Determining Interconnections in Chemical Reaction Networks

Antonis Papachristodoulou and Ben Recht

Abstract— **We present a methodology for robust determination of chemical reaction network interconnections. Given time series data that are collected from experiments and taking into account the measurement error, we minimize the 1-norm of the decision variables (reaction rates) keeping the data in close Euler-fit with a general model structure based on mass action kinetics which models the species' dynamics. We illustrate our methodology on a hypothetical chemical reaction network under various experimental scenarios.**

I. INTRODUCTION

A common problem in engineering and the physical sciences is how to robustly determine the interaction topology of a networked system using time series data collected from experiments. On one hand, such data is often abundant and a high computational effort is required to extract information about the system properties. At the same time, a large amount of data is often necessary to distinguish between different models and decide which components play the most important role in such networks.

Our study focuses on chemical reaction networks where the main objective is to identify the chemical pathways and mechanisms (how the chemical complexes interact within the chemical network) rather than extracting the parameters (the rates of the various reactions) that best fit the particular time series data. The interconnection topology is of greater importance than the particular parameter settings for several reasons. First, the parameters are often collected under noisy experimental conditions and are sensitive to laboratory variations such as temperature and the environment. Second, as is often the case with large networks, persistence of observed phenomena is robust to a large range of parameter values and therefore identifying these parameters exactly is not of great interest. Indeed, networks often have the same functionality in the neighborhood of the nominal reaction rates. But most importantly, networks are rarely robust to the random rewiring of the underlying interconnection structure and hence determining the network structure is much more important than determining the kinetic parameters that fit the particular data.

In this paper we follow a top-down optimization based approach for estimating the connectivity in such networks, proposing models from first principles assuming mass action kinetics and no *a priori* information about the interconnection topology. In particular, we argue that minimizing the one-norm of all possible reaction rates can extract sparse feasible solutions which indicate likely interaction topologies. This optimization can be performed using Linear Programming (LP) for which efficient algorithms exist that can solve large optimization problems. Moreover, the robust counterpart with polytopic uncertainties in the data is also a Linear Program. This is in stark contrast to least-squares where the robust counterpart is at best a second order cone problem (SOCP) [1], which is more computationally expensive than linear programming.

The paper is organized as follows. First, we outline related work in Section II. We describe the system classes under study in Section III. In Section IV we describe our Linear Programming algorithm for determining the reaction network from time series data and discuss the advantages of over a least-squares approach. In Section V we apply our algorithm to a hypothetical chemical reaction network and discuss performance under various scenarios.

II. RELATED WORK

Determining the pathways in chemical reaction networks from time series data has been an active area of research for over a decade. A recent review of available techniques can be found in [2] or [3], but earlier articles, such as [4], also list several approaches to this network identification problem.

We summarize the available techniques as follows. One class of techniques that fall under the rubric of 'stationary-state Jacobian Matrix Elements' analyze the system behavior when it is perturbed locally from the steady state and look at whether the concentration of one species is increased or decreased when another species' concentration is increased. Building on this approach, Kholodenko *et al* [5], [6] have provided an approach to determine the functional interactions in cellular signaling and gene networks, taking into account the 'modular' structure of the networks in question. Alternatively, inferences about the topology of the network can be made by introducing pulse changes in concentration of a chemical species in the network, and observing the network's response, concluding causal chemical connectivities [7].

A variety of data-driven approaches attempt to extract structure from existing experimental data without the ability to tailor experiments to the modeling task. For example, researchers have used time series measurements of concentrations to construct correlation functions of concentrations [8]. An approach using Artificial Neural Networks [9] tries to 'learn' patterns from the complicated and noisy data and to detect trends in the chemical reaction pathways. Related to this is a genetic algorithm approach to study the evolutionary development of a reaction mechanism [10].

A. Papachristodoulou is with the Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, U.K. antonis@eng.ox.ac.uk.

B. Recht is with the Center for the Mathematics of Information, California Institute of Technology, 1200 E. California Blvd., MC 136-93, Pasadena CA 91125, U.S.A.. brecht@caltech.edu.

