To appear in the Wiley International Journal of Robust and Nonlinear Control, 2022

Robustness Guarantees for Structured Model Reduction of Dynamical Systems with Applications to Biomolecular Models

Ayush Pandey and Richard M. Murray

Abstract

Model reduction methods usually focus on the error performance analysis; however, in presence of uncertainties, it is important to analyze the robustness properties of the error in model reduction as well. This problem is particularly relevant for engineered biological systems that need to function in a largely unknown and uncertain environment. We give robustness guarantees for structured model reduction of linear and nonlinear dynamical systems under parametric uncertainties. We consider a model reduction problem where the states in the reduced model are a strict subset of the states of the full model, and the dynamics for all of the other states are collapsed to zero (similar to quasi-steady-state approximation). We show two approaches to compute a robustness guarantee metric for any such model reduction — a direct linear analysis method for linear dynamics and a sensitivity analysis based approach that also works for nonlinear dynamics. Using the robustness guarantees with an error metric and an input-output mapping metric, we propose an automated model reduction method to determine the best possible reduced model for a given detailed system model. We apply our method for the (1) design space exploration of a gene expression system that leads to a new mathematical model that accounts for the limited resources in the system, and (2) model reduction of a population control circuit in bacterial cells.

I. Introduction

For applications of control theory to physical system design, a reduced model is commonly used that describes the dynamics of interest in lower dimensions to simplify the design process. Reduced models are used to specify the desired objectives or the performance specifications of a system. To meet these objectives, the designer needs to map the reduced models to the level of system design and also mathematically characterize this mapping to understand and analyze system performance. For biological systems, this is a challenge that hinders the use of mathematical models in experimental designs and analysis to some extent.

For model-based design of biological circuits, we need to develop mathematical models that map system design specifications to mechanistic details. Commonly used phenomenological models are based on empirical information and their parameters describe lumped properties of the system that are effective in explaining the observed experimental data [1]–[5] but have not been readily used for forward engineering of biological circuits. Towards that end, to explore different design possibilities one needs to carefully justify the validity of the underlying assumptions for each model [6]. Time-scale separation is one of the most common properties exhibited by biological systems. For example, the half-life of an mRNA in *E. coli* is around 100 seconds whereas the average half-life for a protein is of the order of a few hours. Hence, it is important to consider the effects of this assumption on biological system models.

Singular perturbation theory [7] is the formal way of deriving mathematical models for system dynamics with time-scale separation. For such systems, it is common to separate the dynamics into fast and slow modes. Then, singular perturbation theory can be used to derive reduced models that accurately represent the dynamics of interest. A key feature of singular perturbation theory is that the states of the reduced model are a subset of the states of the full model. In other words, the

The authors are with the Control and Dynamical Systems department at California Institute of Technology, Pasadena, CA, USA. Email: apandey@caltech.edu

structure of the model and the meaning of the states and parameters is conserved by construction in any reduced model obtained using singular perturbation theory. This is not automatically the case in other model reduction techniques where transformations are introduced [8], [9] and hence in such techniques, the meanings of the states may not be preserved. We define structured model reduction in this paper as the set of model reduction methods where the states of the reduced models are a strict subset of the states of the full model.

The advantage of structured model reduction techniques is that an explicit mapping between the full and the reduced model is readily available [10]. Moreover, since the parameters and the states in the reduced model have the same meaning as in the full model, the design outputs and analysis results obtained using the reduced model can easily be given context and compared with the full model [11]. This is especially relevant for biological system modeling as it is often important to map the reduced dynamics to the mechanistic details. However, due to the strict condition on the possible reduced model states, structured model reduction methods suffer from the limitation that the choice of reduced models is restricted and dependent on the modeling details of the full system. That is, for a given full model it may not always be possible to analytically derive a reduced model. Other model reduction methods that are projection-based or those which preserve the input-output mapping are better in that respect. In this paper, we focus on the former class of model reduction problems that preserve the modeling structure in the reduced models and hence more relevant for biological systems.

The goal with any model reduction problem is to minimize the error in the performance of the reduced model when compared to the full model. This error performance criterion can be general and depend on trajectories of all states and output variables, or specific, such as minimizing a particular metric of interest. Singular perturbation theory for model reduction and its error analysis is a widely studied topic in the literature [12]–[14] for different system and control design settings. A commonly used method for model reduction that is derived from the singular perturbation concept of time-scale separation is the quasi-steady-state approximation (QSSA) method [15]–[18]. Here, a subset of states is assumed to be at steady-state and hence their dynamics are collapsed to algebraic relationships. Error analysis for QSSA based model reduction [19]–[21] has been studied as well. However, robustness of these model reduction methods is not as widely studied in the literature.

Robust control design is a well-studied problem in control theory. The extensions of robust control theory to singularly-perturbed systems are studied in [22] and [23]. Similarly, robust stability analysis of adaptive control problems, linear time-varying systems, and the general parametric uncertainty problems has been of interest as well [24], [25]. A complementary, although not as widely applicable, approach to study the robustness of systems is to use sensitivity analysis of system variables or derived properties under parameter variations. Due to the success of robust control design methods [24] for different applications, the more holistic approach of sensitivity analysis for robustness estimates has not received much attention. In [26], sensitivities of singular values are used to give estimates for robustness properties of a linear feedback system. The advantage of such a method is that it analyzes the effect of multiple parametric uncertainties and hence can be used to enhance the usual robust stability approaches.

Our problem statement is motivated by this sensitivity analysis approach for robustness and by the lack of existing results for robustness estimates of error in structured model reduction. In particular, we give robustness guarantees for the error in model reduction under parametric uncertainties. Using linear analysis, we give a robustness metric for QSSA-style model reduction of linear dynamical

systems. We present a complementary approach that employs sensitivity analysis of the error in model reduction to estimate the robustness under parametric uncertainties. This approach works for nonlinear dynamical systems as well. We demonstrate our method with the help of common biomolecular system examples. Using our results on the model reduction robustness metric, we derive a new mathematical model for gene expression that accounts for ribosomal loading better than the commonly used models in the literature. Some preliminary versions of results in this paper are in [27]–[29].

II. Preliminaries

A. Notation

We denote an eigenvalue of a matrix P by $\lambda(P)$. The maximum eigenvalue will be denoted by $\lambda_{\max}(P)$. For a state-dependent matrix $P(x)$ we denote the maximum eigenvalue of P over all values of x by $\lambda_{\max_x}(P)$. Throughout this paper, we consider the Euclidean 2-norm for vectors. For example, we use the notation $\|x\|$ for the 2-norm of $x \in \mathbb{R}^n$ and similarly for matrices $\|\cdot\|$ represents the induced 2-norm.

B. Problem Formulation

The full system model is given by the following linear autonomous state-space equation

$$\dot{x} = A(\theta)x, \quad y = Cx, \quad x(0) = x_0, \tag{1}$$

where $x \in \mathbb{R}^n$ are state variables, the output vector is $y \in \mathbb{R}^k$ and $\theta = \left[\theta_1, \theta_2, ..., \theta_p\right]^T$ is the vector of all model parameters. We consider a structured model reduction in this paper where the dynamics of a subset of states (x_c) are collapsed (converted to algebraic relationships) on account of being at quasi-steady-state. The remaining states are the states of the reduced model, $\hat{x}$. The reduced model is given by

$$\dot{\hat{x}} = \hat{A}(\theta)\hat{x}, \quad \hat{y} = \hat{C}\hat{x}, \quad \hat{x}(0) = \hat{x}_0, \tag{2}$$

where $\hat{x} \in \mathbb{R}^{\hat{n}}$ are the reduced state variables and $\hat{y} \in \mathbb{R}^k$ is the output vector. We assume that the full and the reduced model have the same number of outputs but different dynamics. Throughout this paper, we also assume that both the full and the reduced systems are asymptotically stable and observable. This model reduction corresponds to a time-scale separation in the dynamics of the full model where the meaning of all states and parameters is retained in the reduced model. It is a relaxed form of singular perturbation theory based model reduction in that it does not require the system to be in the standard separable form, which is a hard condition to satisfy for general system dynamics, as we will see next.

Singular perturbation theory [7] is the standard way to derive reduced models and bounds on error in model reduction for the problem statement given above. However, to use singular perturbation theory the system dynamics need to be separable according to the different time scales. For the problem formulation above, the requirement would be that we can write the dynamics in the following form,

$$\begin{aligned} \dot{\hat{x}} &= A_{11}\hat{x} + A_{12}x_c, \\ \epsilon\dot{x}_c &= A_{21}\hat{x} + A_{22}x_c, \end{aligned} \tag{3}$$

where $A_{11} \in \mathbb{R}^{\hat{n}\times\hat{n}}, A_{22} \in \mathbb{R}^{(n-\hat{n})\times(n-\hat{n})}$, and similarly we have A_{12} and A_{21}. Now, under the condition that $\epsilon \to 0$ and Hurwitz A_{22}, we get the reduced model as in equation (2) with,

$$\dot{\hat{x}} = (A_{11} - A_{12}A_{22}^{-1}A_{21})\hat{x} := \hat{A}\hat{x}. \tag{4}$$

In Figure 1, the block diagram for a system with time-scale separation is given. Using singular

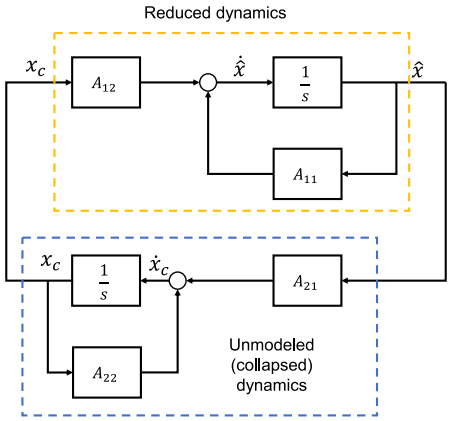

Fig. 1: Structured model reduction of a linear system. The $1/s$ block represents an integrator.

perturbation theory, conditions can be derived under which the error in model reduction converges to zero:

$$\|x - \hat{x}\| \leq O(\epsilon) \tag{5}$$

for a time-scale separation parameter $\epsilon \in [0, \epsilon^*]$. The error in model reduction goes to zero as $\epsilon \to 0$. This is the standard model reduction problem using singular perturbation to separate the time scales of the model where the dynamics of the "fast" states of the system (x_c) are collapsed to zero when $\epsilon \to 0$ and the dynamics of the "slow" states ($\hat{x}$) is the reduced dynamics.

Under uncertainties in the system dynamics, it is important to analyze the robustness of the model reduction. The results on robustness for singular perturbation based model reduction either focus on robust controller design under uncertainties for singularly perturbed systems [22], [23] or analyze the effect on the time-scale separation parameter ϵ due to uncertainties. Recently, a singular perturbation margin [30] (similar to the gain margin and phase margin definitions) has been proposed to assess the robust stability of singularly perturbed systems under uncertainties. It is defined using the ϵ^* given in the error approximation equation (5). This framework can be used to compute a robustness estimate of the model reduction error for the singular perturbation method. An extension of this robust stability margin for nonlinear dynamics is given in [31]. Despite the rich body of literature on singular perturbation theory, the major limitation of such a model reduction approach remains that the system dynamics must be written in the standard form (3). As stated in [32], for physical systems it is usually not straightforward to put a model in the singularly perturbed form since the choice of combination(s) of parameters to be considered small is not always clear. Hence, the relaxed approach of using QSSA is common for various applications.

In QSSA, the dynamics of a set of states are collapsed to zero to get the reduced-order model. The choice of states to be collapsed is usually driven by known time-scale separations in the system model. Although QSSA is a widely used approach it does not necessarily guarantee error performance as in equation (5). The limitation of QSSA is that the mathematical justification and conditions for approximating a variable to be at a steady-state are not always obvious. As a result, there could be many possibilities of reduced models and so it is the designer's task to find a "correct" QSSA based model reduction. Towards that end, in [21], a structured model reduction algorithm is presented that guides the choice of collapsed states so that the error between the output of the full and the reduced models is minimized. Other QSSA error analysis [19] approaches can also be used for this purpose. In this paper, we briefly discuss model reduction error results for QSSA further. Our main focus

is on the problem of robustness of this structured model reduction, that is, how robust a particular model reduction is under parametric uncertainties.

To formulate this problem, we first construct an augmented state-space system that consists of variables of the full and the reduced model together:

$$\bar{x} := \begin{bmatrix} x \\ \hat{x} \end{bmatrix}.$$

We denote all augmented variables similarly with a bar on top of the usual variables. So,

$$\bar{A} := \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}.$$

For the augmented state variable, we can write the following state-space system,

$$\dot{\bar{x}} = \bar{A}(\theta)\bar{x}, \quad \zeta = \bar{C}\bar{x}, \qquad \bar{x}(0) = \begin{bmatrix} x_0 \\ \hat{x}_0 \end{bmatrix}, \tag{6}$$

where ζ is the error in model reduction defined as $\zeta = y - \hat{y}$, hence $\bar{C} = \begin{bmatrix} C & -\hat{C} \end{bmatrix}$.

To study the robustness of the structured model reduction (that is the robustness of deriving the particular reduced model $(\hat{x}, \hat{A})$ under uncertainties in model parameters) we need an upper bound on $\|\zeta\|$ as the parameters θ vary. For the linear augmented system we can write the following by solving for $\zeta(t, \theta)$,

$$\zeta(t, \theta) = \bar{C}e^{\bar{A}t}\bar{x}(0), \tag{7}$$

where $\zeta \in \mathbb{R}^k$.

Lemma 1 (See [33])**.** *For a Hurwitz matrix A, the norm of the matrix exponential is bounded above as*

$$\left\|e^{At}\right\| \leq e^{-|\mu|t}$$

for all $t \geq 0$, where μ is the logarithm norm of A [34]. For the log-norm induced by the 2-norm, we have that

$$\mu(A) = \frac{\lambda_{\max}(A + A^T)}{2},$$

and for Hurwitz A, μ is always negative.

We use this to give an important result on the derivative of the matrix exponential with respect to a parameter.

Lemma 2. *The derivative of the matrix exponential e^{At} with respect to a parameter θ_i is given by*

$$\frac{\partial e^{At}}{\partial \theta_i} = \int_0^t e^{(t-\tau)A} \frac{\partial A}{\partial \theta_i} e^{\tau A} d\tau. \tag{8}$$

If A is Hurwitz, the norm of the derivative of the matrix exponential with θ_i is bounded above by

$$\left\| \frac{\partial e^{At}}{\partial \theta_i} \right\| \leq \left\| \frac{\partial A}{\partial \theta} \right\| t e^{-|\mu|t}. \tag{9}$$

where $|\mu|$ is the absolute value of the log-norm of A as in Lemma 1.

