Model reduction using multiple time scales in stochastic gene regulatory networks

Slaven Peleš,* Brian Munsky,† and Mustafa Khammash‡
Deapartment of Mechanical Engineering, University of California
Santa Barbara, CA 93106-5070
(Dated: August 28, 2006)

Gene network dynamics often involves processes that take place on widely differing time scales – from the order of nanoseconds to the order of several days. Multiple time scales in mathematical models often lead to serious computational difficulties, such as numerical stiffness in the case of differential equations or excessively redundant Monte Carlo simulations in the case of stochastic processes. We present a method that takes advantage of multiple time scales and dramatically reduces the computational time for a broad class of problems arising in stochastic gene regulatory networks. We illustrate the efficiency of our method in two gene network examples, which describe two substantially different biological processes – cellular heat shock response and expression of the *pap* gene in *Escherichia coli* bacteria.

I. INTRODUCTION

Despite the great progress in biosciences in recent years, mathematical modeling of intracellular processes remains in its infancy. The level of complexity found inside the cell is such that intuition obtained from other sciences is often insufficient to develop reliable quantitative models. Many cellular processes take place far from equilibrium and on timescales longer than the cell replication cycle, so they never reach the asymptotic state. Hence, knowing the asymptotic solution alone may not be sufficient to describe the system's dynamics. Furthermore, characteristic time-scales in intracellular processes often differ by several orders of magnitude, which creates additional computational difficulties.

At the essence of the study of gene regulatory networks is identifying specific regulatory mechanisms within a complex structure. One way to do so is to carry out a model reduction with respect to the process of interest. In this paper we propose a model reduction method based on singular perturbation theory that takes an advantage of multiple time scales. It can be applied in the study of asymptotic behavior as well as transient processes. The method is based upon different mathematical formalism[1] and in many aspects is superior to Monte Carlo based approaches[2,3,4], which are typically used to treat these problems. The accuracy of the computation is known *a priori* and can be adjusted before the bulk of calculations is carried out.

We illustrate our method using two examples arising from recent experiments with *Escherichia coli* bacteria: we analyze the *pap* gene regulatory network and cellular heat shock response. The fact that the two processes occur in completely different biological contexts suggests that our method is suitable to study a broad range of problems.

The method can be used as a standalone approach, but its applicability is significantly extended if used in conjunction with the recently proposed Finite State Projection[1]. Model reduction approaches based on singular perturbation theory have been used in various areas of engineering and science[5,6,7,8], but the overwhelming complexity of some biological models has severely limited applicability of these approaches in some areas in biosciences. The Finite State Projection retains an important subset of the state space and projects the remaining part (which can be infinite) onto a single state, while keeping the approximation error strictly within prespecified accuracy. The resulting finite model is given in an analytical form and allows us to implement further reduction techniques, such as singular perturbation, to a broad range of complex systems. As it reduces systems to manageable sizes at little or no cost, the Finite State Projection method, combined with other reduction approaches, has the potential to fundamentally change the computational approach to stochastic biological problems.

This paper is organized as follows: in Section II we give a brief overview of some computational methods that have been used to study stochastic gene network models. In Section III we describe the mathematical details of our method. In Section IV we provide an example of how our method can be applied to a realistic gene network problem, and in Section V we demonstrate how to use time scale separation together with the Finite State Projection method. Finally, in Section VI, we discuss our results and outline prospects for further research.

II. BACKGROUND

Presently, the dominant approach to modeling gene regulatory networks is to describe intracellular processes by a series of chemical reactions involving proteins and RNA molecules. For a system of n chemical species, the state of the system inside the cell is specified by copy numbers of each relevant molecule $\mathbf{X} = (X_1, X_2, \ldots, X_n)$. Often, these numbers are relatively small and reactions take place far from the thermodynamic limit, so that mesoscopic effects, most notably fluctuations, have to be taken into account. The state space of the system is not continuous, but a discrete lattice, where each node corresponds to a different $\mathbf{X}$. The size of the lattice is limited by the maximum pos-

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2006		**00-00-2006 to 00-00-2006**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Model reduction using multiple time scales in stochastic gene regulatory networks	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Department of Electrical and Computer Engineering,University of California,Santa Barbara,CA,93106**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**12**	

sible populations of the n chemical species in the finite volume cell.

At the mesoscopic scale one describes the dynamics of the system in terms of the probability of finding the system in a given state $\mathbf{X}$, rather than in terms of trajectories in the state space. The dynamics of the system can be modeled by the master equation for a Markov process on a lattice[9]. Although respectable attempts have been made to introduce deterministic mesoscopic models for chemical reactions[10], presently stochastic methods are used almost exclusively in the study of intracellular processes at the mesoscopic level.

The master equation describes the time evolution of the probability of finding the system in a particular state $\mathbf{X}$. With an enumeration $\mathbf{X} \to i$, which maps each possible state to a single index, the master equation can be written in a familiar gain-loss form[9]

$$\frac{dp_i(t)}{dt} = \sum_{j \neq i} \left[w_{ij} p_j(t) - w_{ji} p_i(t) \right], \tag{1}$$

where p_i is the probability of finding the system in i^{th} state, while w_{ij} are propensity functions. The propensity functions define the probability $w_{ij}dt$ that the system will transition from the j^{th} to the i^{th} state during an infinitesimal time interval dt. The propensities may be obtained from the chemical reaction rates, which often can be measured experimentally. The first term on the right hand side of the master equation describes an increase in the probability p_i due to transitions to the i^{th} state from all other states j, while the second term describes a decrease in p_i due to transitions from the i^{th} state to other states j. If the system is initially found in a state k, the initial condition for the chemical master equation can be written as $p_i(0) = \delta_{ik}$, where δ_{ik} is the Kronecker delta.

The solution for this problem is the probability $p_i(t)$ that the system initially found in state k will be in state i at the later time t. If we define $A_{ij} = w_{ij} - \delta_{ij} \sum_k w_{ki}$, the chemical master equation can be written in a more compact form

$$\dot{p}_i(t) = \sum_j A_{ij}\, p_j(t). \tag{2}$$

The solution to the chemical master equation generally can be expressed in terms of evolution operator $\mathbf{p}(t) = \mathcal{A}(t,0)\mathbf{p}(0)$, which in case of a finite A can be written as

$$\mathbf{p}(t) = \exp(At)\mathbf{p}(0). \tag{3}$$

Solving the master equation at first seems to be a rather simple problem, as there are many efficient methods for solving systems of linear ordinary differential equations. However, if we consider, for example, a process involving three proteins, where each protein comes in say one thousand copies per cell, that gives us up to a billion of different states and a myriad of possible transitions between them. Carrying out calculations for such system without any insight about its biological structure would be impractical at least.

This problem may be ameliorated by using a Monte Carlo type of computation[11]. The idea behind this approach is to start from some initial probability distribution $p_j(0) = \delta_{jk}$, then choose randomly which reaction will take place next, and compute the new state j where the system will be found at some later time t. The probabilities of picking a particular reaction are given by the propensity functions. The hope is that after sufficiently many calculations like this the histogram containing all outcomes will approximate well the solution of the chemical master equation $\mathbf{p}(t)$. The advantage of this approach is that we do not need to calculate the whole matrix A. Instead, we calculate on the fly only those matrix elements that are required for the computation at hand. Furthermore, this method is broadly applicable as it requires little knowledge about the details of the system under consideration. It has been demonstrated[12] that in the limit case of infinitely many runs the Monte Carlo solution approaches the exact solution to the chemical master equation. Therefore the accuracy of the computation can be increased by simply generating more Monte Carlo simulations.