In [11] the Singular Value Decomposition was used to obtain a family of candidate networks. Since the optimal networks were typically much more dense than would be realistically expected, the sparsest network in the family was identified using robust regression.

Modeling reaction networks using mass action kinetics (described in Section III) results in nonlinear Volterra-type dynamics similar to those studied in [12]. However, since we are interested in determining the 'sparsest' network with mass-action kinetics, our objective function on the kinetic parameters is quite different.

III. CHEMICAL REACTION NETWORKS

Let $\mathbb{R}$ denote the set of real numbers, and $\mathbb{R}^n$ the $n-$dimensional Euclidean space. Similarly, denote by $\mathbb{R}_+$ ($\mathbb{R}_{++}$) the non-negative (positive) real line, and by $\mathbb{R}^n_+$ ($\mathbb{R}^n_{++}$) the $n-$dimensional non-negative (positive) orthant.

By a Chemical Reaction Network we mean a series of elementary reaction steps, i.e., collisions of reactant molecules to form products, which we assume are balanced stoichiometric equations. We follow the modeling framework described in [13].

First, let us explain what we mean by an elementary reaction step. Suppose A, B and C are chemical species, and suppose that the reaction that occurs between them is

$$A + B \xrightarrow{k_1} C \tag{1}$$

For this chemical reaction, which says that a molecule of A reacts with a molecule of B to give a molecule of C with rate k_1, we can define three sets:

- A set of complexes, i.e., whatever appears at the head and tails of the reactions, $\mathcal{C} = \{A + B, C\}$;
- A set of species that participate in the reactions, $\mathcal{S} = \{A, B, C\}$; and
- A set of reactions $\mathcal{R} = \{A + B \to C\}$ and associated with it a vector of reaction rates $k = k_1$.

Suppose now we put together several such elementary reactions, to form the reaction network shown in Figure 1. In this network we have put nodes for complexes and arrows to define reactions, with k_{ij} the rate at which a reaction occurs with reactants complex i and products complex j. What we are interested in, is the structure of the reaction rate dynamics that result from this interconnection, if we assume mass action kinetics. For this network, the set of complexes is $\mathcal{C} = \{A, 2B, A + C, D, B + E\}$, the set of species is $\mathcal{S} = \{A, B, C, D, E\}$ and the set of reactions is $\mathcal{R} = \{A \to B + E, 2B \to A, 2B \to B + E, A + C \to A, A + C \to D, B + E \to A + C\}$ with reaction rates $k = [k_{15}, k_{21}, k_{25}, k_{31}, k_{34}, k_{53}]$. For each of the above sets, we associate a vector field, and, for indexing, we put an arbitrary ordering on the complexes and species as indicated in Figure 1.

The dynamics of a general chemical reaction network, i.e., a network with a species set $\mathcal{S}$, a set of complexes $\mathcal{C}$ and a set of reactions $\mathcal{R}$ with reaction rates k, are assumed to be

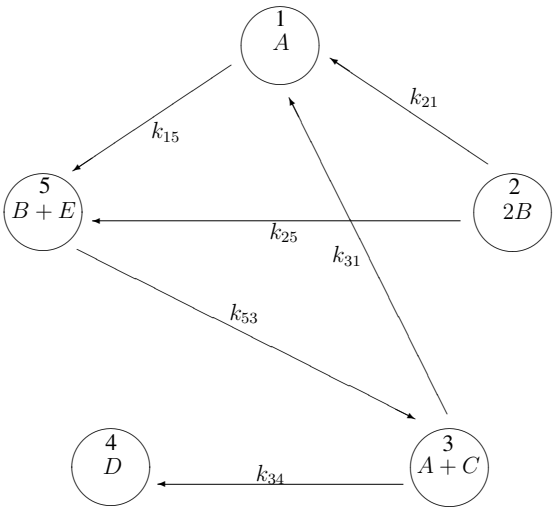

Fig. 1. A chemical reaction network

described by differential equation system in the form:

