Research Statement

Matthew D. Johnston

I am an applied mathematician working in the areas of dynamical systems, stochastic processes, and mathematical biology. My research focuses specifically on the mathematical models of biochemical reaction networks such as metabolic pathways and cellular signaling processes. My professional goal is to build a mathematical research program at the interface of mathematics, biology, and computation, and to establish a network of enthusiastic collaborators in all three fields.

Biochemical reaction networks can be extraordinarily complex. The typical metabolic pathway consists of hundreds of reactions and involves dozens of simultaneously interacting species; furthermore, knowledge about the individual kinetic rates of each reaction is often incomplete or missing entirely. To accommodate this uncertainty, mathematical tools are needed which distill the primary functional components, and which are capable of giving results which are robust to changes in the network’s rate parameters, the reaction rate form, and even the dynamical assumptions of the underlying model (e.g. deterministic vs. stochastic).

My research cuts through the network complexity by identifying patterns and motifs in the underlying network structure itself. The theoretic framework for my results is frequently that of Chemical Reaction Network Theory \[5,7,8\] which represents networks of reactions as graphs with complexes of species as nodes and reactions as edges (see Figure 1). Work in this area is noted for relating the dynamical and steady state properties of a system explicitly to the topological features of the underlying reaction graph, such as the connectivity. In fact, many results hold for all initial conditions and all rate parameters choices, which is surprising given that the models are typically defined by a system of highly nonlinear polynomial equations.

Within the scope of biochemical model dynamics, I have made contributions to the understanding of persistence (i.e. non-extinction) in deterministic models \[12,14,21\], extinction phenomena in stochastic models \[2\], the property of dynamical equivalence in mass-action systems \[13,15,17\], and the ability to characterize the steady state sets of mass-action systems \[9,10\]. I am particularly well-known within the community for my work corresponding networks with disparate topological structure to one another \[10,13\]. This latter work has a significant computational component which has led to several collaborations with researchers in the computational mathematics and optimization communities \[12,15,17\].

My research has expanded the scope of networks which can be analyzed by existing means and also provided critical insight into the underlying biochemical motifs which admit, for instance, bistability-switching and oscillatory behaviors. In addition to the mathematical tools native to dynamical systems theory, research in this area has drawn recent interest from researchers in linear algebra, graph theory, algebraic geometry, and optimization and linear programming. Continuing research of this type is vital to the growing field of systems biology which seeks to understand complex biochemical systems by decomposing them into smaller function-specific motifs and pathways.

In what follows, I briefly summarize the research conducted during my doctoral and postdoctoral work, and outline my plans for future research.

---

**Figure 1:** Network representation of the EnvZ-OmpR mechanism in \*Escherichia coli\* with \(X = \text{EnvZ}\), \(Y = \text{OmpR}\), \(D = \text{ADP}\), \(T = \text{ATP}\), and \(p = \text{phosphate group}\).


Research Contributions

1. Persistence and boundary dynamics of mass-action systems. In many biochemical applications, it is crucial to determine whether a given chemical species will tend toward extinction. For example, this extinction may trigger the shutdown of a metabolic pathway. In large biochemical networks, however, it can be extremely challenging to determine whether the system possesses the property that no initially present reactant will tend toward extinction. The property of non-extinction, which is known as persistence, is also the key to solving the largest open problem in Chemical Reaction Network Theory, the Global Attractor Conjecture.

In my doctoral work, I investigated persistence in the context of deterministically modeled mass-action systems. In this framework, we consider the interaction of \( n \) chemical species involved in \( r \) reactions of the form

\[
\sum_{j=1}^{n} y_{ij} X_j \xrightarrow{k_i} \sum_{j=1}^{n} y'_{ij} X_j, \quad i = 1, \ldots, r,
\]

where \( y_{ij}, y'_{ij} \in \mathbb{Z}_{\geq 0} \) are the stoichiometric coefficients which determine how many of the \( j^{th} \) molecule are lost or gained, respectively, in the \( i^{th} \) reaction. Under the assumption of mass-action kinetics, we may model the dynamics of the network \([1]\) with the system of nonlinear ordinary differential equations

\[
\frac{dx(t)}{dt} = \sum_{i=1}^{r} k_i(y'_i - y_i) \prod_{j=1}^{n} x_j(t)^{y_{ij}}
\]