On the down side Monte Carlo methods are notorious for their slow convergence[11], and the amount of computation necessary to get an accurate result may be too large to be completed in a reasonable amount of time. Also, computers cannot produce truly random numbers, so one has to generate something that is as close as possible. Programs called random number generators[13] create periodic sequences of numbers with a large period, which imitate series of random numbers. If the period is too short, periodic patterns will create numerical artifacts in the calculation. On the other hand, high quality random number generators, such as RANLUX[14], take significantly more computer processing time to execute.

Despite their shortcomings Monte Carlo methods remain an important tool for study of intracellular processes and variety of different Monte Carlo implementations[2,3,4,12,15,16,17] have been successfully used thus far.

An alternative approach known as the Finite State Projection[1,18,19] has been proposed recently by Munsky and Khammash. The method is based on a simple observation that some states are more likely to be reached in a finite time than are others. One can then aggregate all improbable states in (2) into a single sink, and consider all transitions to those states as an irreversible loss. This method automatically provides a guaranteed error bound that may be made as small as desired as it was proved in[1]. With some intuition about the system's dynamics, such as knowing the macroscopic steady state, one can develop an efficient system reduction scheme. It has been demonstrated for a number of biological problems[1,18] that in this way the system (2) can be reduced to a surprisingly small number of ordinary differential equations, thereby dramatically reducing the computation time. The re-

duced system can be treated *analytically*, and the method does not require computationally expensive random number generation.

By discarding unlikely states in a systematic way, the Finite State Projection method provides for a bulk system reduction, but the original Finite State Projection stops far short from what can be achieved. For example, the method does not consider how transitions between the remaining states take place. Transition rates between different states typically vary over several orders of magnitude, and by treating them equally one may waste considerable time performing computations to obtain a precision that far outstrips the model's accuracy.

The low probability transitions occur infrequently, so the processes involving them will take place over long time scales, while high probability transitions correspond to fast intracellular processes. Different time scales can pose computational problems, as the system of ordinary differential equations (2) becomes stiff. On the other hand, depending on the length of observation time, the system can be further simplified. For short times slow processes may be neglected, while for long times the effects of fast processes can be averaged.

In what follows we introduce a computational method that addresses these shortcomings by taking advantage of multiple time scales in the master equation to simplify the system of equations and reduce the computation time. This method is in a sense complementary to the Finite State Projection. It can be used independently, but significant benefits may be achieved when the two methods are combined.

III. TIME SCALE SEPARATION

In order to define a proper chemical master equation matrix A has to satisfy some general properties. Since by definition propensity functions $w_{ij} \geqslant 0$, all off-diagonal elements of A are nonnegative. For the same reason, all diagonal elements of A are nonpositive.

In a closed system the probability has to be conserved, so that $\sum_i p_i(t) = \text{const.}$ for all times. That means

$$\frac{d}{dt}\sum_i p_i(t) = 0 \;\Rightarrow\; \sum_i \sum_j A_{ij} p_j(t) = 0, \tag{4}$$

and hence

$$\sum_j \left(\sum_i A_{ij} \right) p_j(t) = 0, \tag{5}$$

for any probability distribution $\mathbf{p}(t) = (p_1(t), \ldots, p_N(t))$. Here with N we denote the number of all possible states where the system can be found[26].

Therefore it must hold $\sum_i A_{ij} = 0$, i.e. the sum of the elements in each column of A must be zero. In other words vector $\mathbf{1} = (1, 1, \ldots, 1)$ is a left eigenvector of A with associated eigenvalue zero:

$$\mathbf{1}^T A = 0. \tag{6}$$

This further means that for the matrix A there exists at least one right eigenvector $\mathbf{v}$ with the zero eigenvalue,

$$A\mathbf{v} = 0. \tag{7}$$

The eigenvector $\mathbf{v}$ represents the steady state probability distribution for the system, and is the nontrivial solution to (2). Furthermore it can be shown[9] that other eigenvalues of A have negative real parts if the matrix A is irreducible, i.e. it cannot be written in a block diagonal form.

Note that we are interested here in the *nontrivial* solution to (2), which exists since $\det A = 0$. There also exists a trivial solution $\mathbf{p} = \mathbf{0}$, but we can discard it as nonphysical since it does not satisfy the normalization condition $|\mathbf{p}| = \mathbf{1}$.

In gene networks we can often identify clusters of states within which transitions occur quite frequently, while transitions between the clusters are relatively rare. The chemical master equation that corresponds to such a situation has a nearly block diagonal structure, so the matrix A in (2) can be written in the form

$$A = H + \epsilon V, \tag{8}$$

where H is a block diagonal matrix describing transitions within the clusters, matrix V describes transitions from one cluster to another, and $\epsilon > 0$ is a small parameter.

In the limit case, $\epsilon \to 0$, the system remains in one cluster of states for an infinitely long time, i.e. the probability for the system to be found in one of the states within the original cluster is one. Therefore, same as the matrix A, each block of H should define a proper master equation. Each block of H has one zero eigenvalue with associated eigenvector $\mathbf{v}_i$, which describes the steady state probability distribution in the i^{th} cluster, while all other eigenvalues of the block have negative real parts.

It is relatively inexpensive to compute the full eigensystems for the smaller blocks of H. From the eigenvectors for each block one can easily construct a matrix S that diagonalizes H:

$$S^{-1} H S = \Lambda, \quad \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N). \tag{9}$$

Matrix S has the same block diagonal structure as H. This procedure is further simplified if some blocks of H are identical, as is often the case. We label eigenvectors and eigenvalues of H so that $\text{Re}\{\lambda_1\} \geqslant \text{Re}\{\lambda_2\} \ldots \geqslant \text{Re}\{\lambda_N\}$. The first m eigenvalues are then equal to zero ($\lambda_{i \leqslant m} = 0$) and the rest have negative real parts.

In order to keep our presentation streamlined, we shall assume that for all negative eigenvalues $|\text{Re}\{\lambda_{i>m}\}| \gg \epsilon$. This is always satisfied if it is possible to make a clear distinction between fast and slow reactions. This assumption can be relaxed and similar results obtained, as we shall demonstrate later.

By applying now S^{-1} to both sides of (2) we obtain

$$\dot{\mathbf{x}} = \left(\Lambda + \epsilon \tilde{V} \right) \mathbf{x}, \tag{10}$$

where $\mathbf{x} = S^{-1}\mathbf{p}$, and $\tilde{V} = S^{-1}VS$. The equation above can be written in the component form as

$$\dot{x}_i = \lambda_i x_i + \epsilon \sum_{j=1}^{N} \tilde{V}_{ij} x_j. \tag{11}$$

From singular perturbation theory (see Appendix) there exists a near identity transformation

$$\mathbf{x} = (I + \epsilon G)\mathbf{y}, \tag{12}$$

that removes all $O(\epsilon)$ terms, which depend on $x_{i>m}$, from the first m equations ($i \leqslant m$). In other words, equations (11) where $\lambda_i = 0$ can be decoupled from the rest of the system by a coordinate transformation (12) through the order $O(\epsilon)$. In the new coordinates the first m equations become

$$\dot{y}_i = \epsilon \sum_{j=1}^{m} \tilde{V}_{ij} y_j + O(\epsilon^2). \tag{13}$$

By truncating $O(\epsilon^2)$ terms in (13) we reduce our system of equations to an m-dimensional problem. The reduced system is computationally less expensive to solve, while it still approximates well the dynamics of the full system. Because (11) has a stable fixed point solution, if initially $|\mathbf{x}(0) - \mathbf{y}(0)| = O(\epsilon)$, then for all times $t > 0$ it holds $|\mathbf{x}(t) - \mathbf{y}(t)| = O(\epsilon)$.