$$\dot{x} = \sum_{i=1}^{|\mathcal{S}|} \sum_{j=1}^{|\mathcal{S}|} k_{ij} x^{y_i} (y_j - y_i) \tag{2}$$

where x is species concentration vector and y_i is a vector of length $|\mathcal{S}|$ (the cardinality of set $\mathcal{S}$) showing how complex i is comprised of the species in the network, i.e., in our example

$$y_i^T \begin{bmatrix} A \\ \vdots \\ E \end{bmatrix} = c_i$$

where c_i is complex i in the set $\mathcal{C}$. For example $y_3^T = [1, 0, 1, 0, 0]$ in the network in Figure 1. Also, we have used the Laurent-Schwartz notation for x^{y_i} which means

$$x^{y_i} = \prod_{k=1}^{|\mathcal{S}|} x_k^{y_{ik}}$$

The only unknowns in (2) are the parameters k_{ij} which completely determine the network topology. arge-scale networks are seldom heavily connected, but rather enjoy sparsity and modularity [5]. Therefore rather than identifying the 'best fit' of k_{ij}, we seek to establish a sparse interaction topology between the vertices v_i, through the determination of the sparsity structure of the matrix $K = [k_{ij}]$.

IV. PROBLEM STATEMENT AND SOLUTION APPROACH

Suppose time series data has been collected experimentally for the physical system and (for now) that all x_i, $i = 1, \ldots, n$ can be measured. This data takes the form $(t_p, \hat{x}(t_p))$ where $p = 1, \ldots, M$. We assume such experiments yield data at small time steps which may not be regular, and the measurements are contaminated with errors that can be captured as bounds on each measurement

$$x_i^*(t_p) - \epsilon_i^-(p) \leq \hat{x}_i(t_p) \leq x_i^*(t_p) + \epsilon_i^+(p) \tag{3}$$

for $(\epsilon_p^-, \epsilon_p^+) \in \mathbb{R}^2_+$, $p = 1, \ldots, M$ and $i = 1, \ldots, n$.

The problem of nonlinear system identification is generally solved through discretization of the proposed model [14], [15], [12]. Assuming that samples are taken at sufficiently short time intervals, various discretization methods can be applied, the simplest of which is Euler

$$x(t_{k+1}) = x(t_k) + (t_{k+1} - t_k) f(x(t_k)), \quad x(t_0) = x_0 \quad (4)$$

Here $f(x(t_k)$ denotes the right hand side of Equation (2). Each discretization step incurs an error $O((t_{k+1} - t_k)^2)$ locally, and the method is first order globally. Other, more complicated formulae can be proposed to increase the global order of the discretization, such as a modified Euler discretization, taking the form

$$\begin{aligned} x(t_{k+1}) &= x(t_k) + \frac{(t_{k+1} - t_k)}{2} \left(f(x(t_k)) + f(\tilde{x}(t_{k+1})) \right), \\ \tilde{x}(t_{k+1}) &= x(t_k) + (t_{k+1} - t_k) f(x(t_k)), \quad x(t_0) = x_0 \end{aligned}$$

or even Runge-Kutta methods. It was also shown in [14], that including 'extra zero dynamics' in the sampling process can decrease further the local discretization error.

Regardless of the discretization method, the reaction rates appear linearly in these approximations to the ODE (2). In this case, a variety of estimation techniques can be applied to fit the parameters to the observed data. We assume here that we have measurements for all x_i, $i = 1, \ldots, n$, i.e., we have nM measurements. Given a data point $(t_q, \hat{x}(t_q))$ and knowledge of the next time step t_{q+1} yields an estimate for $x_i(t_{q+1})$ of the form

$$x_i(t_{q+1}) = \hat{x}_i(t_q) + (t_{q+1} - t_q) f_i(x(t_q)), \quad (5)$$

for each $i = 1, \ldots, n$ and $q = 0, \ldots, M-1$. A popular criterion for evaluating the quality of a set of parameters is the 2-norm of the error between $x_i(t_{q+1})$ and the measured values $\hat{x}_i(t_{q+1})$. We can write such an error metric as

