

Model Discrimination of Chemical Reaction Networks by Linearization

D. Georgiev, M. Fazel, and E. Klavins

Abstract—Systems biologists are often faced with competing models for a given experimental system. Performing experiments can be time-consuming and expensive. Therefore, a method for designing experiments that, with high probability, discriminate between competing models is desired. In particular, biologists often employ models comprised of polynomial ordinary differential equations that arise from biochemical networks. Unfortunately, the model discrimination problem for such systems is computationally intractable. Here, we examine the *linear discrimination problem*: given two systems of linear differential equations with the same input and output spaces, and uncertain parameters, determine an input that is guaranteed to produce different outputs. In this context, we show that (1) if linearizations of the two nonlinear models can be discriminated, then so can the original nonlinear model; and (2) we show a class of systems for which the linear discrimination problem is convex. The approach is illustrated on a biochemical network with an unknown structure.

I. INTRODUCTION

Systems biologists are often faced with competing dynamical models of experimental systems. For example, a newly discovered regulatory protein identified as an inhibitor may act on one of several possible genes in a pathway, leading to different possible biochemical network models. Frequently, the model structure may be known, but the values of chemical rate constants may be almost entirely unavailable. Furthermore, in many settings the usually high-dimensional full state of the system is not directly observable. Instead, only some low dimensional output such as a few fluorescently tagged protein concentrations, can be measured.

One way to distinguish between different models is to create knockout cell-lines and examine the steady-state behavior of the resulting mutants. However, such experiments are tedious and may be quite difficult depending on how vital the considered network is to the normal operation of the cell. An alternative approach is to run dynamical experiments on cells and interpret their response to perturbations to distinguish between competing models. For example, a nutrient or a chemical signal can be changed in a time-varying manner (i.e. it is an input signal to the system) and the intensity of a fluorescent marker incorporated into the network can be observed, as was recently demonstrated with osmotic pressure regulation in yeast [1].

Dynamical experiments are still expensive to set up, but they are potentially less invasive and much more informative

than static experiments. Within this setting, we address the question: What experiments should be performed on the physical system to ensure that as many candidate models are invalidated by the experimental results as possible? In particular: given a set of candidate models, can we define a probability distribution, which we call a disparity certificate, over the possible input signals that maximally distinguishes the candidate models? If no input signals distinguish the candidate models, the experiments are not worth doing until better candidates are derived. If the candidate models are distinguishable, probing the actual experimental system with the resulting disparity certificate must invalidate at least one of the models (see Figure I).

Invalidation is made difficult by uncertainty in the experimental system. In systems biology, uncertainty arises particularly from unknown reaction rates. Thus, an erroneous model may not be invalidated by an experiment because the uncertain parts of the physical system conspire against the researcher to produce outputs that seem consistent with the model. In our search for a disparity certificate we take a worst-case approach, expecting that uncertainty in the system will work to make candidate models indistinguishable. The resulting problem becomes a min-max game: the input signal tries to force the outputs of two candidate models to be different while the uncertainty tries to make the outputs the same. Invalidation is also made difficult by the fact that models from systems biology are usually nonlinear, either arising from mass action kinetics or enzyme kinetics for example. In previous work [2] we show that the problem of discriminating polynomial systems could be relaxed into a problem that is more tractable, although at the expense of discretizing the system and over-approximating the solution. Here we examine the problem of discriminating linear systems.

In particular, we start with two different candidate nonlinear models that have the same input and output spaces and unknown but bounded parameters. We suppose that the systems are at equilibrium with respect to a given input and that the output of each system (but not the internal state) can be measured. We then linearize the systems about the equilibrium to obtain a linear model discrimination problem. In this context we show the following. First, if we find an input that discriminates the linearized models, this input is guaranteed to discriminate the nonlinear models. Second, we show conditions under which the linear discrimination problem is convex. For this class of systems, we show that the min-max problem described above is equivalent to a max-min problem, which can be solved using standard tools. We have implemented the method in MATLAB and

This work was supported in part by the following grants: NIH/NIDCR, #T32 DE007132-26 and ARO MURI, #W 911 NF 0710287.

M. Fazel, D. Georgiev, and E. Klavins are with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195-2500, USA mfazel@u.washington.edu, dgeorgie@u.washington.edu, klavins@u.washington.edu

demonstrate it on example problems involving candidate parameter regimes in chemical reaction networks. We also show how the method might be iterated when used to design experiments to reduce uncertainty in model parameters.

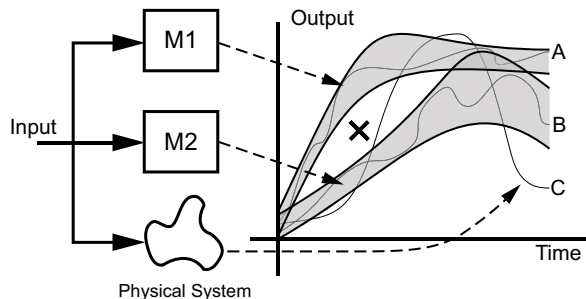


Fig. 1. Model Discrimination for experimental design. X represents the disparity certificate. Shaded regions represent all possible trajectories corresponding to the input u . Trajectories A, B, C are possible outputs of the physical system. For output A model 2 is invalidated, for output B, model 1 is invalidated and for output C both models are invalidated.