Note that if λ_i is smaller or of the same order as ϵ, the near-identity transformation (12) and its inverse introduce corrections to the i^{th} equation that is only of order $O(\epsilon^2)$. Therefore we do not need to find the exact form of the near-identity transformation, we can simply truncate all terms containing $x_{i>m}$ from the system (11).

The first m equations can be solved now independently of the rest of the system, and their solution can be written as

$$y_i(t) = \sum_{j=1}^{m} \left[\exp(\epsilon \tilde{V}' t)\right]_{ij} y_j(0), \tag{14}$$

where $\tilde{V}'$ is $m \times m$ principal submatrix of $\tilde{V}$, with elements $\tilde{V}_{i,j\leqslant m}$. In many cases of interest, solving (13) is a manageable problem, unlike getting general solution for the chemical master equation (2). Since in the long time limit

$$\lim_{t\to\infty} x_{i>m}(t) = O(\epsilon), \tag{15}$$

as we show in the Appendix, we claim that from the solution to the truncated system (14), we can easily construct an approximation to the full solution of the chemical master equation (3). To do so, we first define an evolution operator $\tilde{\mathcal{V}}(t)$ such that $\tilde{\mathcal{V}}_{ij}(t) = [\exp(\epsilon \tilde{V}' t)]_{ij}$ for $i, j \leqslant m$, and $\tilde{\mathcal{V}}_{ij}(t) = 0$ otherwise. In a block matrix form this is

$$\tilde{\mathcal{V}}(t) = \begin{pmatrix} \exp(\epsilon \tilde{V}' t) & 0 \\ 0 & 0 \end{pmatrix}. \tag{16}$$

The price we pay for simplicity here is that $\tilde{\mathcal{V}}(0)$ is not an identity matrix, so the initial condition $y_{i>m}(0)$ also gets truncated, and that truncation error is generally larger than $O(\epsilon)$. The transient time $T(\epsilon)$, after which the solution of the truncated system is $O(\epsilon)$ close to the solution of the full system, can be estimated from the leading nonzero eigenvalue as

$$T(\epsilon) \sim \ln \epsilon / \lambda_{m+1}. \tag{17}$$

If the time scale separation in (8) is done accurately, this transient is negligible for all practical purposes. However, time scale separation in a large system is not always obvious, and may be error prone. We shall further discuss this issue in Section III A.

Finally, by performing the inverse S transformation on $\tilde{\mathcal{V}}(t)$, we obtain

$$\mathcal{V}(t) = S\tilde{\mathcal{V}}(t)S^{-1}, \tag{18}$$

which leads to the $O(\epsilon)$ approximation to the solution of the chemical master equation (2), that is

$$|\mathbf{p}(t) - \mathcal{V}(t)\mathbf{p}(0)| = O(\epsilon), \tag{19}$$

for all times $t > T(\epsilon)$.

Note that neither $\tilde{V}$ nor $\tilde{V}'$ are generators for the evolution operator $\tilde{\mathcal{V}}$, so their eigenvectors cannot be used directly to compute the steady state probability distribution for $\mathbf{p}(t)$.

A. Computational algorithm

One should note that due to the truncation of $\tilde{V}$ only contributions of the first m columns of S and m rows of S^{-1} affect the approximate solution. As a result the computation can be greatly simplified – instead of calculating full eigensystems for each block H_i, it suffices to find only the eigenvectors $\mathbf{v}_i$ associated with the zero eigenvalues. Instead of S we use the $N \times m$ matrix S^R, whose columns are made up of the right eigenvectors of H, while instead of S^{-1} we use the $m \times N$ matrix S^L, whose rows are made up of the left eigenvectors of H. Note that the left eigenvectors are always $\mathbf{1}_i^T$, provided all $|\mathbf{v}_i| = 1$, so the matrix S^L is obtained at no computational cost. The accuracy of the calculation is known *a priori* to be $O(\epsilon)$ for all $t > T(\epsilon)$.

To improve the reliability and robustness of our calculation, we can optionally add a transient time check to our algorithm. To do so we first need to find eigenvalues for all blocks H_i. This comes at a relatively small computational cost, and can be performed before all the other calculations. The transient time needed to obtain the desired accuracy is estimated from the leading negative eigenvalue λ_{m+1} according to (17). If the transient is too long, that can be remedied by expanding matrices S^R and S^L to include the right and left eigenvectors corresponding to λ_{m+1}, respectively. The transient

time is then governed by next negative eigenvalue λ_{m+2}. This procedure can be repeated until the desired accuracy is achieved, thereby sacrificing computational time for precision. Note that in this case the right eigenvectors corresponding to nonzero eigenvalues cannot be obtained trivially.

By performing this test, we also ensure that condition $|\lambda_{i>m}| \gg \epsilon$ is satisfied. Eigenvalues of H that are $O(\epsilon)$ or smaller will result in long transient times. By expanding transformation S to include eigenvectors corresponding to these eigenvalues, we essentially treat them as if they were part of $\tilde{V}$. This procedure adds robustness to the method with respect to separating fast and slow reactions in (8).

Let us summarize how to implement our method:

1. Specify problem parameters. If necessary apply a finite projection to the full state space. Use propensity functions and/or physical intuition to separate H and V.

2. Find the eigenvalues of the uncoupled system, and identify "slow" ones with respect to a preset transient time $T(\epsilon)$.

3. Find right and left eigenvectors corresponding to the slow eigenvalues and construct rectangular matrices S^R and S^L.

4. Calculate $k \times k$ matrix $\tilde{V}' = S^L V S^R$, where k is the number of slow eigenvalues.

5. Compute $k \times k$ matrix $\exp(\epsilon \tilde{V}' t)$.

6. Perform the inverse transformation

$$S^R \exp(\epsilon \tilde{V}' t)\, S^L = \mathcal{V}(t)$$

in order to obtain the approximation to $\exp(At)$ for all times $t > T(\epsilon)$.

B. Example

Let us illustrate this technique with a simple example. We assume two weakly interacting systems that can be found in three different states each. We choose matrices H and V in an arbitrary way, with the only constraint that they define a master equation. In our example

$$H = \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix} \tag{20}$$

is a block diagonal matrix with blocks

$$H_1 = \begin{pmatrix} -4 & 2 & 4 \\ 1 & -2 & 0 \\ 3 & 0 & -4 \end{pmatrix} \quad \text{and} \quad H_2 = \begin{pmatrix} -6 & 3 & 2 \\ 2 & -3 & 0 \\ 4 & 0 & -2 \end{pmatrix}. \tag{21}$$

We find that blocks H_1 and H_2 have one zero eigenvalue each, with corresponding right eigenvectors $\mathbf{v}_1 = (4, 2, 3)$

