d_n.$$

Case 3: $k_1 \neq k_2$ and either $i_1 = i_2, (j_1, j_2) \neq (k_2, k_1)$ or $i_1 \neq i_2$. We claim that the number of such pairs of events is $n(n-1)(n-2)\big((n-1)^2(n-2) - 1\big)$. Indeed, this number is obtained by taking the total number of pairs in Cases 2 and 3 (i.e., all pairs where $k_1 \neq k_2$) – which is readily seen to be $n(n-1)^3(n-2)^2$ – and then subtracting the number in Case 2 (and simplifying). Next, each pair of events in Case 3 occurs when G_n contains the (distinct) reactions $X_{k_1} \leftrightarrows 2X_{k_1}$, $X_{k_2} \leftrightarrows 2X_{k_2}$, $X_{i_1} \leftrightarrows X_{j_1} + X_{k_1}$, $X_{i_2} \leftrightarrows X_{j_2} + X_{k_2}$, and both $G_n^{i_1,j_1}$ and $G_n^{i_2,j_2}$ are connected. Thus for this case

$$\mathbb{P}(A_{k_1,i_1,j_1} \cap A_{k_2,i_2,j_2}) \leq (n^2 p_n)^2 (n p_n)^2 d_n = n^6 p_n^4 d_n.$$

Next, we use equation (17) and the analysis in Cases 1–3 to bound $\mathrm{Var}(T_n)$:

$$\begin{aligned}
\mathrm{Var}(T_n) &= \mathbb{E}(T_n^2) - (\mathbb{E}T_n)^2 \\
&= \sum_{k,i,j} \mathbb{P}(A_{k,i,j}) + \sum_{(k_1,i_1,j_1) \neq (k_2,i_2,j_2)} \mathbb{P}(A_{k_1,i_1,j_1} \cap A_{k_2,i_2,j_2}) - (\mathbb{E}T_n)^2 \\
&\leq (n^3) n^3 p_n^2 d_n + (n^5) n^4 p_n^3 d_n + (n^3) n^5 p_n^3 d_n \\
&\quad + n(n-1)(n-2)\big((n-1)^2(n-2) - 1\big) n^6 p_n^4 d_n - n^8(n-1)^2(n-2)^2 p_n^4 d_n^2 \\
&\leq n^6 p_n^2 d_n + n^9 p_n^3 d_n + n^8 p_n^3 d_n + n^8(n-1)^2(n-2)^2 p_n^4 d_n - n^8(n-1)^2(n-2)^2 p_n^4 d_n^2 \\
(18) \qquad &\leq (n^3 p_n)^2 d_n + (n^3 p_n)^3 d_n + n^8 p_n^3 d_n + (n^3 p_n)^4 d_n(1 - d_n) .
\end{aligned}$$

As noted earlier, inequality (15) implies that $p_n \gg \frac{1}{n^3}$. Hence, certain terms appearing in (18) have the following asymptotic properties: $(n^3 p_n)^2 \ll (n^3 p_n)^4$ and $n^8 p_n^3 \ll (n^3 p_n)^3 \ll (n^3 p_n)^4$. Recall also that $d_n \sim 1$. Thus, from (18), we have $\mathrm{Var}(T_n) \ll (n^3 p_n)^4 = n^{12} p_n^4$. Finally, we showed earlier that $\mathbb{E}T_n \sim n^6 p_n^2$, so we have $\mathrm{Var}(T_n) \ll (\mathbb{E}T_n)^2$, which concludes the proof.   □

Finally, we address the small "window" between the thresholds in Propositions 4.20–4.21. We showed that a random network in this window is nondegenerately multistationary w.h.p., and next we show that it also has ACR w.h.p.

Proposition 4.22 (ACR in window of dense regime). *Consider random reaction networks G_n generated by edge probabilities given by (6). If $\{p_n\}_{n\geq 1}$ satisfies the following:*

$$(19) \qquad\qquad \frac{1}{n^3} \ll p_n \leq \frac{\frac{2}{17}\log(n) - c(n)}{n^3} \quad \text{for some } c(n) \to \infty ,$$

then w.h.p. G_n has unconditional ACR in some species.

