

Optimal Experiment Design and Leveraging Competition for Shared Resources in Cell-Free Extracts

Wolfgang Halter, Frank Allgöwer, Richard M. Murray, and Andras Gyorgy

Abstract—The fact that genes compete for shared cellular resources poses a fundamental challenge when identifying parameters of genetic parts. A recently developed model of gene expression tackles this problem by explicitly accounting for resource competition. In addition to accurately describing experimental data, this model only depends on a small number of easily identifiable parameters with clear physical interpretation. Based on this model, we outline a procedure to select the optimal set of experiments to characterize biomolecular parts in synthetic biology. Additionally, we reveal the role competition for shared resources plays, provide guidelines how to minimize its detrimental effects, and how to leverage this phenomenon to extract the most information about unknown parameters. To illustrate the results, we consider the case of part characterization in cell-free extracts, treat plasmid DNA concentrations as decision variables, and demonstrate the significant performance difference between naïve and optimal experiment design.

I. INTRODUCTION

One of the fundamental goals of synthetic biology is to create large-scale complex systems enabling the precise control of living organisms, both at the cellular and population levels [1]. Today, creating even simple modules requires numerous iterative cycles of designing, building and testing components [2], [3]. Unfortunately, this approach relying on the creation of vast part libraries [4] quickly becomes impractical with increasing circuit complexity. Alternatively, computational tools and control theoretic ideas leveraging predictive models support the development of general design principles that are essential for the rational forward engineering of large-scale synthetic systems [5], [6], [7], [8]. Much of the mystery of today’s genetic engineering stems from the lack of such predictive models and fully characterized parts allowing systems-level designs.

Recently, considerable attention has been focused on various sources of context-dependence that cause unwanted coupling among different genetic modules [9], [10], [11]. One major source of context-dependence is competition for shared resources [12], [13], [14]. This phenomenon causes coupling between virtually any two components that share the same machinery (RNA polymerase, ribosomes, degradation enzymes, etc.), both *in vivo* and *in vitro* [15], [16].

W. Halter and F. Allgöwer are with the Institute for Systems Theory and Automatic Control, University of Stuttgart, 70569 Stuttgart, Germany. halter@ist.uni-stuttgart.de

R.M. Murray is with the Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA. murray@cds.caltech.edu

A. Gyorgy is with the Department of Electrical and Computer Engineering, New York University Abu Dhabi, Saadiyat Island, United Arab Emirates. ag6378@nyu.edu

As a result, part characterization must include quantification of resource sequestration so that the resulting coupling phenomenon can be predicted when parts are co-expressed [15], [17]. This way unwanted coupling effects can be taken into account during systems-level design, thus ensuring correct behavior upon interconnection [18], [19].

To this end, in [15] we developed a mathematical model considering the *in vitro* experimental data originally obtained in [12]. This mechanistic model explicitly accounts for the limited availability of shared transcriptional/translational resources and reveals how the expression of one gene affects that of another via competition for these shared resources. In addition to correctly describing the expression profile of a single gene, we demonstrated that we can successfully predict expression profiles of multiple co-expressed parts characterized separately, thus accounting for loading due to resource competition.

In this paper, we build upon these earlier results in [15] and quantify the information content of candidate experiments in the presence of measurement noise. To this end, we consider parameter estimation from the perspective of both Least Squares Estimation and Maximum Likelihood Estimation. We derive the explicit expression characterizing the information content of a measurement at any given time point. Further, leveraging the analytic nature of this expression, we reveal how competition for shared resources affects the information content of experiments when expressing various copies of a single gene. Following this, we uncover how to extract more information about a part of interest by counter-intuitively co-expressing it with another part. Besides outlining how to select the optimal set of experiments to characterize biomolecular parts in synthetic biology, we also reveal the role competition for shared resources plays in terms of the parameter identification problem. We further provide guidelines how to minimize the detrimental effects of competition, and how to leverage this phenomenon by extracting the most information from potential experiments.

This paper is organized as follows. First, we briefly introduce the mathematical model of gene expression that accounts for the limited availability of shared transcriptional resources. Following this, we quantify the information content of candidate experiments and demonstrate not only how to select the optimal set of experiments yielding the sharpest parameter estimates, but also how the very fact that genes compete for resources can be leveraged to obtain better parameter estimates. Finally, we revisit the experimental data obtained in [12] to compare the performance difference between naïve and optimal experiment design.

II. MODEL AND PROBLEM FORMULATION

Here, we briefly introduce the mathematical model derived in [15] to account for the limited availability of shared cellular resources, together with the problem formulation.

A. System Dynamics

Consider a set of n constitutively expressed genes, with gene i expressing mRNA x_i for $i = 1, \dots, n$, respectively, and introduce $x = (x_1, \dots, x_n)^\top$. To model the limited availability of shared cellular resources, according to [15], the network dynamics are captured by

$$\dot{x}_i = \alpha_i \frac{\frac{d_i}{\kappa_i}}{1 + \sum_{j=1}^n \frac{d_j}{\kappa_j}} - \gamma_i x_i, \quad (1)$$

with $x(0) = 0$, where the parameters are defined as follows. The rate constants α_i and γ_i denote mRNA synthesis and degradation rates, respectively; d_i denotes the DNA concentration encoding gene i ; and κ_i is the dissociation constant of RNA polymerase to the promoter of gene i . For simplicity, we only consider transcriptional dynamics and constitutive genes, but the results can be easily extended to include translational dynamics and regulated genes, as discussed in [15].

Importantly, if we consider the case when $d_i \ll \kappa_i$, the model in (1) simplifies to

$$\dot{x}_i = \frac{\alpha_i}{\kappa_i} d_i - \gamma_i x_i, \quad (2)$$

the standard model of gene expression [20], corresponding to when shared transcriptional resources (RNA polymerase) are available in unlimited amounts, thus competition for these resources is negligible.

B. Experimental Setup and Measurements

Consider J experiments such that the DNA concentrations are $d_i = d_i^{(j)}$ for $i = 1, \dots, n$, and define $d^{(j)} := (d_1^{(j)} \dots d_n^{(j)})^\top \in \mathbb{R}^n$ for $j = 1, \dots, J$. Introducing the vector $\theta := (\alpha_1, \gamma_1, \kappa_1, \dots, \alpha_n, \gamma_n, \kappa_n)^\top$ of unknown parameters, we can rewrite (1) for $i = 1, \dots, n$ as

$$\dot{x} = g(x, \theta, d^{(j)}) \quad \text{with} \quad x(0) = 0. \quad (3)$$

Without measurement noise, let

$$\bar{y}^{(j)}(t) = h(x(t, \theta, d^{(j)})) \in \mathbb{R}^q \quad (4)$$

denote the output of the system in (3), where $h: \mathbb{R}^n \rightarrow \mathbb{R}^q$ is typically such that $h(x) = (x_{i_1} \ x_{i_2} \ \dots \ x_{i_q})^\top$ where $i_j \in \{1, 2, \dots, n\}$, i.e., a subset of the state variables are measured. If we sample the system output $\bar{y}^{(j)}$ at K different time points t_1, \dots, t_K , then by integrating (3) we obtain that $\bar{y}_{j,k}(\theta, d^{(j)}) := \bar{y}^{(j)}(t_k)$ is given by