II. GENERAL PROBLEM SETUP

A. Informal Description

We consider systems of differential equation that correspond to chemical reaction networks. The control inputs are the rates at which certain chemicals are being added to the system and the measured outputs are the concentrations of certain species. The rates of the other reactions are constant over time but unknown. Given a system of differential equations and a set of possible reaction rates, we define a model as a function that maps a control input to the set of measured outputs that can be generated with the possible reaction rates. The following is the problem we want to solve.

Model Discrimination Problem (MDP) (informal). Given a pair of models with the same input and output spaces, find an input, called the disparity certificate, that yields disjoint output sets.

If a disparity certificate exists, we can hope to implement it in the experiment and learn which model does not represent the system. Identifying the inconsistent model is called model invalidation. For work on model invalidation see [3, 4].

B. Formal Description

First define the pair of candidate models that we want to discriminate. For $i = 1, 2$,

$$M_{(i)} : \begin{cases} \dot{x}_{(i)} &= f_{(i)}(x_{(i)}, \kappa_{(i)}) + B_{(i)}u, \\ y_{(i)} &= C_{(i)}x_{(i)}, \quad x_{(i)}(0) = x_{(i)}^e, \quad u(0) = u^e, \\ \kappa_{(i)} &\in [\kappa_{(i)}^-, \kappa_{(i)}^+], \end{cases} \quad (1)$$

where $x_{(i)}^e \in \mathbb{R}^{n_{(i)}}$ is a stable equilibrium point with u^e as the corresponding equilibrium input, $\kappa_{(i)} \in \mathbb{R}^{d_{(i)}}$ is the vector of reaction rates, $u \in \mathbb{R}^r$ is the control input, and $y \in \mathbb{R}^m$ is the measured output. We assume that $f_{(i)}$ is polynomial with respect to the state $x_{(i)}$ and affine with respect

to the reaction rate vector $\kappa_{(i)}$. Note that the equilibrium point is part of the model definition. From experiments, we are likely to get a measurement of the equilibrium output, but not of the entire state. Here the unknown equilibrium states are assumed as part of the candidate model. In the sequel, we will always deal with the pair of models jointly. Therefore, it is more compact to pose all problems in terms of a single composite model M defined as follows.

$$M : \begin{cases} \dot{x} &= f(x, \kappa) + Bu, \\ y &= Cx, \quad x(0) = x^e, \quad u(0) = u^e, \\ \kappa &\in [\kappa^-, \kappa^+], \end{cases} \quad (2)$$

where $x = (x_{(1)}, x_{(2)}) \in \mathbb{R}^n$, $\kappa = (\kappa_{(1)}, \kappa_{(2)}) \in \mathbb{R}^d$, and $y = y_{(1)} - y_{(2)} \in \mathbb{R}^m$. Let the distance between the outputs of the two models be defined as

$$V(t) = \int_0^t \|y_{(1)}(\tau) - y_{(2)}(\tau)\|_1 d\tau = \int_0^t \|y(\tau)\|_1 d\tau. \quad (3)$$

The following is a formal description of the problem we would like to solve.

Model Discrimination Problem (MDP) (formal). Let $\mathcal{U} = \{u \mid \|u - u^e\|_1 < \bar{u}\}$ and let $\Delta(\mathcal{U})$ be the set of all probability distributions over \mathcal{U} . The problem is to

$$\text{maximize} \min_{P \in \Delta(\mathcal{U})} \min_{\kappa^- \leq \kappa \leq \kappa^+} \mathbb{E}_P[V(T)], \quad (4)$$

where the finite T is the time horizon of interest and \mathbb{E}_P is the expected value with respect to the probability distribution P .

The model discrimination problem defined above considers random, open loop control inputs. Random inputs can outperform deterministic inputs as is well known in game theory [5], and have also been used in experiments to provide persistent excitation [6]. In molecular biology, automation has enabled mass scale, parallel experimentation, which lends itself well to the implementation of random inputs. Open loop inputs are of interest, because experimental data is often evaluated after the completion of an experiment.

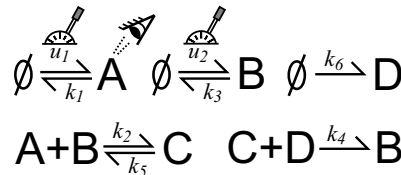


Fig. 2. (a) Chemical reactions where species B can bind to A and thereby enable A to be annihilated by D . The inputs to the system are the rates u_1 and u_2 (indicated by a lever) by which A and B are added to the system, respectively. The output (indicated by the eyeball) is the concentration of A .