Proof. The first part of the lemma (in equation (8)) is a result proven in [35] and a simplified version is given in [36] and [37]. We give an alternative proof in Appendix 1-A. To prove the second part,

given in equation (9), write the norm of the derivative of the matrix exponential with respect to a parameter θ_i as

$$\left\| \frac{\partial e^{At}}{\partial \theta_i} \right\| = \left\| \int_0^t e^{(t-\tau)A} \frac{\partial A}{\partial \theta_i} e^{\tau A} d\tau \right\| \leq \int_0^t \left\| e^{(t-\tau)A} \right\| \left\| \frac{\partial A}{\partial \theta_i} \right\| \left\| e^{\tau A} \right\| d\tau.$$

Now using the result of Lemma 1, since A is Hurwitz, we have,

$$\left\| \frac{\partial e^{At}}{\partial \theta_i} \right\| \leq \left\| \frac{\partial A}{\partial \theta_i} \right\| \int_0^t e^{-(t-\tau)|\mu|} e^{-\tau|\mu|} d\tau.$$

Solving the above integral, we get the desired result:

$$\left\| \frac{\partial e^{At}}{\partial \theta_i} \right\| \leq \left\| \frac{\partial A}{\partial \theta_i} \right\| t e^{-|\mu|t}. \qquad \square$$

With the result from Lemma 1, we can conclude that under our assumption of asymptotically stable full and reduced models, we have that $\bar{A}$ is Hurwitz, and hence the error dynamics given in equation (7) converges to zero at steady-state. In this way, we have set up the problem to focus solely on the analysis of the robustness of model reduction while assuming that the important problem of minimizing the model reduction error has already been addressed. For a given structured model reduction (and hence the corresponding augmented system above), we can get a bound on the error in model reduction as the model parameters vary to give a robustness estimate for this model reduction. We construct a normalized [38] robustness distance estimate for this purpose by computing the change in error with parameter perturbations around any nominal values given by θ_i^*:

$$d_R = \sum_{i=1}^{p} \frac{\theta_i^*}{\|\zeta(t, \theta_i^*)\|} \cdot \left\| \left. \frac{\partial \zeta}{\partial \theta_i} \right|_{\theta_i = \theta_i^*} \right\|, \tag{10}$$

where $\zeta(t, \theta_i^*)$ is the non-zero error in model reduction for $t > 0$ and parameter $\theta_i = \theta_i^*$. We define the sensitivity of the error with respect to a parameter θ_i in the equation above as S_ζ:

$$S_\zeta = \frac{\partial \zeta}{\partial \theta_i},$$

where $S_\zeta \in \mathbb{R}^k$. Using this distance estimate, we propose the following robustness metric to determine the performance of reduced models under parameter uncertainties:

$$R = \frac{1}{1 + d_R} \tag{11}$$

Clearly, for any reduced model we have that $R \in (0, 1)$ with $R \to 0$ corresponding to worst robust performance and $R \to 1$ corresponding to the best robust performance. Hence, our goal in this paper is to compute R (through a bound on the norm of S_ζ) to give a robustness metric for each possible reduced model. The results in the next section give upper bounds to the norm of S_ζ for linear and nonlinear system settings. With the help of some examples, we then demonstrate the computation of the robustness metric as discussed above using these bounds on $\|S_\zeta\|$. Finally, to decide a particular reduced model we may use a combination of the error and the robustness metrics to choose a particular reduced model in a given parameter regime.

III. Results

A. Linear System — Uncertain Initial Conditions

In this section, we consider the uncertainties in the initial conditions — $x(0)$ of a linear system. We can use this result to assess the robust performance of different possible structured model reductions when the initial conditions are dependent on the uncertain model parameters.

Theorem 1. *For the structured model reduction of the autonomous linear system (1) to the reduced form of system (2) by using time-scale separation and quasi-steady-state approximation under uncertain initial conditions, the norm of the sensitivity of the error in model reduction S_ζ is bounded above by*

$$\|S_\zeta\|_2^2 \leq \lambda_{\max}(P) \left\| \frac{\partial \bar{x}(0)}{\partial \theta_i} \right\|_2^2,$$

where P is the Lyapunov matrix that solves the equation $\bar{A}^T P + P\bar{A} = -\bar{C}^T\bar{C}$.

Proof. We have the 2-norm [24] of S_ζ defined as:

$$\|S_\zeta\|_2^2 = \int_0^\infty S_\zeta(t)^T S_\zeta(t) dt.$$

Using equation (7), we can write S_ζ for a parameter θ_i as

$$S_\zeta = \bar{C} e^{\bar{A}t} \frac{\partial \bar{x}(0)}{\partial \theta_i},$$

since we have assumed that the matrices $\bar{A}$ and $\bar{C}$ are not dependent on parameters. The norm of S_ζ then becomes

$$\|S_\zeta\|_2^2 = \int_0^\infty \left(\frac{\partial \bar{x}(0)}{\partial \theta_i} \right)^T e^{\bar{A}^T t} \bar{C}^T \bar{C} e^{\bar{A}t} \left(\frac{\partial \bar{x}(0)}{\partial \theta_i} \right) dt. \tag{12}$$

From [39, Ch.5], we know that for an observable asymptotically stable system, there exists a unique matrix P that solves the Lyapunov equation $\bar{A}^T P + P\bar{A} = -\bar{C}^T\bar{C}$ given by the observability Gramian:

$$P = \lim_{N\to\infty} W_{\mathrm{o}}(N) = \lim_{N\to\infty} \int_0^N e^{\bar{A}^T t} \bar{C}^T \bar{C} e^{\bar{A}t} dt,$$

where $W_{\mathrm{o}}(N)$ is the observability Gramian. Substituting this into equation (12) gives us the desired result:

$$\|S_\zeta\|_2^2 = \left(\frac{\partial \bar{x}(0)}{\partial \theta_i} \right)^T \left[\int_0^\infty e^{\bar{A}^T t} \bar{C}^T \bar{C} e^{\bar{A}t} dt \right] \left(\frac{\partial \bar{x}(0)}{\partial \theta_i} \right)$$

$$\Rightarrow \|S_\zeta\|_2^2 \leq \lambda_{\max}(P) \left\| \frac{\partial \bar{x}(0)}{\partial \theta_i} \right\|_2^2. \square$$

B. Linear System — Uncertain System Dynamics

Now we consider the case where the system dynamics given by the $\bar{A}(\theta)$ matrix is dependent on uncertain parameters. For simplicity we denote $\bar{A}(\theta) := \bar{A}$, noting that it is parameter-dependent.

Theorem 2. *For the structured model reduction of the autonomous linear system (1) to the reduced form of system (2) by using time-scale separation and quasi-steady-state approximation under uncertain system dynamics, the norm of the sensitivity of the error in model reduction S_ζ is bounded above by*

$$\|S_\zeta\|_2^2 \leq \tilde{M} \left\| \frac{\partial \bar{A}}{\partial \theta_i} \right\|_2^2 \left\| \bar{C}^T \bar{C} \right\|_2 \|\bar{x}(0)\|_2^2,$$

where $\tilde{M} = 1/4\,|\mu|^3$ and μ is dependent on $\bar{A}$ as given in Lemma 1.

Proof. Write the norm of S_ζ as

$$\|S_\zeta\|_2^2 = \int_0^\infty S_\zeta(t)^T S_\zeta(t) dt. \tag{13}$$

To derive the bounds, we first write the partial derivative of $\zeta(t, \theta)$ with respect to a parameter θ_i as given in equation (7),

$$S_\zeta = \frac{\partial \zeta}{\partial \theta_i} = \bar{C} \frac{\partial e^{\bar{A}t}}{\partial \theta_i} \bar{x}(0), \tag{14}$$

assuming that the output matrix $\bar{C}$ and the initial conditions are independent of model parameters. We can write the norm of S_ζ as,

$$\|S_\zeta\|^2 = \int_0^\infty \bar{x}(0)^T \left(\frac{\partial e^{\bar{A}t}}{\partial \theta_i} \right)^T \bar{C}^T \bar{C} \left(\frac{\partial e^{\bar{A}t}}{\partial \theta_i} \right) \bar{x}(0) dt \leq \int_0^\infty \left\| \frac{\partial e^{\bar{A}t}}{\partial \theta_i} \right\|^2 \left\| \bar{C}^T \bar{C} \right\| \|\bar{x}(0)\|^2 dt.$$

Using the result from Lemma 2, we can write,

$$\|S_\zeta\|^2 \leq \left\| \frac{\partial \bar{A}}{\partial \theta_i} \right\|^2 \left\| \bar{C}^T \bar{C} \right\| \|\bar{x}(0)\|^2 \int_0^\infty t^2 e^{-2|\mu|t} dt.$$

We can evaluate the integral above by parts:

$$\int_0^\infty t^2 e^{-2|\mu|t} dt = \frac{1}{4\,|\mu|^3}.$$

We get the desired result for the norm of S_ζ by substituting this integral,

$$\|S_\zeta\|^2 \leq \frac{1}{4\,|\mu|^3} \left\| \frac{\partial \bar{A}}{\partial \theta_i} \right\|^2 \left\| \bar{C}^T \bar{C} \right\| \|\bar{x}(0)\|^2.$$

The robustness metric R follows by using the above bound and equation (11). □

Corollary 2.1. *Under simultaneous parametric uncertainties in system dynamics and initial conditions, we can write the norm of S_ζ as*

$$\begin{aligned} \|S_\zeta\|_2^2 \leq \lambda_{\max}(P) \left\| \frac{\partial \bar{x}(0)}{\partial \theta_i} \right\|_2^2 &+ \frac{1}{4\,|\mu|^3} \left\| \frac{\partial \bar{A}}{\partial \theta_i} \right\|_2^2 \left\| \bar{C}^T \bar{C} \right\|_2 \|\bar{x}(0)\|_2^2 \\ &+ \frac{1}{2\,|\mu|^2} \left\| \frac{\partial \bar{A}}{\partial \theta_i} \right\|_2 \left\| \bar{C}^T \bar{C} \right\|_2 \|\bar{x}(0)\|_2 \left\| \frac{\partial \bar{x}(0)}{\partial \theta_i} \right\|_2 \end{aligned} \tag{15}$$

where P is the Lyapunov matrix and μ depends on $\bar{A}$ as in Lemma 1.

Proof. Using the product rule for the derivative of $\zeta(t,\theta)$ in equation (7), we can prove the desired result by combining the results from Theorem 1 and 2 and working out the algebra for other terms that appears in the total derivative equation. □

We present the results for nonlinear dynamical systems next. In Appendix 1-C, we also show that the robustness metric for the nonlinear dynamics simplifies to the same result as given in Corollary 2.1.

C. Nonlinear Dynamics — Model Reduction Robustness

For nonlinear system dynamics, the approach above to derive the bound on the robustness guarantee does not work because we cannot obtain the error dynamics analytically as was possible for linear dynamics in equation (7). An alternate approach for deriving a bound on the sensitivity of the error is using local sensitivity analysis. Consider the following nonlinear dynamics of the full system

$$\dot{x} = f(x,\theta), \quad y = Cx, \quad x(0) = x_0. \tag{16}$$

The reduced nonlinear model is given using similar notation

$$\dot{\hat{x}} = \hat{f}(\hat{x},\theta), \quad \hat{y} = \hat{C}\hat{x}, \quad \hat{x}(0) = \hat{x}_0. \tag{17}$$

For both the full and the reduced dynamics, we assume that $f : E \subset \mathbb{R}^n \to \mathbb{R}^n$, $\hat{f} : \hat{E} \subset \mathbb{R}^{\hat{n}} \to \mathbb{R}^{\hat{n}}$ are locally Lipschitz functions and initial conditions $x(0) \in E$, $\hat{x}(0) \in \hat{E}$. The local Lipschitz continuity gives us that the solutions $x(t)$ and $\hat{x}(t)$ exist and are unique for a finite time interval. We further assume that equilibrium points $x^* \in \mathbb{R}^n$, $\hat{x}^* \in \mathbb{R}^{\hat{n}}$ exist and there is no finite escape time. See [40, Corollary 2.5] for the sufficient smoothness conditions that are needed on the system dynamics for these assumptions to hold. We make these assumptions throughout this paper for any nonlinear function describing the system dynamics.

Theorem 3. *For the structured model reduction of the nonlinear dynamical system (16) to the reduced system (17) by using time-scale separation and quasi-steady-state approximation, the norm of the sensitivity of the error in model reduction, S_ζ is bounded above by*

$$\|S_\zeta\|_2^2 \leq \lambda_{\max_{\bar{x}}}(P(\bar{x}))\left\|\bar{S}_0\right\|_2^2 + 2\int_0^\infty \left\|\bar{Z}^T P(\bar{x})\bar{S}\right\|_2 dt + \lambda_{\max_{\bar{x}}}(\dot{P}(\bar{x}))\int_0^\infty \left\|\bar{S}\right\|_2^2 dt \tag{18}$$

where $P(\bar{x})$ is a matrix that solves the Lyapunov equation $\bar{J}(\bar{x})^T P(\bar{x}) + P(\bar{x})\bar{J}(\bar{x}) = -\bar{C}^T\bar{C}$ at the point $\bar{x}(t) = \bar{x}$ in the augmented nonlinear system trajectory. Here $\bar{J}(\bar{x})$ is the Jacobian matrix at $\bar{x}$, $\bar{Z}$ is the sensitivity to parameter, and $\bar{S}$ is the sensitivity coefficients vector of the augmented system, given by:

$$\bar{J}(\bar{x}) = \begin{bmatrix} \frac{\partial f}{\partial x} & 0 \\ 0 & \frac{\partial \hat{f}}{\partial \hat{x}} \end{bmatrix}, \quad \bar{Z} = \begin{bmatrix} \frac{\partial f}{\partial \theta_i} \\ \frac{\partial \hat{f}}{\partial \theta_i} \end{bmatrix}, \quad \bar{S} = \begin{bmatrix} \frac{\partial x}{\partial \theta_i} \\ \frac{\partial \hat{x}}{\partial \theta_i} \end{bmatrix}.$$

Proof. At the point $\bar{x}(t) = \bar{x}$ in the augmented nonlinear system trajectory, write the sensitivity system equations [38] for a parameter $\theta_i \in \theta$:

$$\dot{\bar{S}} = \bar{J}(\bar{x})\bar{S} + \bar{Z}, \quad S_\zeta = \bar{C}\bar{S}.$$

For the norm of the sensitivity of the error, we can write

$$\|S_\zeta\|_2^2 = \int_0^\infty \bar{S}\bar{C}^T\bar{C}\bar{S}dt.$$

For every $\bar{x}$, given that there exists a matrix $P(\bar{x}) = P^T(\bar{x}) \succ 0$ such that $\bar{J}(\bar{x})^T P(\bar{x}) + P(\bar{x})\bar{J}(\bar{x}) = -\bar{C}^T\bar{C}$, consider a function $V(\bar{S}) = \bar{S}^T P(\bar{x})\bar{S}$. Differentiating this function with respect to time, we have that

$$\frac{dV}{dt} = \bar{S}^T(\bar{J}^T P + P\bar{J})\bar{S} + \bar{S}^T\dot{P}(\bar{x})\bar{S} + (\bar{Z}^T P\bar{S} + \bar{S}^T P\bar{Z}).$$

