

**Fock space, symbolic algebra, and analytical solutions for small stochastic systems**

Fernando A. N. Santos

*Departamento de Matemática Universidade Federal de Pernambuco, 50670-901, Recife, PE, Brazil  
and Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Woodstock Road,  
Oxford OX2 6GG, United Kingdom*

Hermes Gadêlha

*Department of Mathematics, University of York, York YO10 5DD, United Kingdom  
and Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Woodstock Road,  
Oxford OX2 6GG, United Kingdom*

Eamonn A. Gaffney

*Wolfson Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Woodstock Road, Oxford OX2 6GG, United Kingdom*

(Received 21 September 2015; published 28 December 2015)

Randomness is ubiquitous in nature. From single-molecule biochemical reactions to macroscale biological systems, stochasticity permeates individual interactions and often regulates emergent properties of the system. While such systems are regularly studied from a modeling viewpoint using stochastic simulation algorithms, numerous potential analytical tools can be inherited from statistical and quantum physics, replacing randomness due to quantum fluctuations with low-copy-number stochasticity. Nevertheless, classical studies remained limited to the abstract level, demonstrating a more general applicability and equivalence between systems in physics and biology rather than exploiting the physics tools to study biological systems. Here the Fock space representation, used in quantum mechanics, is combined with the symbolic algebra of creation and annihilation operators to consider explicit solutions for the chemical master equations describing small, well-mixed, biochemical, or biological systems. This is illustrated with an exact solution for a Michaelis-Menten single enzyme interacting with limited substrate, including a consideration of very short time scales, which emphasizes when stiffness is present even for small copy numbers. Furthermore, we present a general matrix representation for Michaelis-Menten kinetics with an arbitrary number of enzymes and substrates that, following diagonalization, leads to the solution of this ubiquitous, nonlinear enzyme kinetics problem. For this, a flexible symbolic MAPLE code is provided, demonstrating the prospective advantages of this framework compared to stochastic simulation algorithms. This further highlights the possibilities for analytically based studies of stochastic systems in biology and chemistry using tools from theoretical quantum physics.

DOI: [10.1103/PhysRevE.92.062714](https://doi.org/10.1103/PhysRevE.92.062714)

PACS number(s): 87.10.Mn, 82.20.-w, 87.14.ej

**I. INTRODUCTION**

There is no doubt that the mathematical tools developed to study quantum systems can be applied to other areas of science characterized by stochasticity, even if quantum fluctuations are not influential. While the textbook example is the relation between quantum path integrals and the partition functions of statistical mechanics, there is enormous potential for quantum tools to be exploited in the study of stochasticity in biological and chemical systems. The latter has seen a tremendous surge of interest in recent years, with the method of choice based upon Gillespie's stochastic simulation algorithm and its variants, and to a lesser extent, and for larger systems, the Fokker-Planck equation approximation [1]. In contrast, however, numerous tools from quantum physics are analytical in nature even far from the large-particle number limit.

Hence there is extensive scope for developing our understanding of biological and chemical stochasticity via the exploitation of the mathematical formulations used in the study of quantum physics. This has not escaped attention, though many studies develop the theory, demonstrating that it is universally applicable in principle, rather than applying it in practice.

For instance, during the 1970s it was demonstrated that a master equation describing classical stochastic systems can be

mapped onto the Schrodinger equation, where the probability distribution corresponds to the wave function [2]. In particular, this mapping demonstrated that quantum mechanical tools can be used to investigate systems where species can be created or destroyed, as is the case of chemical reactions [2,3] and birth-death processes on a lattice [4]. This connection between classical systems and quantum mechanical tools was well known to the theoretical physics community at the time [5] and developed further. For instance, a mapping was found between the master equation describing nonequilibrium one-dimensional problems and one-dimensional quantum spin chains [6,7]. In turn this enabled the derivation of exact solutions of several simple reaction models [6], utilizing the fact there are several integrable one-dimensional quantum chains that are exactly understood [8]. However, as was pointed out by Grassberger, Alcaraz, and other physicists [3,6], new approaches allowing the explicit computation of quantities, such as correlation functions, are desirable.