$$\min \|Ak - b\|_2$$

where $k \in \mathbb{R}^{n^2}$ is a vector containing k_{ij}, which we treat as 'decision variables', and $A \in \mathbb{R}^{(M \times n) \times n^2}$ and $b \in \mathbb{R}^{(M \times n)}$ are defined by 'stacking' all such conditions obtained by manipulating the data as per (5). The interesting fact is that the minimum of the 2-norm error has a closed-form solution,

$$k^* = A^\dagger b \triangleq (A^T A)^{-1} A^T b$$

where A^T denotes the transpose of A.

There are a few drawbacks of the above least-squares approach. In the presence of extra constraints on the variables k (constrained regression), the problem does not generally have a closed-form solution. Such constrained minimizations, the simplest of which is Second Order Cone Problems (SOCP), may carry a significant computational cost for large problems. The same is true if the data are contaminated with error which needs to be taken into account when producing k^* [1]. Furthermore, the solution to a least-squares problem will not in general be sparse. Regularized least-squares can be used, in which the solution k^* can be asked to be small, but this solution is also not guaranteed to be sparse [16] and the bi-criterion weights have to be chosen judiciously.

Of course obtaining the sparsest solution k^* so that $\|Ak - b\|$ in some norm is minimized is not computationally tractable as the 0-1 combinatorial element is in the problem statement. A popular relaxation for such combinatorial problems is to minimize the $1-$norm of the solution so that the constraints are satisfied within certain bounds. That is, we propose solving the problem

$$\begin{aligned} \min \quad & \|k\|_1 \\ \text{s.t.} \quad & -l_i^- \leq a_i^T k - b_i \leq l_i^+, i = 1, \ldots, mM \end{aligned} \quad (6)$$

with $(l_i^-, l_i^+) \in \mathbb{R}^2_+$ are small numbers and $a_i's$ are the rows of A and b_is the elements of b. We stress that this is just a heuristic in obtaining a sparse solution. However, the above optimization problem is a Linear Program (LP), for which robust algorithms exist that can solve large problem instances. Furthermore, since the optimization problem is polyhedral, redundant and linearly dependent constraints may allow for a sparse solution. The disadvantage of using this approach to solve the problem at hand, is that the data may not be in close-fit, but of course we would expect some errors to be present, not only because of measurement errors but also because of the discretization approach employed.

A further advantage of this approach is that we may incorporate uncertainties in the measurements with little additional computational complexity. Indeed, the robust counterpart of the above problem is

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax - b \in \mathcal{K}, \quad \forall A \in \mathcal{P} \end{aligned}$$

where $\mathcal{P}$ is a set in which A is allowed to live. If $\mathcal{P}$ can be described by a polytope, whose vertices are known, then the robust counterpart of a conic program is the same conic program. However, seldom do we know the vertices of the polytope, and they may be too many, which makes the robust counterpart a lot harder than the nominal problem. If ellipsoidal uncertainty is used to describe the uncertainty, then the robust counterpart of an LP is an SOCP, which is more complicated to solve than solving an LP. However, if we model the uncertainty in the measurements as in (3) we can use Soyster's framework [17] to formulate the robust counterpart that is still an LP. We note that the framework introduced in [18] gives a series of LPs if not all data are subject to uncertainty. In particular, if the coefficients a_{ij} of matrix A are perturbed so that they take values in $[a_{ij} - \epsilon_{ij}, a_{ij} + \epsilon_{ij}]$ then the following LP is a robust formulation of Program (6):

$$\begin{aligned} \min \quad & \|k\|_1 \\ \text{s.t.} \quad & -l_i^- \leq \sum_{j=1}^{N} a_{ij} k_j + \sum_{j=1}^{N} \epsilon_{ij} \mu_j - b_i \leq l_i^+, \\ & \qquad i = 1, \ldots, mM \\ & -\mu_i \leq k_i \leq \mu_i, \quad i = 1, \ldots, n^2 \\ & \mu_i \geq 0, \quad i = 1, \ldots, n^2 \end{aligned} \quad (7)$$

Using the above ideas, we aim to extract from data the sparsity pattern in the solution K of the optimization problem, which is related to the connectivity of the underlying graph structure, drawing important conclusions on the interaction topologies inside the network. In the next section, we will illustrate this idea, implementing this Linear Programming approach to variety of different network identification scenarios for a particular chemical reaction network.