Example (Mass action model). The concentrations of the species in the chemical reaction network of Figure 2 is

described by the following system of differential equations.

$$\begin{aligned}\dot{X}_A &= u_1 - \kappa_1 X_A - \kappa_2 X_A X_B + \kappa_5 X_C \\ \dot{X}_B &= u_2 - \kappa_3 X_B - \kappa_2 X_A X_B + \kappa_4 X_C X_D + \kappa_5 X_C \\ \dot{X}_C &= \kappa_2 X_A X_B - \kappa_5 X_C - \kappa_4 X_C X_D \\ \dot{X}_D &= \kappa_6 - \kappa_4 X_C X_D,\end{aligned}\quad (5)$$

where the output $y = X_A$. Note that the dependence of the state derivative on the rates κ and the inputs u is linear. For such a network, to find whether A binds to B , we would construct two models with different bounds on κ_2 and try to discriminate them. If MDP returns affirmative, we would apply the input to the experiment and use the output to find our answer.

III. BACKGROUND

Model discrimination is widely studied in the literature. Experimental design based on model discrimination has also received a considerable attention [7–11]. The traditional approach is to apply Bayesian conditioning and discriminate models using criterion such as the maximum likelihood [7, 12, 13]. Methods based on the Bayesian approach assume that the real world system behaves according to a given candidate model with a known probability. As a result, the success of these methods depends on the quality of prior information. More recent methods are based on deterministic models for uncertainty [8–10], as is the case in this paper. In [10], a method is developed that computes the initial state to maximize an upper bound on the distance between outputs of competing models. The models, however, are deterministic and, by maximizing the upper bound, the method cannot guarantee discrimination. In [8] algorithms for efficiently discriminating models from experimental data are developed, however, no inputs are considered. Finally, a great deal of literature on auxiliary input design for fault detection deals with model discrimination [12, 14]. The majority of this work is based on linear models and Bayesian conditioning. In our work, nonlinear (polynomial) systems of difference equations with unstructured uncertainty are considered and input signal distributions guaranteeing discrimination are computed.

IV. MODEL DISCRIMINATION BY LINEARIZATION

A. Linearization Setup

Use of linearization to study properties of equilibria is well documented [15]. In this paper, we show that linearization can be used to study discriminability as well. The corresponding linear discrimination problem is still difficult to solve in general. We study specific instances of the linear problem and provide solutions.

The relationship between discrimination of the nonlinear models and the linearized models is local (as will be shown in the next section). The first n_M (where $n_M = \max(n_{(1)}, n_{(2)})$) derivatives of the linearized models determine the discriminability of the nonlinear models. We use this fact to define the related, linear model discrimination problem. Let the composite system of differential equations

linearized around the equilibrium state and control input be written as

$$\begin{aligned}\delta\dot{x} &= A(\kappa)\delta x + B\delta u, \\ \delta y &= C\delta x, \quad \delta x(0) = 0,\end{aligned}\quad (6)$$

where $A(\kappa) = A_0 + \sum_{i=1}^d \kappa_i A_i$. Note, we cannot linearize with respect to the reaction rates. Deviations of the reactions rates are not necessarily small.

Define the discrete composite model by computing the time derivatives of the linearized systems.

$$M^K : \begin{cases} \bar{x}(k+1) &= A(\kappa)\bar{x}(k) + B\bar{u}(k), \\ \bar{y}(k) &= C\bar{x}(k), \bar{x}(0) = 0 \\ \kappa &\in [\kappa^-, \kappa^+], k \in \{1, \dots, K\} \end{cases}\quad (7)$$

where $\bar{u}(k) = \frac{d^k}{dt^k} \delta u(0)$. Note that we've assumed the inputs are continuously differentiable; this class of inputs is sufficient to discriminate linear systems. As before, let the distance between the outputs be $V_L(k) = \sum_{i=0}^k \|\bar{y}(i)\|_1$. The linear discrimination problem corresponding to the original MDP is defined as follows.

K-LMDP. Let $\mathcal{U}_L = \{u_L \mid \|u_L\|_1 < 1\}$ and let $\Delta(\mathcal{U}_L)$ be the set of all probability distributions over \mathcal{U}_L . The K th order linear model discrimination problem (K-LMDP) is to

$$\text{maximize}_{P \in \Delta(\mathcal{U}_L)} \min_{\kappa^- \leq \kappa \leq \kappa^+} \mathbb{E}_P [V_L(K)].\quad (8)$$

Remark 1. Although one can solve K-LMDP for any $K > 0$, discriminability of the models is determined for $K = \max(n_{(1)}, n_{(2)})$.

The remainder of the paper is concerned with the K-LMDP. First, we show how it is related to the MDP. Then we show that the minimization and maximization operations can be switched if V_L is convex with respect to κ . Finally, we use the minimax-maximin equivalence to derive a solution algorithm for the convex version of K-LMDP.