where \( x(t) = (x_1(t), \ldots, x_n(t)) \in \mathbb{R}_{\geq 0}^n \) and \( y_i = (y_{i1}, \ldots, y_{in}) \in \mathbb{Z}_{\geq 0}^n \). The question of persistence is the question of whether, along any time sequence \( t_i \to \infty \), we have \( x_j(t_i) \to 0 \). In \([14]\) and \([21]\), we presented two novel conditions on the reaction graph of a network which are sufficient to guarantee the persistence of all chemical species. We also showed how the conditions could be verified computationally and applied them to prove special cases of the Global Attractor Conjecture.

I have additionally worked on problems within the scope of endotactic networks \([4, 6, 19]\). Endotacticity is roughly a measure of how “inward-point” the reaction vectors \( y'_i - y_i \in \mathbb{Z}^n \) are (see Figure 2). The condition is thought to lead to persistence of \([2]\) based on geometric considerations. In collaboration with Casian Pantea and Pete Donnell, I presented a mixed-integer linear programming algorithm capable of verifying the endotacticity of chemical reaction networks for networks of arbitrary dimension \([12]\). The algorithm is the first computational implementation of work in this area and is also the first capable of verifying endotacticity for networks with more than two constituent species. We have implemented the algorithm in the web-based open-source software program CoNtRol and successfully classified over 400 networks from the BioModels Database.

2. Extinction in stochastically-modeled reaction networks. In many biochemical reaction systems the number of interacting molecules is on the order of tens or hundreds. For such
networks, it is inappropriate to model the dynamics deterministically over a continuous state space; rather, it is suitable to model it stochastically over discrete molecular counts as a continuous-time Markov chain. In this framework, we have

\[ X(t) = X(0) + \sum_{i=1}^{r} (y'_i - y_i)Y_i \left( \int_{0}^{t} \lambda_i(X(s)) \, ds \right) \]  

(3)

where \( \{Y_i\} \) are unit rate Poisson processes and \( \lambda_i(\cdot) \) are rate intensities which are often chosen to be of mass-action form. The chain (3), however, may exhibit “extinction events” not permitted by the deterministic model (2) (see Figure 3). To accurately capture the qualitative behavior of the biochemical system being modeled, therefore, it is important to understand which systems exhibit this distinction in long-term behavior when modeled stochastically versus deterministically.

In my postdoctoral work, I investigated this stochastic extinction phenomenon from the perspective of Chemical Reaction Network Theory. In collaboration with David F. Anderson and Germán Enciso, I identified a broad class of networks for which extinction events not present in the deterministic model necessarily arise when modeled stochastically [2]. The class we identified contains, for instance, all of the systems with so-called absolute concentration robustness identified by Guy Shinar and Martin Feinberg in their recent Science paper [20]. Although this disparity in long-term behavior had been studied previously (for example, Keizer’s paradox), our work was the first to identify a general class of networks for which this key distinction holds.

3. Dynamical equivalence and linear conjugacy of mass-action systems. Under mass-action kinetics, two chemical reaction networks with distinct underlying structure may generate the same governing set of ordinary differential equations (2). For example, it can be easily seen that the networks

\[ 2X \xrightarrow{1} 2Y \xrightarrow{2} X + Y \]

and

\[ 2X \xrightleftharpoons{1}{2} 2Y \]

both generate the mass-action system \( \dot{x} = -\dot{y} = -2x^2 + 2y^2 \). It follows that, for a known dynamics [2], we may not uniquely infer the network structure. This is a problem in the study of structural and parameter identifiability in chemical reaction networks.

We may also, however, use this observation as a tool for expanding the scope of known dynamical results within the study of biochemical reaction networks. In particular, if two networks are dynamically equivalent (i.e. have the same governing systems [2]), any known dynamical properties of one system may trivially be transferred to the other, where the property may be
otherwise unknown. Building upon this intuition, in [13] we extend dynamical equivalence of systems to linear conjugacy of systems. For linearly conjugate systems, the dynamical flows are related by a non-trivial linear transformation in the reactant concentrations (see Figure 1). In subsequent collaborations with Gábor Szederkényi, we addressed the question of determining, from within the class of dynamically equivalent or linearly conjugate systems, ones which possess network properties such as weak reversibility [17], detailed or complex balancing [16], and a minimal possible deficiency, which is a key network parameter in the study of system analysis [15]. We approached these problems through novel application of the well-established and portable mixed-integer linear programming framework.