Proof. Assume (19). As $\frac{1}{n^3} \ll p_n$, the random network G_n contains all reactions in $E_{0,1}$, namely, $0 \leftrightarrows X_k$ and $0 \leftrightarrows 2X_k$, for all $1 \leq k \leq n$ (Remark 4.8). For $1 \leq k \leq n$, let $B_k = B_{k,n}$ denote the event that, in all other reactions of G_n, the species X_k appears as a catalyst-only species. We claim that the event B_k implies that G_n has unconditional ACR in X_k. Indeed, in this event, the mass-action ODE for species X_k (for any choice of positive rate constants) has the following form:

$$\frac{dx_k}{dt} \;=\; c_2 x_k^2 + c_1 x_k + c_0 ,$$

where $c_0 > 0$ and $c_2 < 0$ (and $c_1 \in \mathbb{R}$). This quadratic polynomial has a unique positive root (by Descartes' rule). Additionally, G_n necessarily admits a positive steady state (Remark 2.5). We conclude that G_n has unconditional ACR in X_k when B_k occurs.

It therefore suffices to show that, for the random variables $T_n := \sum_{k=1}^{n} 1_{B_k}$, the following limit holds: $\lim_{n\to\infty} \mathbb{P}(T_n > 0) = 1$. Hence, by Lemma 4.18, it is enough to prove $\mathrm{Var}(T_n) \ll (\mathbb{E}T_n)^2$.

We begin with computing $\mathbb{E}T_n$. For fixed k, let $\overline{B}_{0,2}, \overline{B}_{1,1}, \overline{B}_{1,2}, \overline{B}_{2,2}$ be the sets of edges (reactions) in $E_{0,2}, E_{1,1}, E_{1,2}, E_{2,2}$, respectively, in which species X_k appears as a non-catalyst-only species. So, by construction, B_k occurs if and only if G_n contains no reaction from the sets $\overline{B}_{i,j}$:

$$\overline{B}_{0,2} \;=\; \{0 \leftrightarrows X_k + X_j \mid j \neq k, \ 1 \leq j \leq n\}$$

$$\overline{B}_{1,1} \;=\; \{aX_k \leftrightarrows bX_j \mid a,b = 1,2, \ j \neq k, \ 1 \leq j \leq n\} \ \cup \ \{X_k \leftrightarrows 2X_k\}$$

$$\overline{B}_{1,2} \;=\; \{X_k \leftrightarrows X_j + X_\ell \mid j \neq \ell, \ j,\ell \neq k, \ 1 \leq j,\ell \leq n\} \ \cup \ \{2X_k \leftrightarrows X_j + X_\ell \mid j \neq \ell, \ 1 \leq j,\ell \leq n\}$$
$$\qquad\qquad \cup \ \{aX_j \leftrightarrows X_K + X_\ell \mid j,\ell \neq k, \ 1 \leq j,\ell \leq n\}$$

$$\overline{B}_{2,2} \;=\; \{X_k + X_i \leftrightarrows X_j + X_\ell \mid i,j,\ell \neq k, \ j \neq \ell, \ 1 \leq i,j,\ell \leq n\} .$$

It is then straightforward to compute the cardinalities of the sets $\overline{B}_{i,j}$:

$$|\overline{B}_{0,2}| = n-1, \quad |\overline{B}_{1,1}| = 4n-3, \quad |\overline{B}_{1,2}| = (n-1)(3n-3), \quad |\overline{B}_{2,2}| = \frac{(n-1)^2(n-2)}{2}.$$