$$\bar{y}_{j,k}(\theta, d^{(j)}) = h\left(\int_0^{t_k} g(x, \theta, d^{(j)}) dt\right) \in \mathbb{R}^q, \quad (5)$$

and introduce $\bar{y} := (\bar{y}_{1,1}^\top \ \dots \ \bar{y}_{J,K}^\top)^\top$. In the presence of additive measurement noise the output becomes

$$Y_{j,k} = \bar{y}_{j,k}(\theta, d^{(j)}) + Z_{j,k} \quad (6)$$

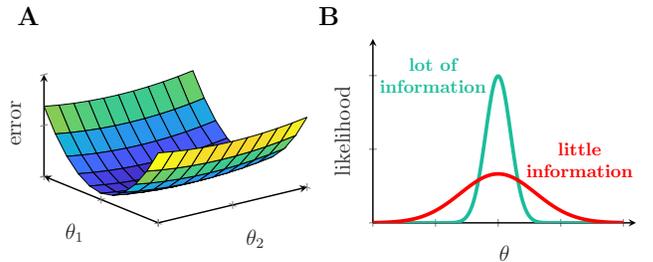


Fig. 1. (A) An optimization algorithm converges to the minimum $\hat{\theta}_{\text{LSE}}$ in θ_1 faster than in θ_2 . (B) The same small change in θ yields a greater decrease in case of the narrower curve than in case of the wider curve.

for $j = 1, \dots, J$ and $k = 1, \dots, K$, where the random variables $Z_{j,k} \sim \mathcal{N}_q(0, \Sigma)$ follow an independent and identically distributed (i.i.d.) q -dimensional multivariate normal distribution with zero mean and symmetric covariance matrix $\Sigma \in \mathbb{R}^{q \times q}$, assumed to be known (e.g., from equipment calibration).

C. Problem Formulation

The measurements $y := (y_{1,1}^\top \ \dots \ y_{J,K}^\top)^\top \in \mathbb{R}^{qJK}$ correspond to the J experiments with DNA concentrations $d := ((d^{(1)})^\top \ \dots \ (d^{(J)})^\top)^\top \in \mathbb{R}^{nJ}$ and K time points. Based on this, the two most common estimators are the Least Square Estimator $\hat{\theta}_{\text{LSE}}$ and the Maximum Likelihood Estimator $\hat{\theta}_{\text{MLE}}$. In particular, we have that

$$\hat{\theta}_{\text{LSE}} = \arg \min_{\theta} L(y, \theta, d), \quad (7)$$

where

$$L(y, \theta, d) := \frac{1}{2} [y(d) - \bar{y}(\theta, d)]^\top [y(d) - \bar{y}(\theta, d)] \quad (8)$$

quantifies the total error between measurements and model predictions. Similarly, we have that

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta} f_Y(y|\theta, d), \quad (9)$$

where $f_Y(y|\theta, d)$ is the probability density function of observing the samples $Y := \{Y_{1,1}, \dots, Y_{J,K}\}$. In what follows, we omit the subscript of $\hat{\theta}$ as under a Gaussian model with i.i.d. noise, as in our case, we have that $\hat{\theta}_{\text{LSE}} = \hat{\theta}_{\text{MLE}}$ [21].

Here, we are interested in the quality of the estimator $\hat{\theta}$. To illustrate this, consider first the error function depicted in Fig. 1A, where the error is less sensitive to changes in θ_2 than in θ_1 , yielding significant uncertainty about θ_2 , thus an optimization algorithm converges slower in θ_2 than in θ_1 . Similarly, considering the two likelihood functions in Fig. 1B, although they share the same $\hat{\theta}$, the parameter uncertainty is smaller corresponding to the narrower one (green). In the following, we focus on how the error/likelihood surface is shaped in the vicinity of $\hat{\theta}$, thus studying local observability properties using a linearization approach. A more general description of this problem for a wider model class is provided in [22].

Importantly, the surfaces $L(y, \theta, d)$ and $f_Y(y|\theta, d)$ in Fig. 1 around the estimator $\hat{\theta}$ depend on the experimental

conditions, i.e., our choice of DNA concentrations d . Therefore, we seek to answer the following questions. How do we quantify the quality of parameter estimates? What is the set of optimal decision variables d ? Most importantly: how can we leverage competition for shared resources to obtain optimal error/likelihood landscapes?

III. RESULTS

We first briefly review the results concerned about the local description of $L(y, \theta, d)$ and $f_Y(y|\theta, d)$. After providing a simple procedure to quantify the information content of candidate experiments, thus finding the optimal set, we reveal the role that competition for shared resources plays. Note that while the dynamics of x_i and x_j in (3) are decoupled, identification of the unknown parameters in θ are coupled.

A. Local Behavior of the Error and Likelihood Functions

Assume that there exists a parameter vector $\hat{\theta}$ such that

$$\bar{y}_{j,k}(\hat{\theta}, d^{(j)}) = y_{j,k} \quad \forall j \in [1, J], k \in [1, K], \quad (10)$$

so that $L(y, \hat{\theta}, d) = 0$ together with

$$\left. \frac{\partial L(y, \theta, d)}{\partial \theta_1} \right|_{\hat{\theta}} = \left. \frac{\partial L(y, \theta, d)}{\partial \theta_2} \right|_{\hat{\theta}} = \dots = \left. \frac{\partial L(y, \theta, d)}{\partial \theta_{3n}} \right|_{\hat{\theta}} = 0.$$

In this case, we can approximate $L(y, \theta, d)$ around $\hat{\theta}$ as

$$\begin{aligned} L(y, \theta, d) &\approx L(y, \hat{\theta}, d) + \left. \frac{\partial L(y, \theta, d)}{\partial \theta} \right|_{\hat{\theta}} [\theta - \hat{\theta}] + \\ &[\theta - \hat{\theta}]^\top \left. \frac{\partial^2 L(y, \theta, d)}{\partial \theta^2} \right|_{\hat{\theta}} [\theta - \hat{\theta}] \\ &= [\theta - \hat{\theta}]^\top H(\hat{\theta}, d) [\theta - \hat{\theta}] \end{aligned} \quad (11)$$

with the Hessian matrix

$$H(\hat{\theta}, d) := \left. \frac{\partial^2 L(y, \theta, d)}{\partial \theta^2} \right|_{\hat{\theta}}. \quad (12)$$

As the eigenvalues and eigenvectors of $H(\hat{\theta}, d)$ determine the form of the paraboloid (Fig. 1A), they strongly influence the convergence properties of any gradient based optimization method.