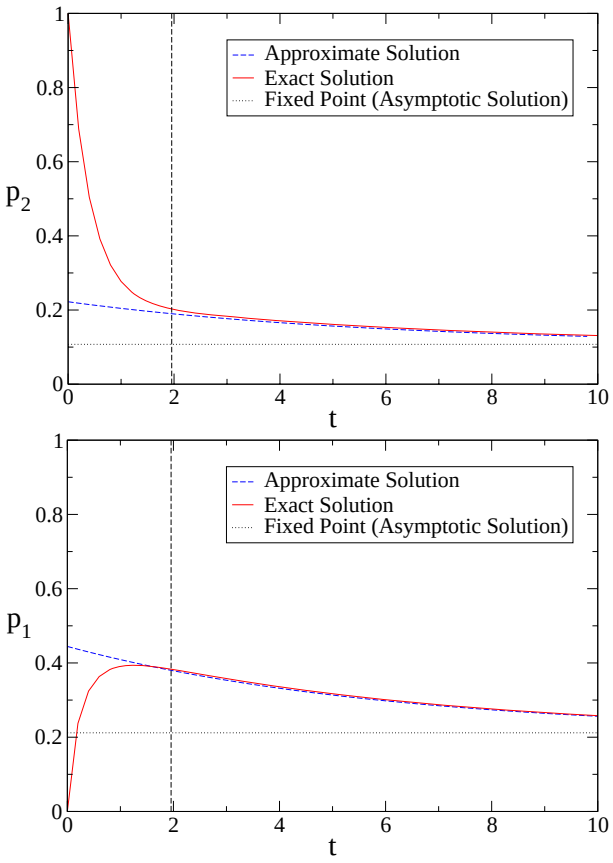

FIG. 1: Comparison of the approximate and the exact solution to the master equation in Section III B. The initial probability distribution is $p_i(0) = \delta_{2i}$. The transient time is estimated to be $T(\epsilon) = \ln \epsilon / \lambda_3 = 1.96$ for $\epsilon = 0.01$, and is denoted by the vertical line on the graph.

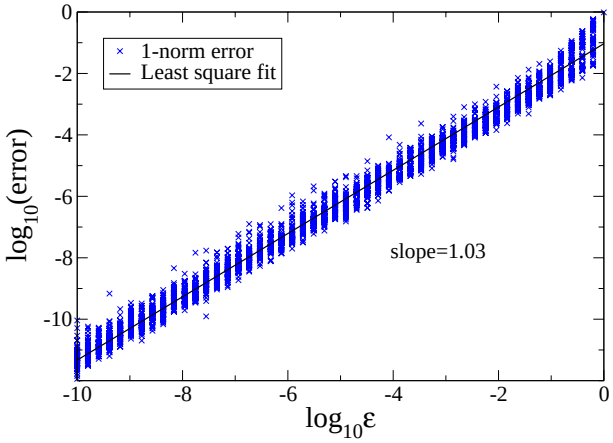

FIG. 2: 1-norm error in probability distribution for the truncated solution versus ϵ. For each value of ϵ we have randomly generated 50 matrices H and V, so that every $H + \epsilon V$ defines a proper master equation. Each matrix H has between 2 and 6 blocks and each block has size between 2 and 21. The elements of H and V are randomly generated from a uniform distribution between 0 and 1. The probability distributions were calculated after time $t = 2T(\epsilon) = 2 \ln \epsilon / \lambda_{m+1}$.

and $\mathbf{v}_2 = (3, 2, 6)$. From these eigenvectors, we assemble the matrix S^R,

$$S^R = \begin{pmatrix} 4/9 & 0 \\ 2/9 & 0 \\ 3/9 & 0 \\ 0 & 3/11 \\ 0 & 2/11 \\ 0 & 6/11 \end{pmatrix}. \tag{22}$$

The matrix composed of left eigenvectors of H_1 and H_2 is similarly used to form S^L,

$$S^L = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}. \tag{23}$$

In our example the coupling matrix is

$$V = \begin{pmatrix} -8 & 0 & 0 & 5 & 3 & 2 \\ 0 & -5 & 0 & 2 & 3 & 1 \\ 0 & 0 & -12 & 4 & 6 & 2 \\ 4 & 2 & 3 & -11 & 0 & 0 \\ 1 & 2 & 5 & 0 & -12 & 0 \\ 3 & 1 & 4 & 0 & 0 & -5 \end{pmatrix}. \tag{24}$$

To get the equations for the slowly changing variables (13), we calculate

$$\tilde{V}' = S^L V S^R = \begin{pmatrix} -87/11 & 78/11 \\ 29/3 & -26/3 \end{pmatrix}. \tag{25}$$

Next, we calculate the evolution operator for the truncated system, $\exp(\epsilon \tilde{V}' t)$, and perform the inverse S-transformation to obtain

$$\mathcal{V}(t) = S^R \exp(\epsilon \tilde{V}' t) S^L. \tag{26}$$

Finally, we obtain the approximate solution as

$$\mathbf{p}(t) = \mathcal{V}(t)\, \mathbf{p}(0). \tag{27}$$

As an illustration, in Figure 1 we show components $p_1(t)$ and $p_2(t)$ of the solution above for the initial condition $p_i(0) = \delta_{2i}$, and $\epsilon = 0.01$. We can see that after the transient time (17) has elapsed we obtain a good agreement between the exact and the approximate solution to this example problem.

To further support our results, we randomly generate a large number of master equations with similar near block-diagonal structure and compare their exact solutions to the approximate solutions obtained using our approach. The numerical results presented in Figure 2 show that the approximation error is controlled by the small parameter ϵ.

IV. PAP SWITCH

Pili are small hair-like structures that enable bacteria to bind to epithelial cells and thereby significantly increase the bacteria's ability to infect host organisms. However, pili expression comes at cost to the bacteria, as the production of pili requires a large portion of the cellular energy. Whether or not *E. coli* are piliated depends upon the regulation of genes such as the pyelonephritis-associated pili (*pap*) genes. Here we study a simplified version of the PAP switch model[18], which analyzes the regulatory network responsible for controlling one type of pili.

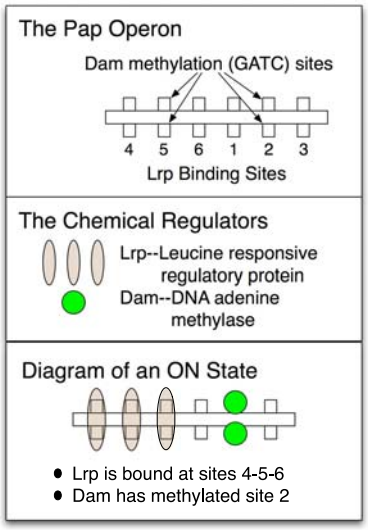

FIG. 3: Schematic of the *pap* operon (top), key regulatory components of the Pap switch (middle) and diagram of the operon in its on state (bottom).

Recent experiments[20,21] have identified two transcription factors that affect expression of the *pap* gene, and six binding sites for the two. The transcription factors are DNA adenine methylase (Dam) and leucine responsive regulatory protein (Lrp). Dam binds and applies methyl groups to GATC sites at 2 and 5, as shown in Figure 3. This Dam methylation is an irreversible and relatively slow process. On the other hand, Lrp binds cooperatively to three adjacent sites at a time, either 1-2-3 or 4-5-6 (Figure 3). These reactions are fast and reversible. Lrp binding also inhibits Dam methylation. Altogether this makes 16 possible states in which the PAP switch can be found. These are schematically described by the network model shown in Figure 4. In this model *pap* transcription occurs only in the state 11 when Lrp is bound to sites 1-2-3 and site 5 is methylated. We assume that cell replication always resets the system to state 1. A solution to the chemical master equation for this problem gives the time evolution of the probability of finding the system in each state including state 11, which is proportional to the probability of *pap* gene expression.