Thus, using the edge probabilities (6) and the hypothesis $p_n < \frac{1}{n^2}$, we get $\mathbb{P}(B_k)$ and hence $\mathbb{E}T_n$, as follows:

$$(20) \quad \mathbb{P}(B_k) \;=\; (1 - n^2 p_n)^{n-1+4n-3}(1 - np_n)^{(n-1)(3n-3)}(1 - p_n)^{(n-1)^2(n-2)/2}$$

$$\qquad\qquad =\; (1 - n^2 p_n)^{5n-4}(1 - np_n)^{(n-1)(3n-3)}(1 - p_n)^{(n-1)^2(n-2)/2},$$

$$(21) \qquad \mathbb{E}T_n \;=\; \sum_{k=1}^{n} \mathbb{P}(B_k) = n\mathbb{P}(B_k) = n(1 - n^2 p_n)^{5n-4}(1 - np_n)^{(n-1)(3n-3)}(1 - p_n)^{(n-1)^2(n-2)/2}.$$

We recall the following, which is well known:

Fact: For a sequence $\lambda(n)$ and $q \geq 1$, if $\lambda(n) \ll n$, then $(1 - \frac{\lambda(n)}{n^q})^{n^q} \sim e^{-\lambda(n)}$.

To apply this fact, we use (19) to obtain $n^3 p_n \leq (\log(n) - c(n)) \ll n$. We can now apply the fact (with $q = 1,2,3$) to the following three factors in (21): $(1 - n^2 p_n)^{5n-4}$, $(1 - np_n)^{(n-1)(3n-3)}$,

and $(1 - p_n)^{(n-1)^2(n-2)/2}$. One of these analyses is shown below (and the other two are similar):

$$(1 - np_n)^{(n-1)(3n-3)} = \left(\left(1 - \frac{n^3 p_n}{n^2}\right)^{n^2} \right)^{\frac{(n-1)(3n-3)}{n^2}} \sim \left(e^{-n^3 p_n}\right)^{\frac{(n-1)(3n-3)}{n^2}} = e^{-p_n n(n-1)(3n-3)} .$$

The resulting three limits combine to yield the first limit here:

$$\mathbb{E}T_n \sim n e^{-p_n(n^2(5n-4)+n(n-1)(3n-3)+(n-1)^2(n-2)/2)} \geq n e^{-\frac{17}{2}n^3 p_n} \geq n e^{-\log(n)+\frac{17}{2}c(n)} = e^{\frac{17}{2}c(n)} \gg 1 ,$$

and the remaining inequalities are direct computations or come from (19).

Next, we compute $\mathbb{P}(B_k \cap B_h)$. To that end, for fixed k, h (with $1 \leq k, h \leq n$ and $k \neq h$), let $\overline{A}_{0,2}, \overline{A}_{1,1}, \overline{A}_{1,2}, \overline{A}_{2,2}$ denote the sets of edges in $E_{0,2}, E_{1,1}, E_{1,2}, E_{2,2}$, respectively, in which species X_k or X_h (or both) appear as a non-catalyst-only species. By construction, $B_k \cap B_h$ occurs if and only if G_n contains no reaction from the sets $\overline{A}_{i,j}$. Also, $\overline{A}_{i,j}$ is the union of two sets of the form $\overline{B}_{i,j}$, one for k and one for h. Thus, the cardinalities of $\overline{A}_{i,j}$ are computed (in a straightforward way) using the inclusion-exclusion principle:

$$\begin{aligned}
|\overline{A}_{0,2}| &= 2|\overline{B}_{0,2}| - 1 , & |\overline{A}_{1,1}| &= 2|\overline{B}_{1,1}| - 4 , \\
|\overline{A}_{1,2}| &= 2|\overline{B}_{1,2}| - 4(n-2) , & |\overline{A}_{2,2}| &= 2|\overline{B}_{2,2}| - (n-2)(3n-7)/2 .
\end{aligned}$$

The "exclusion" terms above yield:

$$(22) \qquad \mathbb{P}(B_k \cap B_h) = \mathbb{P}(B_k)^2 (1 - n^2 p_n)^{-5} (1 - np_n)^{-4(n-2)} (1 - p_n)^{-(n-2)(3n-7)/2}.$$