For the likelihood function in (9), we follow a similar approach and first define the log-likelihood function $l(y|\theta, d) := \log f_Y(y|\theta, d)$. Due to the monotonicity of $\log(\cdot)$, note that $\hat{\theta} = \arg \max_{\theta} l(y|\theta, d)$. For $\theta \in \mathbb{R}^r$ we have from [23] that $\hat{\theta}$ approximately follows the r -dimensional multivariate normal distribution $\hat{\theta} \sim \mathcal{N}_r(\hat{\theta}, \mathcal{I}^{-1}(\hat{\theta}, d))$, where $\mathcal{I}(\hat{\theta}, d) \in \mathbb{R}^{r \times r}$ is the Fisher information matrix defined as

$$\mathcal{I}(\theta, d) := \int \left[\frac{\partial l(y|\theta, d)}{\partial \theta} \right]^\top \frac{\partial l(y|\theta, d)}{\partial \theta} f_Y(y|\theta, d) dy \quad (13)$$

evaluated at $\hat{\theta}$. Just like the properties of $H(\hat{\theta}, d)$ characterize the behavior of measurement error in a small neighborhood, $\mathcal{I}(\hat{\theta}, d)$ provides a measure of the amount of information that the observable random variable Y contains about the unknown parameters in θ : the ‘‘greater’’ it is, the more

information the experiments contain about $\hat{\theta}$ (i.e., ‘‘smaller’’ covariance, thus less uncertainty).

Both with $H(\hat{\theta}, d)$ and $\mathcal{I}(\hat{\theta}, d)$, to determine the optimal set of experiments, we need to consider some real-valued summary statistics. Among popular choices, the most common in terms of the Fisher information matrix are the A-optimal, E-optimal and D-optimal designs, see e.g. [24], [25], [26], yielding the optimization problems

$$d_A^* = \arg \min_{d \in \mathcal{D}} \text{tr} \left(\mathcal{I}^{-1}(\hat{\theta}, d) \right), \quad (14)$$

$$d_E^* = \arg \min_{d \in \mathcal{D}} \lambda_{\max} \left(\mathcal{I}^{-1}(\hat{\theta}, d) \right), \quad (15)$$

$$d_D^* = \arg \min_{d \in \mathcal{D}} \det \left(\mathcal{I}^{-1}(\hat{\theta}, d) \right), \quad (16)$$

respectively, where \mathcal{D} is the set of possible d . While d_A^* and d_E^* minimize the average variance and maximum error, respectively, d_D^* maximizes the differential Shannon information content

$$- \int f(y|\theta, d) l(y|\theta, d) dy,$$

extending the concept of Shannon information [27] to random variables with continuous probability distributions. Similarly, with λ_{\min} and λ_{\max} denoting the smallest and largest eigenvalues of $H(\hat{\theta}, d)$ in (12), respectively, one might seek to minimize the condition number of $H(\hat{\theta}, d)$:

$$d^* = \arg \min_{d \in \mathcal{D}} \left(\frac{\lambda_{\max}}{\lambda_{\min}} \right), \quad (17)$$

thus avoiding flat error surfaces in any direction [28]. Note that in this case, we assume structural identifiability of the problem, thus $\lambda_{\min} \neq 0$.

B. Decomposition of the Error and Likelihood Functions

Next, we present two key lemmas to easily compute $H(\hat{\theta}, d)$ and $\mathcal{I}(\hat{\theta}, d)$. To simplify notation, we do not denote the dependence on d and $d^{(j)}$ in this subsection.

Lemma 1. $H(\hat{\theta})$ from (12) can be computed as

$$H(\hat{\theta}) = \sum_{j=1}^J \sum_{k=1}^K H_{j,k}(\hat{\theta}) \quad (18)$$

where

$$H_{j,k}(\theta) = \left[\frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta} \right]^\top \frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta}. \quad (19)$$

Proof. From $\frac{\partial L(y, \theta)}{\partial \theta_a} = \sum_{j=1}^J \sum_{k=1}^K \frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta_a} [\bar{y}_{j,k}(\theta) - y_{j,k}]$, considering (10) we obtain that

$$\frac{\partial^2 L(y, \theta)}{\partial \theta_a \partial \theta_b} = \sum_{j=1}^J \sum_{k=1}^K \frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta_a} \frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta_b},$$

thus yielding (18) with (19). \square

Lemma 2. $\mathcal{I}(\hat{\theta})$ from (13) can be computed as

$$\mathcal{I}(\hat{\theta}) = \sum_{j=1}^J \sum_{k=1}^K \mathcal{I}_{j,k}(\hat{\theta}), \quad (20)$$

where

$$\mathcal{I}_{j,k}(\theta) = \left[\frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta} \right]^\top \Sigma^{-1} \frac{\partial \bar{y}_{j,k}(\theta)}{\partial \theta}. \quad (21)$$

Proof. Combining (6) with the fact that $Z_{1,1}, \dots, Z_{J,K}$ are independent and identically distributed random variables yields that $f_Y(y|\theta) = \prod_{j=1}^J \prod_{k=1}^K f_{Y_{j,k}}(y_{j,k}|\theta)$ with

$$f_{Y_{j,k}}(y_{j,k}|\theta) = \frac{\exp \left\{ [y_{j,k} - \bar{y}_{j,k}(\theta)]^\top \frac{\Sigma^{-1}}{2} [y_{j,k} - \bar{y}_{j,k}(\theta)] \right\}}{\sqrt{(2\pi)^q |\Sigma|}}.$$

Introduce the indexing $i := (j-1)J + k$ to simplify notation. With this, we obtain that $l(y|\theta) = \sum_{i=1}^{JK} l_i(y_i|\theta)$ where $l_i(y_i|\theta) := \log f_{Y_i}(y_i|\theta)$. Next, we show that with

$$\mathcal{I}_i(\theta) := \int \left[\frac{\partial l_i(y_i|\theta)}{\partial \theta} \right]^\top \frac{\partial l_i(y_i|\theta)}{\partial \theta} f_{Y_i}(y_i|\theta) dy_i$$

we obtain $\mathcal{I}(\theta) = \sum_{i=1}^{JK} \mathcal{I}_i(\theta)$. To see this, note that $f_a(y_a|\theta) = e^{l_a(y_a|\theta)}$, thus $\int f_{Y_a}(y_a|\theta) dy_a = 1$ yields

$$\begin{aligned} 0 &= \frac{\partial}{\partial \theta_p} \int f_{Y_a}(y_a|\theta) dy_a = \int \frac{\partial e^{l_a(y_a|\theta)}}{\partial \theta_p} dy_a \\ &= \int \frac{\partial l_a(y_a|\theta)}{\partial \theta_p} e^{l_a(y_a|\theta)} dy_a = \int \frac{\partial l_a(y_a|\theta)}{\partial \theta_p} f_{Y_a}(y_a|\theta) dy_a. \end{aligned}$$

Therefore, if $a \neq b$ then we obtain

$$\int \int \frac{\partial l_a(y_a|\theta)}{\partial \theta_p} \frac{\partial l_b(y_b|\theta)}{\partial \theta_q} f_{Y_a}(y_a|\theta) f_{Y_b}(y_b|\theta) dy_a dy_b = 0,$$

so that $\mathcal{I}(\theta) = \sum_{i=1}^{JK} \mathcal{I}_i(\theta)$. Finally, as Σ does not depend on θ , from [29] it follows that

$$[\mathcal{I}_i(\theta)]_{a,b} = \frac{\partial \bar{y}_i^\top(\theta)}{\partial \theta_a} \Sigma^{-1} \frac{\partial \bar{y}_i(\theta)}{\partial \theta_b},$$

yielding (21) with $i = (j-1)J + k$. \square

The results in Lemmas 1–2 can be readily leveraged to characterize the set of optimal experiments according to some real-valued summary statistics, e.g., those in (14)–(17). Additionally, these results are equally well-suited for precisely quantifying how competition for shared cellular resources affects parameter estimation. Thus, they can guide the experiment design not only by providing guidelines how to minimize its detrimental effects, but also by highlighting how to leverage this phenomenon, detailed next.