Since the two transcription factors bind at significantly different rates, a smart bookkeeping practice would be to record Lrp binding propensities in the matrix H and methylation propensities in V as defined in (8). With a convenient labeling scheme, as shown in Figure 4, we can

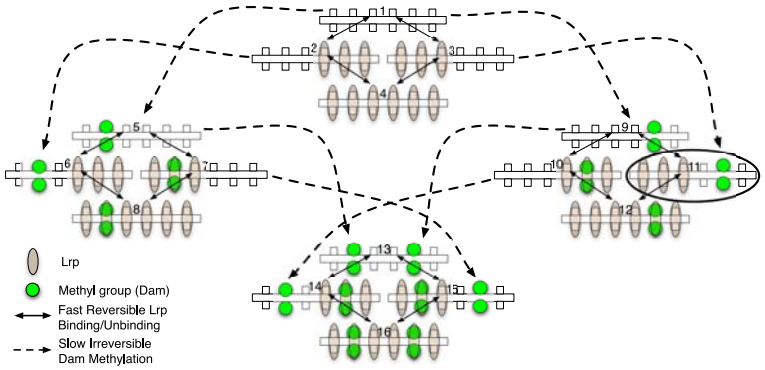

FIG. 4: PAP switch schematic diagram.

express H in a simple block diagonal form,

$$H = \begin{pmatrix} H_1 & 0 & 0 & 0 \\ 0 & H_2 & 0 & 0 \\ 0 & 0 & H_3 & 0 \\ 0 & 0 & 0 & H_4 \end{pmatrix}. \tag{28}$$

Recent experimental data[21] reveals that the propensities of Lrp binding at sites 4-5-6 depend strongly on the methylation pattern of site 5, while propensities of Lrp at sites 1-2-3 do not significantly depend upon the methylation pattern of site 2. Thus we find that there are only two distinct blocks as

$$H_1 = H_3 = \begin{pmatrix} -9500 & 6.8 & 0.09 & 0 \\ 9270 & -18.4 & 0 & 0.09 \\ 230 & 0 & -463.29 & 6.79 \\ 0 & 11.6 & 463.2 & -6.88 \end{pmatrix} \tag{29}$$

and

$$H_2 = H_4 = \begin{pmatrix} -9500 & 62 & 0.09 & 0 \\ 9270 & -73 & 0 & 0.09 \\ 230 & 0 & -463.29 & 61.76 \\ 0 & 11 & 463.2 & -61.85 \end{pmatrix}. \tag{30}$$

Leading eigenvalues for both, H_1 and H_2, are zero, while the next largest eigenvalue is of order $\lambda_5 \sim -10$. On the other hand we estimate that all methylation propensities have the same value $\epsilon = 0.17$. Following our labeling scheme (Figure 4) the nonzero entries of matrix V are then: $V_{1,1} = -2$, $V_{2,2} = V_{3,3} = V_{5,5} = V_{7,7} = V_{9,9} = V_{10,10} = -1$, and $V_{5,1} = V_{9,1} = V_{6,2} = V_{11,3} = V_{13,5} = V_{13,9} = V_{14,10} = V_{15,7} = 1$. Therefore, all we need to construct the matrix S^R are the right eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ that correspond to the zero eigenvalues of H_1 and H_2, respectively. Following the footsteps outlined in Section III we reduce the PAP switch model to a 4-dimensional system and carry out calculation for the probability p_{11}, which is proportional to the PAP transcription probability.

The PAP switch model we presented here is simple enough to be integrated directly so we can compare results for the full system and the reduced system. As we show in the Figure 5, all the important information about the system's behavior is preserved in the reduced model.

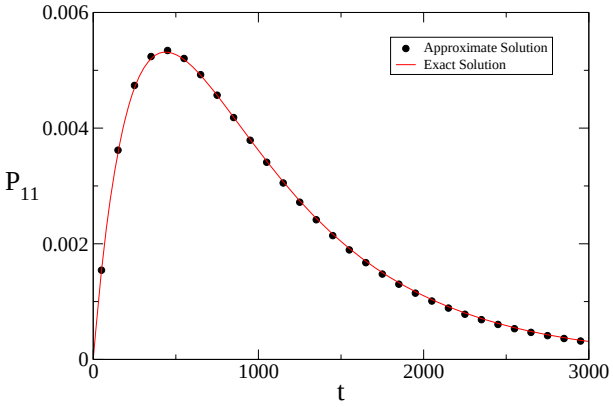

FIG. 5: Time evolution of pap gene expression probability. Initially no transcription factors are bound the pap operon, so the initial condition is $p_1(0) = 1$ and $p_{i\neq 1} = 0$. The transient time 17 is less than 1 in our time units.

This model predicts a short time lag between replication and Pap production, since methylation of site 2 must occur before *pap* expression. Further, since Dam methylation at 5 prohibits expression, if the cell waits too long to decide to switch "on", it will most probably miss its chance and remain "off". Thus, a newly created *E. coli* cell will most likely express the pap gene at some point shortly after replication. Probability of expressing pili drops significantly at later times and cell resources are used for other functions, such as initiating the next replication cycle.

V. HEAT SHOCK SYSTEM

Through evolution all living organisms have developed mechanisms for dealing with environmental stress. One such mechanism is cellular heat shock response. Increased temperature causes proteins inside the cell to misfold and thereby loose their functionality. In response, the cell produces heat shock proteins, most no-

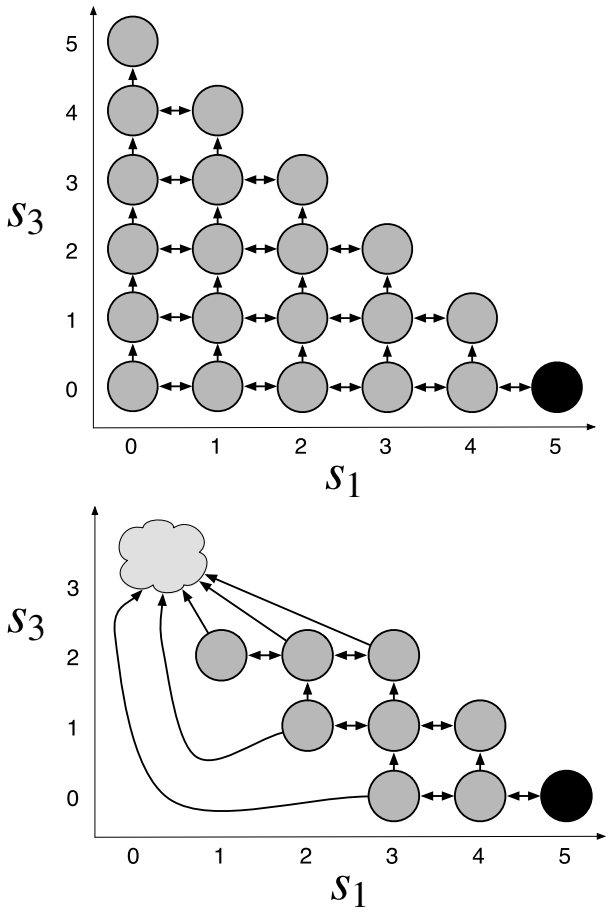

FIG. 6: (a) Two dimensional lattice representing possible states and transitions in the heat shock model. Here s_1 and s_3 are populations of σ_{32}-DnaK and σ_{32}-RNAP compounds, respectively, while s_2 is the population of free σ_{32} molecules. Reactions $s_1 \rightleftharpoons s_2$, are represented by bidirectional horizontal arrows and reactions $s_2 \rightarrow s_3$ is represented with vertical arrows. The total number of σ_{32} is constant (in this example $s_1 + s_2 + s_3 = 5$), so the chemical state of the system is uniquely defined by s_1 and s_3 alone. (b) The same lattice after applying the Finite State Projection. Unlikely states have been aggregated into a single sink state.

tably molecular chaperones and proteases. Molecular chaperones help refold deformed proteins and restore their original function. Proteases, on the other hand, help degrade damaged proteins, hopefully before they trigger unwanted chemical reactions in the cell. At the crux of the heat shock response mechanism in *E. coli* is the synthesis of σ_{32}-RNAP complex[22]. Here we study a simplified model for σ_{32}-RNAP synthesis using our time scale separation method in conjunction with a Finite State Projection. The analysis of the entire heat shock response mechanism in *E. coli* is rather complicated, and is beyond the scope of this paper.