V. NUMERICAL EXPERIMENTS

Consider a network with 5 species $\mathcal{S} = \{A, B, C, D, E\}$ and 5 complexes, $\mathcal{C} = \{A, 2B, A + C, D, B + E\}$. Suppose we are given time series data for all the species this system, but we do not know the topology of the interconnection. The first experiment examines the recovered interconnection diagram using the (non-robust) linear program (6). Later on, we will consider the same problem with missing data on one species and a robust network determination problem.

We have implemented the network shown in Figure 1 with a K matrix of the form:

$$K = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.8492 \\ 0.3386 & 0 & 0 & 0 & 0.4290 \\ 0.8244 & 0 & 0 & 0.0563 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.7364 & 0 & 0 \end{bmatrix} \tag{8}$$

and have simulated the system with uniformly distributed initial conditions. The data sets were obtained by simulating the above set of nonlinear equations using SIMULINK. Ten such data sets were generated and incorporated in the linear program.

Since we do not know how the chemical network is connected, and we cannot even speculate how part of it may be connected, we need to assume a general structure for it and write the dynamics for the complete network. The structure of all possible interconnections that we could have with these complexes is shown in Figure 2. A least-squares approach, would yield the following array of reaction rates:

$$\begin{bmatrix} 0 & 0.1 & 1 & 0.1 & 1 \\ 1 & 0 & 0.8 & 0.2 & 1 \\ 1 & 0.6 & 0 & 1 & 0.9 \\ 0.1 & 0.8 & 0.9 & 0 & 0 \\ 1 & 0.9 & 1 & 0.9 & 0 \end{bmatrix}$$

and essentially the only zero element predicted is k_{45}. Note that the diagonal of this matrix does not enter into our optimization. We write these entries as zero, but this is merely a convenient notational place holder.

The resulting K from our linear programming approach is given by

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0.5 & 0.4 & 0.9 \\ 1 & 0 & 0 & 1 & 0.2 \\ 0 & 0 & 0.8 & 0 & 0 \\ 0.1 & 0 & 1 & 0.2 & 0 \end{bmatrix}$$

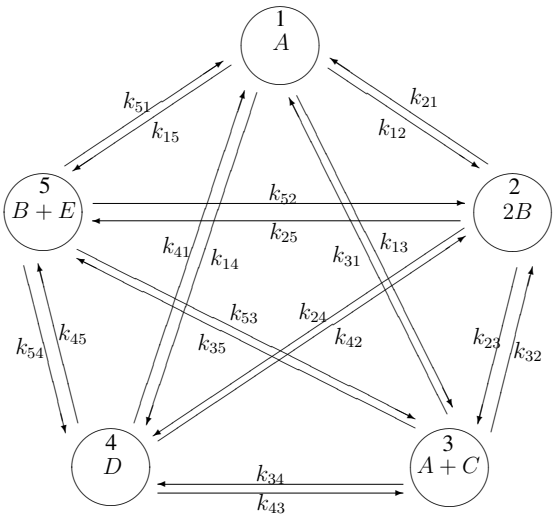

Fig. 2. The complete network

Observe that the second column is equal to zero which implies that the second complex is not the product of any reaction. Having determined this sparse structure for the system, we can repeat the same LP optimization, but now impose the new information about the sparse structure obtained in the new linear program, i.e. that $k_{12} = 0$ etc. Iterating once on this data, we get the following results:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0.5 & 0.3 & 0.7 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0.8 & 0 & 0 \\ 0 & 0 & 0.8 & 0 & 0 \end{bmatrix}$$