Using (22) and other expressions found above, we compute $\mathrm{Var}(T_n)$:

$$\begin{aligned}
\mathrm{Var}(T_n) &= \mathbb{E}(T_n^2) - (\mathbb{E}T_n)^2 \\
&= \sum_{k=1}^{n} \mathbb{P}(B_k) + \sum_{h \neq k} \mathbb{P}(B_k \cap B_h) - (\mathbb{E}T_n)^2 \\
&= n\mathbb{P}(B_k) + n(n-1)\mathbb{P}(B_k)^2 (1 - n^2 p_n)^{-5}(1 - np_n)^{-4(n-2)}(1 - p_n)^{-(n-2)(3n-7)/2} - n^2\mathbb{P}(B_k)^2 \\
&= n\mathbb{P}(B_k) - n\mathbb{P}(B_k)^2 g(n) + n^2\mathbb{P}(B_k)^2(g(n) - 1) ,
\end{aligned}$$

where $g(n) := (1 - n^2 p_n)^{-5}(1 - np_n)^{-4(n-2)}(1 - p_n)^{-(n-2)(3n-7)/2}$. We claim that $\lim_{n\to\infty} g(n) = 1$. In fact, since $n^2 p_n \ll 1$, we have (for n sufficiently large) the inequalities $1 \leq (1 - n^2 p_n)^{-5}$ and $\log(1 - n^2 p_n) \geq -2n^2 p_n$, the second of which further implies $(1 - n^2 p_n)^{-5} \leq e^{10n^2 p_n}$. Applying similar inequalities for the remaining two factors of $g(x)$, we obtain:

$$(23) \qquad 1 \leq g(n) \leq e^{p_n\left(10n^2 + 8(n-2)n + (n-2)(3n-7)\right)} \qquad \text{for } n \text{ sufficiently large.}$$

By (19), the exponent appearing in (23) limits to 0, as $n \to \infty$, so indeed $g(n) \to 1$. Hence,

$$\frac{\mathrm{Var}(T_n)}{(\mathbb{E}T_n)^2} = \frac{1}{n\mathbb{P}(B_k)} - \frac{g(n)}{n} + g(n) - 1 = \frac{1}{\mathbb{E}T_n} - \frac{g(n)}{n} + g(n) - 1 \to 0 + 0 + 1 - 1 = 0 ,$$

as $n \to \infty$, where we also use $\mathbb{E}T_n \gg 1$. Thus, $\mathrm{Var}(T_n) \ll (\mathbb{E}T_n)^2$, which completes the proof. $\qquad \square$

Remark 4.23 (Decoupling in window of dense regime). The proof of Proposition 4.22 shows that, if $\frac{1}{n^3} \ll p_n \leq \frac{\frac{2}{17}\log(n) - c(n)}{n^3}$, then w.h.p. a random network G_n contains a subnetwork of the form $\{0 \leftrightarrows X_k, \ 0 \leftrightarrows 2X_k\}$ for some species X_k that is a catalyst-only species in all other reactions. This highlights the fact that unconditional ACR arises because G_n is a union of two "almost decoupled" subnetworks, one with ACR and the other with multistationarity (by Proposition 4.20) w.h.p.

5. Discussion

We have shown that it is highly atypical for multistationarity and ACR to coexist in certain random reaction networks. In particular, for the type-homogeneous stochastic block model, the window for co-existence is relatively small: It corresponds to when the expected number of edges is approximately between n and $\frac{2}{17}n\log(n)$, where n is the number of species. This window does not even exist unless n is quite large (Remark 4.15). Moreover, when this window exists, the resulting random networks exhibit multistationarity and ACR simply as a result of nearly decoupling into two subnetworks, one with ACR and the other with multistationarity (Remark 4.23).