For simplicity of exposition, we denote $P(\bar{x})$ as P noting that this is a state-dependent Lyapunov matrix, similarly $\bar{J}$ is a state-dependent Jacobian matrix of the augmented nonlinear system evaluated at $\bar{x}(t) = \bar{x}$. Integrating the expression above from 0 to ∞ and then substituting the expression for $\|S_\zeta\|_2^2$, we get

$$\|S_\zeta\|_2^2 = \int_0^\infty -\frac{dV}{dt}dt + \int_0^\infty \bar{S}^T\dot{P}\bar{S}dt + \int_0^\infty (\bar{Z}^T P\bar{S} + \bar{S}^T P\bar{Z})dt,$$

$$\|S_\zeta\|_2^2 = -\lim_{t\to\infty} V(\bar{S}(t)) + V(\bar{S}(0)) + \int_0^\infty \bar{S}^T\dot{P}\bar{S}dt + \int_0^\infty (\bar{Z}^T P\bar{S} + \bar{S}^T P\bar{Z})dt,$$

Since, P is a positive semidefinite matrix, V will be a non-negative function [32]. Using this fact and denoting $\bar{S}(0) := \bar{S}_0$, we have the inequality

$$\|S_\zeta\|_2^2 \leq \bar{S}_0^T P\bar{S}_0 + \int_0^\infty \bar{S}^T\dot{P}\bar{S}dt + \int_0^\infty (\bar{Z}^T P\bar{S} + \bar{S}^T P\bar{Z})dt. \tag{19}$$

For the first part of this equation, we can write

$$\left\|\bar{S}_0^T P\bar{S}_0\right\| \leq \lambda_{\max_{\bar{x}}}(P)\left\|\bar{S}_0\right\|^2,$$

where we compute the maximum eigenvalue of P over all points $\bar{x}$. Note that if the initial conditions are independent of all model parameters then $\bar{S}_0$ = 0. Similarly, we get,

$$\left\|\int_0^\infty \bar{S}^T\dot{P}\bar{S}dt\right\| \leq \lambda_{\max_{\bar{x}}}(\dot{P})\int_0^\infty \left\|\bar{S}\right\|^2 dt.$$

Evaluating the last part in equation (19) proves the theorem, combined with the above results:

$$\int_0^\infty \left(\bar{Z}^T P\bar{S} + \bar{S}^T P\bar{Z}\right)dt \leq 2\int_0^\infty \left\|\bar{Z}^T P\bar{S}\right\|_2 dt.$$

□

Remark. For computational purposes, we may modify the bound above as

$$\|S_\zeta\|_2^2 \leq \lambda_{\max_{\bar{x}}}(P)\left\|\bar{S}(0)\right\|_2^2 + 2N\sup_t\left\|\bar{Z}^T P\bar{S}\right\|_2 + N\lambda_{\max_{\bar{x}}}(\dot{P})\sup_t\left\|\bar{S}\right\|_2^2 \tag{20}$$

where $N > 0$ is a time at which the system solution is arbitrarily close to the equilibrium point, that is, $\|\bar{x}(N) - \bar{x}^*\| < \epsilon$ for some $\epsilon > 0$ and $\bar{x}^*$ is the equilibrium point for the augmented nonlinear dynamical system. Since we compute $P(\bar{x})$ at every time step, we can also compute $\dot{P}(\bar{x})$ numerically. Further, a direct computation for S_ζ is also possible but the bounds that we give may be used as interpret-able metrics for system performance analysis.

Next, we consider the structured model reduction of a controlled nonlinear dynamical system. The nonlinear system dynamics are then given by,

$$\begin{aligned} \dot{x} &= f(x,\theta) + g(x,\theta)u, \\ y &= Cx, \quad x(0) = x_0. \end{aligned} \tag{21}$$

The reduced nonlinear model is given using similar notation

$$\begin{aligned}\dot{\hat{x}} &= \hat{f}(\hat{x},\theta) + \hat{g}(\hat{x},\theta)u,\\ \hat{y} &= \hat{C}\hat{x}, \quad \hat{x}(0) = \hat{x}_0.\end{aligned} \tag{22}$$

Here, we have assumed a scalar input u for simplicity of exposition. The results that follow can be derived for systems with multiple inputs as well but with more complicated algebra.

Theorem 4. *For the structured model reduction of the nonlinear controlled dynamics (21) to the reduced dynamics (22), the norm of the sensitivity of the error S_ζ is bounded above by,*

$$\begin{aligned}\|S_\zeta\|_2^2 \leq \lambda_{\max_{\bar{x}}}(P)\left\|\bar{S}_0\right\|^2 + 2\int_0^\infty \left(\left\|\bar{Z}_f^T P\bar{S}\right\| + \left\|u^T\bar{S}^T\bar{J}_g^T P\bar{S}\right\| + \left\|\bar{Z}_g^T P\bar{S}u\right\|\right)dt\\ + \lambda_{\max_{\bar{x}}}(\dot{P})\int_0^\infty \left\|\bar{S}\right\|^2 dt\end{aligned}$$

if there exists $P(\bar{x}) = P(\bar{x})^T \succ 0$ such that $P(\bar{x})\bar{J}_f(\bar{x}) + \bar{J}_f(\bar{x})^T P(\bar{x}) = -\bar{C}^T\bar{C}$ at the point $\bar{x}(t) = \bar{x}$ in the augmented nonlinear system trajectory. Here $\bar{J}_f(\bar{x}), \bar{J}_g(\bar{x})$ and $\bar{Z}_f, \bar{Z}_g$ are the Jacobian and parameter sensitivity matrices of augmented nonlinear functions $\bar{f}$ and $\bar{g}$ respectively.

Proof. The proof follows similar to the proof of the previous theorem by defining a function $V(\bar{S}) = \bar{S}^T P(\bar{x})\bar{S}$ at every point $\bar{x}$ in the system trajectory and calculating the bound for $\|S_\zeta\|$. See Appendix 1-B for full proof. □

The results above assumed that the outputs of the system are linearly related to the states $y = Cx$, however, we can derive similar results even without this assumption.

Theorem 5. *For the system dynamics in equation (21) with output dynamics given by $y = h(x,\theta)$ and the reduced model dynamics given in equation (22) with output dynamics given by $\hat{y} = \hat{h}(\hat{x},\theta)$, the norm of the sensitivity of error is bounded above by,*

$$\begin{aligned}\|S_\zeta\|_2^2 \leq \lambda_{\max_{\bar{x}}}(P)\left\|\bar{S}(0)\right\|^2 + 2\int_0^\infty \left(\left\|\bar{Z}_f^T P\bar{S}\right\|_2 + \left\|u^T\bar{S}^T\bar{J}_g^T P\bar{S}\right\| + \left\|\bar{Z}_g^T P\bar{S}u\right\|\right)dt\\ + 2\left\|\bar{Z}_h\right\|_2^2 + \lambda_{\max_{\bar{x}}}(\dot{P})\int_0^\infty \left\|\bar{S}\right\|^2 dt\end{aligned}$$

if there exists $P(\bar{x}) = P(\bar{x})^T \succ 0$ such that $\bar{J}_f(\bar{x})^T P(\bar{x}) + P(\bar{x})\bar{J}_f(\bar{x}) = -\bar{C}_1^T\bar{C}_1$, where

$$\bar{C}_1 = \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}J_h(\bar{x}) & 0\\ 0 & \hat{J}_h(\bar{x})\end{bmatrix} \triangleq C_e\bar{J}_h(\bar{x}),$$

and $\bar{J}_*(\bar{x}), \bar{Z}_*$ refer to the Jacobian and the parameter sensitivity matrices for each nonlinear function respectively at a point $\bar{x}(t) = \bar{x}$ in the system trajectory given by:

$$\bar{J}_*(\bar{x}) = \begin{bmatrix}J_*(x) & 0\\ 0 & \hat{J}_*(\hat{x})\end{bmatrix}, \quad \bar{Z}_* = \begin{bmatrix}Z_*\\ \hat{Z}_*\end{bmatrix}, \quad * : f, g, h.$$

Proof. For every point $\bar{x}(t) = \bar{x}$, we write the sensitivity of the error in model reduction for a parameter $\theta_i \in \theta$ using chain rule:

$$S_\zeta = \frac{d(h-\hat{h})}{d\theta_i} = \left(\frac{\partial h}{\partial x}\right)\left(\frac{\partial x}{\partial \theta_i}\right) + \left(\frac{\partial h}{\partial \theta_i}\right) - \left[\left(\frac{\partial \hat{h}}{\partial \hat{x}}\right)\left(\frac{\partial \hat{x}}{\partial \theta_i}\right) + \left(\frac{\partial \hat{h}}{\partial \theta_i}\right)\right].$$

Define the Jacobians $J_h(\bar{x}), \hat{J}_h(\bar{x})$, parameter sensitivity matrices $Z_h, \hat{Z}_h$ for h and $\hat{h}$ respectively and substitute back to write,

$$S_\zeta = \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} J_h(x) & 0 \\ 0 & \hat{J}_h(\hat{x}) \end{bmatrix} \begin{bmatrix} S \\ \hat{S} \end{bmatrix} + \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} Z_h \\ \hat{Z}_h \end{bmatrix},$$
$$S_\zeta = \bar{C}_1 \bar{S} + C_e \bar{Z}_h.$$

Now, consider a function $V(\bar{S}) = \bar{S}^T P(\bar{x}) \bar{S}$ and proceed in a similar way as in the proof of previous results to write $\|S_\zeta\|_2^2$,

$$\begin{aligned} \|S_\zeta\|_2^2 &= \int_0^\infty \bar{S}^T \bar{C}_1^T \bar{C}_1 \bar{S} dt + \int_0^\infty \bar{Z}_h^T C_e^T C_e \bar{Z}_h dt \\ &= \int_0^\infty \bar{S}^T \bar{C}_1^T \bar{C}_1 \bar{S} dt + 2 \left\| \bar{Z}_h \right\|_2^2. \end{aligned}$$

Using the result from Theorem 4 and $\bar{J}_f(\bar{x})^T P(\bar{x}) + P(\bar{x}) \bar{J}_f(\bar{x}) = -\bar{C}_1^T \bar{C}_1$, we get the desired result. □

Remark. For all results on robustness estimate, it is possible to further simplify the bounds using the method in [21] to express this bound only in terms of $\hat{x}$, the reduced state variables and x_c, the collapsed state variables.

D. Input-Output mapping

For forced nonlinear systems, in addition to the error and the robustness metric, it is also important to consider the input-output mapping so that the response of a reduced model for an input is similar to that of the full model. If the mapping from $u \mapsto y$ is linear, the computation of induced system norms is a well studied topic, see for example [41]. However, for general nonlinear systems, this is still an active research area [42] with results only available under certain structural conditions on the system dynamics. Similar to the gap-metric for linear systems [43], [44], there have been a few results on computation of the gap metric for nonlinear systems [45]. We can use similar computational results in our model reduction procedure by assigning the gap metric to each nonlinear reduced-order model.

On the other hand, if we are only interested in the response to an input at steady-state or at a fixed number of points in the response, then we can linearize the dynamics at these points and assess the induced system norm for each reduced model and use this as a metric while choosing a reduced-order model. For the system operator $H : u \mapsto y$, we define,

$$\gamma(H) := \sup_{u \neq 0} \frac{\|Hu\|}{\|u\|}. \tag{23}$$

E. Automated Model Reduction

Given a well-behaved full system model (linear or nonlinear) under the stability assumptions, we develop an automated model reduction pipeline based on QSSA and our results on the robustness of model reduction. In the first step, we solve the time-scale separation problem to obtain all possible reduced models. Then, we may numerically bound the error in the model reduction for each reduced model as shown in [21] or compute an error metric ($\|\zeta\| = \|y - \hat{y}\|$) by directly simulating the system. The latter method also works for nonlinear dynamical systems. We may reject any reduced model at this stage that does not exhibit the desired level of error performance. However, to make a clear choice of a reduced model we compute a robustness metric for each reduced model using the results in the previous section. Finally, if the system is controlled then we may use the input-output mapping metric as well for the linearized system dynamics to determine the performance of

a reduced model in terms of the input-output response. Note that it is not required to explore the space of all possible reduced models to compute any of the proposed metrics. This is especially important for large system models where symbolically computing all possible reduced models would be computationally infeasible. Hence, for large system models, we can compute a heuristically chosen set of reduced models. Then, we can compute the error, robustness, and input-output mapping metrics as desired for the reduced models of interest.

We provide an implementation of these model reduction tools in a package called `AutoReduce`. This package is available as an open-source project on GitHub [46]. The software is based on Python Sympy [47] and works by loading a Systems Biology Markup Language (SBML) [48] model to import any biological system model. The following important tools are available in this software:
1) Time-scale separation: To solve the time-scale separation problem for a given model, the package methods can be used to set the dynamics of the given states to collapse (x_c) to zero and to automatically substitute back into the dynamics for the reduced state variables ($\hat{x}$). This method automates the QSSA procedure and can be used to compute QSSA based reduced models without specifying parameter values.
2) Conservation laws: Using conservation laws, we can eliminate states that are conserved from the imported SBML model. We integrate this approach in our method similar to the method in [49]. The conservation laws can be explicitly defined and these are used in the software package accordingly to eliminate state variables and compute the reduced-order models.
3) Comparison metrics: The performance metrics discussed above are implemented in this package as well. For any pair of full model and reduced model, metrics such as the norm of the error ($\|\zeta\|$) or the robustness metric ($\|S_\zeta\|$) can be computed.

IV. Examples

In this section, we apply our model reduction robustness estimate results to biomolecular systems to demonstrate the utility of our approach.

A. Enzymatic reaction dynamics

For the enzymatic reaction system, we can write a chemical reaction network model using mass-action kinetics for the following reactions:

$$E + S \underset{d}{\overset{a}{\rightleftharpoons}} C \xrightarrow{k} P + E \tag{24}$$

The full mass-action kinetics based model with four species is given by:

$$\frac{dS}{dt} = -aES + dC, \quad \frac{dC}{dt} = aES - (d+k)C$$
$$\frac{dE}{dt} = -aES + dC + kC, \quad \frac{dP}{dt} = kC$$

For this example, we discuss the standard model reduction approach using the singular perturbation method along with our automated model reduction approach. Recall that for singular perturbation theory approach it is important to separate the dynamics analytically according to the time scales, which might not be possible for general system dynamics. But for this example, it is possible to derive the reduced model using singular perturbation theory as well.

We write the conservation laws for this system as $E = E_{\text{tot}} - C$ and $S + C + P = S_{\text{tot}}$, where E_{tot} and S_{tot} are the total enzyme and substrate concentrations respectively. We also specify that the output of interest for this system is concentration of P, the product species.

