

# A reduced model solution for the chemical master equation arising in stochastic analyses of biological networks.

Brian Munsky and Mustafa Khammash

**Abstract**—This article introduces the Observability Aggregated Finite State Projection (OAFSP) method for use in the stochastic analysis of biological systems. The small chemical populations of such systems have probability distributions that evolve according to a set of linear, time-invariant, ordinary differential equations known as the Chemical Master Equation (CME). The original FSP algorithm directly approximates the full CME solution to within a prespecified error. However, one may be interested only in certain portions of the distribution or certain statistical quantities such as mean or variance, and the full FSP method may provide an excess of information. In these cases, one can define a linear output signal and extract only the reachable and observable regions from the full distribution state space. The unobservable regions of the distribution can be aggregated with no accuracy loss but with less computational cost. This paper presents the resulting OAFSP algorithm and illustrates its benefits on a simple chemical reaction.

## I. INTRODUCTION

In many situations, especially in the case of cell fate decisions, a single gene-protein interaction may dramatically affect subsequent biological processes. Deterministic models fail to capture the inherent randomness of these system, and stochastic models are necessary [1]. The dynamics of a well-mixed, fixed volume, constant temperature chemically reacting system can be described by a Markovian process, whose probability distribution evolves according to the chemical master equation (CME) [2], [3]. In most cases the CME involves a very large, or infinite, number of linear time invariant (LTI) ordinary differential equations (ODEs), and has typically been assumed to be intractable to direct analysis [4]. Therefore, the dynamics of the CME have almost invariably been studied using Monte Carlo (MC) simulation algorithms such as Gillespie’s Stochastic Simulation Algorithm (SSA) [5], which simulates every individual reactive event in the chemical process. Although simple, the SSA can be extremely computationally expensive. As a result, various approximations have been considered to speed up the computational time of the SSA. These approximations come in two varieties: time leaping methods [6], [7], [8], [9] and system partitioning methods [10], [11], [12], [13].

For any MC method, computing a statistical description of system dynamics requires many individual realizations. Furthermore, even after a large number of MC realizations,

there is no clear guarantee as to how well the generated statistical information corresponds to the actual CME solution. To address this issue, we recently introduced the Finite State Projection (FSP) method and the FSP algorithm for approximating the CME solution [14]. Unlike MC analyses, the FSP directly computes the system’s probability distribution at a given time without requiring large numbers of realizations. Furthermore, the FSP approximation guarantees upper and lower bounds on the true CME solution. The FSP algorithm provides a systematic means of increasing the size of the finite state projection until these bounds are within any pre-specified error tolerance.

By approximating the CME as a finite dimensional linear system, the FSP opens the field of stochastic analysis to a myriad of common tools from the established fields of controls and dynamics. For example, several readily available tools facilitate lower order approximations of larger systems and promise significant reductions in computational cost. In this paper, we will illustrate how even the most basic elements of modern controls theory—minimal realizations and simple model reduction techniques—can significantly improve the FSP.

The remainder of this paper is organized into four sections. The next section briefly reviews the basic theory of the FSP. The third section applies elementary concepts related to reachability and observability to formulate the observability aggregated FSP algorithm (OAFSP). Next, a simple example will illustrate the usefulness of the new aggregate FSP algorithm. The final section concludes with a discussion of the advantages of applying controls understanding to chemical kinetics.

## II. THE FINITE STATE PROJECTION METHOD

Consider a well-mixed, fixed-temperature, and fixed volume system of  $N$  distinct reacting chemical species. At any time, one can use an integer vector  $\mathbf{x} := [\xi_1 \ \xi_2 \ \dots \ \xi_N]^T$  in the non-negative set  $\mathcal{N}^N$  to describe the discrete populations of the  $N$  molecular species. We will *a priori* fix a sequence of configurations  $\mathbf{x}_1, \mathbf{x}_2, \dots$  and define  $\mathbf{X} := [\mathbf{x}_1, \mathbf{x}_2, \dots]^T$  as the *ordered configuration set*. Suppose that there can be a maximum of  $M$  possible chemical reactions, where each  $\mu^{th}$  reaction takes the system from one configuration to another:  $\mathbf{x}_i \rightarrow \mathbf{x}_j = \mathbf{x}_i + \nu_\mu$ . For each reaction,  $\nu_\mu$  is the associated stoichiometry (directional transition on the configuration set), and  $a_\mu(\mathbf{x}_i)$  is the associated propensity function. Define  $p_i(t)$  as the probability that the system will be have the  $i^{th}$  configuration at time  $t$ . At an incrementally later time,  $p_i(t + dt)$  is equal to the sum

B. Munsky is with the Department of Mechanical and Environmental Engineering, University of California, Santa Barbara, CA 93106-5070 [brianem@engr.ucsb.edu](mailto:brianem@engr.ucsb.edu)