C. Expression of a single part

To illustrate how the results in Lemmas 1–2 can be used to better understand the effects of competition for shared resources, we first consider the transcription of a single gene ($n = 1$) into mRNA x both when competition for shared resources is accounted for like in (1), and when competition is neglected, see (2).

The solution of (1) is given by

$$x(t) = \frac{\alpha}{\gamma} \frac{d}{1 + \frac{d}{\kappa}} (1 - e^{-\gamma t}), \quad (22)$$

whereas integrating (2) yields

$$x(t) = \frac{\alpha}{\gamma} \frac{d}{\kappa} (1 - e^{-\gamma t}). \quad (23)$$

Claim 1. Let $H^0(\hat{\theta}, d)$ be like in (12) corresponding to model (2) with $n = 1$, output $y = x$ and $\theta = (\alpha, \gamma, \kappa)^\top$. With this, for any $i \in \{1, \dots, J\}$ we have that

$$H^0(\hat{\theta}, d) = H^0(\hat{\theta}, d^{(i)}) \sum_{j=1}^J \left(\frac{d^{(j)}}{d^{(i)}} \right)^2. \quad (24)$$

Proof. From (23) we obtain that

$$\begin{aligned} \frac{\partial \bar{y}_{j,k}}{\partial \alpha} &= \frac{d^{(j)}}{\kappa \gamma} (1 - e^{-\gamma t_k}), \\ \frac{\partial \bar{y}_{j,k}}{\partial \gamma} &= \frac{\alpha d^{(j)}}{\kappa \gamma^2} (e^{-\gamma t_k} - 1 + \gamma t_k e^{-\gamma t_k}), \\ \frac{\partial \bar{y}_{j,k}}{\partial \kappa} &= \frac{\alpha d^{(j)}}{\kappa^2 \gamma} (e^{-\gamma t_k} - 1), \end{aligned}$$

so that $\frac{\partial \bar{y}_{j,k}}{\partial \alpha}, \frac{\partial \bar{y}_{j,k}}{\partial \gamma}, \frac{\partial \bar{y}_{j,k}}{\partial \kappa} \propto d^{(j)}$, yielding (24). \square

It follows from Claim 1 that experiments with different DNA concentrations do not change the shape of the error function $L(y, \theta, d)$ when competition for shared cellular resources is neglected. Conversely, when considering (1) instead of (2), the result in (24) does not hold, thus the shape of the paraboloid in (11) changes with different DNA concentrations, accounting for loading effects due to the limited availability of shared resources. To further reveal the role that competition plays, we next focus on it from the perspective of information content when comparing (1) with (2). Without loss of generality, assume that $\kappa = 1$.

Claim 2. With $\theta = (\alpha, \gamma)^\top$, let $\mathcal{I}(\hat{\theta}, d)$ and $\mathcal{I}^0(\hat{\theta}, d)$ denote the Fisher information matrices corresponding to (1) and (2), respectively. With this, we have that

$$\det \left(\mathcal{I}(\hat{\theta}, d) \right) < \det \left(\mathcal{I}^0(\hat{\theta}, d) \right). \quad (25)$$

Proof. Considering (23), from Lemma 2 we obtain that

$$\mathcal{I}_{j,k}^0(\hat{\theta}, d^{(j)}) = \begin{bmatrix} A_{j,k}^2 & A_{j,k} B_{j,k} \\ A_{j,k} B_{j,k} & B_{j,k}^2 \end{bmatrix}$$

with

$$\begin{aligned} A_{j,k} &= \frac{1}{\sqrt{|\Sigma|}} \frac{d^{(j)}}{\gamma} (1 - e^{-\gamma t_k}), \\ B_{j,k} &= \frac{1}{\sqrt{|\Sigma|}} \left[\frac{\alpha d^{(j)}}{\gamma^2} (e^{-\gamma t_k} - 1) + \frac{\alpha d^{(j)}}{\gamma} t_k e^{-\gamma t_k} \right]. \end{aligned}$$

Similarly, from (2) we obtain that $\mathcal{I}_{j,k}(\hat{\theta}, d^{(j)}) = \mathcal{I}_{j,k}^0(\hat{\theta}, d^{(j)}) / (1 + d^{(j)})^2$. Therefore, it follows from Lemma 2 with $\Omega = \{1, \dots, J\} \times \{1, \dots, K\}$ that

$$\begin{aligned} \det \left(\mathcal{I}^0(\hat{\theta}, d) \right) &= \sum_{(j,k) \in \Omega} \sum_{(a,b) \neq (j,k)} (A_{j,k} B_{a,b} - A_{a,b} B_{j,k})^2, \\ \det \left(\mathcal{I}(\hat{\theta}, d) \right) &= \sum_{(j,k) \in \Omega} \sum_{(a,b) \neq (j,k)} \frac{(A_{j,k} B_{a,b} - A_{a,b} B_{j,k})^2}{(1 + d^{(j)})^2 (1 + d^{(a)})^2}, \end{aligned}$$

yielding (25) since $d^{(j)}, d^{(a)} \geq 0$ for $j, a = 1, \dots, J$. \square

In Claim 1 we saw that competition for shared resources was essential for taking advantage of multiple experiments, thus it was advantageous. Conversely, Claim 2 reveals that the information content is decreased by competition for shared resources.

D. Co-expression of two parts

After demonstrating how competition for shared cellular resources affects the characterization of a single part, here we demonstrate how it can be leveraged. To this end, we consider the expression of two genes.

The first gene of interest expresses mRNA x_1 with parameters α_1, γ_1 , and κ_1 , whereas the second gene expresses mRNA x_2 with the corresponding parameters α_2, γ_2 , and κ_2 . We consider two cases: first, only x_1 is expressed, so that from (1) its dynamics are given by

$$\dot{x}_1 = \alpha_1 \frac{\frac{d_1}{\kappa_1}}{1 + \frac{d_1}{\kappa_1}} - \gamma_1 x_1, \quad (26)$$

whereas in the second case both genes are expressed:

$$\dot{x}_i = \alpha_i \frac{\frac{d_i}{\kappa_i}}{1 + \frac{d_1}{\kappa_1} + \frac{d_2}{\kappa_2}} - \gamma_i x_i, \quad i = 1, 2. \quad (27)$$

Assume that we have already characterized the second part, thus α_2, γ_2 , and κ_2 are known parameters, and that the above two parts only differ in their promoters, captured by $\kappa_1 \neq \kappa_2$, but are otherwise identical, thus $\alpha_1 = \alpha_2$ and $\gamma_1 = \gamma_2$. Therefore, our task is to estimate $\theta = \kappa_1$.

Let $\mathcal{I}^S(\hat{\theta}, d_1)$ denote the Fisher information matrix corresponding to (26) when $y = x_1$. Similarly, with $d = (d_1, d_2)^\top$ let $\mathcal{I}^{C1}(\hat{\theta}, d)$ and $\mathcal{I}^{C2}(\hat{\theta}, d)$ correspond to (27) when $y = x_1$ and when $y = (x_1, x_2)^\top$, respectively.