B. Sufficiency of Linear Discrimination

In this section we will motivate solving K-LMDP by showing that if $V_L(k)$ is positive for some $k \geq 0$, then $V(k)$ will be positive as well. In the remainder of this section, without loss of generality, we make the following assumptions:

- 1) $u(t)$ is constant (constant input is sufficient for linear system), and a part of the state vector,
- 2) $x_e = 0, u_e = 0$,
- 3) The unknown parameters vector κ is a constant.

Assumption 1 is without loss of generality. To discriminate a linear system, only constant inputs need to be considered. Moreover, even if one discriminates the linear system using a higher class of n -differentiable inputs, the results hold since such inputs reduce to constant inputs by appending their derivatives to the state. Assumption 2 simply means we've translated the equilibrium to zero. There is no loss of generality since this translation preserves the polynomial/affine system structure. Assumption 3 is valid for a large class of biological systems, e.g., all systems governed by mass

action kinetics. Under the above assumptions, the nonlinear composite system in Equation 2 has the following form:

$$\begin{aligned} \dot{x} &= Ax + h(x), \\ y &= Cx, \quad x(0) = (0, u(0)). \end{aligned} \quad (9)$$

Lemma 1. $\lim_{x \rightarrow 0} \frac{|h^{(\ell)}(x)|}{|x|} \rightarrow 0$, where $h^{(\ell)} = \frac{\partial^\ell}{\partial t^\ell} h$.

Proof: Show, by induction that $h_i^{(\ell)}(x)$ has the form $x^T Q_i^\ell(x) x$, where all entries of $Q_i^\ell(x)$ are polynomials.

- 1) True for $h^{(0)}$ by assumption on h .
- 2) Suppose $h_i^{(j)}(x)$ has the form $x^T Q_i^j x$, for $j \leq \ell$. Write $h^{(j)}(x) = Q^j(x) x$, where entries of $Q^j(x)$ are polynomials of order 1 or greater. Then

$$\begin{aligned} h_i^{(\ell+1)}(x) &= 2(Ax + Q^0(x)x)^T Q_i^\ell(x) x \\ &+ x^T \frac{\partial Q_i^\ell}{\partial x} x (Ax + Q^0(x)x) = x^T \tilde{Q}(x) x. \end{aligned}$$

Lemma 2. If $\|CA^\ell \nu\| > 0$, for some $\ell \leq n$, then there exist constants $\rho^+ > 0, \tau^+ > 0$ such that $\|y^\ell(\tau)\| > 0$ for all $\tau < \tau^+$ and $x(0) = \rho \nu, \rho < \rho^+$.

Proof: Compute output derivatives for $\tau = 0$ and $x(0) = \rho \nu$. Throughout the proof we omit the dependence on τ whenever appropriate.

$$\begin{aligned} y^{(\ell)} &= Cx^{(\ell)} \\ &= C \left(A^\ell x + A^{\ell-1} h(x) + \dots + h^{(\ell-1)}(x) \right). \end{aligned}$$

From Lemma 1, we know that $\lim_{\rho \rightarrow 0} \frac{h^{(k)}(\rho \nu)}{|\rho|} \rightarrow 0$. So choose $\rho^+ > 0$ to satisfy

$$\|CA^\ell \nu\| > \|C(A^{\ell-1} h(\rho \nu) + \dots + h^{(\ell-1)}(\rho \nu))\| / \rho,$$

for all $\rho \leq \rho^+$. It follows that $\|y^{(\ell)}(0)\| > 0$ when $x(0) = \rho \nu$. Finally, we know that there exists a time interval $[0, \tau_M], \tau_M > 0$, in which $y^{(\ell)}(\tau)$ is continuous with respect to τ . Therefore, there exists a constant τ^+ such that $\|y^{(\ell)}(\tau)\| > 0$ for all $\tau < \tau^+$. ■

Theorem 1. If $\|CA^\ell \nu\| > 0$, for some $\ell \leq n$, and $x(0) = \rho \nu$, then there exists a constant $\rho^+ > 0$ such that, for all $t > 0$ and for all $\rho < \rho^+, V(t) > 0$.

Proof: The proof is by contradiction. Choose ρ^+ according to Lemma 2. Let $u = \rho u_L, 0 < \rho \leq \rho^+$, and suppose that $V(\tau) = 0$ for all $\tau < t$. This implies that $y(\tau) = 0$ for all $0 \leq \tau \leq t$. Since all functions that are zero on an open interval must also have zero derivatives on that interval, this is a contradiction to Lemma 2. ■

The condition of Theorem 1 is equivalent to the input ν discriminating the linear system. If $\|CA^\ell \nu\| = 0$ for all $\ell \leq n$, then the output of the linear system remains zero for all time. Therefore, the theorem implies that if an input discriminates the linear system, then the same input scaled will discriminate the nonlinear system.

C. Maximin and Minimax Equivalence

To solve K -LM DP, it is useful to be able to switch the minimization and maximization operations without changing the optimal value. In this section we first rewrite the problem in its static form.