M. Khammash is with the Department of Mechanical and Environmental Engineering, University of California, Santa Barbara, CA 93106-5070 [khammash@engr.ucsb.edu](mailto:khammash@engr.ucsb.edu)

of (i) the probability that the system begins in  $\mathbf{x}_i$  at  $t$  and remains there until  $t + dt$ , and (ii) the probability that the system is in  $\mathbf{x}_j \neq \mathbf{x}_i$  at  $t$  and will transition to  $\mathbf{x}_i$  during the time step,  $dt$ . This probability can be written as:

$$p_i(t + dt) = p_i(t) \left( 1 - \sum_{\mu=1}^M a_{\mu}(\mathbf{x}_i) dt \right) + \sum_{\mu=1}^M p(\mathbf{x}_i - \nu_{\mu}, t) a_{\mu}(\mathbf{x}_i - \nu_{\mu}) dt. \quad (1)$$

From (1) it is easy to derive the differential equation known as the Chemical Master Equation, or CME [5] which we write in vector form as:

$$\dot{\mathbf{p}}_i(t) = \begin{bmatrix} -\sum_{\mu=1}^M a_{\mu}(\mathbf{x}_i) \\ a_1(\mathbf{x}_i - \nu_1) \\ a_2(\mathbf{x}_i - \nu_2) \\ \vdots \\ a_M(\mathbf{x}_i - \nu_M) \end{bmatrix}^T \begin{bmatrix} p(\mathbf{x}_i, t) \\ p(\mathbf{x}_i - \nu_1, t) \\ p(\mathbf{x}_i - \nu_2, t) \\ \vdots \\ p(\mathbf{x}_i - \nu_M, t) \end{bmatrix}. \quad (2)$$

Based upon our enumeration of  $\mathbf{X}$ , we can combine the equations for every  $p_i(t)$  into a single linear expression:

$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t), \quad (3)$$

where  $\mathbf{P}(t)$  is the probability density state vector at time  $t$ , and the generator matrix  $\mathbf{A}$  is the *state reaction matrix*. The elements of  $\mathbf{A}$  are defined by the reaction stoichiometries and propensities and our chosen enumeration of  $\mathbf{X}$ ;  $\mathbf{A}$  is independent of  $t$ ; all of its diagonal elements are non-positive; all its off-diagonal elements are non-negative; and all its columns sum to zero. The solution to the linear ODE beginning at  $t = 0$  and ending at  $t = t_f$  in (3) is the expression:  $\mathbf{P}(t_f) = \Phi(0, t_f)\mathbf{P}(0)$ . When  $\mathbf{X}$  is finite dimensional, the operator,  $\Phi(0, t_f) = \exp(\mathbf{A}t_f)$ , and the solution is simply:  $\mathbf{P}(t_f) = \exp(\mathbf{A}t_f)\mathbf{P}(0)$ . However, when  $\mathbf{X}$  is infinite dimensional or extremely large, the corresponding analytic solution is unclear or vastly difficult to compute. In these cases, one may devise a systematic means of approximating the full system using finite dimensional sub-systems. This systematic truncation approach is the Finite State Projection method [14].

We must introduce some convenient notation. Let  $J = \{j_1, j_2, j_3, \dots\}$  denote an index set. If  $\mathbf{X}$  is an enumerated set  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots\}$ , then  $\mathbf{X}_J$  denotes the subset  $\{\mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \mathbf{x}_{j_3}, \dots\}$ . Furthermore, let  $\mathbf{v}_J$  denote the subvector of  $\mathbf{v}$  whose elements are chosen according to  $J$ , and let  $\mathbf{A}_{IJ}$  denote the submatrix of  $\mathbf{A}$  such that the rows have been chosen according to  $I$  and the columns have been chosen according to  $J$ . For example, if  $I$  and  $J$  are defined as  $\{3, 1, 2\}$  and  $\{1, 3\}$ , respectively, then:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}_{IJ} = \begin{bmatrix} g & k \\ a & c \\ d & f \end{bmatrix}.$$

For convenience let  $\mathbf{A}_J := \mathbf{A}_{JJ}$ . For any two index sets  $I$  and  $J$ , let  $J \subseteq I$  denote that  $I$  contains every element from  $J$ , and let  $J \cup I$  and  $J \cap I$  be the union and intersection,

respectively, of  $J$  and  $I$ . Let  $J'$  denote the complement of the set  $J$ , where unless stated otherwise the complement is taken with respect to the set of all non-negative integers. In addition,  $\mathbf{1}$  will denote a column of ones: the operation  $\mathbf{1}^T \mathbf{v}$  sums each of the columns in  $\mathbf{v}$ . With this notation we can state the following theorem:

**Theorem 2.1.** If  $\mathbf{A} \in \mathcal{R}^{n \times n}$  has no negative off-diagonal elements, then for any index set,  $J$ ,

$$[\exp(\mathbf{A})]_J \geq \exp(\mathbf{A}_J) \geq \mathbf{0}. \quad (4)$$

Ref. [14] provides a detailed proof of this theorem. Since  $\mathbf{A}$  in (3) has no negative off-diagonal terms, if  $J_2 \supseteq J_1$ , Theorem 2.1 assures that:

$$[\exp(\mathbf{A}_{J_2 t_f})]_{J_1} \mathbf{P}_{J_1}(0) \geq \exp(\mathbf{A}_{J_1 t_f}) \mathbf{P}_{J_1}(0).$$

where  $\mathbf{P}_{J_1}$  is the probability distribution of the elements of  $\mathbf{X}$  indexed by  $J_1$ . This result guarantees that as one increases the finite configuration subset, the approximate solution increases monotonically for each element in the configuration subset.

In addition to being non-negative,  $\mathbf{P}(t)$  sums to exactly one. These properties and the nonnegativity of the off-diagonal elements of  $\mathbf{A}$  allow one to state a second theorem (a sketch of the proof is given below for Theorem 3.1. See Ref [14] for the full proof).

**Theorem 2.2.** Consider a Markov process in which the probability distribution evolves according to the linear ODE,  $\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t)$ , where  $\mathbf{A}$  has no negative off-diagonal entries. If for some finite index set  $J$ ,  $\varepsilon > 0$ , and  $t_f \geq 0$ ,

$$\mathbf{1}^T \exp(\mathbf{A}_J t_f) \mathbf{P}_J(0) \geq 1 - \varepsilon, \quad (5)$$

then

$$\exp(\mathbf{A}_J t_f) \mathbf{P}_J(0) \leq \mathbf{P}_J(t_f), \text{ and} \quad (6)$$

$$\|\mathbf{P}_J(t_f) - \exp(\mathbf{A}_J t_f) \mathbf{P}_J(0)\|_1 \leq \varepsilon. \quad (7)$$

While Theorem 2.1 guarantees that as we add points to the finite configuration subset, the approximate solution monotonically increases, Theorem 2.2 provides a certificate of how close the approximation is to the true solution. Together the two theorems suggest the FSP algorithm presented in Ref [14]. The following section will illustrate how basic modern controls theory may build upon these tools to improve the FSP algorithm.

### III. THE OBSERVABILITY AGGREGATED FSP

We begin with the CME written in the form (3). Suppose that at time  $t = 0$ , the initial pdv is supported only on the set indexed by  $U$ ; in other words  $p_i(0) = 0$  for all  $i \notin U$ . We write the initial value problem in (3) as the equivalent impulse response problem:

$$\dot{\mathbf{P}}(t) = \mathbf{A}\mathbf{P}(t) + \mathbf{b}\delta(t), \text{ where } \mathbf{b} = \mathbf{P}(0).$$

Suppose that we wish only to compute the statistical quantity  $\mathbf{y}(t) = \mathbf{C}\mathbf{P}(t)$ . As one example, if one were interested in estimating the mean or variance of the population of the  $m^{\text{th}}$  molecular species, then  $\mathbf{C}$  would simply be the row vector  $\mathbf{C}_{\text{mean}} = [\mathbf{x}_{1m}, \mathbf{x}_{2m}, \dots]$  or

$\mathbf{C}_{var} = [ \mathbf{x}_{1m}^2 - \mathbf{x}_{1m}, \mathbf{x}_{2m}^2 - \mathbf{x}_{2m}, \dots ]$ , respectively, where  $\mathbf{x}_{im}$  is the  $m^{\text{th}}$  component of the integer vector  $\mathbf{x}_i$ . Alternatively, as in the next subsection, we may choose the output to correspond to the probability density on a portion of the configuration set. For any  $\mathbf{C}$  the resulting problem now takes on a familiar form:

$$\begin{aligned} \dot{\mathbf{P}}(t) &= \mathbf{A}\mathbf{P}(t) + \mathbf{b}\delta(t); \\ \mathbf{y} &= \mathbf{C}\mathbf{P}(t). \end{aligned} \quad (8)$$

For systems on a finite configuration set, or for systems that have been projected onto a finite configuration set, this standard representation is open to a host of computational tools. In the following subsection we illustrate how one may use concepts of observability and reachability to improve upon the efficiency of the FSP.