Nonetheless, the exploitation of techniques from physics for problems directly motivated from other sciences, rather than mapping exact solutions that have been found in quantum systems, has been very limited. That said, notable examples have been published such as the application of renormalization group procedures to studies of morphogen gradient formation

[9] and the application of many-body theory in the analysis of stochastic gene expression, stochastic models in population biology, and stem-cell differentiation [10–12]. For instance, a mapping between a spin-boson system and a model of gene expression has been constructed [10], which can be analyzed via a variational approach analogous to that used for nonequilibrium statistical dynamics [13,14]. This approach in turn provides a spectral method for the study of stochastic systems [10,14,15], which is numerically efficient and yields a good approximation for steady-state solutions, while avoiding cumbersome noncommutative operations inherent to the quantum-field-theoretical mapping [2]. Here, in contrast, we make use of symbolic algebra to circumvent the latter issue and look for exact solutions for small stochastic systems.

In particular, we consider the Fock space of quantum mechanics and show how its representation allows the study of small biological or chemical stochastic systems. Effectively, we demonstrate how the quantum Fock space can provide a suitable representation for the study and solution of the chemical master equation [1], which is the fundamental governing equation across a wide of array of studies in stochastic systems. The basic concept of solving the system of linear ordinary differential equations associated with the chemical master equation has been suggested previously in the context of the finite state projection algorithm [16]. However, the methodology here provides a means of automatically generating the matrix representation of the time evolution operator in Fock space associated with these ordinary differential equations, as provided within the MAPLE code in the Supplemental Material [17], as well as numerous solution techniques. In particular, this representation may provide computational advantages over the analytically equivalent generating function method [14] since the latter typically requires one to represent exponentials in terms of their power series in symbolic computations. Furthermore, the method presented below allows us to automate the route from the chemical master equation to its direct solution for small reactant number stochastic law of mass action systems, as well as exploiting tools designed for quantum and statistical physics in exploring these systems.

In addition, the subject of our study, small stochastic systems are highly nontrivial in general. For instance, they can have relatively large parameter spaces, leading to difficulties of inference, parameter estimation and model identifiability [18], and hence alternative techniques to Gillespie's stochastic simulation algorithm, which do not require the simulation of ensembles across high-dimensional parameter space, are likely to find generic application. In addition, ensemble averaging entails that the modeling of stochastic systems where rare events are important to final outcomes, for instance, extinctions in ecology or cell mutations in carcinogenesis or drug resistance, is particularly difficult to study with stochastic simulation algorithms, in contrast to the techniques presented here.

To demonstrate the practical application of the Fock space framework for stochastic law-of-mass-action systems, we consider a small copy number example of arguably the most famous exemplar of a biochemical system, Michaelis-Menten kinetics, which is a paradigm for enzyme-catalyzed biochemistry and has been studied for over a century, as reviewed for single-molecule enzymology by Grima *et al.* [19].

We focus on the regime of both low numbers of enzymes and substrates, noting that systems with an infinite reservoir of substrate are well classified due to the simplicity of not being required to track substrate levels [19,20]. The case with one enzyme and small numbers of substrate has been mathematically analyzed by Aranyi and Toth [29], who produced analytical results in terms of summations of marginal generating functions. Here we parenthetically recover a simpler representation of Aranyi and Toth's presented solution for one enzyme and one substrate, as well as analytical solutions for more than one substrate.

However, the real power and purpose of the presented framework lies in its flexibility and its physical interpretation. In particular, given the chemical master equation, the combination of symbolic algebra and Fock space techniques allows the rapid generation of a matrix representation for the time evolution operator of the system, as illustrated in the Supplemental Material [17], which makes extensive use of Fock space manipulations. In turn, this enables either an analytical solution via a single symbolic calculation, as with the single-enzyme stochastic Michaelis-Menten system, or simple solution schemes utilizing numerical linear algebra. In particular, and in contrast to stochastic simulations, ensembles are not required to generate variable means, standard deviations, or higher moments for the study of stochastic effects. However, the increase in combinatorial complexity of the Fock space dimension with system size implies the solution technique presented here is most appropriate for small systems, where large-system-size simplifications are not valid. Hence, the methodologies presented here can be productively applied to any small stochastic system of chemical, biological, or biochemical interactions utilizing either analytical methods or personal computing. Prospective application examples include simple stochastic gene expression models [10,21], especially in the context of single-molecular techniques and data [19,22], and thus there is a generality that extends far beyond enzyme kinetics.