Observation 1. Problem has the following form.

$$\text{maximize} \min_{P \in \Delta(\mathcal{U}_L)} \mathbb{E}_{L \in \mathcal{L}} [\|Lu\|_1]. \quad (10)$$

We proceed under the assumption that \mathcal{L} is closed, convex, and bounded. This will always be true for $K \leq 2$. The set is not in general convex for $K > 2$; exploring convex relaxations for this set is a direction for future work. Under this assumption on \mathcal{L} we have the following result.

Theorem 2. Let \mathcal{U}_{LV} be the finite set of vertices of \mathcal{U}_L in the positive orthant. Then

$$\begin{aligned} \min_{L \in \mathcal{L}} \max_{u \in \mathcal{U}_L} \|Lu\|_1 &= \max_{P \in \Delta(\mathcal{U}_L)} \min_{L \in \mathcal{L}} \mathbb{E}_P [\|Lu\|_1], \\ &= \max_{P \in \Delta(\mathcal{U}_{LV})} \min_{L \in \mathcal{L}} \mathbb{E}_P \|Lu\|_1. \end{aligned} \quad (11)$$

Proof: The maximization problem in the left hand side of Equation 11 can be solved explicitly and we can rewrite the left hand side as

$$\text{minimize} \max_{L \in \mathcal{L}} \sum_j |L_{ij}|. \quad (12)$$

This shows that the maximization in the minimax problem only needs to be carried out over the vertices of \mathcal{U}_L . If we restrict \mathcal{U} to \mathcal{U}_L , the domain of the cost function $\mathbb{E}_P [\|Lu\|_1]$ becomes finite dimensional as well as compact. Since the cost function is also convex-concave with respect to L and P , it is well known from the theory of zero sum games that it has a saddle point in the set $\mathcal{L} \times \Delta(\mathcal{U}_L)$, Theorem 4.6 in [16]. Thus it holds that

$$\min_{L \in \mathcal{L}} \max_j \sum_i |L_{ij}| = \max_{P \in \Delta(\mathcal{U}_{LV})} \min_{L \in \mathcal{L}} \mathbb{E}_P \|Lu\|_1. \quad (13)$$

The desired result follows.

$$\begin{aligned} \max_{P \in \Delta(\mathcal{U}_{LV})} \min_{L \in \mathcal{L}} \mathbb{E}_P \|Lu\|_1 &\leq \max_{P \in \Delta(\mathcal{U}_L)} \min_{L \in \mathcal{L}} \mathbb{E}_P \|Lu\|_1 \\ &\leq \min_{L \in \mathcal{L}} \max_j \sum_i |L_{ij}| \\ &= \max_{P \in \Delta(\mathcal{U}_{LV})} \min_{L \in \mathcal{L}} \mathbb{E}_P \|Lu\|_1. \end{aligned}$$

Thus we only need to consider the control inputs that are the positive orthant vertices of \mathcal{U}_L . We develop this idea in the next section to construct a solution algorithm.

D. 2nd Order Solution Algorithm

In this section we derive a solution algorithm for K -LM DP when $K = 2$. First, we rewrite the full problem, explicitly listing all the variables and constraints. Note that this problem includes an optimization variable that belongs to an infinite dimensional set of probability distributions.

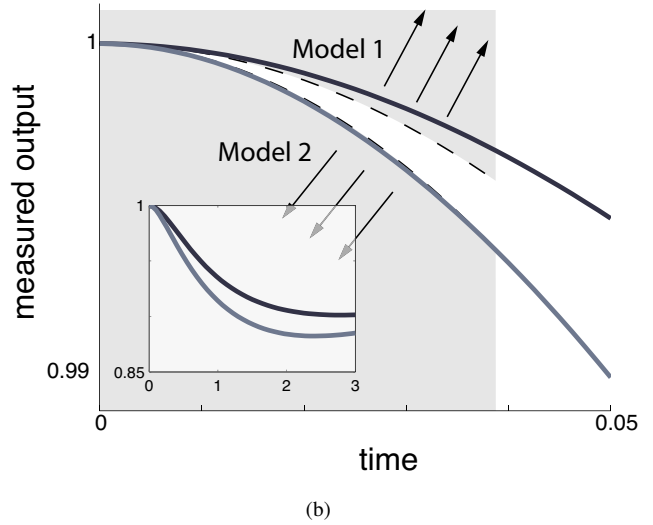
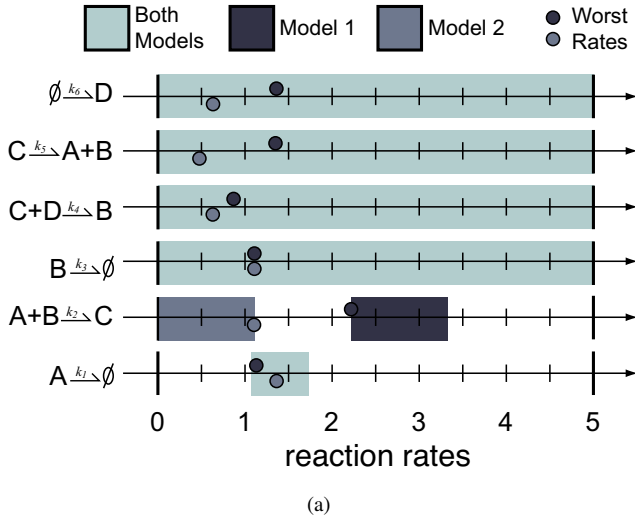