This experiment reveals that the sparsity structure can be further reduced by an iterative procedure. One could also use the above as a 'probability' lookup table, and investigate sparsity structures, such as setting k_{23} and k_{24} equal to zero. Indeed this solution is also feasible, which reveals additional structure in the matrix K. Working this way, we have found that the following non-zero matrix results in feasible LPs, with $l = 0.01$:

$$K_{\text{nom}} = \begin{bmatrix} 0 & 0 & k_{13} & 0 & k_{15} \\ k_{21} & 0 & 0 & 0 & k_{25} \\ k_{31} & 0 & 0 & k_{34} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_{53} & 0 & 0 \end{bmatrix}$$

which is the same as the network shown in Figure 1, but for a link between complex 1 and complex 3. At this point we should recall that there is no unique reaction mechanism that can fit a data set; in fact, we can only hope to invalidate a postulated reaction mechanism using data. We will return to this issue in the concluding section.

The next experiment we performed was to assume that some of the species could not be observed in the experiments for technical reasons. In particular, we assumed that the concentration of species A could not be measured. This

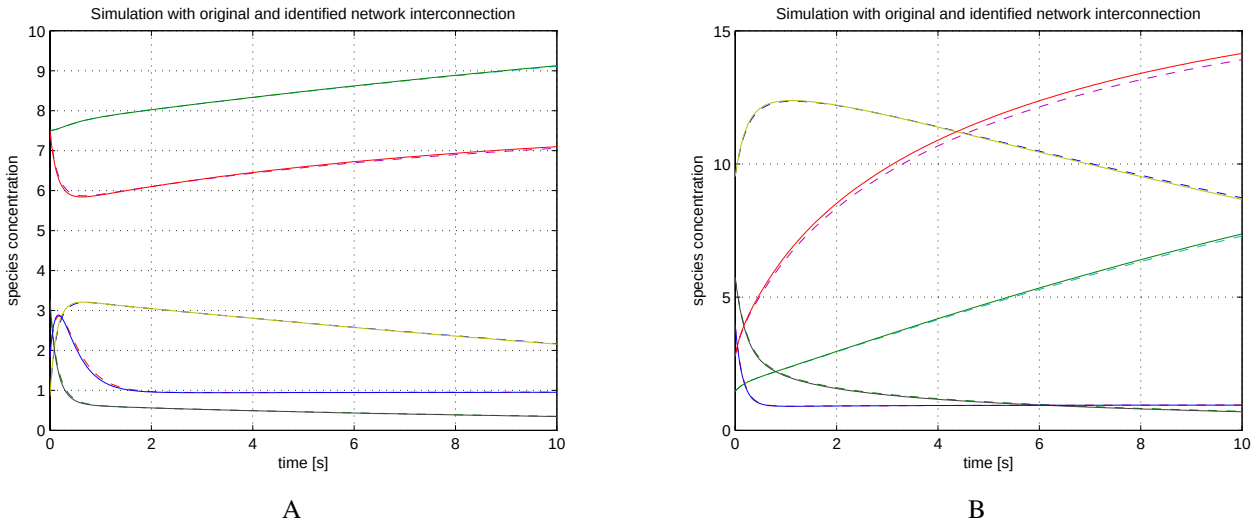

Fig. 3. Simulation of network shown in Figure 1 with reaction rates (8) (solid line) and the network shown in Figure 2 with reaction rates given by (9) (dashed line) from two initial conditions.

does not pose significant problems, as we can replace the occurrences of the terms in the vector field involving the variable x_1 with a vector of new variables q which we also ask to be 'sparse', through minimization of the sum of q_i. Eight such substitutions need to be made; the result is a matrix of the form:

$$K_{\text{miss}} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ k_{21} & 0 & k_{23} & 0 & k_{25} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_{43} & 0 & 0 \\ 0 & 0 & k_{53} & k_{54} & 0 \end{bmatrix}$$

and a $q = [q_1, \ldots, q_8]$ which corresponds to nonzero entries for k_{31}, k_{34}, k_{35} k_{13} and k_{15}. Therefore in this case too, a sparse topology interconnection is obtained, but the matrix in this case is not as sparse as before.