#### A. Estimating the pdv of important states

Suppose that we begin with the known population vector,  $\mathbf{x}_u$ , and we want the probability distribution on the configuration subset  $\mathbf{X}_K = \{\mathbf{x}_{k_1}, \mathbf{x}_{k_2}, \dots\}$ . In other words, we wish to compute  $\mathbf{y}(t) = \mathbf{P}_K(t)$ . For example,  $\mathbf{X}_K$  may correspond to configurations that exhibit a specific biological trait, such as the expression of a certain gene. As above, define the vector  $\mathbf{b} = \mathbf{P}(0) = \{\mathbf{b}_i\}_{i=1}^\infty$ . In this case  $\mathbf{b}_i = 1$  for  $i = u$  and zero otherwise. For this  $\mathbf{b}$  and the impulse response in (8), let  $\mathbf{X}_R$  be the subset of all configuration points  $\mathbf{x}_i$  such that  $p_i(t) > 0$  at any  $t \geq 0$ . This subset is indexed by  $R$  to denote that it is the *reachable configuration subset*; its complement  $\mathbf{X}_{R'}$  is the *unreachable configuration subset*. Define the *observable configuration subset*,  $\mathbf{X}_O$ , as the set of all  $\mathbf{x}_i$  such that  $p_i(t_0) > 0$  at time  $t_0$  guarantees that  $|\mathbf{y}| > \mathbf{0}$  at some  $t \geq t_0$ . We will call the complement,  $\mathbf{X}_{O'}$ , the *unobservable configuration subset*. Note that our definitions of reachability and observability are slightly less restrictive than the traditional usage. While using the usual concepts of observability and reachability would often allow bigger reductions in the order of the problem, it is often much easier—and less computationally intensive—to categorize the system as shown here.

Now that the configuration set has been decomposed into subsets, we can introduce the following theorem:

**Theorem 3.1.** Consider a process whose distribution evolves according to the linear ODE:

$$\begin{bmatrix} \dot{\mathbf{P}}_{I_1} \\ \dot{\mathbf{P}}_{I_2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{I_1} & \mathbf{0} \\ \mathbf{A}_{I_2 I_1} & \mathbf{A}_{I_2} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{I_1} \\ \mathbf{P}_{I_2} \end{bmatrix} \quad (9)$$

where  $I_1$  and  $I_2$  are disjoint index sets.

If for some finite index set  $J \subseteq I_1$ ,  $\varepsilon > 0$ , and  $t_f \geq 0$ ,

$$\mathbf{1}^T \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ \mathbf{1}^T \mathbf{P}_{I_2}(0) \end{bmatrix} \geq 1 - \varepsilon, \quad (10)$$

then

$$\exp(\mathbf{A}_J t_f) \mathbf{P}_J(0) \leq \mathbf{P}_J(t_f), \quad \text{and} \quad (11)$$

$$\|\mathbf{P}_J(t_f) - \exp(\mathbf{A}_J t_f) \mathbf{P}_J(0)\|_1 \leq \varepsilon. \quad (12)$$

*Proof.* (For the reader's convenience, this proof includes an outline of the proof of Theorem 2.2—see also [14]. Note

that all probability distributions are nonnegative and sum exactly to one). We begin by proving (11). Let  $J'$  denote the complement of  $J$  on the set  $I_1$ . The evolution of the full probability density vector is governed by the permuted ODE:

$$\begin{bmatrix} \dot{\mathbf{P}}_J \\ \dot{\mathbf{P}}_{J'} \\ \dot{\mathbf{P}}_{I_2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_J & \mathbf{A}_{JJ'} & \mathbf{0} \\ \mathbf{A}_{J'J} & \mathbf{A}_{J'} & \mathbf{0} \\ \mathbf{A}_{I_2 J} & \mathbf{A}_{I_2 J'} & \mathbf{A}_{I_2} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J \\ \mathbf{P}_{J'} \\ \mathbf{P}_{I_2} \end{bmatrix}, \quad (13)$$

where the submatrices  $\mathbf{A}_{JJ'}$  and  $\mathbf{A}_{I_2 J'}$  are nonnegative since  $\mathbf{A}$  has no negative off-diagonal terms. We now sum all of the rows corresponding to the set  $I_2$ :

$$\begin{bmatrix} \dot{\mathbf{P}}_J \\ \dot{\mathbf{P}}_{J'} \\ \mathbf{1}^T \dot{\mathbf{P}}_{I_2} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_J & \mathbf{A}_{JJ'} & \mathbf{0} \\ \mathbf{A}_{J'J} & \mathbf{A}_{J'} & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} & \mathbf{1}^T \mathbf{A}_{I_2 J'} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J \\ \mathbf{P}_{J'} \\ \mathbf{P}_{I_2} \end{bmatrix}, \quad (14)$$

where we have used the fact that all columns of  $\mathbf{A}$ , particularly those indexed by  $I_2$ , sum to zero:  $\mathbf{1}^T \mathbf{A}_{I_2} = \mathbf{0}$ .