For singular perturbation approach, we manually simplify [50, Ch.3] the system dynamics to write the system with conservation laws in the required form (3):

$$\epsilon \frac{dC}{dt} = \frac{k}{K_d}\left(E_{\text{tot}} - C\right)\left(S_{\text{tot}} - C - P\right) - kC - \epsilon kC,$$

$$\frac{dP}{dt} = kC,$$

where $K_d = d/a$ and the small parameter $\epsilon := k/d$. When $\epsilon \to 0$, we obtain the reduced model. For singular perturbation theory based model reduction, a singular perturbation margin (SPM) has been proposed in the literature to assess the robustness of model reduction [51]. The SPM evaluates the maximum value of ϵ such that the singularly perturbed system dynamics are unstable. But for biologically relevant parameters the dynamics given above do not become unstable and hence we do not have a metric to compute the robustness of model reduction directly.

The automated model reduction package first eliminates the two state-variables E and S based on the conservation laws. In the next step, it solves for time-scale separation to obtain a possible reduced-order model. For this reduced model, we computed the error metric ($\|\zeta\| = 4.2 \times 10^{-3}$) as well as the robustness metric over the model initial conditions and parameters [46]. If the system performance is satisfactory, then we can conclude that the final reduced model is given by:

$$\frac{dP}{dt} = K_L \bar{C},$$

where $K_L = kE_{\text{tot}}$ is the lumped parameter and $\bar{C}$ is a function of $E_{\text{tot}}, S_{\text{tot}}$ and model parameters given by

$$\bar{C} = \frac{S_{\text{tot}} - P}{S_{\text{tot}} + K_m}$$

where $K_m = (d+k)/a$. All of the simulations, and code required for computations for this example is available at [46]. An important distinction with our automated computational method to derive reduced models is that any spurious conditions are easier to catch. For this example, it is shown in [52, Ch.3] that the commonly used model derived above fails to capture the true dynamics under certain parameter regimes. The robustness properties of the model reduction would inform the analysis whenever such a condition may occur.

B. Gene Expression — Design Space Exploration

With this example, we demonstrate that a two-state ribosome and protein model can robustly capture the chemical reaction dynamics of gene expression. We analytically derive phenomenological models for gene expression with exactly known mappings to the mechanistic details. We explore the modeling assumptions of time-scale separation [53]–[55], conservation laws [49], [52], [56] and prove the robustness of various models under certain parametric conditions.

For gene expression, a two-state model is commonly used [1], [2], [57]–[59] in the literature that models the dynamics of the mRNA transcript (T) and the protein concentration (X) as a function of the DNA copy number (G) and regulatory effects:

$$\frac{dT}{dt} = k_{tx} f_{tx}(G, \cdot) - d_T T$$

$$\frac{dX}{dt} = k_{tl} f_{tl}(T) - d_X X,$$

where k_{tx} is defined as the transcription rate and k_{tl} is the translation rate. Similarly, d_T and d_X are the degradation and dilution parameters for the transcript and the protein respectively. The function $f_{tx}(\cdot)$ is usually a Hill function dependent on the mechanism of transcriptional regulation (activation or repression). For constitutive expression, this is assumed to be a constant function of the DNA copy number, $f_{tx}(G) = kG$. Similarly, $f_{tl}(\cdot)$ could be a constant or a Hill function dependent on the transcriptional regulation mechanism [52]. Clearly, the parameters in this model and any parameters in the Hill functions all have empirical meanings but an analytical relationship with the mechanistic reaction rates is usually obscured [60]. Moreover, a closer analysis would show that such phenomenological models are only valid under certain assumptions and parameter regimes.

The full CRN model: The full chemical reaction network (CRN) model for the expression of protein X from a single gene G is described in Table I. In this CRN, the gene G is transcribed by RNA polymerase (P) to an mRNA transcript T via a complex (C_1) formation reaction. Then, the transcript T binds to the ribosome R to form the second complex C_2, which then translates to express the protein X. Under the assumption of mass-action kinetics for all reactions, the ordinary differential equation (ODE) model can be derived as shown on the right in Table I. We refer to this as the full CRN model for the rest of this paper.

Reduced-order modeling: For the CRN model, the first step is to solve for the conserved quantities in the model. We assume that the total RNA polymerase in the system and the total ribsomes remain conserved. From Table I, observe that

$$\frac{dP}{dt} + \frac{dC_1}{dt} = 0, \quad \frac{dR}{dt} + \frac{dC_2}{dt} = 0.$$

Hence, for constants P_{tot} and R_{tot}, we have that $P_{\text{tot}} = P + C_1$ and $R_{\text{tot}} = R + C_2$. These conservation laws can be used to eliminate C_1 and C_2. Using `AutoReduce`, we obtained the following reduced-order model under conservation laws,

$$\begin{aligned}
\frac{dP}{dt} &= (k_{tx} + k_{up})(P_{\text{tot}} - P) - k_{bp} G P, \\
\frac{dT}{dt} &= k_{tx}(P_{\text{tot}} - P) + (k_{tl} + k_{ur})(R_{\text{tot}} - R) - k_{br} R T - d_T T, \\
\frac{dR}{dt} &= (k_{tl} + k_{ur})(R_{\text{tot}} - R) - k_{br} R T, \\
\frac{dX}{dt} &= k_{tl}(R_{\text{tot}} - R) - d_X X.
\end{aligned} \tag{25}$$

Next, we obtain various reduced models under different time-scale separation assumptions. Recall that we denote all reduced-order model variables with a hat to differentiate the corresponding variable in the full model, for example, in a reduced model the protein species will be represented as $\hat{X}$ and the corresponding species in the full model is denoted by X. All variables denote the concentrations for each species and parameters take appropriate units. Furthermore, we define the following lumped parameter notations which appear as Hill function activation parameters in the reduced model expressions that we derive next.

$$K_0 := \frac{k_{tl} + k_{ur}}{k_{br}}, \qquad K_1 := \frac{k_{tx} + k_{up}}{k_{bp}}, \qquad K_d := \frac{k_{tl} + k_{ur}}{d_T} \tag{26}$$

The mRNA transcript ($\hat{T}$) and protein ($\hat{X}$) model: Under the assumption that the free ribosomes and the RNA polymerase dynamics are at QSS (that is $\dot{R} = 0, \dot{P} = 0$), we obtain the following model

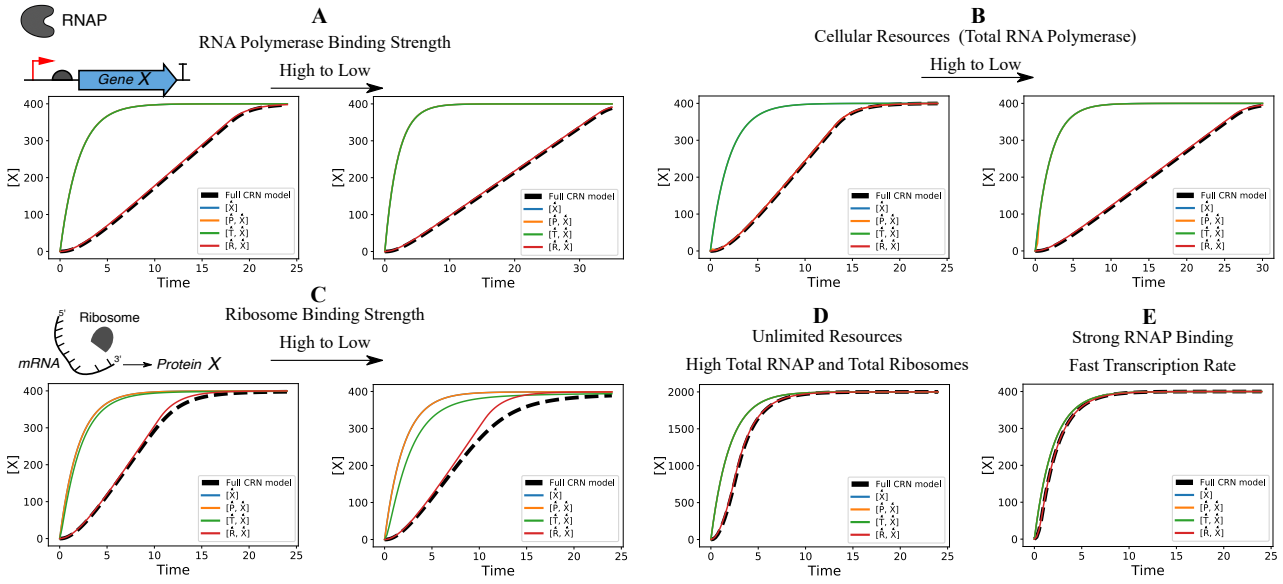

Fig. 2: Performance of the gene expression models under different biological conditions. **(A)** We observe that for weaker binding of RNA polymerase to the promoter region of the DNA, the time-response of the full CRN model is slower as it takes a longer time to reach steady-state. The mathematical model with only the mRNA transcript and protein dynamics is unable to capture this effect since this binding reaction is assumed to be at quasi-steady-state in this model. On the other hand, the $\hat{R}_\Delta$ model, which describes the dynamics of the free ribosome and the protein is able to capture the effect. Note that the performance of the $[\hat{R}, \hat{X}]$ and the $\hat{R}_\Delta$ model are equivalent. **(B)** With decreasing total RNA polymerase, the time-response of the full model is slower and only the ribosome-protein ($[\hat{R}, \hat{X}]$) model is able to account for this effect. **(C)** With decreasing ribosome binding strength, none of the reduced models perfectly capture the full CRN dynamics, but still, the ribosome-protein model is the closest in error performance to the full model. **(D)** With a very high amount of total RNA polymerase and the total ribosome count in the system, all models reach steady-state faster. **(E)** Under strong RNA polymerase binding to the DNA and high transcription rate, we see that all models exhibit good error performance. These can be understood as the ideal conditions under which using a one-state protein dynamics model is also justified. Python code used for this analysis is available publicly on GitHub [46] and can also be run online using this link.

with only the mRNA transcript and the protein dynamics as a function of the DNA copy number G:

$$\begin{aligned} \frac{d\hat{T}}{dt} &= k_{tx} P_{\text{tot}} \left(\frac{G}{K_1 + G} \right) - d_T \hat{T} \\ \frac{d\hat{X}}{dt} &= k_{tl} R_{\text{tot}} \frac{\hat{T}}{K_0 + \hat{T}} - d_X \hat{X}. \end{aligned} \tag{27}$$

The free ribosome ($\hat{R}$) and protein ($\hat{X}$) model: Under the assumption that the mRNA transcript and the RNA polymerase dynamics are at QSS (that is $\dot{T} = 0, \dot{P} = 0$), we obtain the following model with only the free ribosome and the protein dynamics:

$$\begin{aligned} \frac{d\hat{R}}{dt} &= \frac{d_T (R_{tot} - \hat{R})}{K_0^{-1} \hat{R} + K_d^{-1}} - k_{tx} P_{tot} \left(\frac{G}{K_1 + G} \right) \left(\frac{\hat{R}}{\hat{R} + \frac{K_0}{K_d}} \right) \\ \frac{d\hat{X}}{dt} &= k_{tl} \left(R_{\text{tot}} - R \right) - d_X \hat{X}. \end{aligned} \tag{28}$$

Similar to the ribosome-protein ($[\hat{R}, \hat{X}]$) and the mRNA transcript-protein models ($[\hat{T}, \hat{X}]$), it is possible to derive the polymerase-protein ($[\hat{P}, \hat{X}]$) and the only protein model ($[\hat{X}]$). The detailed equations for these two models are not presented here for brevity but their performance is shown in Figures 2 and 3.

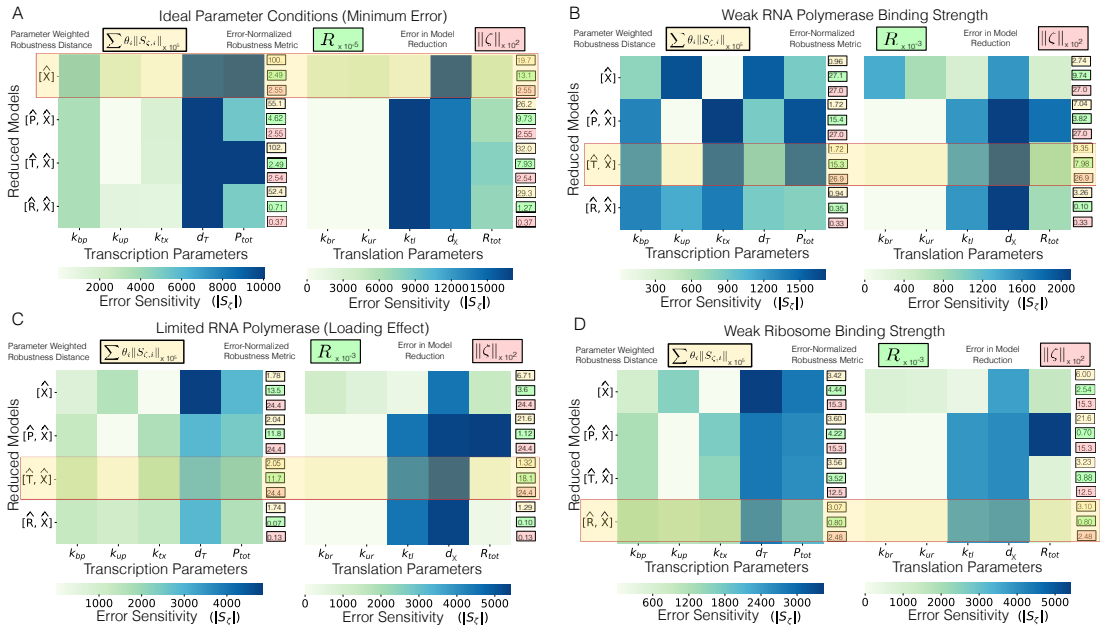

Fig. 3: Robustness analysis of gene expression models. The figure shows the fragility of different models for all model parameters as computed by the sensitivity of the error in protein X concentration between the full model and the reduced models under different parameter conditions. The weighted robustness distance, the normalized robustness metric, and the normed error in model reduction are labeled for each reduced model in each parameter condition. The combined decision metric r, with weights $w_1 = 1, w_2 = 0.3$, is used to choose one final reduced model in each of the four conditions as shown by the red box. **(A)** Ideal parameter conditions under which all reduced models have similar error performance that is close to the full model. The one-state model with only the protein dynamics has the best performance for the combined metric on error and robustness. **(B)** Under the condition of weak RNA polymerase binding strength, we get that the mRNA transcript-protein model has the best performance. **(C)** Under the condition of limited RNA polymerase resources, we get that the mRNA transcript-protein model has the best performance similar to (B).**(D)** Under the condition of weak ribosome binding strength, we get that the ribosome-protein model ($[\hat{R}, \hat{X}]$) has the best performance as all other models have a much higher error and lower robustness. Python code used for this analysis is available publicly on GitHub [46].