Fig. 3. (a) Allowable reaction rate sets for models 1 and 2. Each uncontrolled reaction is listed along the y-axis and the interval containing the corresponding reaction rate is shaded on the axis to the right. The reaction rates of model 1 can take any values in light or dark grey regions. The reaction rates of model 2 can take any values in light or medium grey regions. Values of the worst case rates from 2-LMDP are also shown. (b) The affirmative solution to the model discrimination problem predicts that (initially) the output output of model 1 (model 2) will lie in the upper (lower) shaded region. The outputs of the nonlinear models generated by the worst rates (see (a)) are also shown. The inset shows the responses on a longer time-scale.

2-LMDP (infinite dimensional). The 2nd order Linear Model Discrimination Problem is to

$$\text{maximize}_P \min_{\kappa} \mathbb{E}_P [\|Lu\|_1], \quad (14)$$

subject to

$$P \in \Delta(\mathcal{U}_L), \quad L = L_0 + \sum_i \kappa_i L_i, \quad \kappa \in [\kappa^-, \kappa^+]$$

where

$$L_0 = \begin{pmatrix} C^T B & 0 \\ C^T A_0 B & C^T B \end{pmatrix}, \quad L_i = \begin{pmatrix} 0 & 0 \\ C^T A_i B & 0 \end{pmatrix},$$

for $i = 1, \dots, d$.

Next, we apply Theorem 2 to derive an equivalent finite linear program. Let $P^*(e_i) = p_i^*$ where e_i is the i th vector of the standard Euclidean basis, and form the following dual problem of the finite, minimax reformulation of 2-LMDP.

2-LMDP (finite dimensional).

$$\text{maximize}_{p, \lambda_{\ominus}, \lambda_{\oplus}, \Lambda_{\oplus}, \Lambda_{\ominus}} \lambda_{\ominus}^T \kappa^- - \lambda_{\oplus}^T \kappa^+ + \text{tr} [(\Lambda_{\oplus} - \Lambda_{\ominus})^T L_0],$$

subject to

$$\begin{aligned} p &\geq 0, & \sum p_i &= 1, & p_j - (\Lambda_{\oplus} + \Lambda_{\ominus})_{ij} &= 0 \\ \lambda_{\ominus} &\geq 0, & \lambda_{\oplus} &\geq 0, & (\Lambda_{\oplus})_{ij} &\geq 0, & (\Lambda_{\ominus})_{ij} &\geq 0 \\ \text{tr} (\Lambda_{\oplus}^T L_i) - \text{tr} (\Lambda_{\ominus}^T L_i) + (\lambda_{\oplus} - \lambda_{\ominus})_i &= 0 \end{aligned}$$

Example (Model Discrimination). A solution of 2-LMDP (finite dimensional) was used to discriminate candidate models for the reaction network in Figure 2. Two candidate models were generated. The model dynamics are given by Equation 5. The equilibrium concentrations are all equal to 1. Figure 3(a) shows the bounds on the reaction rates of both

models. The models differ in the bounds on κ_2 , the rate at which A binds to B . Model 1 imposes a higher rate on κ_2 increasing the annihilation of A . The two models were found to be disparate, with the optimal cost of 2-LMDP equal to $1/\frac{1}{9}$. The optimal distribution P^* over the control inputs has the following form,

$$P^* \left(u_1(0) = 1, u_2(0) = \frac{du_1}{dt}(0) = \frac{du_2}{dt}(0) = 0 \right) = 1.$$

This same input, with some scaling, discriminates the nonlinear models (see Section IV-B). Figure 3(b) illustrates the predicted disjoint sets of the nonlinear model outputs. The worst case rates κ^* (see Figure 3(a) for numeric values) were also computed by minimizing the expected cost in Equation 14 with P equal to P^* . Figure 3(b) illustrates the nonlinear model outputs that are generated by κ^* .

The next section builds on this example to give an iterative model generation and discrimination algorithm that can be used to reduce a wide range of candidate models to a more informative family.

V. REPEATED DISCRIMINATION IN EXPERIMENTS

The discrimination problem solved in Section IV-D is used to construct a system identification algorithm. The purpose of the algorithm is the invalidate as many members of a given family of candidate models. A viable candidate must 1) yield the correct equilibrium, and 2) generate the experimental output within a given tolerance. The algorithm is outlined below.

Algorithm (Automated Model Invalidation).