These results suggest that reaction networks that combine multistationarity and ACR in a non-trivial way require specialized architecture, and the properties do not occur together coincidentally. Of course, real biochemical networks are far from random and exist only when they offer a selective advantage to the organism in its environment. It is a reasonable speculation that combining the two seemingly opposite properties may be favorable. A biochemical network may require robustness in its internal operation while maintaining flexibility as a signal-response mechanism. Said differently, such a network may operate through an essential combination of ACR with multistability.

These ideas raise a natural question: Which special structures, even if statistically rare, can produce ACR and multistability in networks of biochemically reasonable size and complexity? In future work, we will report on such mechanisms and their underlying principles. Interestingly, we find families of biochemical networks with ACR and multistationarity that employ ubiquitous designs such as enzyme-catalyzed reactions, lock-and-key mechanisms for enzyme binding, and redundancy through parallel pathways.

Returning to the current work, we gave asymptotic results on multistationarity when n (the number of species) is large. We are also interested in multistationarity when n is of medium size (say, $n = 10$ to 30). We would like to investigate, by generating random such networks (at various edge-probabilities), what fraction are multistationary. Although checking multistationarity is generally difficult, an approach used here – namely, finding a small multistationary motif (ours had only 3 species) and then lifting it – can be applied. For performing this task, note that certain classes of small multistationary networks have been established [24, 26, 31], as have various criteria for lifting multistationarity (surveyed in [8]).

Going forward, it would be interesting to discover more small multistationary motifs. Are there more multistationary networks with only 3 species that are well suited for lifting to larger networks? Establishing such networks might aid in analyzing the prevalence of multistationarity – with or without ACR – in random reaction networks generated by stochastic block models besides the type-homogeneous one we focused on here.

Another important direction is to consider more biologically realistic models. For instance, we could require our networks to satisfy mass conservation. As another example, we could relax the assumption of reversibility. While the existence of a positive steady state is not guaranteed in this setting (recall Remark 2.5), we nevertheless could exploit new archetypes of reaction networks having both multistationarity and ACR. One such network is the union of the well-known ACR network $\{A + B \to 2B, \ B \to A\}$ (Example 2.1) with a multistationary network in which A is a catalyst-only species. It would be interesting to compute concrete thresholds based on such motifs, and then to compare them to the thresholds found in the current work.

A final promising direction is to study the prevalence and thresholds of additional reaction-network properties. In particular, properties that can be lifted from small networks to larger ones – such as periodic orbits [8, 9, 15, 30] – can also be analyzed in our random-network framework. Do periodic orbits co-exist with ACR in random networks? If so, then, as is the case for multistationarity and ACR, the window of co-existence is likely very small.

REFERENCES

[1] Noga Alon and Joel H Spencer. *The probabilistic method*. John Wiley & Sons, 2016.

[2] David F Anderson, Daniele Cappelletti, and Thomas G Kurtz. Finite time distributions of stochastically modeled chemical systems with absolute concentration robustness. *SIAM Journal on Applied Dynamical Systems*, 16(3):1309–1339, 2017.

[3] David. F. Anderson and Simon. L. Cotter. Product-form stationary distributions for deficiency zero networks with non-mass action kinetics. *Bulletin of Mathematical Biology*, 78(12), 2016.

[4] David F. Anderson, Gheorghe Craciun, and Thomas G. Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bulletin of Mathematical Biology*, 72(8), 2010.

[5] David F. Anderson and Tung D. Nguyen. Results on stochastic reaction networks with non-mass action kinetics. *Mathematical Biosciences and Engineering*, 16(4):2118–2140, 2019.

[6] David F Anderson and Tung D Nguyen. Deficiency zero for random reaction networks under a stochastic block model framework. *Journal of Mathematical Chemistry*, 59(9):2063–2097, 2021.

[7] David F. Anderson and Tung D. Nguyen. Prevalence of deficiency zero reaction networks in an Erdos-Renyi framework. *Journal of Applied Probability*, 59(2):384–398, 2022.

[8] Murad Banaji. Splitting reactions preserves nondegenerate behaviours in chemical reaction networks. *Preprint,* `arXiv:2201.13105`, 2022.