At normal physiological temperatures σ_{32} protein is found almost exclusively in a complex σ_{32}-DnaK. As the temperature increases this complex becomes less stable

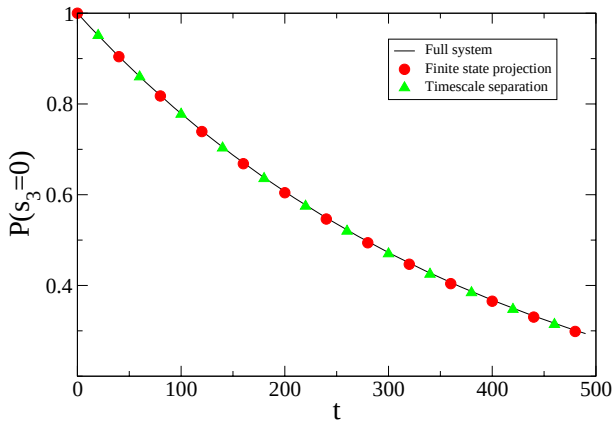

FIG. 7: Probability that no σ_{32}-RNAP molecule has been sythesized in the heat shock toy model.

and there is a non-negligible probability of finding free σ_{32} inside the cell. The free σ_{32} then can combine with RNA polymerase through what can be considered an irreversible reaction to form a σ_{32}-RNAP complex. In turn, σ_{32}-RNAP initiates transcription of genes encoding heat shock proteins. This regulatory mechanism can be summarized in a simple set of reactions,

$$s_1 \rightleftharpoons s_2 \rightarrow s_3, \tag{31}$$

where s_1, s_2 and s_3 correspond to the σ_{32}-DnaK complex, the σ_{32} heat shock regulator and the σ_{32}-RNAP complex, respectively. This model of the heat shock subsystem has been analyzed before using various computational methods including Monte Carlo implementations[3,23].

In the biological system, the relative rates of the reactions are such that the reaction from s_2 to s_1 is by far the fastest, and σ_{32} molecules infrequently escape from DnaK long enough to form the σ_{32}-RNAP complex. The purpose of this mechanism is to strike a balance between fixing the damage produced by heat and saving the cell's resources, as a significant portion of cell energy is consumed when producing heat shock proteins. The optimal response to the heat shock is not massive, but measured production of heat shock proteins, which leaves sufficient resources for other cellular functions. We use the following set of parameters values for the reaction rates[3,23]:

$$c_1 = 10, \quad c_2 = 4 \times 10^4, \quad c_3 = 2. \tag{32}$$

For simplicity, in our model we assume that the total number of σ_{32} – free or in compounds – is constant, so that $s_1 + s_2 + s_3 = \text{const}$. With this constraint the reachable states of this three species problem can be represented on a two dimensional lattice.

For illustrative purposes, Figure 6a, shows one such lattice for an initial condition of $s_1 = 5$ and $s_2 = s_3 = 0$. Here, the total population is fixed at five, and there is a total of 21 reachable states.

We first apply the Finite State Projection. We estimate that all states where $s_2 > 2$ or $s_3 > 2$ are unlikely

to be reached in a short time, so we aggregate them into a sink node as shown in Figure 6b, thereby reducing this to a ten state problem. From the transitions to the aggregated state, we find a strict upper bound on the error introduced by such an approximation. For our set of parameters the approximation the error is guaranteed to be below 8% for any time $t \leqslant 500$.

Next, we further reduce this system by applying time scale separation. Elements of the matrix A_{FSP}, which defines the master equation for the system obtained after the Finite State Projection, can be read off of Figure 6b. In accordance with our bookkeeping practice we write $A_{FSP} = H + V$ and record all reversible reactions $s_1 \rightleftharpoons s_2$ in the matrix H, and all other reactions, including $s_2 \to s_3$ and transitions to the aggregated state, in the matrix V. By doing so we ensure that all fast reactions are contained in H. Note that there is no unique way to separate fast and slow reactions and we chose this one for its simplicity. Matrix H has a block diagonal structure

$$H = \begin{pmatrix} H_3 & & & \\ & H_2 & & \\ & & H_1 & \\ & & & 0 \end{pmatrix}, \tag{33}$$

where each block

$$H_k = \begin{pmatrix} -(k+2)c_1 & c_2 & 0 \\ (k+2)c_1 & -(k+1)c_1 - c_2 & 2c_2 \\ 0 & (k+1)c_1 & -2c_2 \end{pmatrix} \tag{34}$$

corresponds to a row of states in Figure 6b. The zero in the last row is just a scalar, and it corresponds to the aggregated state. The matrix ϵV is made up of irreversible reactions (vertical transitions in Figure 6b) and therefore has a lower triangular form:

$$\epsilon V = \begin{pmatrix} 0 & & & & & & & & & \\ 0 & -c_3 & & & & & & & & \\ 0 & 0 & -3c_1 - 2c_3 & & & & & & & \\ 0 & c_3 & 0 & 0 & & & & & & \\ 0 & 0 & 2c_3 & 0 & -c_3 & & & & & \\ 0 & 0 & 0 & 0 & 0 & -2c_1 - 2c_3 & & & & \\ 0 & 0 & 0 & 0 & c_3 & 0 & 0 & & & \\ 0 & 0 & 0 & 0 & 0 & 2c_3 & 0 & -c_3 & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -c_1 - 2c_3 & \\ 0 & 0 & 3c_1 & 0 & 0 & 2c_1 & 0 & c_3 & c_1 + 2c_3 & 0 \end{pmatrix}.$$

For the reaction rates above, the first four eigenvalues of H are zero, and the rest have negative real parts each with magnitude of order 10^4 or larger, suggesting that the truncation in (16) is indeed valid for this problem. Therefore the dynamics of 9-dimensional system obtained by the Finite State Projection can be well approximated by a system of only four linear ordinary differential equations. By applying algorithm from Section III A we find that the time scale separation introduces error of order 10^{-3} with respect to the solution obtained by Finite State Projection alone. The transient time (17) is estimated to be 2×10^{-4}, and is negligible considering the time interval of interest.

In Figure 7 we compare results obtained by solving the full system directly, using Finite State Projection alone and using Finite State Projection and time scale separation combined. The figure shows how probability of having no σ_{32}-RNAP complex in the cell decreases with time. All three results are in a good agreement as our calculations predicted. Note that the Finite State Projection error in this example is much smaller than the estimated upper bound.