Claim 3. Define $d_1^* := \arg \max_{d_1} \mathcal{I}^S(\hat{\theta}, d_1)$ together with $\mathcal{I}^* := \max_{d_1} \mathcal{I}^S(\hat{\theta}, d_1)$. With this, we obtain that $d_1^* = \kappa_1$ and $\mathcal{I}^{C1}(\hat{\theta}, d) \leq \mathcal{I}^*$. Furthermore, $\mathcal{I}^{C2}(\hat{\theta}, d) \geq \mathcal{I}^*$ when

$$\frac{d_1}{\kappa_1} \sqrt{1 + 2 \frac{d_2}{\kappa_2} + 2 \left(\frac{d_2}{\kappa_2} \right)^2} \geq \left(1 + \frac{d_1}{\kappa_1} + \frac{d_2}{\kappa_2} \right)^2. \quad (28)$$

Proof. Integrating (26), we obtain that

$$x_1(t) = \frac{\frac{d_1}{\kappa_1}}{1 + \frac{d_1}{\kappa_1}} \frac{\alpha_1}{\gamma_1} (1 - e^{-\gamma_1 t}),$$

so that Lemma 2 yields $\mathcal{I}^S(\hat{\theta}, d_1) = \sum_{k=1}^K \mathcal{I}_k^S(\hat{\theta}, d_1)$, where

$$\mathcal{I}_k^S(\hat{\theta}, d_1) = \left[\frac{\alpha_1 (1 - e^{-\gamma_1 t_k})}{\kappa_1 \gamma_1} \right]^2 \left(\frac{d_1}{\kappa_1} \right)^2 \left(1 + \frac{d_1}{\kappa_1} \right)^{-4},$$

thus $d_1^* = \kappa_1$. Similarly, from (27) we obtain that

$$x_i(t) = \frac{\frac{d_i}{\kappa_i}}{1 + \frac{d_1}{\kappa_1} + \frac{d_2}{\kappa_2}} \frac{\alpha_i}{\gamma_i} (1 - e^{-\gamma_i t}) \quad i = 1, 2,$$

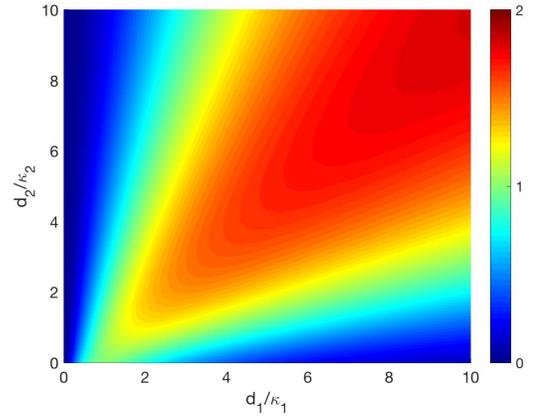


Fig. 2. The information $\tilde{\mathcal{I}}(\theta, d)$ exceeds $\mathcal{I}^* = \max_{d_1} \mathcal{I}(\theta, d_1)$ when the condition in (28) is satisfied. For instance, when $d_1/\kappa_1, d_2/\kappa_2 \approx 10$, the information content is approximately doubled. Note that when $d_1/\kappa_1 \approx 10$, the information content expressing and measuring only x_1 is $\approx \mathcal{I}^*/10$.

thus Lemma 2 yields that $\mathcal{I}^{C1}(\hat{\theta}, d) = \sum_{k=1}^K \mathcal{I}_k^{C1}(\hat{\theta}, d)$ and $\mathcal{I}^{C2}(\hat{\theta}, d) = \sum_{k=1}^K \mathcal{I}_k^{C2}(\hat{\theta}, d)$ with $\mathcal{I}_k^{C1}(\hat{\theta}, d) = \mathcal{I}_k^S(\hat{\theta}, d_1) s\left(\frac{d_1}{\kappa_1}, \frac{d_2}{\kappa_2}\right)$, $\mathcal{I}_k^{C2}(\hat{\theta}, d) = \mathcal{I}_k^{C1}(\hat{\theta}, d) q\left(\frac{d_1}{\kappa_1}, \frac{d_2}{\kappa_2}\right)$, and

$$s(a, b) = \frac{(1+a)^4 (1+b)^2}{(1+a+b)^4}, \quad (29)$$

$$q(a, b) = \frac{(1+a)^4 [(1+b)^2 + b^2]}{(1+a+b)^4}. \quad (30)$$

Consequently, we have that $\mathcal{I}^{C1}(\hat{\theta}, d) = \mathcal{I}^S(\hat{\theta}, d_1) s\left(\frac{d_1}{\kappa_1}, \frac{d_2}{\kappa_2}\right)$, so that from (29) it follows that $\mathcal{I}^{C1}(\hat{\theta}, d) \leq \mathcal{I}^*$. Similarly, we obtain that $\mathcal{I}^{C2}(\hat{\theta}, d) = \mathcal{I}^{C1}(\hat{\theta}, d) q\left(\frac{d_1}{\kappa_1}, \frac{d_2}{\kappa_2}\right)$, so that $\mathcal{I}^{C2}(\hat{\theta}, d) \geq \mathcal{I}^*$ follows from (30) when (28) is satisfied. \square

The above result can be interpreted as follows. First, $d_1^* = \kappa_1$ means that the system (26) is “most excited” when $d_1/(d_1 + \kappa_1)$ is steepest, a widely known fact in biology [20]. Next, Claim 3 states that $\mathcal{I}^{C1}(\hat{\theta}, d) \leq \mathcal{I}^*$, that is, without measuring the expression of the known part, it is not possible to exceed the maximal information content \mathcal{I}^* about the unknown part. This is not surprising, as comparing (26) and (27), the dynamics of x_1 are identical if

$$\frac{\frac{d_1}{\kappa_1}}{1 + \frac{d_1}{\kappa_1}} = \frac{\frac{\tilde{d}_1}{\kappa_1}}{1 + \frac{\tilde{d}_1}{\kappa_1} + \frac{\tilde{d}_2}{\kappa_2}} \rightarrow \frac{d_1}{\kappa_1} = \frac{\tilde{d}_1}{\kappa_1 + \frac{\tilde{d}_2}{\kappa_2}},$$

where \tilde{d}_i denotes the values of d_i in (27). Although we have information about the expression of x_2 via x_1 , we do not possess any extra information about the expression of x_1 .

Conversely, once we measure both x_1 and x_2 , we also have access to the effect of x_1 expression on that of x_2 as a result of competition for shared resources. Thus, there are conditions under which the overall information, direct from measuring x_1 and indirect from measuring x_2 , exceeds \mathcal{I}^* (Fig. 2). As a general design principle, when characterizing promoters it is thus beneficial to co-express an unknown part with a known reporter and to use generous concentrations of DNA for both parts (top right corner in Fig. 2).