- **Inputs**

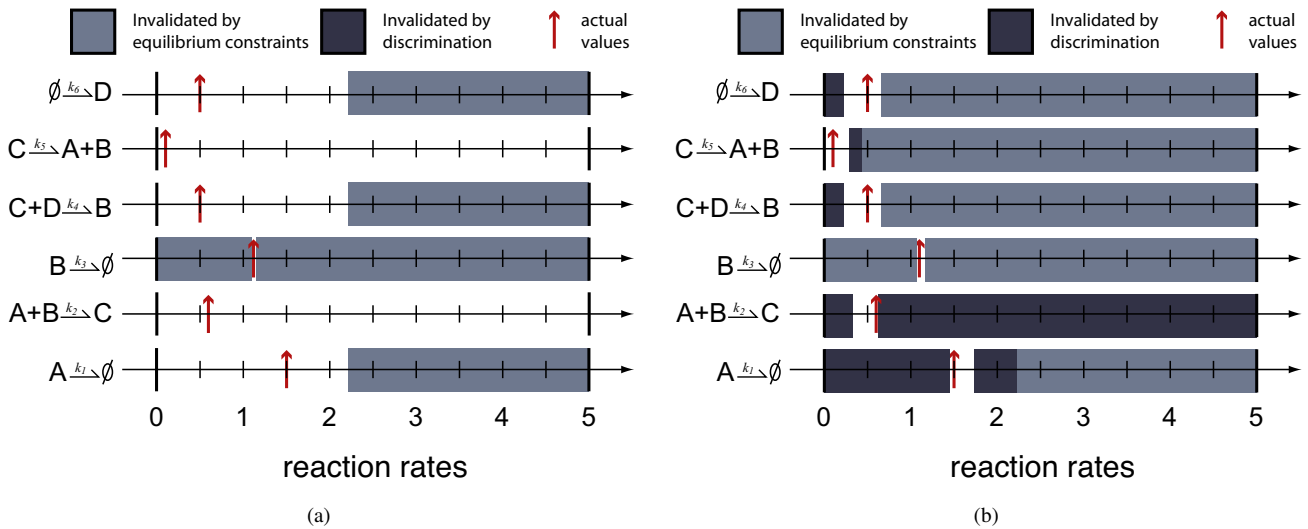


Fig. 4. Outcomes of the Automated Model Invalidation Algorithm. Each uncontrolled reaction is listed along the y-axis. The regions to the right of the reactions indicate whether the reaction rate value was invalidated using equilibrium constraints, invalidated using a discriminating experiment, or remains viable. The actual reaction rates (shown for reference only) that define the physical system are shown by an arrow in the corresponding row. (a) Outcome of the AMIA that is based only on equilibrium constraints. (b) Outcome of the AMIA that uses model discrimination.

- 1) *Dynamics: A nonlinear system of ordinary differential equations*

$$\begin{aligned} \dot{x} &= f(x, \kappa) + Bu \\ y &= Cx, x(0) = x^e, u(0) = u^e, \end{aligned} \quad (15)$$

where f is a polynomial function that is affine with respect to κ .

- 2) *Bounding set: A set \mathcal{K} that contains all possible values of κ .*

• Iteration Steps

- 1) *Generate candidate models: for a given constant $\rho > 0$, partition \mathcal{K} into blocks $\{\mathcal{K}_1, \dots, \mathcal{K}_N\}$ with norm at most ρ . For all $i \in \{1, \dots, N\}$, construct a model $M_{(i)}$ from the dynamics in Equation 15 and the set \mathcal{K}_i .*
- 2) *Invalidate inconsistent models: a model is invalidated if for all possible values of κ , x_e and u_e do not generate an equilibrium point.*
- 3) *Find disparate models: any two models are disparate if they can be discriminated using 2-LMDP.*
- 4) *Execute experiment: for any two disparate models, run the discriminating experiment and invalidate either model if it is inconsistent with the experimental output.*
- 5) *Reduce ρ , redefine \mathcal{K} to be the union of all valid blocks, and return to Step 1. Stop when ρ is sufficiently small.*

The guiding philosophy of the Automated Model Invalidation Algorithm is that there exists a single differential equation system (Equation 15) that parametrizes all candidate models of interest. Such a system may be formed from a single chemical reaction network that includes all possible reactions.

To demonstrate the value of model discrimination for model invalidation, we've implemented a simple version of the algorithm above that makes the following simplifications. The set \mathcal{K} is restricted to be an interval $[\kappa^-, \kappa^+]$. For Step 1, the set \mathcal{K} is partitioned into blocks that are themselves intervals. For Steps (2-3), we considered only those blocks that result in an interval when subtracted from \mathcal{K} . For Step 4, we replaced the physical experiment with an oracle that knows the actual values of κ and reveals an invalid model given a disparate pair of models.