The advantage of combining the Finite State Projection and time scale separation becomes obvious if we consider a more realistic and much larger problem with the same reactions but with initial conditions $s_1 = 2000$ and $s_2 = s_3 = 0$. In this case there are 2,001,000 reachable states, and the full chemical master equation is too large to be tackled directly. However, by applying the Finite State Projection we find that truncating every state where $s_3 \leqslant 340$ and $s_2 \leqslant 15$ introduces 1-norm error that is less than 10^{-3} for times $t \leqslant 300s$. The resulting matrix A is of size 5457×5457, and has near block diagonal form (8) similar to the example in Figure 6. Its block diagonal part H contains 341 irreducible blocks each with 16 rows and columns. Same as in the previous example the leading nonzero eigenvalue of H has a negative real part of magnitude 10^4, so the system can be reduced to a 342 state model using the time scale separation algorithm. Should we apply time scale separation directly to the full system, not only the amount of the computation would be significantly larger, but there would be no simple way to obtain the near block diagonal structure as it was the case in the previous example.

The solution to this problem shows how the number of compounds σ_{32}-RNAP grows in time if the temper-

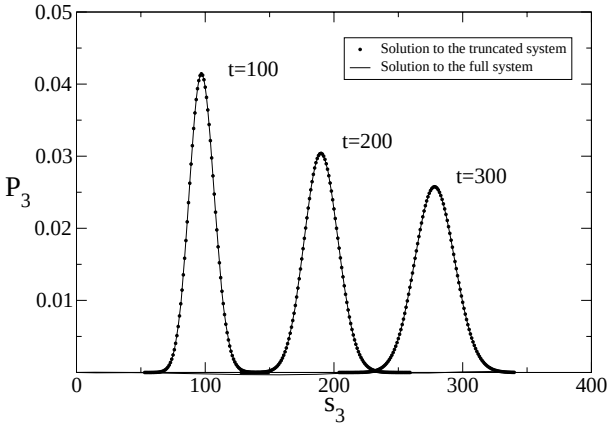

FIG. 8: Probability distribution for s_3 calculated at three different times. The truncated solution (dots) approximates well the solution to the full system (solid lines).

ature is constant and slightly above normal physiological level. This number is proportional to the number of heat shock proteins produced in the cell. With the Finite State Projection solution, we have computed the probability distribution for s_3 at three times $t = 100$, 200, and 300. Figure 8 illustrates the probability distribution of species σ_{32}-RNAP (s_3) at these instances in time. The Finite State Projection solution, shown with solid lines, takes approximately 220 seconds to compute. The dots represent the probability distribution of s_3 as computed using our time scale separation method applied atop of the Finite State Projection. The difference between the two results is indistinguishable, but the total computational effort for the reduced, 342 state model is 200-fold less; the truncated model takes only about a second to solve the problem.

VI. CONCLUSION

Until recently, it was thought that the chemical master equation could not be analytically solved except for the most trivial systems. Previous work on the Finite State Projection demonstrated that for many biological systems, bulk system reductions could bring models closer into the fold of solvable problems. Here we have shown that the Finite State Projection method can be further enhanced when solving the chemical master equation for systems involving multiple time scales. In combination with Finite State Projection method, we have shown that our algorithm, based upon singular perturbation theory, provides a powerful computational tool for studying intracelular processes and gene regulatory networks.

Similar problems were studied earlier with specially designed Monte Carlo implementations[2,3] or hybrid methods[4]. In contrast to these, our method does not require random number generation, and its accuracy is given *a priori*. A further advantage of our method is its ease of implementation and the speed of computations. The proposed algorithm is particularly fast when implemented on systems for which there are strict means of separating slow and fast reactions.

The Finite State Projection and our time scale separation approach also provide valuable insight as to how one may further deal with the bewildering complexity that intracellular processes exhibit. First, cellular processes are limited by cell size and available energy. It is then plausible that the main features of intracellular dynamics can be captured in a relatively small subset of the state space, as the results obtained by Finite State Projection suggest. Another typical feature of intracellular processes is that they are composed of reactions that take place on different timescales. Depending on the observation time of interest, some of these reactions can be neglected, while some will contribute only through their averages. Preliminary success with our approach gives us a hope that relatively simple models for intracellular processes can be tailored when a region in the state space and observation time of interest are known.

Of course, one can easily envision that additional model reductions may be possible to even further enhance the power of both the Finite State Projection and the time scale separation approach. Indeed some reductions based upon control theory[19] are already becoming apparent. Also, in our computations we have used off the shelf numerical routines for eigensystem calculations and matrix exponentiation. Further improvements in computational speed can be achieved if these routines are optimized for matrices which define master equations and their special properties. We intend to investigate these possibilities in the future.

Acknowledgments

This material is based upon work supported by National Science Foundation grant CCF-0326576, Institute of Collaborative Biotechnologies grant DAAD19-03-D-0004 from the US Army Research office. S. P. thanks Krešimir Josić of University of Houston for a productive discussion and useful comments.

APPENDIX

Singular perturbation theory has been extensively studied in various literature. However, most of the literature in this area is of wide scope and often very technical. In order to spare the reader some time, here we present a heuristic argument, which provides a mathematical justification for our method, while keeping technicalities at minimum. For rigorous proofs interested reader may want to consult for example[24,25].

Consider a weakly perturbed linear N-dimensional sys-

tem described by (11)

$$\dot{x}_i = \lambda_i x_i + \epsilon \sum_{j=1}^{N} V_{ij} x_j, \tag{A.1}$$

where $\lambda_i = 0$ for $i \leqslant m$, and λ_i has negative real part for $i > m$. We want to find a near identity coordinate transformation (12) that would remove as many $O(\epsilon)$ terms as possible from (A.1) and "push" them to higher orders in ϵ. After substituting $\mathbf{x} = (I + \epsilon G)\mathbf{y}$ in (A.1) we get

$$\begin{aligned} \dot{y}_i \;=\; & \lambda_i y_i + \epsilon \sum_{j=1}^{N} \tilde{V}_{ij} y_j - \epsilon \sum_{j=1}^{N} G_{ij} \lambda_j y_j \\ & + \epsilon \lambda_i \sum_{j=1}^{N} G_{ij} y_j + O(\epsilon^2) \end{aligned} \tag{A.2}$$

By equating all $O(\epsilon)$ terms to zero we find

$$\sum_{j=1}^{N} \left(V_{ij} - G_{ij}\lambda_j + \lambda_i G_{ij} \right) y_j = 0, \tag{A.3}$$

and by solving for G_{ij} we obtain

$$G_{ij} = \frac{V_{ij}}{\lambda_j - \lambda_i}. \tag{A.4}$$

Therefore, we can always find G_{ij} except when $\lambda_i = \lambda_j$. In other words, all nonresonant terms can be removed through $O(\epsilon)$ from (A.1) by a near identity transformation (12). In our method we are interested in separating slow and fast processes in the system, so we shall define matrix G in (12) as follows:

$$G_{ij} = \begin{cases} \dfrac{V_{ij}}{\lambda_j - \lambda_i} & i \leqslant m < j \\ 0 & \text{otherwise} \end{cases} \tag{A.5}$$

By substituting this expression for G in (A.2) we find that

$$\begin{aligned} \dot{y}_i &= \epsilon \sum_{j=1}^{m} V_{ij} y_j + O(\epsilon^2) && i \leqslant m \\ \dot{y}_i &= \lambda_i y_i + \epsilon \sum_{j=1}^{N} V_{ij} y_j + O(\epsilon^2) && i > m \end{aligned}$$

We observe that first m equations decouple from the rest of the system through $O(\epsilon)$, and can be solved independently after truncating higher order terms. Furthermore, the near identity transformation (12) does not introduce any new $O(\epsilon)$ terms to the first m equations, so it is essentially just a truncation of all $x_{i>m}$ terms from (A.1). We do not need to calculate G and perform transformation (12) as such transformation is guaranteed to exist.