Let  $p_{agg} := \mathbf{1}^T \mathbf{P}_{I_2}$ . The aggregated probability density is now governed by the finite linear ODE:

$$\begin{bmatrix} \dot{\mathbf{P}}_J \\ \dot{p}_{agg} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_J & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J \\ \mathbf{P}_{I_2} \end{bmatrix} + \begin{bmatrix} \mathbf{A}_{JJ'} \\ \mathbf{1}^T \mathbf{A}_{I_2 J'} \end{bmatrix} \mathbf{P}_{J'}.$$

The solution of this forced ODE is

$$\begin{bmatrix} \mathbf{P}_J(t_f) \\ p_{agg}(t_f) \end{bmatrix} = \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ p_{agg}(0) \end{bmatrix} + \int_0^{t_f} \exp \begin{bmatrix} \mathbf{A}_J(t_f - \tau) & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J}(t_f - \tau) & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{JJ'} \\ \mathbf{1}^T \mathbf{A}_{I_2 J'} \end{bmatrix} \mathbf{P}_{J'}(\tau) d\tau.$$

Since  $\mathbf{A}_{JJ'}$ ,  $\mathbf{A}_{I_2 J'}$ ,  $\mathbf{P}_{J'}(t)$ , and  $\exp \begin{bmatrix} \mathbf{A}_J t & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t & \mathbf{0} \end{bmatrix}$  are all nonnegative for  $t \geq 0$ , we obtain the inequality in (11) as the top part of

$$\begin{bmatrix} \mathbf{P}_J(t_f) \\ p_{agg}(t_f) \end{bmatrix} \geq \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ p_{agg}(0) \end{bmatrix}. \quad (15)$$

Using (9) and the fact that the probability distribution on the  $J$  and  $I_2$ -indexed sets must be nonnegative and have a combined sum of no more than one we get:

$$\begin{aligned} & \left\| \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ \mathbf{1}^T \mathbf{P}_{I_2}(0) \end{bmatrix} \right\|_1 \\ & \geq \left\| \begin{bmatrix} \mathbf{P}_J(t_f) \\ p_{agg}(t_f) \end{bmatrix} \right\|_1 - \varepsilon, \end{aligned} \quad (16)$$

Finally, applying (15) and rearranging terms yields:

$$\left\| \begin{bmatrix} \mathbf{P}_J(t_f) \\ p_{agg}(t_f) \end{bmatrix} - \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{I_2 J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ \mathbf{1}^T \mathbf{P}_{I_2}(0) \end{bmatrix} \right\|_1 \leq \varepsilon, \quad (17)$$

and completes the proof.

By our definition of reachable, the probability density vector on the configuration subset  $\mathbf{X}_{R'}$  is zero, and a permutation can reorder the remaining rows of (3) as:

$$\begin{bmatrix} \dot{\mathbf{P}}_{RO} \\ \dot{\mathbf{P}}_{RO'} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{RO} & \mathbf{A}_{RO'RO'} \\ \mathbf{A}_{RO'RO} & \mathbf{A}_{RO'} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{RO} \\ \mathbf{P}_{RO'} \end{bmatrix}, \quad (18)$$

where  $RO := R \cap O$  indexes the reachable/observable configuration subset, and  $RO' := R \cap O'$  indexes the reachable/unobservable configuration subset. Also by definition, no configuration in  $\mathbf{X}'_O$  can transition into the configuration subset  $\mathbf{X}_O$ , which results in the identity:  $\mathbf{A}_{RORO'} = \mathbf{0}$ , and the system reduces to:

$$\begin{bmatrix} \dot{\mathbf{P}}_{RO} \\ \dot{\mathbf{P}}_{RO'} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{RO} & \mathbf{0} \\ \mathbf{A}_{RO'RO} & \mathbf{A}_{RO'} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{RO} \\ \mathbf{P}_{RO'} \end{bmatrix}. \quad (19)$$

Applying Theorem 3.1 yields the following corollary:

**Corollary 3.2.** Consider any Markov process in which the probability density state vector evolves according to (19). Let  $J$  be a finite subset of the index set  $RO$ . If for  $\varepsilon > 0$ , and  $t_f \geq 0$

$$\mathbf{1}^T \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{RO'J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_J(0) \\ \mathbf{1}^T \mathbf{P}_{RO'}(0) \end{bmatrix} \geq 1 - \varepsilon, \quad (20)$$

then

$$\exp(\mathbf{A}_J t_f) \mathbf{P}_J(0) \leq \mathbf{P}_J(t_f), \quad \text{and} \quad (21)$$

$$\|\mathbf{P}_J(t_f) - \exp(\mathbf{A}_J t_f) \mathbf{P}_J(0)\|_1 \leq \varepsilon. \quad (22)$$