Example (Repeated Discrimination). The simple version of the algorithm was applied to the chemical reaction network in Figure 2. The general differential equation system (Input 1) is given by Equation 5 and the bounding set \mathcal{K} (Input 2) is given by the interval $[0, 5]^6$.

We ran the algorithm with and without model discrimination (Steps 3 and 4). The results are shown in Figure (a) and (b), respectively. Unshaded regions indicate viable values of κ , lightly shaded regions indicate values of κ invalidated based on the equilibrium constraints, and dark shaded regions indicate values of κ invalidated by running a discriminating experiment given by 2-LMDP. With the exception of κ_3 , the degradation rate of B , equilibrium constraints alone do not invalidate significant portions of \mathcal{K} . With model discrimination, we can identify viable regions of \mathcal{K} arbitrarily close to the actual values of κ . The accuracy is only limited by the discrimination threshold (a positive threshold on the optimal cost of 2-LMDP below which models are not considered disparate).

VI. CONCLUSION

We considered the problem of model discrimination of polynomial systems with unknown parameters that lie in a given set, and proposed a linearization-based method. We

showed the method gives a sufficient condition for discrimination of the nonlinear system. The problem of finding inputs that guarantee discrimination is in general computationally intractable, but for the case of $K = 2$ (which corresponds to discrimination based on the value of the output and its first derivative only) the feasible set becomes convex. If we use ℓ_1 norm to measure the distance between the model outputs, the problem is shown to reduce to a linear program. This method can be coupled with simple branch-and-bounds over the parameter space to help with automated model invalidation, as shown for a simple illustrative example. Note that it appears that the affine structure with respect to input u is not needed. One can linearize around the equilibrium input as well as the equilibrium state and then proceed in Section IV-B (the only related section) in the same way.

Future work will explore efficient convex relaxations to extend the method for $K > 2$, allowing discrimination based on higher order derivatives and side information. Relaxations that find tight bounds on the feasible parameter region will be very helpful in adapting this approach for model invalidation. The work here takes the first step towards these goals.

REFERENCES

- [1] J. T. Mettetal, D. Muzzey, C. Gmez-Urbe, and A. van Oudenaarden, "The frequency dependence of osmo-adaptation in *saccharomyces cerevisiae*," *Science*, vol. 319, no. 5862, 2008.
- [2] D. Georgiev and E. Klavins, "Model discrimination of polynomial systems via stochastic inputs," in *Proc. IEEE Conf. on Decision and Control*, 2008.
- [3] S. Prajna, "Barrier certificates for nonlinear model validation," *Automatica*, vol. 42, no. 1, 2006.
- [4] X. L. T. O. J. G. A. P. S. P. T.-M. Yi, M. Fazel and J. Doyle, "Application of robust model validation using sostoools to the study of g-protein signaling in yeast," in *Proc. Foundations of Systems Biology and Engineering (FOSBE)*, August 2005.
- [5] R. B. Myerson, *Game Theory, Analysis of Conflict*. Harvard University Press, 1991.
- [6] R. D. Nowak and B. D. V. Veen, "Random and pseudorandom excitation sequences for nonlinear system identification," in *IEEE Winter Workshop on Nonlinear Digital Signal Proc.*, 1993.
- [7] R. Horn, "Statistical methods for model discrimination. Applications to gating kinetics and permeation of the acetylcholine receptor channel," *Biophysical Journal*, vol. 51, no. 2, 1987.
- [8] R. Feely, M. Frenklach, M. Onsum, T. Russi, A. Arkin, and A. Packard, "Model discrimination using data collaboration," *Journal of Physical Chemistry A*, vol. 110, 2006.
- [9] J. F. Apgar, J. E. Toettcher, D. Endy, F. M. White, and B. Tidor, "Stimulus design for model selection and validation in cell signaling," *PLoS Comp. Biol.*, vol. 4, no. 2, 2008.
- [10] A. Papachristodoulou and H. El-Samad, "Algorithms for discriminating between biochemical reaction network models: Towards systematic experimental design," in *Proc. of the American Control Conf.*, 2007.
- [11] A. Kremling, S. Fischer, K. Gadkar, F. J. Doyle, T. Sauter, E. Bullinger, F. Allgower, and E. D. Gilles, "A benchmark for methods in reverse engineering and model discrimination: Problem formulation and solutions," *Genome Research*, vol. 14, 2004.
- [12] L. Blackmore and B. Williams, "Finite horizon control design for optimal model discrimination," in *Proc. IEEE Conf. Decision & Control*, 2005.
- [13] K. Felsenstein, "Optimal Bayesian design for discrimination among rival models," *Comp. Stat. and Data Anal.*, vol. 14, no. 4, 1992.
- [14] T. Hatanaka and K. Uosaki, "Optimal auxiliary input for fault detection of systems with model uncertainty," in *Int. Conf. on Contr. App.*, 1999.
- [15] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 1995.
- [16] T. Basar and G. Olsder, *Dynamic noncooperative game theory*. SIAM Classics in Applied Mathematics, second ed., 1999.