[9] Murad Banaji and Balázs Boros. The smallest bimolecular mass action reaction networks admitting Andronov-Hopf bifurcation. *Preprint,* `arXiv:2207.04971`, 2022.

[10] Murad Banaji and Casian Pantea. The inheritance of nondegenerate multistationarity in chemical reaction networks. *SIAM Journal on Applied Mathematics*, 78(2):1105–1130, 2018.

[11] Béla Bollobás and Andrew Thomason. Threshold functions. *Combinatorica*, 7:35–38, 1987.

[12] Balázs Boros. Existence of positive steady states for weakly reversible mass-action systems. *SIAM Journal on Mathematical Analysis*, 51(1):435–449, 2019.

[13] Corentin Briat, Ankit Gupta, and Mustafa Khammash. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell systems*, 2(1):15–26, 2016.

[14] Jian Deng, Martin Feinberg, Chris Jones, and Adrian Nachman. On the steady states of weakly reversible chemical reaction networks. Preprint, `arXiv:1111.2386`.

[15] Radek Erban and Hye-Won Kang. Chemical systems with limit cycles. *Preprint,* `arXiv:2211.05755`, 2022.

[16] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. In *Publication of the Mathematical Institute of the Hungarian Academy of Sciences*, pages 17–61, 1960.

[17] Martin Feinberg. Complex balancing in general kinetic systems. *Archive for Rational Mechanics and Analysis*, 49:187–194, 1972.

[18] Martin Feinberg. *Foundations of chemical reaction network theory*. Springer, 2019.

[19] Paul W Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.

[20] Fritz Horn. Necessary and sufficient conditions for complex balancing in chemical kinetics. *Archive for Rational Mechanics and Analysis*, 49:172–186, 1972.

[21] Fritz Horn and Roy Jackson. General mass action kinetics. *Archive for Rational Mechanics and Analysis*, 47:187–194, 1972.

[22] Badal Joshi and Gheorghe Craciun. Reaction network motifs for static and dynamic absolute concentration robustness. *to appear in SIAM Journal on Applied Dynamical Systems*.

[23] Badal Joshi and Gheorghe Craciun. Foundations of static and dynamic absolute concentration robustness. *Journal of Mathematical Biology*, 85(53), 2022.

[24] Badal Joshi and Anne Shiu. Atoms of multistationarity in chemical reaction networks. *Journal of Mathematical Chemistry*, 51:153–178, 2013.

[25] Badal Joshi and Anne Shiu. A survey of methods for deciding whether a reaction network is multistationary. *Math. Model. Nat. Phenom., special issue on "Chemical dynamics"*, 10(5):47–67, 2015.

[26] Badal Joshi and Anne Shiu. Which small reaction networks are multistationary? *SIAM Journal on Applied Dynamical Systems*, 16(2):802–833, 2017.

[27] Jinsu Kim and German Enciso. Absolutely robust controllers for chemical reaction networks. *Journal of the Royal Society Interface*, 17(166):20200031, 2020.

[28] Nicolette Meshkat, Anne Shiu, and Angelica Torres. Absolute concentration robustness in networks with low-dimensional stoichiometric subspace. *Vietnam Journal of Mathematics*, 50:623–651, 2022.

[29] Guy Shinar and Martin Feinberg. Structural sources of robustness in biochemical reaction networks. *Science*, 327(5971):1389–1391, 2010.

[30] Xiaoxian Tang and Kaizhang Wang. Hopf bifurcations of reaction networks with zero-one stoichiometric coefficients. *Preprint,* `arXiv:2208.04196`, 2022.

[31] Xiaoxian Tang and Hao Xu. Multistability of small reaction networks. *SIAM Journal on Applied Dynamical Systems*, 20(2):608–635, 2021.

[32] John J Tyson, Reka Albert, Albert Goldbeter, Peter Ruoff, and Jill Sible. Biological switches and clocks. *J. R. Soc. Interface*, 5:S1–S8, 2008.