To further emphasize how to leverage the co-expression of a known reporter with an unknown part, we next reveal how the choice of $H_{j,k}$ contributes to the eigenvalues and eigenvectors of H . This question is not trivial and has been studied in form of the Horn conjecture [30], [31]. While previous results focused on deriving conditions on the sequence of eigenvalues of $H_{j,k}$ for solving this problem, we study the eigenvectors instead. Here, we consider the combination of two observations (either about the same part or about co-expressed unknown and reporter parts, see Claim 3), yielding H_1 and H_2 , respectively. Note that both H_1 and H_2 have a single non-zero eigenvalue (follows from the diadic decomposition in (19)). Let $\langle a, b \rangle$ denote the standard scalar product.

Lemma 3. *Let $H_1, H_2 \in \mathbb{R}^{n \times n}$ be real symmetric matrices, both with precisely one non-zero eigenvalue, denoted by λ_1 and λ_2 , respectively. Further, let h_1 and h_2 denote the corresponding normalized eigenvectors, respectively. Let η_1 and η_2 denote the two non-zero eigenvalues of $H = H_1 + H_2$ with the corresponding eigenvectors v_1 and v_2 , respectively. If $\langle h_1, h_2 \rangle = 0$, then $\eta_i = \lambda_i$ and $v_i = h_i$ for $i = 1, 2$. Otherwise, $v_1 = a_1 h_1 + h_2$ and $v_2 = a_2 h_1 + h_2$ with $\eta_1 = \lambda_2(1 + a_1 \langle h_1, h_2 \rangle)$ and $\eta_2 = \lambda_2(1 + a_2 \langle h_1, h_2 \rangle)$, where*

$$a_{1,2} = \frac{(\lambda_1 - \lambda_2) \pm \sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_1\lambda_2\langle h_1, h_2 \rangle^2}}{2\lambda_2\langle h_1, h_2 \rangle}.$$

Proof. With $a \in \mathbb{R}$ and the ansatz $v = ah_1 + bh_2$ we want to find a and b such that $Hv = \eta v$. As for both H_1 and H_2 , only one eigenvalue is non-zero, it suffices to consider the orthogonal projection of v on the eigenvectors of H_1 and H_2 , respectively, thus

$$Hv = (a + \langle h_1, h_2 \rangle)\lambda_1 h_1 + (a\langle h_1, h_2 \rangle + 1)\lambda_2 h_2. \quad (31)$$

When $\langle h_1, h_2 \rangle = 0$, the Lemma follows directly. Otherwise, without loss of generality we set $b = 1$ and with the substitution $h_2 = v - ah_1$, we obtain that all terms in h_1 vanish in the above equation if a satisfies $\lambda_2\langle h_1, h_2 \rangle a^2 + (\lambda_2 - \lambda_1)a - \lambda_1\langle h_1, h_2 \rangle = 0$. The solutions of this quadratic formula are a_1 and a_2 defined in the claim. With this, the eigenvalues η_1 and η_2 of H from (31) are given by $\eta_{1,2} = \lambda_2(1 + a_{1,2}\langle h_1, h_2 \rangle)$. \square

Therefore, when combining two observations (either about the same part or about co-expressed unknown and reporter parts), from Lemma 3 we obtain that the eigenvectors/eigenvalues of the cumulative Hessian $H = H_1 + H_2$ are convex combinations of those of the Hessians H_1 and H_2 corresponding to the two observations. Most importantly, the difference between eigenvalues η_1 and η_2 of H is

$$\eta_1 - \eta_2 = \sqrt{(\lambda_1 - \lambda_2)^2 + 4\lambda_1\lambda_2\langle h_1, h_2 \rangle^2}. \quad (32)$$

Therefore, in order to have $\eta_1 \approx \eta_2$ for an optimal shape of $L(y, \theta, d)$ (Fig. 1A), one should choose the two experiments such that the non-zero eigenvalues of the constituent Hessians H_1 and H_2 are comparable, and also their eigenvectors are approximately orthogonal.

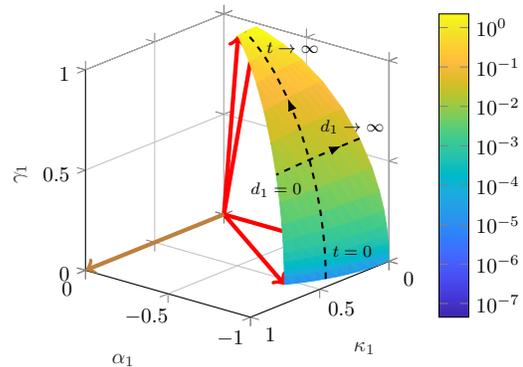


Fig. 3. Eigenvectors corresponding to the non-zero eigenvalue of H_1 (colored surface) and H_2 (brown arrow). Color shading indicates value of associated eigenvalue. Black dashed lines indicate the levels of constant DNA amount (vertical) and constant time point (horizontal).

IV. APPLICATION EXAMPLES

We consider two application examples to illustrate the implications of our results. We first focus on how competition between an unknown part and a reporter can be leveraged to obtain a better-conditioned error surface $L(y, \theta, d)$, then we revisit the experiments in [12] from the perspective of information content and sharpness of parameter estimates.

A. Least Squares Estimation

Consider the case of co-expression of an unknown part with output x_1 together with a known reporter with output x_2 . The unknown parameters are $\theta = (\alpha_1, \gamma_1, \kappa_1)$, and let H_1 and H_2 denote the Hessians corresponding to the observations $x_1(t_k)$ and $x_2(t_k)$. The eigenvector corresponding to the only non-zero eigenvalue of H_2 is given by $(0, 0, 1)^\top$ (brown in Fig. 3), since the output x_2 only depends on κ_1 and is independent of α_1 and γ_1 , see (27). It can be shown that the eigenvector of H_1 corresponding to the only non-zero eigenvalue lies in the polyhedral cone $C = \{a_1 q_1 + a_2 q_2 + a_3 q_3 + a_4 q_4\}$ (red in Fig. 3) with $q_1 = (-\frac{\gamma_1}{\alpha_1}, \frac{\gamma_1}{\kappa_1}, 0)^\top$, $q_2 = (-\frac{\gamma_1}{\alpha_1}, \frac{\gamma_1}{\kappa_1}, 1)^\top$, $q_3 = (-\frac{\gamma_1}{\alpha_1}, 0, 0)^\top$ and $q_4 = (-\frac{\gamma_1}{\alpha_1}, 0, 1)^\top$.

From Lemma 3 we know that the eigenvector of H_1 should be orthogonal to that of H_2 , depicted in brown in Fig. 3. Therefore, the optimal choice lies in the (α_1, γ_1) -plane, corresponding to large $d_1^{(j)}$. Furthermore, Lemma 3 also reveals that the non-zero eigenvalues of the constituent Hessians H_1 and H_2 should be comparable, therefore, the optimal choice of time point t_k to take the measurement x_2 is when the corresponding eigenvalue is approximately equal to the non-zero eigenvalue of H_1 .