The proof of Corollary 3.2 follows directly from Theorem 3.1 where  $I_1 = RO$  and  $I_2 = RO'$ . To illustrate the underlying intuition of Corollary 3.2, Fig. 1(top) illustrates a two dimensional state lattice for a two chemical reacting system. The system begins with an initial configuration  $\mathbf{x}_u$  at time  $t = 0$ , and we are interested in calculating the probability that the system has configuration,  $\mathbf{x}_y$ , at the time  $t = t_f \geq 0$ . The configuration set can be separated into three disjoint subsets: the unreachable region,  $\mathbf{X}_{R'}$ ; the unobservable region,  $\mathbf{X}_{O'}$ ; and the reachable/observable region  $\mathbf{X}_{RO}$ . Using the OAFSP, we remove the  $\mathbf{X}_{R'}$  from the system and aggregate  $\mathbf{X}_{O'}$  to a single point, as shown in Fig. 1(bottom left). We then project  $\mathbf{X}_{RO}$  onto a finite configuration subset  $\mathbf{X}_J$ . The projected system is shown in Fig. 1(bottom right), where the subsets  $\mathbf{X}_{J'}$  and  $\mathbf{X}_{RO'}$  have each been aggregated to a single point. Because the projected system is finite dimensional, its solution can be computed using the matrix exponential function or by using a standard ODE solver. Theorem 2.1 shows that as the subset  $\mathbf{X}_J$  increases, fewer trajectories are lost to  $\mathbf{X}_{J'}$  and the probability of remaining in  $\mathbf{X}_J \cup \mathbf{X}_{RO'}$  increases. Corollary 3.2 shows that the probability that the system is currently in  $\mathbf{X}_J \cup \mathbf{X}_{RO'}$  must be at least as large as the probability that the system has been in  $\mathbf{X}_J \cup \mathbf{X}_{RO'}$  for all times  $t = 0$  to  $t = t_f$ .

### B. The OAFSP Algorithm

The results above and our previous work on the FSP [14] suggest a systematic procedure for solving the chemical kinetic problem as posed in (8). This algorithm, which we refer to as the Observability Aggregated FSP algorithm, can be stated as follows:

**Step 0** Define reaction propensities and stoichiometry.

Choose the initial pdv,  $\mathbf{P}(0)$ .

Choose the final time of interest,  $t_f$ .

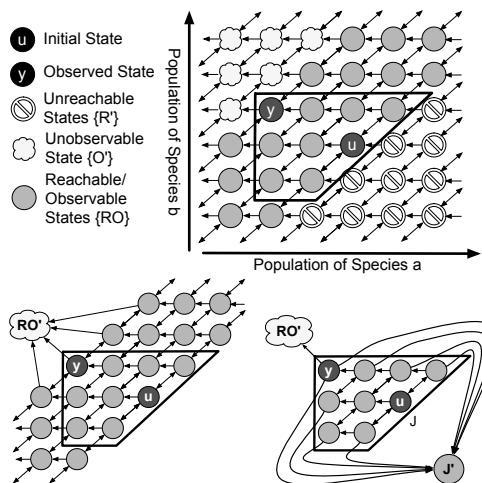


Fig. 1. Top: schematic of a two dimensional integer lattice representing the configuration set of a two species chemical reaction. Each configuration point  $[a,b]$  is represented by a circle and transitions (reactions) are shown by the connecting arrows. Bottom: aggregation of the unobservable configuration subset (left), and projection of the observable/reachable configuration subset onto a finite configuration subset:  $\mathbf{X}_J \in \mathbf{X}_{RO}$  (right).

Specify the total acceptable error,  $\varepsilon > 0$ .

Define configuration subsets:  $\mathbf{X}_{RO}$  and  $\mathbf{X}_{RO'}$ .

Choose initial finite index set,  $J_o \subseteq RO$ .

Initialize a counter,  $i = 0$ .

**Step 1** Use propensities and stoichiometry to compute

$$\Gamma_{J_i} = \mathbf{1}^T \exp \begin{bmatrix} \mathbf{A}_J t_f & \mathbf{0} \\ \mathbf{1}^T \mathbf{A}_{RO'J} t_f & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{J_i}(0) \\ \mathbf{1}^T \mathbf{P}_{RO'}(0) \end{bmatrix}.$$

**Step 2** If  $\Gamma_{J_i} \geq 1 - \varepsilon$ , **Stop**.

$\exp(\mathbf{A}_{J_i} t_f) \mathbf{P}_{J_i}(0)$  is within  $\varepsilon \mathbf{1}$  error from  $\mathbf{P}_{J_i}(t_f)$ .

**Step 3** Add more configurations to find  $\mathbf{X}_{J_{i+1}}$ .

Increment  $i$  and return to **Step 1**.

In the following section, we will illustrate a simple example, where we apply the OAFSP algorithm.