The available free ribosome ($\hat{R}_\Delta$) and protein ($\hat{X}$) model: We define the available free ribosomes in the gene expression system as $\hat{R}_\Delta = R_{\text{tot}} - \hat{R}$. Substituting $\hat{R}$ for $R_{\text{tot}} - \hat{R}_\Delta$, we can derive a new reduced-order model from the $\begin{bmatrix}\hat{R} & \hat{X}\end{bmatrix}$ model — the $\hat{R}_\Delta$ model:

$$\frac{d\hat{R}_\Delta}{dt} = k_{tx} P_{tot} \left(\frac{G}{K_1 + G}\right)\left(\frac{R_{\text{tot}} - \hat{R}_\Delta}{R_{\text{tot}} - \hat{R}_\Delta + \frac{K_0}{K_d}}\right) - \left(\frac{d_T}{K_0^{-1}\left(R_{\text{tot}} - \hat{R}_\Delta\right) + K_d^{-1}}\right)\hat{R}_\Delta,$$

$$\frac{d\hat{X}}{dt} = k_{tl}\hat{R}_\Delta - d_X\hat{X}. \tag{29}$$

This model simplifies the robustness analysis that follows. Note that the $\hat{R}_\Delta$ model closely resembles the commonly used gene expression model and all of its terms are exactly similar to the mRNA transcript and protein model but scaled by a ribosome count factor. We give two statements regarding the error performance analysis of the reduced models that we have shown. A justification of these statements is given in detail in Appendix 2-A and 2-B.

Statement 1. *The error between the $\hat{R}_\Delta$ model and the full CRN model is robust to perturbations in the binding/unbinding of ribosome to the mRNA transcript (k_{br}, k_{ur}), and also total resources (both RNA polymerase, P_{tot} and ribosome, R_{tot}). Only the perturbations in the translation rate (k_{tl}) and the protein degradation (d_X) parameters significantly affect the error performance of the $\hat{R}_\Delta$ model.*

Figure 2 shows the error performance of the $\hat{R}_\Delta$ model under various biologically plausible parameter conditions and assumptions. The Euclidean norm of S_ζ is plotted in Figure 3 to compare the robustness of the $\hat{R}_\Delta$ model alongside other models in different biologically plausible parameter conditions. Using $\|S_\zeta\|$ we can obtain a weighted sum to compute the robustness distance d_R for each reduced model using equation (10). A normalized robustness metric R is also shown for the reduced models. To choose a particular reduced model for each parameter condition, we use a linear combination of the error and the robustness distance defined as:

$$r := w_1 \|\zeta\| + w_2 d_R, \tag{30}$$

where w_1 and w_2 are free parameters that can be chosen to weigh the error and the robustness performance respectively. This metric can be used as a single scalar to compare the reduced models, lower r implies better performance. Similar to the result above, we give the following statement for the transcript and the protein model.

Statement 2. *The mathematical model with the mRNA transcript and protein dynamics, given in equation (27), captures the full CRN model dynamics under the assumption of unlimited ribosomal resources and fast binding reactions. As a result, the error performance of the transcript-protein model is directly dependent on the mRNA-ribosome binding/unbinding parameters and the translation parameter. Hence, unlike the $\hat{R}_\Delta$ model, this model is not robust to perturbations in k_{br}, k_{ur}, and R_{tot} in addition to k_{tl} and d_X.*

We scale up the above analysis by using a gene expression model that consists of endonuclease mediated mRNA degradation as well. The results for this analysis are given in Appendix 2-C.

C. Population Control — Non-Identifiability Analysis

We discuss a population control synthetic biological circuit example to demonstrate the application of model reduction to improve parameter inference. Using this example, we also show how our approach can be used to analytically obtain the non-identifiable manifold description for a system. We consider a synthetic circuit that controls the population density and composition of a two-member bacterial consortium [4], [61]. In this circuit, under the control of two inducer input signals, each cell kills itself by expressing a toxin protein (ccdB). Each cell type rescues the other cell type by producing a signal that activates the expression of an anti-toxin (ccdA) in its partner cell. There are two different fluorescent signal readouts corresponding to the population of each cell type, given by L_1 and L_2 in the model. The circuit schematic is shown in Figure 4-A, B. For the mathematical model, we denote the average concentration of the toxin protein (ccdB) in the cell population with N_i, $i = 1, 2$. Similarly, the anti-toxin (ccdA) is denoted by A_i. Finally, the concentration of AHL signals in the consortia is given by R_i for each signal. The description of the model parameters and their values are given in Table II. For more details on the parameter values and system description, the reader is referred to [4].

In this mathematical model, an intrinsic assumption is that the signal transport is instantaneous as the dynamics of transport of signals across cell membranes is not explicitly modeled. The inducer signals, denoted by L and I, are the two inputs to the system. The mathematical model is given by

the following ODEs:

$$\begin{aligned}
\frac{dN_1}{dt} &= \beta_{R_1}\left(l_{R_1} + \frac{R_1^2}{K_{R_1} + R_1^2}\right) - k_b A_1 N_1 - d_T N_1 \\
\frac{dA_1}{dt} &= K_r \beta_{R_2}\left(l_{R_2} + \frac{R_2^2}{K_{R_2} + R_2^2}\right) - k_b A_1 N_1 - d_T A_1 \\
\frac{dR_1}{dt} &= \beta_{tac}\left(l_{tac} + \frac{I^2}{K_{tac} + I^2}\right) L_1 - d_S R_1 \\
\frac{dR_2}{dt} &= \beta_{sal}\left(l_{sal} + \frac{L^2}{K_{sal} + L^2}\right) L_2 - d_S R_2 \\
\frac{dN_2}{dt} &= \beta_{R_2}\left(l_{R_2} + \frac{R_2^2}{K_{R_2} + R_2^2}\right) - k_b A_2 N_2 - d_T N_2 \\
\frac{dA_2}{dt} &= K_r \beta_{R_1}\left(l_{R_1} + \frac{R_1^2}{K_{R_1} + R_1^2}\right) - k_b A_2 N_2 - d_T A_2 \\
\frac{dL_1}{dt} &= k_C\left(1 - \frac{L_1 + L_2}{C_{max}}\right) L_1 - d_c L_1 \frac{N_1}{K_{tox} + N_1} - dL_1 \\
\frac{dL_2}{dt} &= k_C\left(1 - \frac{L_1 + L_2}{C_{max}}\right) L_2 - d_c L_2 \frac{N_2}{K_{tox} + N_2} - dL_2.
\end{aligned} \tag{31}$$

We cannot use singular perturbation theory to reduce this system model since it is not clear how this model can be expressed in the standard form for singular perturbation framework. Using our automated model reduction method, we obtain various possible reduced models for this system and choose the "best" reduced model based on the performance metrics we discussed. Since the full model has two output variables (L_1 and L_2), any reduced model for this system must have atleast these two output variables so that the reduced model may be used effectively for parameter inference.

With our approach, we obtain four different reduced models each with four states. Using the performance metrics discussed earlier, we choose the reduced-order model in equation (32) since it is the most robust reduced-model (with the least robustness metric for the given parameters). The advantages of using this reduced model for parameter identification are clear since the reduced models only have 13 parameters (compared to 24 in the full model). Further, since an analytical mapping between the full and the reduced models is available, we can determine the non-identifiable manifold that could assist in the identifiability analysis for this model [62].

The error and robustness performance for all the reduced models is given in Figure 4-D–F. The decision metric r with weights $w_1 = 1, w_2 = 0.5$, is lowest for the reduced model with both of the toxin-states. Hence, we choose this reduced model. The dynamics for this model are given by:

$$\begin{aligned}
\hat{f}_1 &= \left(l_{R_1} + \frac{\beta_{R_1} x_7^2}{x_7^2 + K_{I0}}\right) - d_T x_1 - \frac{\beta_{R_2} k_b x_1 \left(K_{a0} l_{R_2} + x_8^2\right)}{k_b x_1 x_8^2 + K_{a0} k_b x_1 + d_T x_8^2 + K_{a0} d_T}, \\
\hat{f}_2 &= \left(l_{R_2} + \frac{\beta_{R_2} x_8^2}{x_7^2 + K_{a0}}\right) - d_T x_5 - \frac{\beta_{R_1} k_b x_5 \left(K_{I0} l_{R_1} + x_7^2\right)}{k_b x_5 x_7^2 + K_{I0} k_b x_5 + d_T x_7^2 + K_{I0} d_T}, \\
\hat{f}_3 &= k_c\left(1 - \frac{x_7 + x_8}{C_{max}}\right) x_7 - \frac{d_c x_1 x_7}{x_1 + K_{tox}} - d x_7, \\
\hat{f}_4 &= k_c\left(1 - \frac{x_7 + x_8}{C_{max}}\right) x_8 - \frac{d_c x_5 x_8}{x_5 + K_{tox}} - d x_8.
\end{aligned} \tag{32}$$

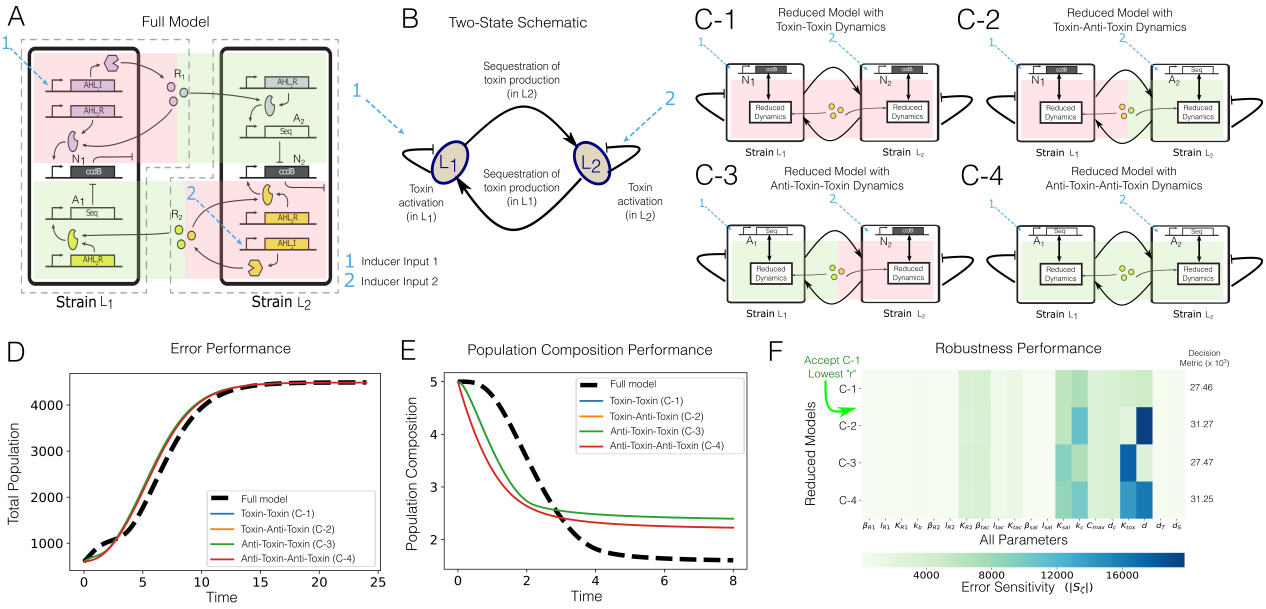

Fig. 4: **(A)**: Two-member population and composition control circuit schematic. The two inducer inputs (labeled 1 and 2 in cyan) activate the ccdB expression in L_1 and L_2 cell strains respectively. AHL signals diffuse out of each cell type to signal the expression of anti-toxin (ccdA) in the other cell type. The anti-toxin sequesters away the toxin protein to rescue the cell population. **(B)**: A simple schematic demonstrating the proliferation and death of each cell type under input signals. **(C-1 – C-4)**: Reduced models obtained using our automated model reduction approach. Each reduced model has four states. Two of these states are the output signals which are never reduced in our method. The other two states in each reduced model are labeled in the Figure. **(D)**: Total population (L_1 + L_2) obtained on simulating each reduced model and the full model. As we can see, the error performance for all of these four reduced models is satisfactory. **(E)**: Population composition (L_1/L_2) for the reduced models and the full model. Composition control is a feature of this circuit that the reduced models demonstrate as well, although the dynamics of population composition are not fully captured. **(F)**: Robustness analysis for each reduced model gives us a guide to choose a particular reduced model given the parameters in the full model. The heatmap shows the norm of the sensitivity of error in model reduction $\|S_\zeta\|$ with respect to the model parameters. The decision metric $r = w_1 \|\zeta\| + w_2 d_R$ is then used to choose the reduced model given in (C-1) since it has the least r.

Here $\hat{x} = \begin{bmatrix} x_1 & x_5 & x_7 & x_8 \end{bmatrix}^T$ and the exact expressions for the lumped parameters are available part of our software package, AutoReduce, on GitHub [46].

V. Discussion

Our main result in this paper gives a closed-form expression for the robustness guarantee of structured model reduction of linear dynamical systems. We show two different methods to derive this result — a direct linear analysis approach for the linear systems and a sensitivity analysis based approach that also works for nonlinear dynamics. The advantage of our method is that the system does not need to be in the standard form as in singular perturbation theory and that we can compute the robustness of the model reduction error with respect to each model parameter individually for a holistic analysis of different possible model reductions.

We demonstrated the applications of our model reduction approach to biological system examples. The exploration of design space by changing experimentally tunable parameters in the models as shown in Figure 2 is an important step towards using mathematical models for biological circuit design. There are two main advantages to this approach:

1) For any complex biological circuit, a single gene expression would usually form a small part of the design. Hence, interpreting the key states and parameters involved in tuning its dynamics is important for the modular design of the complete circuit.

2) For parameter identification, reduced models are commonly used but it is important to verify the correctness of these reduced models. The automated model reduction method discussed in this paper is a step towards that goal as it provides a mapping of the reduced models with the full model alongside the performance guarantees. Further, due to parameter lumping, the reduced models have fewer number of parameters, improving the parameter identifiability of the system given measurement data [62]. The non-identifiable manifolds are analytically known as well.

An interesting future direction would be to extend the results in this paper for the expression of multiple genes together to explore retroactivity [63] and its effects on various phenomenological models used for the design of such systems. This study would be a step in building towards a modular design framework [64] that considers the design of multiple modules together with their context-dependence. Representing large system models in phenomenological terms is a challenging task which would require going beyond QSSA based model reduction presented in this paper. Hence, derivation of reduced models by introducing defined coordinate transformations for states and parameters might be worth investigating as well. Heuristic guidance is another approach that may be used for model reduction of large system models. The greedy algorithm in [21] is a possible direction to apply our results to large biological network models. More research on similar lines could improve the scalability of the tools discussed in this paper. Finally, we would also like to note that to compute the robustness metric, new theoretical results on sensitivity of the model reduction error with respect to simultaneous multiple parameter co-variations could give new insights. Using sensitivities for robustness analysis is not as widely explored for nonlinear system analysis and hence building on the results in this paper could lead to alternative robustness analysis methods.