4. Characterization of steady states of mass-action systems. To determine the steady state sets of mass-action systems, one must solve the following system of polynomial equations

\[ \sum_{i=1}^{r} k_i(y'_i - y_i) \prod_{j=1}^{n} x_j(t)^{y_{ij}} = 0. \]

This solution must then be intersected with an affine invariant space known as a stoichiometric compatibility class. This task is complicated by the observation that, in application, the rate parameters are often unknown and so must be left as symbolic parameters. The study of these steady state sets has attracted the attention of many leading algebraists in recent years, including Bernd Sturmfels, Alicia Dickenstein, Ezra Miller, and the late Karin Gatermann, among others.

I have recently introduced a network-based method for determining these steady states. The method, which is called network translation, cuts through the inherent algebraic complexity of (4) by associating the network to a generalized one with the same reaction vectors but different underlying network structure (especially strongly connected) [10] (see Figure 5). The method of network translation has the advantage of graphically identifying the underlying modes of metabolic balance which may be obscured by traditional steady state analysis. Network translation has furthermore been recently applied by Carsten Conradi and Anne Shiu to characterize the steady state set of the processive multisite phosphorylation system [3]. I have recently shown that network translation has wider applicability within models of MAPK networks [11] and also developed an algorithm capable of implementing the process of network translation [9].
Future Research

1. **Generalized mass-action systems and network translation.** Stefan Müller and Georg Regensburger have recently initiated the study of generalized mass-action systems. In such systems, the mass-action monomials in (2) are allowed to take powers which do not necessarily correspond to the stoichiometry of the underlying reaction network [18]. My work on network translation meanwhile has opened the door for using structural properties of such networks to characterize the steady states of classical mass-action systems [10]. The property of a system having toric steady states, which is of significant interest in the algebraic and biochemical communities, can often be established directly through network translation.

The study of network translation and generalized mass-action systems in general, however, is in its very early stages. A primary goal moving forward is to develop these theories in the same way that Chemical Reaction Network Theory itself was developed throughout the 1970s. Among the next aspects to study are:

1. **Multistationarity.** Multistationarity is the property that the system permits multiple stoichiometrically compatible steady states and is thought to underlie hysteresis and switching behavior in biochemical models. A suitable next step in the theory of network translation is to determine conditions on the network translation process itself which guarantee, or preclude, the capacity for the underlying system to exhibit multistationarity. This challenge is especially tantalizing given recent work on multistationarity in generalized polynomial systems.

2. **Dynamical behavior.** Very little work has been conducted to date on characterizing the dynamics (as opposed to steady state properties) of generalized mass-action systems. Such systems, however, are not static objects; they evolve just as the classical systems. A primary focus moving forward will be adapting the classical tools such as Lyapunov stability theory and linear stability analysis to these generalized mass-action systems. This work will continue to include collaboration with the algebraic community.

2. **Extending network theory to stochastic models.** Sparse work has been conducted extending the classical results of Chemical Reaction Network Theory from the deterministic to the stochastic setting. The most comprehensive results to date are contained in the papers [1], where the authors show dynamical agreement between the deterministic and stochastic models for the class of “complex-balanced systems,” and in my own previously mentioned paper [2] where we show necessary differences in long-term behavior. Future research will include:

1. **Extinction sets.** It is still unclear which general conditions are sufficient to guarantee extinction in stochastic models of biochemical reaction networks and how these extinction sets are structured within the discrete state space. These questions have received surprisingly little attention in the modeling literature. Future work will focus on adapting results from the well-established Petri net and queueing literature to biochemical models. This work will necessitate collaboration within the stochastic and combinatorics research communities.

2. **Positive recurrence and stationary distributions.** Research into the property of positive recurrence of continuous-time Markov chain models of biochemical reaction networks has started to gain traction within the stochastic modeling community; however, work still lags significantly behind that of the corresponding question of long-term dynamical behavior for deterministic models. Future work will seek to establish necessary and sufficient conditions on the underlying network structure under which positive recurrence holds and, where possible, characterize the resulting stationary distributions.
References