## IV. APPLYING THE OAFSP

Here we consider an example chemical reaction involving three molecular species,  $a$ ,  $b$  and  $c$ , the configuration set is the set of non-negative integer vectors:  $\mathbf{x}_i = [a_i \ b_i \ c_i]^T$ . At time  $t = 0$ , the system begins at a known configuration point:

$$\mathbf{x}_u^T = [a_u, b_u, c_u] = [a, b, c]_{t=0} = [70, 50, 80],$$

and the molecular species interact via the chemical reactions:  $(a + b \rightleftharpoons c + c)$  and  $(a \rightarrow c)$ . Since none of these reactions change the total number of molecules in the system, for a given initial condition, it is sufficient to know the populations  $a_i$  and  $b_i$  in order to uniquely determine  $\mathbf{x}_i$ , and the entire configuration set can be represented on a 2-dimensional integer lattice. Fig. 1 illustrates such a lattice where the initial condition is shaded in dark gray, and the directionality of the reactions are as indicated. From the stoichiometry, one can also show that all points in the reachable configuration

TABLE I

COMPARISON OF THE COMPUTATIONAL EFFORT REQUIRED FOR THE OAFSP, THE FSP AND THE SSA FOR THREE DIFFERENT SIMULATION TIME INTERVALS. ALL SIMULATIONS HAVE BEEN CONDUCTED ON A 3.0 GHZ INTEL PENTIUM 4 PROCESSOR RUNNING A LINUX ENVIRONMENT. FOR THE SSA, WE HAVE USED THE STOCHKIT SOFTWARE PACKAGE [15], FOR THE FSP AND THE OAFSP, THE CALCULATIONS ARE PERFORMED IN USING MATHWORKS MATLAB.

Final time, $t_f$	Computational Time (s)		
	OAFSP	FSP	$10^4$ SSA runs. <sup>a</sup>
10s	44	41	460 (460)
20s	89	121	900 (870)
30s	108	331	1400 (1100)

<sup>a</sup>Second value is for simulations that are stopped as soon as  $[a] \leq [b]$

subset,  $\mathbf{x}_i \in \mathbf{X}_R$ , satisfy the inequalities:  $a_i \leq a_u + b_i - b_u$  and  $a_i + b_i \leq a_u + b_u + c_u$ , where the subscript  $u$  refers to the initial configuration as defined above. The rest of the configuration set may be immediately excluded in the FSP description. The dimension of the  $\mathbf{X}_R$  is about  $1.2 \times 10^4$ .

Suppose that we are interested in computing the probability density at time  $t_f$  of all of the configurations for which the population of  $b$  is less than that of  $a$ . Let  $K$  be the subset of all indices  $k$  such that  $b_k < a_k$ , then the output is  $\mathbf{P}_K(t)$ . For this output and the system’s stoichiometry, one can show that the observable subset and the output subset are identical:  $O = K$ . While the dimension of  $\mathbf{X}_O$  is infinite, the intersection of  $\mathbf{X}_R$  and  $\mathbf{X}_O$  includes fewer than  $2 \times 10^3$  states. Thus, by aggregating the unobservable/reachable subset, one is able to reduce the order of the problem by a factor of six before making any approximation.

Once the observability aggregation has been applied, the remaining problem may be analyzed with the FSP and OAFSP algorithms; in each case the total error tolerance has been set to  $\varepsilon = 10^{-3}$ . Fig. 2 illustrates the joint probability distribution of the species  $a$  and  $b$  at three different times—each plot is exactly the same whether one uses the original FSP algorithm or the OAFSP algorithm. Table I provides the computational time required to achieve these results with the FSP and OAFSP algorithms. For an additional comparison, Table I also provides the time required to simulate the system  $10^4$  times using the basic SSA as well as a version of the SSA in which the simulation is interrupted as soon as the system leaves the observable space. In the three cases considered here, the FSP and OAFSP are both much faster than the SSA. Furthermore, for the situation in which a significant portion of the probability distribution lies outside the observable space (Fig. 2, right), the OAFSP method dramatically outperforms the original FSP algorithm (Table I, bottom row).

## V. CONCLUSIONS

This paper introduced the Observability Aggregated Finite State Projection (OAFSP) method and algorithm for the solution of the chemical master equation. While the original FSP

directly computes the system’s full probability distribution, the OAFSP concentrates only on the reachable/observable portion of the configuration set. By intelligently aggregating unimportant regions of the configuration set, the order of the CME can be significantly reduced. Like the FSP, the OAFSP solution guarantees upper and lower bounds on the solution of the true system and provides a systematic means of increasing the size of the finite state projection until these bounds are within any pre-specified error tolerance. The OAFSP method was effectively demonstrated on an example problem. Given a known initial state, we have used the OAFSP to find the desired portion of the distribution at a later time of interest. We have compared the efficiency of the OAFSP with the original FSP and the SSA for three different simulation time intervals. In this example, we have shown that the OAFSP and FSP algorithms outperform the SSA. Further, we have shown that the aggregation can significantly decrease computational cost with no loss in accuracy.

While the FSP is still not feasible for all classes of chemical systems, this study has shown one example of how readily available tools may facilitate lower order approximations of the CME and significantly reduce computational cost. Additional state aggregation, and multi-scale partitioning methods will also provide further benefit over the original FSP and are currently being developed to be presented elsewhere. Like the OAFSP, these major gains in the performance of the FSP originate in the well-developed theory of controls and dynamics. In this manner, new implementations will continue to expand the class of problems for which the FSP and its derivatives are efficient and versatile tools for stochastic analysis.