ACKNOWLEDGMENT

We would like to thank Chelsea Hu for her help in reviewing the examples presented in this paper and Ankita Roychoudhury for her contributions to the AutoReduce software. This research is sponsored in part by the National Science Foundation under grant number: CBET-1903477 and the Defense Advanced Research Projects Agency (Agreement HR0011-17-2-0008). The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred.

REFERENCES

[1] M. B. Elowitz and S. Leibler, "A synthetic oscillatory network of transcriptional regulators," *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[2] T. S. Gardner, C. R. Cantor, and J. J. Collins, "Construction of a genetic toggle switch in escherichia coli," *Nature*, vol. 403, no. 6767, pp. 339–342, 2000.

[3] C. Y. Hu, J. D. Varner, and J. B. Lucks, "Generating effective models and parameters for rna genetic circuits," *ACS synthetic biology*, vol. 4, no. 8, pp. 914–926, 2015.

[4] R. D. McCardell, A. Pandey, and R. M. Murray, "Control of density and composition in an engineered two-member bacterial community," *bioRxiv*, 2019.

[5] A. Swaminathan, V. Hsiao, and R. M. Murray, "Quantitative modeling of integrase dynamics using a novel python toolbox for parameter inference in synthetic biology," *bioRxiv*, p. 121152, 2017.

[6] A. Gyorgy and D. Del Vecchio, "Limitations and trade-offs in gene expression due to competition for shared cellular resources," in *53rd IEEE Conference on Decision and Control*. IEEE, 2014, pp. 5431–5436.

[7] P. Kokotović, H. K. Khalil, and J. O'reilly, *Singular perturbation methods in control: analysis and design*. SIAM, 1999.

[8] S. Z. Rizvi, J. Mohammadpour, R. Tóth, and N. Meskin, "A kernel-based pca approach to model reduction of linear parameter-varying systems," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 5, pp. 1883–1891, 2016.

[9] B. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, no. 1, pp. 17–32, 1981.

[10] M. K. Transtrum and P. Qiu, "Bridging mechanistic and phenomenological models of complex biological systems," *PLoS computational biology*, vol. 12, no. 5, p. e1004915, 2016.

[11] S. Rao, A. Van der Schaft, K. Van Eunen, B. M. Bakker, and B. Jayawardhana, "A model reduction method for biochemical reaction networks," *BMC systems biology*, vol. 8, no. 1, pp. 1–17, 2014.

[12] A. Gorban, "Model reduction in chemical dynamics: slow invariant manifolds, singular perturbations, thermodynamic estimates, and analysis of reaction graph," *Current Opinion in Chemical Engineering*, vol. 21, pp. 48–59, 2018.

[13] P. Kokotovic and P. Sannuti, "Singular perturbation method for reducing the model order in optimal control design," *IEEE Transactions on Automatic Control*, vol. 13, no. 4, pp. 377–384, 1968.

[14] R. Marino and S. Nicosia, "Singular perturbation techniques in the adaptive control of elastic robots," *IFAC Proceedings Volumes*, vol. 18, no. 16, pp. 95–100, 1985.

[15] K. R. Schneider and T. Wilhelm, "Model reduction by extended quasi-steady-state approximation," *Journal of Mathematical Biology*, vol. 40, no. 5, pp. 443–450, 2000.

[16] N. Vora and P. Daoutidis, "Nonlinear model reduction of chemical reaction systems," *AIChE Journal*, vol. 47, no. 10, pp. 2320–2332, 2001.

[17] F. G. Helfferich, "Systematic approach to elucidation of multistep reaction networks," *The Journal of Physical Chemistry*, vol. 93, no. 18, pp. 6676–6681, 1989.

[18] J. A. Christiansen, "The elucidation of reaction mechanisms by the method of intermediates in quasi-stationary concentrations," in *Advances in Catalysis*. Elsevier, 1953, vol. 5, pp. 311–353.

[19] T. Turanyi, A. Tomlin, and M. Pilling, "On the error of the quasi-steady-state approximation," *The Journal of Physical Chemistry*, vol. 97, no. 1, pp. 163–172, 1993.

[20] A. Yannacopoulos, A. Tomlin, J. Brindley, J. Merkin, and M. Pilling, "The error of the quasi steady-state approximation in spatially distributed systems," *Chemical Physics Letters*, vol. 248, no. 1-2, pp. 63–70, 1996.

[21] A. Papachristodoulou, Y.-C. Chang, E. August, and J. Anderson, "Structured model reduction for dynamical networked systems," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 2670–2675.

[22] R. Babaghasabha, M. A. Khosravi, and H. D. Taghirad, "Adaptive robust control of fully constrained cable robots: singular perturbation approach," *Nonlinear Dynamics*, vol. 85, no. 1, pp. 607–620, 2016.

[23] L. A. Belhaj, M. Ait-Ahmed, and M. F. Benkhoris, "Embarked electrical network robust control based on singular perturbation model," *ISA transactions*, vol. 53, no. 4, pp. 1143–1151, 2014.

[24] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*. Dover, 2009.

[25] M. K. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of joint parametric uncertainty and unmodeled dynamics," in *1988 American control conference*. IEEE, 1988, pp. 1195–1200.

[26] J. Freudenberg, D. Looze, and J. Cruz, "Robustness analysis using singular value sensitivities," *International Journal of Control*, vol. 35, no. 1, pp. 95–116, 1982.

[27] A. Pandey and R. M. Murray, "An automated model reduction tool to guide the design and analysis of synthetic biological circuits," *bioRxiv*, p. 640276, 2019.

[28] ——, "Model reduction tools for phenomenological modeling of input-controlled biological circuits," *bioRxiv*, 2020. [Online]. Available: https://www.biorxiv.org/content/early/2020/02/20/2020.02.15.950840

[29] ——, "A two-state ribosome and protein model can robustly capture the chemical reaction dynamics of gene expression," *bioRxiv*, 2020.

[30] X. Yang, J. J. Zhu, and A. S. Hodel, "Singular perturbation margin and generalised gain margin for linear time-invariant systems," *International Journal of Control*, vol. 88, no. 1, pp. 11–29, 2015.

[31] X. Yang and J. J. Zhu, "Singular perturbation margin and generalised gain margin for nonlinear time-invariant systems," *International Journal of Control*, vol. 89, no. 3, pp. 451–468, 2016.

[32] H. K. Khalil and J. W. Grizzle, *Nonlinear Systems*. Prentice Hall, Upper Saddle River, NJ, 2002, vol. 3.

[33] J. Schmidt, "G. dahlquist, stability and error bounds in the numerical integration of ordinary differential equations. 85 s. stockholm 1959. k. tekniska högskolans handlingar," 1961.

[34] T. Ström, "On logarithmic norms," *SIAM Journal on Numerical Analysis*, vol. 12, no. 5, pp. 741–753, 1975.

[35] R. F. Snider, "Perturbation variation methods for a quantum boltzmann equation," *Journal of Mathematical Physics*, vol. 5, no. 11, pp. 1580–1587, 1964.

[36] R. M. Wilcox, "Exponential operators and parameter differentiation in quantum physics," *Journal of Mathematical Physics*, vol. 8, no. 4, pp. 962–982, 1967.

[37] H. Tsai, K. Chan *et al.*, "A note on parameter differentiation of matrix exponentials, with applications to continuous-time modelling," *Bernoulli*, vol. 9, no. 5, pp. 895–919, 2003.

[38] R. P. Dickinson and R. J. Gelinas, "Sensitivity analysis of ordinary differential equation systems - a direct method," *Journal of Computational Physics*, vol. 21, no. 2, pp. 123–143, 1976.

[39] P. J. Antsaklis and A. N. Michel, *Linear Systems*. Springer Science & Business Media, 2006.

[40] W. M. Haddad and V. Chellaboina, *Nonlinear dynamical systems and control: a Lyapunov-based approach*. Princeton university press, 2011.

[41] V. Chellaboina, W. M. Haddad, D. S. Bernstein, and D. A. Wilson, "Induced convolution operator norms of linear dynamical systems," *Mathematics of Control, Signals and Systems*, vol. 13, no. 3, pp. 216–239, 2000.

[42] V. Zahedzadeh, H. J. Marquez, and T. Chen, "Upper bounds for induced operator norms of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 54, no. 5, pp. 1159–1165, 2009.

[43] T. T. Georgiou, "On the computation of the gap metric," *Systems & Control Letters*, vol. 11, no. 4, pp. 253–257, 1988.

[44] B. D. Anderson, T. S. Brinsmead, and F. De Bruyne, "The vinnicombe metric for nonlinear operators," *IEEE Transactions on Automatic Control*, vol. 47, no. 9, pp. 1450–1465, 2002.

[45] V. Zahedzadeh, H. J. Marquez, and T. Chen, "On the computation of an upper bound on the gap metric for a class of nonlinear systems," in *2008 American Control Conference*. IEEE, 2008, pp. 1917–1922.

[46] A. Pandey, "Auto-reduce : Python based automated model reduction package," [Online].
[47] A. Meurer, C. P. Smith, M. Paprocki, O. Čertík, S. B. Kirpichev, M. Rocklin, A. Kumar, S. Ivanov, J. K. Moore, S. Singh, T. Rathnayake, S. Vig, B. E. Granger, R. P. Muller, F. Bonazzi, H. Gupta, S. Vats, F. Johansson, F. Pedregosa, M. J. Curry, A. R. Terrel, v. Roučka, A. Saboo, I. Fernando, S. Kulal, R. Cimrman, and A. Scopatz, "Sympy: symbolic computing in python," *PeerJ Computer Science*, vol. 3, p. e103, Jan. 2017. [Online]. Available: https://doi.org/10.7717/peerj-cs.103
[48] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden *et al.*, "The systems biology markup language (sbml): a medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, no. 4, pp. 524–531, 2003.
[49] T. J. Snowden, P. H. Van Der Graaf, and M. J. Tindall, "A combined model reduction algorithm for controlled biochemical systems," *BMC Systems Biology*, vol. 11, no. 1, p. 17, 2017.
[50] D. Del Vecchio and R. M. Murray, *Biomolecular Feedback Systems*. Princeton University Press Princeton, NJ, 2015.
[51] X. Yang and J. J. Zhu, "Singular perturbation margin assessment of linear time-invariant systems via the bauer-fike theorems," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012, pp. 6521–6528.
[52] K. J. Åström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2008.
[53] A. N. Tikhonov, "Systems of differential equations containing small parameters in the derivatives," *Matematicheskii Sbornik*, vol. 73, no. 3, pp. 575–586, 1952.
[54] A. Goeke, S. Walcher, and E. Zerz, "Determining "small parameters" for quasi-steady state," *Journal of Differential Equations*, vol. 259, no. 3, pp. 1149–1180, 2015.
[55] T. J. Snowden, P. H. van der Graaf, and M. J. Tindall, "Methods of model reduction for large-scale biological systems: a survey of current methods and trends," *Bulletin of mathematical biology*, vol. 79, no. 7, pp. 1449–1486, 2017.
[56] M. Gasparyan, A. Van Messem, and S. Rao, "An automated model reduction method for biochemical reaction networks," *Symmetry*, vol. 12, no. 8, p. 1321, 2020.
[57] O. Buse, R. Pérez, and A. Kuznetsov, "Dynamical properties of the repressilator model," *Physical Review E*, vol. 81, no. 6, p. 066206, 2010.
[58] C. Y. Hu and R. M. Murray, "Design of a genetic layered feedback controller in synthetic biological circuitry," *bioRxiv*, p. 647057, 2019.
[59] A. J. Meyer, T. H. Segall-Shapiro, E. Glassey, J. Zhang, and C. A. Voigt, "Escherichia coli "marionette" strains with 12 highly optimized small-molecule sensors," *Nature chemical biology*, vol. 15, no. 2, pp. 196–204, 2019.
[60] L. Pasotti, M. Bellato, D. De Marchi, and P. Magni, "Mechanistic models of inducible synthetic circuits for joint description of dna copy number, regulatory protein level, and cell load," *Processes*, vol. 7, no. 3, p. 119, 2019.
[61] X. Ren, A.-A. Baetica, A. Swaminathan, and R. M. Murray, "Population regulation in microbial consortia using dual feedback control," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 5341–5347.
[62] F. P. Davidescu and S. B. Jørgensen, "Structural parameter identifiability analysis for dynamic reaction networks," *Chemical Engineering Science*, vol. 63, no. 19, pp. 4754–4762, 2008.
[63] S. Jayanthi, K. S. Nilgiriwala, and D. Del Vecchio, "Retroactivity controls the temporal dynamics of gene transcription," *ACS synthetic biology*, vol. 2, no. 8, pp. 431–441, 2013.
[64] A. Gyorgy and D. Del Vecchio, "Modular composition of gene transcription networks," *PLoS Comput Biol*, vol. 10, no. 3, p. e1003486, 2014.