We next focus on a series of measurements and the role competition between an unknown part and a known reporter plays by comparing 6 different experiment designs, summarized in Table I. A standard design usually involves a series of experiments with the DNA concentration and time points both spaced linearly between some minimum and maximum value, without additional reporter DNA (scenario #1). As seen before, adding reporter DNA (scenario #2) can be

TABLE I

LOAD DNA CAN BE LEVERAGED TO YIELD STEEPER ERROR SURFACE

| Scenario | d_1 ($j \in [1, 10]$) | d_2 | t ($k \in [1, 10]$) | $\frac{\lambda_{\max}}{\lambda_{\min}}$ |
|----------|---------------------------|-------|-------------------------|---|
| 1 | $10 \cdot j$ | 0 | $k/10$ | 1855 |
| 2 | $10 \cdot j$ | 100 | $k/10$ | 212 |
| 3 | 100 | 100 | $k/10$ | 232 |
| 4 | 100 | 100 | $[0.1, 1]$ | 500 |
| 5 | 100 | 100 | $10^{k/10-1}$ | 169 |
| 6 | 14.6 | 200 | see (33) | 24.5 |

beneficial: with the same measurement time points and DNA concentrations, the fraction $\lambda_{\max}/\lambda_{\min}$ decreases about 10-fold. Removing 9 of the 10 experiments yields almost no increase in the fraction (scenario #3), and further removing 8 of the 10 time points only increases the fraction $\lambda_{\max}/\lambda_{\min}$ about 2-fold (scenario #4). Importantly, this increase can be recovered by spreading the time points logarithmically (scenario #5), significantly outperforming the original naïve experiment design with only 1 of the 10 experiments and a set of wisely chosen time points. Finally, in scenario #6, keeping the number of DNA concentrations and time points constant, we apply a pattern search optimization algorithm to find the optimal values $d_1^* = 14.6$, $d_2^* = 200$ and measurement time points

$$t^* = \{0.075, 0.078, 0.082, 0.083, 0.084, 0.086, 0.088, 0.090, 0.093, 0.307\} \quad (33)$$

minimizing the condition number $\lambda_{\max}/\lambda_{\min}$ under the constraints $0 \leq d_i \leq 200$ and $0 \leq t_k \leq 1$.

B. Maximum Likelihood Estimation

In [12] there were $J = 6$ DNA concentrations ($d = (0.2, 0.5, 1, 2, 5, 20)^\top$) and measurements every minute, thus $K = 60$ time points over the first hour, considered in [15]. In Table II, we consider 18 setups and the information content (log-scale) of the corresponding D-optimal solution (16) relative to the information content \mathcal{I}_0 corresponding to the original setup in [12]. In particular, we consider possible DNA concentrations from the set $\mathcal{D} = \{0.1, 0.2, \dots, 1, 2, 4, \dots, 20\}$ of which we chose $J \in \{1, 2, 3, 4, 5, 6\}$, with $K \in \{6, 20, 60\}$ measurements (every 10, 3, and 1 minute, respectively). From Table II, we conclude that (i) the information content could have been increased over 100-fold (red in Table II) with the optimal choice of experiments compared to \mathcal{I}_0 obtained in the original setup in [12]; and (ii) we could have obtained the same information content \mathcal{I}_0 with only 20 time points instead of 60 (blue in Table II).

We next focus more on the widely used logarithmic distribution of probing DNA concentrations used in [12]. In particular, we consider the set \mathcal{D} of possible DNA concentrations with $J = 4$ choices and $K = 20$ time points. With this, we have approximately 5,000 possible combinations, among which about 200 yield parameter estimates with smaller average variance than the original combination of 6 DNA concentrations (Fig. 4A). Furthermore, while in the original combination 4 out of 6 DNA concentrations are no greater than 2nM, only 6% of the DNA concentrations in the best 50

TABLE II

D-OPTIMAL INFORMATION CONTENT

| | | J | | | | | |
|---|----|-------|------|------|------|------|------|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| K | 6 | -16.8 | -5.1 | -3.3 | -2.9 | -2.7 | -2.5 |
| | 20 | -15.0 | -2.7 | -0.9 | -0.4 | -0.2 | 0.0 |
| | 60 | -13.2 | -0.4 | 1.4 | 1.9 | 2.1 | 2.3 |

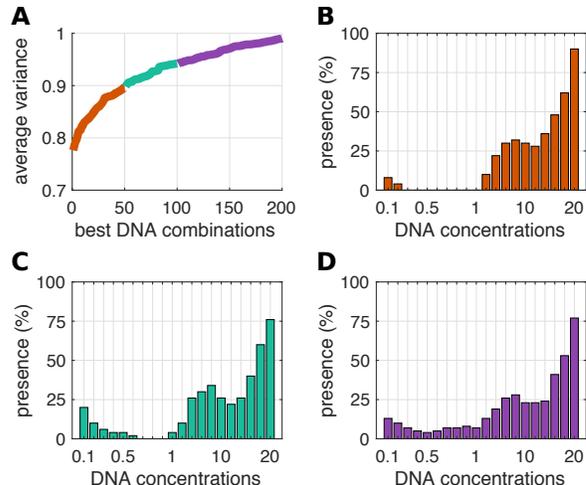


Fig. 4. (A) Average variance (weighted average considering nominal parameter values in [15]) of the best 200 combinations normalized to that in [12] with 6 DNA concentrations. (B) DNA distribution corresponding to the best 50 combinations. (C) DNA distribution corresponding to the 50 combinations following the 50 best. (D) DNA distribution corresponding to the 100 combinations following the 100 best.

combinations belong to this range (Fig. 4B), and it increases to 15% and 22% for the following 50 and 100 combinations (Fig. 4C–D), respectively. This suggests that choosing higher DNA concentrations in [12] would have yielded parameter estimates of reduced variance, thus of less uncertainty.

V. CONCLUSION

A system is said to be globally structurally identifiable if its output is unique for every admissible parameter vector θ . This property has been studied extensively [32], [33], [34], and can be rigorously verified by studying the Lie derivatives of the output function. While structural identifiability is a necessary condition for identifying all parameters, it is not sufficient. Even if a parameter is found to be structurally identifiable, the likelihood/error function around the true parameter value may be flat in some directions, leading to a practically unidentifiable parameter.

In this paper, we focused on two complementary aspects of practical identifiability of unknown parameters in the context of cell-free extracts. We considered the Maximum Likelihood Estimator $\hat{\theta}_{MLE}$ in case the noise distribution is known, and the Least Squares Estimator $\hat{\theta}_{LSE}$ otherwise. Both estimators give rise to a local description of error/likelihood surfaces, where the appropriate Hessian matrix characterizes the curvature of the error/likelihood surface. By studying these Hessian matrices, we can thus quantify the information content of candidate experiments.

Leveraging this, it is possible to design the optimal set of experiments that minimizes some real-valued summary statistics, leading to sharper parameter estimates. Additionally, the novelty of the results presented here stems from the fact that we revealed the role that competition for shared resources plays. In particular, we first illustrated that it is essential for taking advantage of different plasmid DNA concentrations. Second, we demonstrated that when characterizing an unknown part, the information content can be significantly increased by co-expressing the unknown part with a known reporter, thus extracting additional information via loading, leading to reduced parameter uncertainty.

Finally, we revisited the original experiments in [12] to compare the information content of naïve and optimized experiments. Since the results show that the information content can be increased with even less experiments, we next seek to experimentally verify these findings and demonstrate that optimizing experiments indeed yield appreciably sharper parameter estimates and more accurate predictions. Additionally, we seek to provide an experiment design algorithm which suggests future experiments based on already given experimental data and explore both the theoretical and experimental aspects of considering time varying inputs (decision variables) of different signal shapes (i.e., administering inducer/inhibitor molecules continuously).