Suppose now that data are uncertain, and we want to search for *robust* sparse structures for the K matrix. Using the framework developed in [18], we set $\epsilon_i^+(p) = 0.0004$ and $\epsilon_i^-(p) = 0.0005$ for $i = 1, \ldots, 5$ and all data points p — such uncertainty could be due to roundoff errors (see Equation (3)). A robust LP can be formulated, as discussed in section IV, and the resulting optimization results in a network with a richer sparsity structure:

$$K_{\text{rob}} = \begin{bmatrix} 0 & 0 & k_{13} & 0 & k_{15} \\ k_{21} & 0 & k_{23} & 0 & k_{25} \\ k_{31} & 0 & 0 & k_{34} & 0 \\ 0 & 0 & k_{43} & 0 & 0 \\ 0 & 0 & k_{53} & 0 & 0 \end{bmatrix}$$

Finally, we note that once a candidate network is determined, we can perform a least-squares minimization to obtain the best k values for a particular sparsity structure. For example, if we choose K_{nom} as the sparsity structure and fit the least squares error over all 10 experiments, we get the following K matrix:

$$K = \begin{bmatrix} 0 & 0 & 0.0364 & 0 & 0.7721 \\ 0.3295 & 0 & 0 & 0 & 0.3999 \\ 0.7804 & 0 & 0 & 0.0553 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6668 & 0 & 0 \end{bmatrix} \quad (9)$$

In figures 3A and 3B we show how the nominal system, with the K matrix given by Equation (8) compares in simulation with the K matrix given by Equation (9) for different initial conditions. We see that some initial conditions have better behavior for the two parameter sets than others. There is hope, however, that using other methods and through choice of a particular initial condition we can eliminate a parameter set, as the initial condition in Figure 3B shows some deviation in the dynamics of x_1.

VI. DISCUSSION AND CONCLUSIONS

We have presented a system identification heuristic for recovering the structure of chemical reaction networks from noisy and incomplete measurements using linear programming. Our method couples simulation with experimental data and attempts to find a sparse set of parameters consistent with the measurements. We have shown how to add parameter uncertainty at little computational cost. And we have demonstrated numerically that this technique is able to recover reasonable approximations to true network topologies in the presence of noise and missing data.

Even though our final least-squares fit recovers reaction rates close to the true rates, our heuristic does not recover the true network structure. Indeed proving that a particular reaction mechanism is correct is a misnomer. This is because there may be another reaction mechanism that may be consistent with the available data, within experimental error etc. Disproving that a mechanism could ever be represented by data is possible, and new techniques exist in that direction [19]. Furthermore, it is of great interest to characterize

the networks for which our heuristic would recover the true sparsity pattern.

A model that has been developed using a top-down approach is sometimes too complicated for analysis. Even though for particular network structures there are important results that establish stability, existence of equilibria etc [13], in general this task is difficult and reduced order models need to be constructed for analysis.

Our heuristic technique is aimed specifically in obtaining robust, sparse, network interconnections to approximately fit data rather than in fitting the 'best' K which would vary for different experimental conditions. Specific *a priori* knowledge for the system structure, such as modularity etc., will definitely reduce significantly the computational burden. However, we insist that our problem stays within the class of linear programs, for which reliable algorithms exist of complexity of order n^4M (where n is the number of species and M is the number of data points) not taking into account a sparse structure implied from an *a priori* knowledge of the system interconnection.

Finally, we note that even though mass action kinetics are common in chemical reaction networks modeling, there are many situations where some other model description such as the Generalized Mass Action model or the S-system model would be preferred [20], [21], [22]. The above technique can be used under dynamics with other 'canonical form' descriptions, as long as Equations (5) are affine in the unknown parameters. Otherwise, a linearization approach can be employed, or some other nonlinear methods can be used. The effectiveness of our scheme in such modeling schemes is a promising direction left for future study.