APPENDIX 1 — PROOFS

A. *Linear Dynamics — Derivative of Matrix Exponential*

Here we prove equation (8) as given in Lemma 2:

Proof. For a linear system, $\dot{x} = Ax$, we can write the solution $x(t) = e^{At}x(0)$, where e^{At} is the matrix exponential. Now, for a parameter θ_i, we can write

$$\frac{\partial x(t)}{\partial \theta_i} = e^{At}\frac{\partial x(0)}{\partial \theta_i} + \frac{\partial e^{At}}{\partial \theta_i}x(0),$$

using the product rule of differentiation. Define

$$S(t) := \frac{\partial x(t)}{\partial \theta_i}, \tag{33}$$

so we have,

$$S(t) = e^{At}S(0) + \frac{\partial e^{At}}{\partial \theta_i}x(0), \tag{34}$$

and write $\dot{S}(t)$ using equation (33) and $\dot{x} = Ax$ as

$$\frac{dS}{dt} = A\frac{\partial x}{\partial \theta_i} + \frac{\partial A}{\partial \theta_i}x = AS + \frac{\partial A}{\partial \theta_i}x.$$

Solving for $S(t)$, we get,

$$S(t) = e^{At}S(0) + \int_0^t e^{A(t-\tau)}\frac{\partial A}{\partial \theta_i}x(\tau)d\tau.$$

Since $x(\tau) = e^{A\tau}x(0)$, we can simplify the above equation and write

$$S(t) = e^{At}S(0) + \left[\int_o^t e^{A(t-\tau)}\frac{\partial A}{\partial \theta_i}e^{A\tau}d\tau\right]x(0).$$

Comparing this with equation (34), we get the desired result for the derivative of the matrix exponential

$$\frac{\partial e^{At}}{\partial \theta_i} = \int_o^t e^{A(t-\tau)}\frac{\partial A}{\partial \theta_i}e^{A\tau}d\tau. \qquad \square$$

B. Robustness Metric — Nonlinear Controlled Dynamics

Proof of Theorem 4:

Proof. For the augmented system we can write the sensitivity coefficients as

$$\bar{S} = \frac{\partial \bar{x}}{\partial \theta_i},$$

where $\theta_i \in \theta$. Using chain rule at point $\bar{x}(t) = \bar{x}$ in the system trajectory, we can derive the sensitivity system equation given by

$$\dot{\bar{S}} = \left(\bar{J}_f(\bar{x})\bar{S} + \bar{Z}_f\right) + \left(\bar{J}_g(\bar{x})\bar{S} + \bar{Z}_g\right)u,$$

for a scalar u and Jacobian matrices J_f and J_g are state-dependent. Consider a function $V(\bar{S}) = \bar{S}^T P(\bar{x})\bar{S}$ for $P(\bar{x}) = P(\bar{x})^T \succ 0$ that satisfies the conditions given in the theorem statement. Note that we drop the $\bar{x}$ notation from $\bar{J}_f, \bar{J}_g, P$ for simplicity. Taking the derivative of V with respect to time, we can write,

$$\frac{dV}{dt} = \dot{\bar{S}}^T P\bar{S} + \bar{S}^T P\dot{\bar{S}} + \bar{S}^T \dot{P}\bar{S}.$$

Substituting for $\dot{\bar{S}}$ we get

$$\frac{dV}{dt} = \bar{S}^T\left(P\bar{J}_f + \bar{J}_f^T P\right)\bar{S} + u^T\bar{S}^T\bar{J}_g^T P\bar{S} + \bar{S}^T P\bar{J}_g\bar{S}u + \left(\bar{S}^T P\bar{Z}_f + \bar{Z}_f^T P\bar{S}\right) + \left(\bar{S}^T P\bar{Z}_g u + u^T\bar{Z}_g^T P\bar{S}\right) + \bar{S}^T\dot{P}\bar{S}$$

Now if there exists a matrix $P = P^T$ such that

$$P\bar{J}_f + \bar{J}_f^T P = -\bar{C}^T\bar{C}$$

then we get the following bound by manipulating the 2-norm of S_ζ and denoting $\bar{S}(0) := \bar{S}_0$,

$$\|S_\zeta\|_2^2 \leq \bar{S}_0^T P\bar{S}_0 + 2\int_0^\infty \left(\left\|u^T\bar{S}^T\bar{J}_g^T P\bar{S}\right\| + \left\|\bar{Z}_f^T P\bar{S}\right\| + \left\|u^T\bar{Z}_g^T P\bar{S}\right\|\right)dt + \int_0^\infty \bar{S}^T\dot{P}\bar{S}dt$$

We can simplify the above to get the desired result,

$$\|S_\zeta\|_2^2 \leq \lambda_{\max_{\bar{x}}}(P)\left\|\bar{S}_0\right\|^2 + 2\int_0^\infty \left(\left\|u^T\bar{S}^T\bar{J}_g^T P\bar{S}\right\| + \left\|\bar{Z}_f^T P\bar{S}\right\| + \left\|u^T\bar{Z}_g^T P\bar{S}\right\|\right)dt + \lambda_{\max_{\bar{x}}}(\dot{P})\int_0^\infty \left\|\bar{S}\right\|^2 dt. \qquad \square$$

C. Equivalence of the two results for linear dynamics

A direct comparison of the results in Corollary 2.1 (robustness estimate for linear dynamics) and Theorem 3 (robustness estimate for nonlinear dynamics) is not evident. But for the special case of linear dynamics, we have the closed-form solutions for $\bar{S}(t)$ and $\bar{Z}(t)$. Using these we can evaluate the bound given in Theorem 3 further.

Claim. The bound on the sensitivity of the error in model reduction when obtained using sensitivity analysis approach (as in equation (18)) is same as the bound obtained using direct linear analysis approach (as given in equation (15)). In particular, we have that,

$$2\int_0^\infty \left\| \left(\frac{\partial \bar{A}}{\partial \theta_i}\bar{x}\right)^T P\bar{S} \right\| dt = \frac{1}{4\left|\mu\right|^3}\left\|\frac{\partial \bar{A}}{\partial \theta_i}\right\|^2 \left\|\bar{C}^T\bar{C}\right\| \left\|\bar{x}(0)\right\|^2 + \frac{1}{2\left|\mu\right|^2}\left\|\frac{\partial \bar{A}}{\partial \theta_i}\right\| \left\|\bar{C}^T\bar{C}\right\| \left\|\bar{x}(0)\right\| \left\|\frac{\partial \bar{x}(0)}{\partial \theta_i}\right\| \tag{35}$$

Note that the first term in equations (15) and (18) is the same and hence we have removed that term in the equation above. This term corresponds to the parametric uncertainty in the initial conditions.

Proof. To prove the above claim, we start by evaluating the different parts of the left hand side expression using the closed-form solutions for linear dynamics. First, note that from the sensitivity equation [38] for a linear system we have that

$$\dot{\bar{S}} = \bar{A}\bar{S} + \frac{\partial \bar{A}}{\partial \theta_i}\bar{x}(t).$$

Solving the equation above for $\bar{S}(t)$ and taking the norm we get

$$\begin{aligned} \left\|\bar{S}(t)\right\| &\leq \left\|e^{\bar{A}t}\bar{S}(0)\right\| + \left\|\int_0^t e^{\bar{A}(t-\tau)}\frac{\partial \bar{A}}{\partial \theta_i}\bar{x}(\tau)d\tau\right\| \\ &\leq e^{-|\mu|t}\left\|\bar{S}(0)\right\| + \left\|\frac{\partial \bar{A}}{\partial \theta_i}\right\| \left\|\bar{x}(0)\right\| te^{-|\mu|t} \end{aligned} \tag{36}$$

using Lemma 1 and 2. Similarly, for $\left\|\bar{Z}\right\|$, we have

$$\left\|\bar{Z}\right\| = \left\|\frac{\partial \bar{A}}{\partial \theta_i}\bar{x}\right\| \leq \left\|\frac{\partial \bar{A}}{\partial \theta_i}\right\| \left\|\bar{x}(0)\right\| e^{-|\mu|t}. \tag{37}$$

Finally, for the Lyapunov matrix, we know that

$$\left\|P\right\| = \left\|\int_0^\infty e^{\bar{A}^T t}\bar{C}^T\bar{C}e^{\bar{A}t}dt\right\|,$$

using the observability Gramian. Using Lemma 1, we have

$$\left\|P\right\| \leq \int_0^\infty e^{-|\mu|t}\left\|\bar{C}^T\bar{C}\right\| e^{-|\mu|t}dt.$$

So,

$$\left\|P\right\| \leq \left\|\bar{C}^T\bar{C}\right\| \int_0^\infty e^{-2|\mu|t}dt,$$

which gives us that

$$\left\|P\right\| \leq \frac{1}{2\left|\mu\right|}\left\|\bar{C}^T\bar{C}\right\|. \tag{38}$$

Substituting the equations (36), (37), and (38) into the left hand size of equation (35), we get the desired result that proves our claim. □

Although the two approaches give equivalent results for linear dynamics, the advantage with the sensitivity analysis approach is that it is a general method that can be used for nonlinear dynamical systems as well.

APPENDIX 2 — EXAMPLES

A. Gene Expression — Robustness Statement 1

Justification. To demonstrate robustness of the $\hat{R}_\Delta$ model, we can look at the sensitivity of the error between the $\hat{R}_\Delta$ model and the full CRN model to various parameter perturbations. Define the error as $\zeta := X - \hat{X}$, where X and $\hat{X}$ represent the protein concentration in the full CRN model and the $\hat{R}_\Delta$ model respectively. In the case where the error ζ between the two models is within acceptable bounds, we additionally desire that this error is not sensitive to any of the parameters. The sensitivity of the error to perturbation in a parameter θ_i is given by $S_\zeta = \partial\zeta/\partial\theta_i$. Hence, this "fragility metric" S_ζ must be minimized to achieve higher robustness. Analyzing the rate of change of S_ζ with time for different parameters can justify the statement.

To analytically derive $\dot{S}_\zeta$, we use the equation $S_\zeta = \bar{C}\bar{S}$ to write

$$\dot{S}_\zeta = \left(\begin{bmatrix} C & -\hat{C} \end{bmatrix}\right) \left(\begin{bmatrix} J & 0 \\ 0 & \hat{J} \end{bmatrix} \begin{bmatrix} S \\ \hat{S} \end{bmatrix} + \begin{bmatrix} Z \\ \hat{Z} \end{bmatrix} \right).$$

For the model parameters $\theta_i \in \{k_{tx}, k_{bp}, k_{up}, k_{br}, k_{ur}, P_{\text{tot}}, R_{\text{tot}}\}$, we have the dynamics of S_ζ given by

$$\dot{S}_\zeta = k_{tl} \frac{\partial \left(\hat{R}_\Delta - \hat{R}_\Delta \right)}{\partial \theta_i} - d_X S_\zeta. \tag{39}$$

For the protein degradation rate $\theta_i = d_X$, we have

$$\dot{S}_\zeta = k_{tl} \frac{\partial \left(\hat{R}_\Delta - \hat{R}_\Delta \right)}{\partial d_X} - \zeta - d_X S_\zeta. \tag{40}$$

For the translation rate $\theta_i = k_{tl}$, we have

$$\dot{S}_\zeta = k_{tl} \frac{\partial \left(\hat{R}_\Delta - \hat{R}_\Delta \right)}{\partial k_{tl}} + \left(\hat{R}_\Delta - \hat{R}_\Delta \right) - d_X S_\zeta. \tag{41}$$

From equations (39) – (41), we have that for all positive parameter values and stable models, the dynamics of S_ζ are convergent to a fixed point. More importantly, the fragility metric, S_ζ directly depends on the translation rate k_{tl} and the protein degradation rate d_X, proving the assertion of the statement above. Also, for k_{tl} and d_X, we see that an extra term appears in the error sensitivity dynamics implying higher fragility of the $\hat{R}_\Delta$ model under perturbations to the translation rate and the protein degradation rate. □

B. Gene Expression — Robustness Statement 2

Justification. Similar to the justification of Statement 1, we can analyze the dynamics of the sensitivity of the error to parameter perturbations for the mRNA transcript and protein dynamical model. For the translation rate, that is $\theta_i = k_{tl}$,

$$\dot{S}_\zeta = k_{tl} \frac{\partial \hat{R}_\Delta}{\partial k_{tl}} - \frac{k_{tl} R_{\text{tot}} K_0}{\left(K_0 + \hat{T} \right)^2} \frac{\partial \hat{T}}{\partial k_{tl}} + \hat{R}_\Delta - \frac{R_{\text{tot}} \hat{T}}{K_0 + \hat{T}} + \frac{k_{tl} R_{\text{tot}} \hat{T}}{k_{br} \left(K_0 + \hat{T} \right)^2} - d_X S_\zeta. \tag{42}$$

For the protein degradation parameter $\theta_i = d_X$,

$$\dot{S}_e = k_{tl}\frac{\partial \hat{R}_\Delta}{\partial d_X} - \frac{k_{tl}R_{\text{tot}}K_0}{\left(K_0+\hat{T}\right)^2}\frac{\partial \hat{T}}{\partial d_X} - \zeta - d_X S_\zeta. \tag{43}$$

For the total ribosome count, $\theta_i = R_{\text{tot}}$, we have,

$$\dot{S}_e = k_{tl}\frac{\partial \hat{R}_\Delta}{\partial R_{\text{tot}}} - \frac{k_{tl}R_{\text{tot}}K_0}{\left(K_0+\hat{T}\right)^2}\frac{\partial \hat{T}}{\partial R_{\text{tot}}} - \frac{k_{tl}\hat{T}}{K_0+\hat{T}} - d_X S_\zeta. \tag{44}$$

For the ribsome-transcript binding parameter, $\theta_i = k_{br}$, we have,

$$\dot{S}_\zeta = k_{tl}\frac{\partial \hat{R}_\Delta}{\partial k_{br}} - \frac{k_{tl}R_{\text{tot}}K_0}{\left(K_0+\hat{T}\right)^2}\frac{\partial \hat{T}}{\partial k_{br}} - \frac{k_{tl}K_0^2R_{\text{tot}}\hat{T}}{\left(K_0+\hat{T}\right)^2(k_{tl}+k_{ur})} - d_X S_\zeta. \tag{45}$$

For the ribsome-transcript unbinding parameter, $\theta_i = k_{ur}$, we have,

$$\dot{S}_\zeta = k_{tl}\frac{\partial \hat{R}_\Delta}{\partial k_{ur}} - \frac{k_{tl}R_{\text{tot}}K_0}{\left(K_0+\hat{T}\right)^2}\frac{\partial \hat{T}}{\partial k_{ur}} + \frac{k_{tl}R_{\text{tot}}K_0\hat{T}}{\left(K_0+\hat{T}\right)^2(k_{tl}+k_{ur})} - d_X S_\zeta. \tag{46}$$

For all other parameters, we have,

$$\dot{S}_\zeta = k_{tl}\frac{\partial \hat{R}_\Delta}{\partial k_{ur}} - \frac{k_{tl}R_{\text{tot}}K_0}{\left(K_0+\hat{T}\right)^2}\frac{\partial \hat{T}}{\partial k_{ur}} - d_X S_\zeta. \tag{47}$$

From equations (42) – (47), we can conclude that S_ζ directly depends on the total ribosome count, R_{tot}, the binding/unbinding parameters of the transcript with ribosome, k_{br}, k_{ur}, the translation rate, k_{tl}, and the protein degradation rate, d_X. This is in contrast with the results for the $\hat{R}_\Delta$ model where the dynamics of S_ζ only depend on the translation rate and the protein degradation rate. Hence, the $\hat{R}_\Delta$ model is robust to all other parameters whereas the transcript-protein model is fragile to all of these parameters, proving the statement assertion. □

C. Gene Expression with Endonuclease Mediated mRNA Degradation

To investigate the scalability of this approach, we expand on the gene expression example by including enzymatic degradation of the mRNA transcript mediated by endonucleases. Here we have this enzymatic degradation in addition to the basal degradation rate d_T. The CRN and the corresponding mass-action ODE model is given in Table III.

Observe that in this model we have the following conservation law relationships,

$$\frac{dP}{dt} + \frac{dC_1}{dt} = 0, \quad \frac{dR}{dt} + \frac{dC_2}{dt} = 0, \quad \frac{dE}{dt} + \frac{dC_3}{dt} = 0.$$

Hence, for constants P_{tot}, R_{tot}, and E_{tot}, we can write

$$P + C_1 = P_{\text{tot}}, \;\; R + C_2 = R_{\text{tot}}, \;\; E + C_3 = E_{\text{tot}}.$$

Using these algebraic relationships, we can eliminate the complexes C_1, C_2, and C_3 to obtain a reduced ODE model. Next, we explore various time-scale separation assumptions that may be used to get further reduced models and discuss their performance and robustness with respect to the model parameters.