It remains to show that the solution to truncated system (13) will be $O(\epsilon)$ close to solution to (A.1) on a time interval of interest. These equations are linear and hence can be solved analytically, but let us take an extra step here and expand the solution to (A.1) in powers of ϵ

$$x_i(t) = x_i^{(0)}(t) + \epsilon x_i^{(1)}(t) + \epsilon^2 x_i^{(2)}(t) + \dots \tag{A.6}$$

By substituting this expression into (A.1) and grouping same orders in ϵ we get series of equations

$$\begin{aligned} \epsilon^0 &: \; \dot{x}_i^{(0)}(t) = \lambda_i x_i^{(0)}(t) \\ \epsilon^1 &: \; \dot{x}_i^{(1)}(t) = \lambda_i x_i^{(1)}(t) + \sum_{j=1}^{N} V_{ij} x_i^{(0)}(t) \\ \epsilon^2 &: \; \dot{x}_i^{(2)}(t) = \lambda_i x_i^{(2)}(t) + \sum_{j=1}^{N} V_{ij} x_i^{(1)}(t) \\ \dots & \quad \dots \end{aligned}$$

which we can solve in a straightforward way to obtain

$$\begin{aligned} & x_i^{(0)}(t) = e^{\lambda_i t} x_i^{(0)}(0) \\ & x_i^{(1)}(t) = e^{\lambda_i t} x_i^{(1)}(0) + e^{\lambda_i t} \sum_{j=1}^{N} V_{ij} x_i^{(0)}(0) \int_0^t e^{(\lambda_j - \lambda_i)s} ds \\ & \dots \end{aligned}$$

Let us first consider equations $i \leqslant m$. The solution to (A.1) through $O(\epsilon)$ then can be written as

$$\begin{aligned} x_i(t) \;=\; & x_i(0) + \epsilon \sum_{j=1}^{m} V_{ij} x_j(0) t \\ & + \epsilon \sum_{j=m+1}^{N} \frac{e^{\lambda_j t} - 1}{\lambda_j} V_{ij} x_j(0) + O(\epsilon^2), \end{aligned}$$

where we substituted $x_i(0) = x_i^{(0)}(0) + \epsilon x_i^{(1)}(0) + O(\epsilon^2)$ and $\lambda_i = 0$. Since the system has one stable steady state solution the series above must converge for all times. The first two terms in the expansion above are equal to the first two terms in the expansion of the solution (14) for the truncated system. Therefore, for $y_i(0) = x_i(0)$ it is

$$|x_i(t) - y_i(t)| = \epsilon \left| \sum_{j=m+1}^{N} \frac{e^{\lambda_j t} - 1}{\lambda_j} V_{ij} x_j^{(0)}(0) \right| + O(\epsilon^2).$$

In the expression above all $\lambda_j < 0$, therefore $|x_i(t) - y_i(t)| = O(\epsilon)$ holds for all $t > 0$. Since the expression for $x_i(t)$ is convergent series and $x_i(0)$ are linearly independent, we conclude that $y_i(t)$ must also have a fixed point solution, which is $O(\epsilon)$ close to the solution of the full system.

Next we consider equations in (A.1) where $i > m$. The solution to these can be expanded in terms of ϵ as

$$\begin{aligned} x_i(t) \;=\; & e^{\lambda_i t} x_i(0) + \epsilon \frac{e^{\lambda_i t} - 1}{\lambda_i} \sum_{j=1}^{m} V_{ij} x_j(0) \\ & + \epsilon t e^{\lambda_i t} \sum_{j=m+1}^{N} V_{ij} x_j(0) + O(\epsilon^2) \end{aligned}$$

Our truncation algorithm (Sec. III A) sets all $y_i(t) \equiv 0$, so initially the difference between full and truncated solution is whatever the initial condition $x_i(0)$ is, and it can be larger than $O(\epsilon)$. However, in the limit case

$$\lim_{t\to\infty} |x_i(t)| = \epsilon \left| \frac{1}{\lambda_i} \sum_{j=1}^{m} V_{ij} x_j^{(0)}(0) \right| + O(\epsilon^2) \qquad \text{(A.7)}$$

That means the truncation introduces $O(\epsilon)$ error to the asymptotic solution. Larger errors may occur only during the finite transient time $0 < t < T(\epsilon)$, where $T(\epsilon)$ is given in (17). One can verify this by substituting right hand side of (17) for time in the solution $x_{i>m}(t)$ above.

* URL: http://www.engineering.ucsb.edu/~peles/
† Electronic address: brianem@engineering.ucsb.edu
‡ Electronic address: khammash@engineering.ucsb.edu

[1] B. Munsky and M. Khammash, Journal of Chemical Physics **124**, 044104 (2006).
[2] C. V. Rao and A. P. Arkin, Journal of Chemical Physics **118**, 4999 (2003).
[3] Y. Cao, D. T. Gillespie, and L. R. Petzold, Journal of Chemical Physics **122** (2005).
[4] Y. N. Kaznessis, Chemical Engineering Science **61**, 940 (2006).
[5] H. A. Simon and A. Ando, Econometrica **29**, 111138 (1961).
[6] R. G. Phillips and P. V. Kokotovic, IEEE Transactions on Automatic Control **26**, 1087 (1981).
[7] S. Dey and I. Mareels, IEEE Transactions on Signal Processing (2004).
[8] M. A. Gallivan and R. M. Murray, International Journal of Robust and Nonlinear Control (2004).
[9] N. G. V. Kampen, *Stochastic Processes in physics and chemistry* (Elsevier, 2001).
[10] L. A. Bunimovich and M. F. Demers, Journal of Statistical Physics **120** (2005).
[11] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2005), 2nd ed.
[12] D. T. Gillespie, Journal of Computational Physics **22**, 403 (1976).
[13] P. L'Ecuyer, in *Handbook on Simulation*, edited by J. Banks (Wiley, 1998), chap. 4, pp. 93–137.
[14] M. Lüscher, Computer Physics Communications **79**, 100 (1994).
[15] D. T. Gillespie, Journal of Chemical Physics **115**, 1716 (2001).
[16] D. T. Gillespie and L. R. Petzold, Journal of Chemical Physics **119**, 8229 (2003).
[17] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, Journal of Chemical Physics **119**, 12784 (2003).
[18] B. Munsky, A. Hernday, D. Low, and M. Khammash, in *Proceedings of FOSBE* (2005), pp. 145–148.
[19] B. Munsky and M. Khammash, in *Proceedings of the 45th IEEE Conference on Decision and Control* (2006), to be published.
[20] A. D. Hernday, M. Krabbe, B. A. Braaten, and D. A. Low, Proceedings of the National Academy of Sciences of the United States of America **99**, 16470 (2002).
[21] A. D. Hernday, B. A. Braaten, and D. A. Low, Molecular Cell **12**, 947 (2003).
[22] H. E. Samad, H. Kurata, J. C. Doyle, C. A. Gross, and M. Khammash, Proceedings of the National Academy of Sciences of the United States of America **102**, 2736 (2005).
[23] H. E. Samad, M. Khammash, L. Petzold, and D. T. Gillespie, International Journal of Robust Nonlinear Control (2005).
[24] T. J. Kaper, in *Analyzing multiscale phenomena using singular perturbation methods* (American Mathematical Society, 1999), pp. 85–131.
[25] T. Kato, *Perturbation Theory for Linear Operators* (Springer-Verlag, Berlin, Heidelberg, New York, 1980).

[26] We shall assume through the rest of the paper that N is finite. For systems in which N is infinite, the Finite State Projection could be applied to find a finite system, which approximates the original to an arbitrary degree of accuracy.