VI. ACKNOWLEDGMENTS

The authors are grateful for Vipul Singhal and William Poole for the insightful comments and thought provoking discussions.

REFERENCES

- [1] S. A. Benner and A. M. Sismour, "Synthetic biology," *Nature Reviews Genetics*, vol. 6, no. 7, pp. 533–543, 2005.
- [2] E. Andrianantoandro, S. Basu, D. K. Karig, and R. Weiss, "Synthetic biology: new engineering rules for an emerging discipline," *Molecular Systems Biology*, vol. 2, 2006.
- [3] J. A. J. Arpino, E. J. Hancock, J. Anderson, M. Barahona, G.-B. V. Stan, A. Papachristodoulou, and K. Polizzi, "Tuning the dials of Synthetic Biology," *Microbiology*, vol. 159, no. 7, pp. 1236–1253, 2013.
- [4] M. J. Smanski, S. Bhatia, D. Zhao, Y. Park, L. B. A. Woodruff, G. Giannoukos, D. Ciulla, M. Busby, J. Calderon, R. Nicol, D. B. Gordon, D. Densmore, and C. A. Voigt, "Functional optimization of gene clusters by combinatorial design and assembly," *Nature biotechnology*, vol. 32, no. 12, pp. 1241–1249, 2014.
- [5] J. T. MacDonald, C. Barnes, R. I. Kitney, P. S. Freemont, and G.-B. V. Stan, "Computational design approaches and tools for synthetic biology," *Integrative Biology*, vol. 3, no. 2, pp. 97–108, 2011.
- [6] N. Crook and H. S. Alper, "Model-based design of synthetic, biological systems," *Chemical Engineering Science*, vol. 103, pp. 2–11, 2013.
- [7] T. P. Prescott and A. Papachristodoulou, "Synthetic biology: A control engineering perspective," in *Proc. European Control Conference (ECC)*, 2014, pp. 1182–1186.
- [8] D. Del Vecchio, A. J. Dy, and Y. Qian, "Control theory meets synthetic biology," *Journal of The Royal Society Interface*, vol. 13, no. 120, 2016.
- [9] R. J. R. Algar, T. Ellis, and G.-B. Stan, "Modelling essential interactions between synthetic genes and their chassis cell," in *Proc. 53rd IEEE Conference on Decision and Control (CDC)*, 2014, pp. 5437–5444.
- [10] A. Y. Weiße, D. A. Oyarzún, V. Danos, and P. S. Swain, "Mechanistic links between cellular trade-offs, gene expression, and growth," *Proceedings of the National Academy of Sciences*, vol. 112, no. 9, pp. E1038–E1047, 2015.
- [11] O. Borkowski, F. Ceroni, G. Stan, and T. Ellis, "Overloaded and stressed: whole-cell considerations for bacterial synthetic biology," *Current Opinion in Microbiology*, vol. 33, pp. 123–130, 2016.
- [12] D. Siegal-Gaskins, Z. A. Tuza, J. Kim, V. Noireaux, and R. M. Murray, "Gene Circuit Performance Characterization and Resource Usage in a Cell-Free Breadboard," *ACS Synthetic Biology*, vol. 3, no. 6, pp. 416–425, 2014.
- [13] A. Gyorgy, J. I. Jiménez, J. Yazbek, H.-H. Huang, H. Chung, R. Weiss, and D. Del Vecchio, "Isocost Lines Describe the Cellular Economy of Genetic Circuits," *Biophysical Journal*, vol. 109, no. 3, pp. 639–646, 2015.
- [14] F. Ceroni, R. Algar, G.-B. Stan, and T. Ellis, "Quantifying cellular capacity identifies gene expression designs with reduced burden," *Nature Methods*, vol. 12, no. 5, pp. 415–418, 2015.
- [15] A. Gyorgy and R. M. Murray, "Quantifying resource competition and its effects in the TX-TL system," in *Proc. 55th IEEE Conference on Decision and Control (CDC)*, 2016, pp. 3363–3368.
- [16] W. Halter, J. M. Montenbruck, and F. Allgöwer, "Systems with integral resource consumption," in *Proc. 56th IEEE Conference on Decision and Control (CDC)*, 2017, pp. 2667–2673.
- [17] W. Halter, J. M. Montenbruck, Z. A. Tuza, and F. Allgöwer, "A resource dependent protein synthesis model for evaluating synthetic circuits," *Journal of Theoretical Biology*, vol. 420, pp. 267–278, 2017.
- [18] Y. Qian, H.-H. Huang, J. I. Jiménez, and D. Del Vecchio, "Resource Competition Shapes the Response of Genetic Circuits," *ACS Synthetic Biology*, vol. 6, no. 7, pp. 1263–1272, 2017.
- [19] V. H. Nagaraj, J. M. Greene, A. M. Sengupta, and E. D. Sontag, "Translation inhibition and resource balance in the TX-TL cell-free gene expression system," *Synthetic Biology*, vol. 2, no. 1, pp. 1–7, 2017.
- [20] D. D. Vecchio and R. M. Murray, *Biomolecular Feedback Systems*. Princeton University Press, 2014.
- [21] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, 2009.
- [22] S. Zeng, "Observability measures for nonlinear systems," in *Proc. 57th IEEE Conference on Decision and Control (CDC)*, 2018, to appear.
- [23] B. Efron and T. Hastie, *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge University Press, 2016.
- [24] P. Weber, A. Kramer, C. Dingler, and N. Radde, "Trajectory-oriented Bayesian experiment design versus Fisher A-optimal design: an in depth comparison study," *Bioinformatics (Oxford, England)*, vol. 28, no. 18, pp. i535–i541, 2012.
- [25] E. Balsa-Canto, J. Banga, and A. Alonso, "Computational procedures for optimal experimental design in biological systems," *IET Systems Biology*, vol. 2, no. 4, pp. 163–172, 2008.
- [26] F. Pukelsheim, *Optimal design of experiments*. SIAM, 1993, vol. 50.
- [27] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [28] L. N. Trefethen and D. Bau III, *Numerical Linear Algebra*. SIAM, 1997, vol. 50.
- [29] B. Porat and B. Friedlander, "Computation of the exact information matrix of gaussian time series with stationary random components," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 1, pp. 118–130, 1986.
- [30] H. Weyl, "Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)," *Mathematische Annalen*, vol. 71, no. 4, pp. 441–479, 1912.
- [31] A. Horn, "Eigenvalues of sums of Hermitian matrices," *Pacific Journal of Mathematics*, vol. 12, no. 1, pp. 225–241, 1962.
- [32] H. Pohjanpalo, "System identifiability based on the power series expansion of the solution," *Mathematical Biosciences*, vol. 41, no. 1-2, pp. 21–33, 1978.
- [33] E. Sontag and Y. Wang, "I/O equations for nonlinear systems and observation spaces," in *Proc. 30th IEEE Conference on Decision and Control (CDC)*, 1991, pp. 720–725.
- [34] G. Margaria, E. Riccomagno, M. J. Chappell, and H. P. Wynn, "Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences," *Math. Biosci.*, vol. 174, pp. 1–26, 2001.