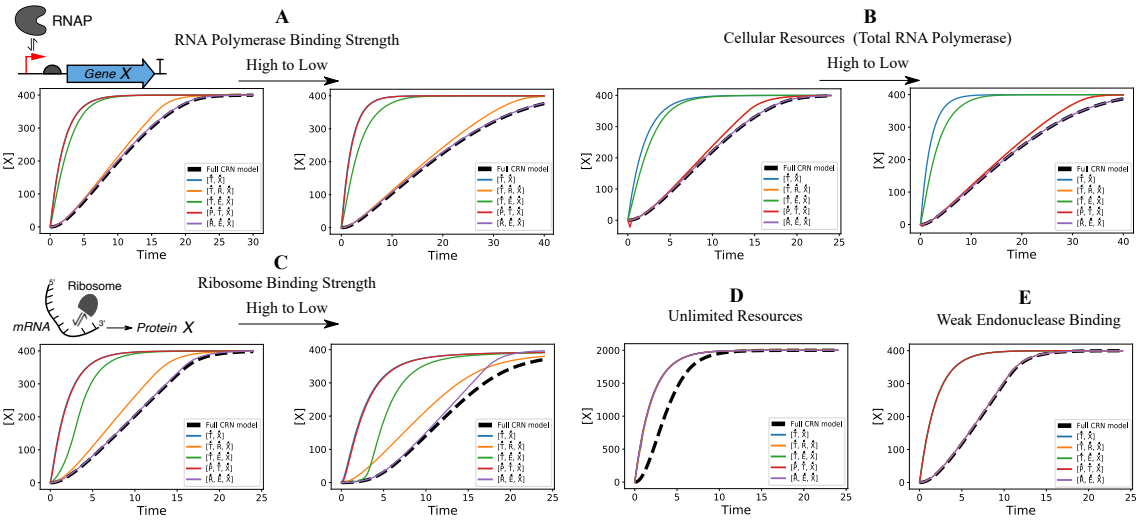

Fig. 5: Performance of the gene expression with endonuclease mediated mRNA degradation models under different biological conditions. **(A)** We observe that for weaker binding of RNA polymerase to the promoter region of the DNA implies a slower time-response which is only captured perfectly by the free ribosome model that also models the endonuclease dynamics. The mathematical model with only the mRNA transcript and protein dynamics is unable to capture this effect since this binding reaction is assumed to be at quasi-steady-state in this model, but it is the only two-state model for this system with a satisfactory steady-state performance. **(B)** We observe a similar effect with decreasing total RNA polymerase in the system. **(C)** With decreasing ribosome binding strength, none of the reduced models perfectly capture the full CRN dynamics, but still, the ribosome-protein model with endonuclease dynamics ($[\hat{R}, \hat{E}, \hat{X}$) is the closest in error performance to the full model. **(D)** With a very high amount of cellular resources, all models reach steady-state faster. **(E)** Under weak endonuclease binding, we observe that only the model that explicitly models the endonuclease dynamics gives good performance. Python code used for this analysis is available publicly on GitHub [46].

1) The mRNA transcript and protein dynamical model: Assuming that the dynamics of all species in the model other than the mRNA transcript T and the protein X are at quasi-steady-state, we get the following model. Observe that a new degradation Hill function term appears in the dynamics of the mRNA transcript that is dependent on the endonuclease binding parameters.

$$\frac{d\hat{T}}{dt} = k_{tx}P_{\text{tot}}\frac{G}{K_1 + K_1 G} - d_E E_{\text{tot}}\frac{\hat{T}}{K_2 + \hat{T}} - d_T\hat{T}$$
$$\frac{d\hat{X}}{dt} = k_{tl}R_{\text{tot}}\frac{\hat{T}}{K_0 + \hat{T}} - d_X\hat{X} \tag{48}$$

where K_2 is a new lumped parameter that is the Hill activation parameter for the endonuclease binding,

$$K_2 = \frac{d_E + k_{ue}}{k_{be}}.$$

2) Modeling the dynamics of available free ribosomes: As before, we define the available free ribosomes in the system as $\hat{R}_\Delta = R_{\text{tot}} - \hat{R}$. Using this definition and assuming that the dynamics of all species in the model other than the free ribosomes, the mRNA transcript, and the protein are at quasi-steady-state we get the following model.

$$\frac{d\hat{T}}{dt} = k_{tx}P_{\text{tot}}\frac{G}{K_1 + G} + (k_{tl} + k_{ur})\hat{R}_\Delta - d_E E_{\text{tot}}\frac{\hat{T}}{K_2 + \hat{T}} - k_{br}(R_{\text{tot}} - \hat{R}_\Delta)\hat{T} - d_T\hat{T}$$
$$\frac{d\hat{R}_\Delta}{dt} = k_{br}(R_{\text{tot}} - \hat{R}_\Delta)\hat{T} - (k_{tl} + k_{ur})\hat{R}_\Delta, \quad \frac{d\hat{X}}{dt} = k_{tl}\hat{R}_\Delta - d_X\hat{X} \tag{49}$$

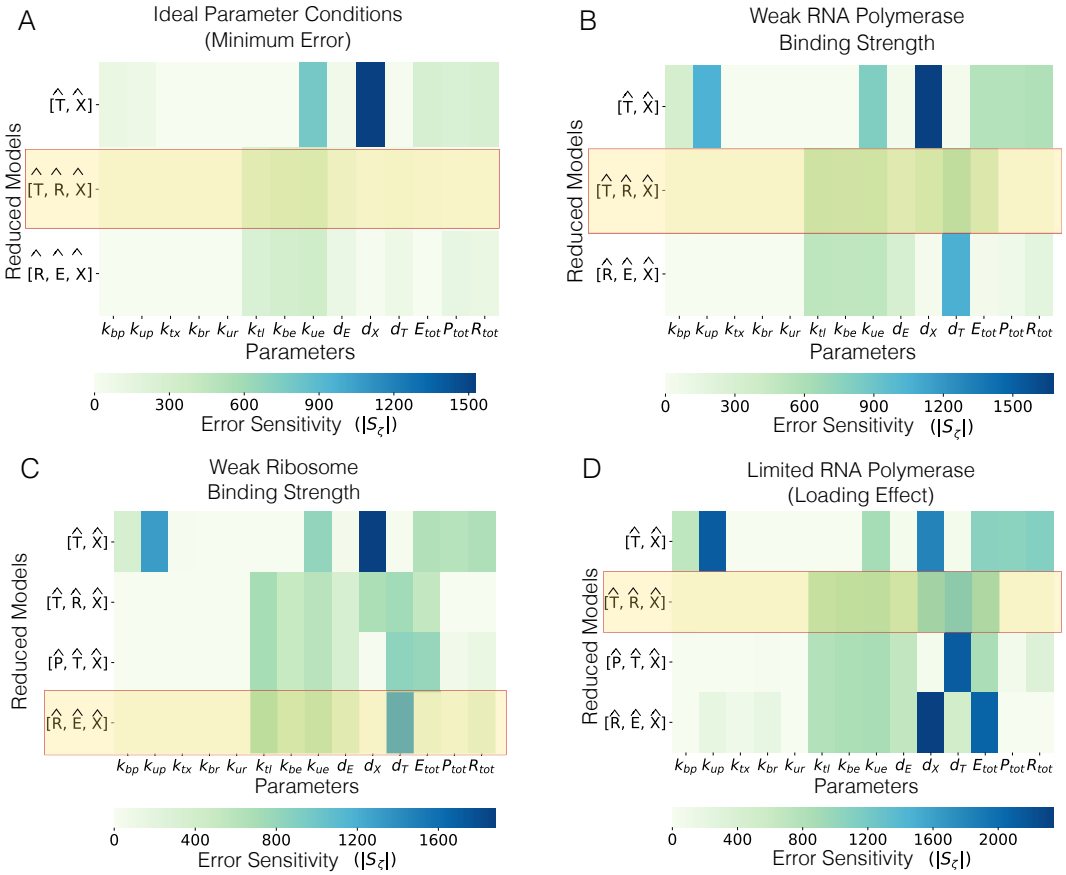

Fig. 6: Robustness analysis of gene expression models with endonuclease mediated degradation of mRNA. The heatmaps show the norm of S_ζ for each model parameter. Lower S_ζ implies better robust performance. **(A)**: Robustness of the reduced models under ideal parameter conditions corresponding to unlimited resources as shown in Figure 5-D where the error for all reduced models is the minimum. Using the combined decision metric $r = w_1 \|\zeta\| + w_2 d_R$, we get that the reduced model with the states $\hat{T}, \hat{R}, \hat{X}$ is the best choice. **(B)**: For weak RNA polymerase binding strength, we obtain that the $\hat{T}, \hat{R}, \hat{X}$ model performs better compared to the other reduced models. **(C)**: Under weak ribosome binding strength condition, we get that the reduced model consisting of $\hat{R}, \hat{E}, \hat{X}$ as its states has the best performance. **(D)**: Under limited resources, the $\hat{T}, \hat{R}, \hat{X}$ model is again able to capture the full CRN dynamics with the least error and most robustness.

Note that the model with only $\hat{R}_\Delta$ and the protein $\hat{X}$ dynamics does not work in this case since the enzymatic degradation reactions for the mRNA transcript are significant for the overall dynamics. So, either T or E is necessary in the ribosome and protein model to get satisfactory performance. As a result, we can also obtain a reduced model with $\begin{bmatrix} \hat{R} & \hat{E} & \hat{X} \end{bmatrix}$ as the states. The performance of all of the possible reduced models is shown in Figure 5.

Although a similar robustness analysis as shown in Statements 1 and 2 can be done for this system dynamics it would be easier to numerically compute a bound on S_ζ for all reduced models. The heatmap of the Euclidean norm of this metric is shown in Figure 6. We use the upper bound for the norm of S_ζ derived in this paper to compute the robustness estimate for each model parameter. The `AutoReduce` software can be used to compute the model equations and these metrics.

TABLE I: Gene expression model.

CRN	ODE model	Nominal Parameter Values
$G + P \underset{k_{up}}{\overset{k_{bp}}{\rightleftharpoons}} C_1$	$\frac{dP}{dt} = (k_{up} + k_{tx})\, C_1 - k_{bp} GP$	$k_{bp} = 80, k_{up} = 2$
$C_1 \xrightarrow{k_{tx}} G + P + T$	$\frac{dC_1}{dt} = k_{bp} GP - (k_{up} + k_{tx})\, C_1$	$k_{tx} = 0.5$
$T + R \underset{k_{ur}}{\overset{k_{br}}{\rightleftharpoons}} C_2$	$\frac{dT}{dt} = k_{tx} C_1 + (k_{ur} + k_{tl})\, C_2 - k_{br} TR - d_T T$	$k_{br} = 80, k_{ur} = 2$
$C_2 \xrightarrow{k_{tl}} T + R + X$	$\frac{dR}{dt} = (k_{ur} + k_{tl})\, C_2 - k_{br} TR$	$k_{tl} = 0.5$
$T \xrightarrow{d_T} \varnothing$	$\frac{dC_2}{dt} = k_{br} TR - (k_{ur} + k_{tl})\, C_2$	$d_T = 0.5$
$X \xrightarrow{d_X} \varnothing$	$\frac{dX}{dt} = k_{tl} C_2 - d_X X$	$d_X = 0.01$

TABLE II: Model parameters

S.no.	Parameters	Description	Unit	Guess
1	β_{R_1}	Max transcription rate of inducible promoter (for N_1 and A_2)	con./hr	6
2	l_{R_1}	Leak constant of inducible promoter (for N_1 and A_2)	N/A	2e-3
3	K_{R_1}	Activation constant of inducible promoter (for N_1 and A_2)	con.	430
4	k_b	Binding rate between toxin and anti-toxin	1/con.hr	30
5	β_{R_2}	Max transcription rate of inducible promoter (for N_2 and A_1)	con./hr	6
6	l_{R_2}	Leak constant of inducible promoter (for N_2 and A_1)	N/A	2e-3
7	K_{R_2}	Activation constant of inducible promoter (for N_2 and A_1)	con.	190
8	β_{tac}	Max transcription rate of inducible promoter (for R_1)	con./hr	19.8e-3
9	l_{tac}	Leak constant of inducible promoter (for R_1)	N/A	1.5e-3
10	K_{tac}	Activation constant of inducible promoter (for R_1)	con.	1.4e5
11	β_{sal}	Max transcription rate of inducible promoter (for R_2)	con./hr	14.4e-3
12	l_{sal}	Leak constant of inducible promoter (for R_2)	N/A	2.1e-4
13	K_{sal}	Activation constant of inducible promoter (for R_2)	con.	13
14	k_C	Cell division rate	1/hr	0.6
15	C_{max}	Population cap	conc.	5500
16	d_c	Cell death rate	1/conc.hr	0.8
18	I	Max induced I (input 1) concentration	con.	1e6
20	L	Max induced L (input 2) concentration	con.	324
21	K_{tox}	Repression coefficient of toxin to proliferation	con.	1
22	d_S	Degradation constant of AHL signals	1/hr	0.5
23	d	Basal degradation of each cell	1/hr	0.1
24	d_T	Basal degradation of toxins and antitoxins	1/hr	1.5
25	K_r	Ribosome scaling factor	N/A	5

TABLE III: Gene expression with endonuclease mediated mRNA degradation

CRN	ODE model	Nominal Parameter Values
$G + P \underset{k_{up}}{\overset{k_{bp}}{\rightleftharpoons}} C_1$	$\frac{dP}{dt} = (k_{up} + k_{tx})\, C_1 - k_{bp} GP$	$k_{bp} = 80, k_{up} = 2$
$C_1 \xrightarrow{k_{tx}} G + P + T$	$\frac{dC_1}{dt} = k_{bp} GP - (k_u p + k_t x)\, C_1$	$k_{tx} = 0.5$
$T + R \underset{k_{ur}}{\overset{k_{br}}{\rightleftharpoons}} C_2$	$\frac{dT}{dt} = k_{tx} C_1 + (k_{ur} + k_{tl})\, C_2$	
	$-k_{br} TR + k_{ue} C_3 - k_{be} TE - d_T T$	$k_{br} = 80, k_{ur} = 2$
$C_2 \xrightarrow{k_{tl}} T + R + X$	$\frac{dR}{dt} = (k_{ur} + k_{tl})\, C_2 - k_{br} TR$	$k_{tl} = 0.5$
$T + E \underset{k_{ue}}{\overset{k_{be}}{\rightleftharpoons}} C_3$	$\frac{dC_2}{dt} = k_{br} TR - (k_{ur} + k_{tl})\, C_2$	$k_{be} = 10, k_{ue} = 2$
$C_3 \xrightarrow{d_E} E$	$\frac{dE}{dt} = (k_{ui} + d_E) C_3 - k_{be} TE$	$d_E = 0.1$
$T \xrightarrow{d_T} \varnothing$	$\frac{dC_3}{dt} = k_{be} TE - (k_{ue} + d_E) C_3$	$d_T = 0.5$
$X \xrightarrow{d} \varnothing$	$\frac{dX}{dt} = k_{tl} C_2 - d_X X$	$d_X = 0.01$