REFERENCES

[1] L. El Ghaoui and H. Lebret, "Robust solutions to least-squares problems with uncertain data," *SIAM J. MATRIX ANAL. APPL.*, vol. 18, no. 4, pp. 1035–1064, 1997.

[2] E. J. Crampin, S. Schnell, and P. E. McSharry, "Mathematical and computational techniques to deduce complex biochemical reaction mechanisms," *Progress in Biophysics and Molecular Biology*, vol. 86, pp. 77–112, 2004.

[3] J. Ross, I. Schreiber, and M. O. Vlad, *Determination of Complex Reaction Mechanisms: Analysis of Chemical, Biological and Genetic Networks*. Oxford University Press, Inc., 2006.

[4] T. Chevalier, I. Schreiber, and J. Ross, "Toward a systematic determination of complex reaction mechanisms," *J. Phys. Chem.*, vol. 97, no. 26, pp. 6776–6787, 1993.

[5] B. N. Kholodenko, A. Kiyatkin, F. J. Bruggeman, E. D. Sontag, H. V. Westerhoff, and J. B. Hoek, "Untangling the wires: A strategy to trace functional interactions in signalling and gene networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 20, pp. 12841–12846, 2002.

[6] B. N. Kholodenko and E. D. Sontag, "Determination of functional network structure from local parameter dependence data," 2002. oai:arXiv.org:physics/0205003.

[7] W. Vance, A. Arkin, and J. Ross, "Determination of causal connectivities of species in random networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 5816–5821, 2002.

[8] A. Arkin, P. D. Shen, and J. Ross, "A test case of correlation metric construction of a reaction pathway from measurements," *Science*, vol. 277, pp. 1275–1279, 1997.

[9] J. S. Almeida, "Predictive non-linear modeling of complex data by artificial neural networks," *Current Option in Biotechnology*, vol. 13, pp. 72–76, 2002.

[10] M. Wahde and J. Hertz, "Coarse-grained reverse engineering of genetic regulatory networks," *Biosystems*, vol. 55, pp. 129–136, 2000.

[11] M. K. Stephen Yeung, J. Tegnér, and J. J. Collins, "Reverse engineering gene networks using singular value decomposition and robust regression," *Proceedings of the National Academy of Sciences*, vol. 99, no. 9, pp. 6163–6168, 2002.

[12] F. J. Doyle III, R. K. Pearson, and B. Ogunnaike, *Identification and Control of Process Systems Using Volterra Models*. Springer, 2001.

[13] M. Feinberg, "Lectures on chemical reaction networks," 1979. Lectures delivered at the Mathematics Research Center, University of Wisconsin-Madison.

[14] J. I. Yuz and G. C. Goodwin, "On sampled-data models for nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 10, pp. 1477–1485, 2005.

[15] L. Ljung, *System Identification: Theory for the User*. Information and System Sciences, Prentice Hall, Inc., second ed., 1999.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[17] A. L. Soyster, "Convex programming with set-inclusive constraints and applications to inexact linear programming," *Operations Research*, vol. 21, pp. 1154–1157, 1973.

[18] D. Bertsimas and M. Sim, "The price of robustness," *Operations Research*, vol. 52, no. 1, pp. 35–53, 2004.

[19] S. Prajna, "Barrier certificates for nonlinear model validation," *Automatica*, vol. 42, no. 1, pp. 117–126, 2006.

[20] M. A. Savageau, "Biochemical systems analysis: I. some mathematical properties of the rate law for the component enzymatic reactions," *Journal of Theoretical Biology*, vol. 25, pp. 365–369, 1969.

[21] M. A. Savageau, "Biochemical systems analysis: I. the steady-state solutions for an n-pool system using a power law approximation," *Journal of Theoretical Biology*, vol. 25, pp. 370–379, 1969.

[22] E. O. Voit, *Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists*. Cambridge University Press, 2000.