## VI. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant NSF-ITR CCF-0326576 and the Institute for Collaborative Biotechnologies through Grant DAAD19-03-D-0004 from the U.S. Army Research Office.

## REFERENCES

- [1] M. Khammash and H. El Samad, “Stochastic modeling and analysis of genetic networks,” *Proc. 45th IEEE CDC*, Dec. 2006.
- [2] D. McQuarrie, “Stochastic approach to chemical kinetics,” *J. Applied Probability*, vol. 4, pp. 413–478, 1967.
- [3] D. T. Gillespie, “A rigorous derivation of the chemical master equation,” *Physica A*, vol. 188, pp. 404–425, 1992.
- [4] H. El Samad, M. Khammash, L. Petzold, and D. Gillespie, “Stochastic modeling of gene regulatory networks,” *Int. J. Robust Nonlin.*, vol. 15, pp. 691–711, 2005.
- [5] D. T. Gillespie, “Exact stochastic simulation of coupled chemical reactions,” *J. Phys. Chem.*, vol. 81, no. 25, pp. 2340–2360, May 1977.
- [6] —, “Approximate accelerated stochastic simulation of chemically reacting systems,” *J. Chem. Phys.*, vol. 115, no. 4, pp. 1716–1733, Jul. 2001.
- [7] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, “Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method,” *J. Chem. Phys.*, vol. 119, no. 24, pp. 12 784–12 794, Dec. 2003.

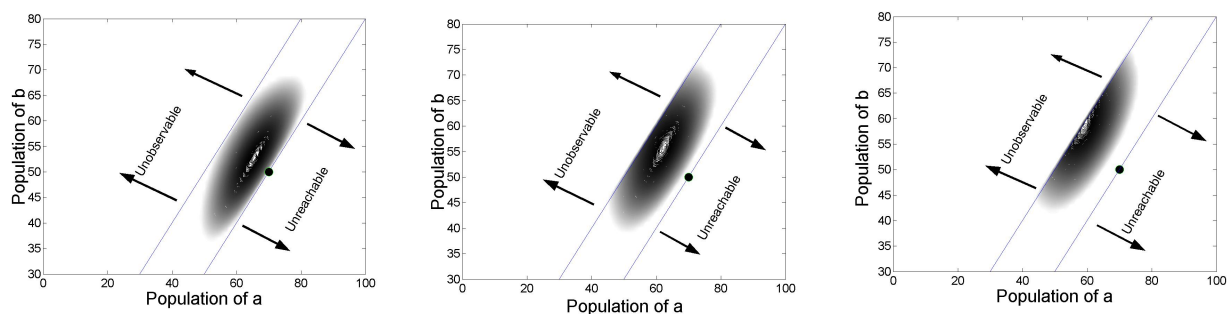


Fig. 2. The probability density on the reachable and observable portion of the configuration space at three different times:  $t=10s$  (left),  $t=20s$  (middle), and  $t=30s$  (right). In each plot, the initial condition is shown by a black circle. The contrast of the probability density contour lines is defined using a logarithmic scale where white corresponds to a probability density less than  $10^{-6}$ , and black corresponds to a density of  $10^{-1}$ .

- [8] D. T. Gillespie and L. R. Petzold, "Improved leap-size selection for accelerated stochastic simulation," *J. Chem. Phys.*, vol. 119, no. 16, pp. 8229–8234, Oct. 2003.
- [9] T. Tian and K. Burrage, "Binomial leap methods for simulating stochastic chemical kinetics," *J. Chem. Phys.*, vol. 121, no. 21, pp. 10 356–10 364, Dec. 2004.
- [10] C. V. Rao and A. P. Arkin, "Stochastic chemical kinetics and the quasi-steady-state assumption: Application to the gillespie algorithm," *J. Chem. Phys.*, vol. 118, no. 11, pp. 4999–5010, Mar. 2003.
- [11] E. Haseltine and J. Rawlings, "Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics," *J. Chem. Phys.*, vol. 117, no. 15, pp. 6959–6969, Jul. 2002.
- [12] Y. Cao, D. Gillespie, and L. Petzold, "The slow-scale stochastic simulation algorithm," *J. Chem. Phys.*, vol. 122, no. 014116, Jan. 2005.
- [13] H. Salis and Y. Kaznessis, "Accurate hybrid stochastic simulation of a system of coupled chemical or biological reactions," *J. Chem. Phys.*, vol. 112, no. 054103, 2005.
- [14] B. Munsky and M. Khammash, "The finite state projection algorithm for the solution of the chemical master equation," *J. Chem. Phys.*, vol. 124, no. 044104, 2006.
- [15] L. Petzold and coworkers, "StochKit Beta Version," Nov. 2004, downloaded from the Internet at <http://www.engineering.ucsb.edu/cse/StochKit/index.html>.