In addition, continuum-scale Michaelis-Menten kinetics are known to be extremely stiff, with a dramatic ratio of time scales that becomes apparent in the deterministic limit via the initial transient boundary layer, which is characterized in terms of the ratio of substrate to enzyme molecules in textbook studies [23]. Hence we explore whether stiffness is present in small Michaelis-Menten systems, despite the fact there is no generic analog of a large substrate-to-enzyme ratio for small systems. Furthermore, stiffness is well known as a source of difficulty in stochastic simulations [24], and thus this particular investigation also highlights how the presented techniques can be applied to studies that exacerbate weaknesses in stochastic simulation algorithms. More generally, this study additionally highlights that there is still enormous scope for a large-scale objective of further developing the tools of quantum physics to the extent that exploitation in biological and chemical systems can be readily implemented.

## II. THE FOCK SPACE APPROACH TO MASS ACTION SYSTEMS

In this section we demonstrate how the Fock-space framework for quantum physics can be combined with symbolic and linear algebra tools for the study of small stochastic systems

characterized by the law of mass action, which thus finds application in the modeling of biological, biochemical, and chemical systems. In particular, a typical law-of-mass-action reaction with  $N$  reactants,  $R_i$ , producing  $M$  products,  $P_i$ , at a rate  $k_{\pm}$ , can be described via

$$\sum_{i=1}^N r_i R_i(t) \xrightleftharpoons[k_-]{k_+} \sum_{j=1}^M p_j P_j(t), \quad (1)$$

where  $r_i, p_j$  are stoichiometric coefficients. Furthermore, the probability of finding the system in a given state

$$\mathbf{N}(t) = (R_1(t), \dots, R_N(t), P_1, \dots, P_M(t))$$

at time  $t$  is denoted by  $\mathcal{P}(\mathbf{N}, t)$ . By simple balances and the inherent Markovian nature of law-of-mass-action reactions, the associated (chemical) master equation is given by [25]

$$\frac{\partial \mathcal{P}(\mathbf{N}, t)}{\partial t} = \sum_{\mathbf{N}'} [T_{\mathbf{N}' \rightarrow \mathbf{N}} \mathcal{P}(\mathbf{N}', t) - T_{\mathbf{N} \rightarrow \mathbf{N}'} \mathcal{P}(\mathbf{N}, t)], \quad (2)$$

where the transition rates,  $T_{\mathbf{N}' \rightarrow \mathbf{N}}$ , between the configurations  $\mathbf{N}'$  and  $\mathbf{N}$  are independent of the time.

To start exploiting Fock space tools, a given configuration state of the system is represented in the Fock space by a direct product of the Hilbert space  $\mathcal{S}_i$  of each species  $k$  in the chemical reaction. That is, with  $\mathcal{S}_i = \{1, \dots, N_i\}$ , where  $N_i$  is the maximal number of the reactant  $R_i$  (or the product  $P_i$ ), we have  $\mathcal{F} = \mathcal{S}_1 \otimes \dots \otimes \mathcal{S}_k$ . Hence with  $s_i \in \mathcal{S}_i$ , a state with  $s_i$  molecules from species  $i \in \{1, \dots, k\}$  is an element of the Fock space and can be represented in Dirac's bra-ket notation via  $|n\rangle = |s_1 \dots s_k\rangle$  and is referred to as a pure Fock state, with the set of all pure Fock states providing a basis for the Fock space. With the probability of being in a state  $\mathbf{N}$  at time  $t$ , from the master equation,  $\mathcal{P}(\mathbf{N}, t)$ , rewritten in the new notation as  $P(n, t)$ , the stochastic system time  $t$  can be fully characterized by  $|\Psi(t)\rangle$ , which is defined via the linear summation of pure Fock states,

$$|\Psi(t)\rangle := \sum_n P(n; t) |n\rangle. \quad (3)$$

As an aside note that in Ref. [14] this wave function maps onto a generating function representation via

$$|\Psi(t)\rangle = \sum_n P(n; t) |n\rangle \Leftrightarrow |G(t)\rangle = \sum_n P(n; t) x^n, \quad (4)$$

where  $x^n$  is a shorthand for a multinomial  $x_1^{s_1} x_2^{s_2} \dots x_k^{s_k}$ , though we pursue a route that allows convenient symbolic algebra. Hence working with Eq. (3), we introduce the creation and annihilation operators for each species, which act on the pure Fock states via

$$\begin{aligned} \alpha_i^\dagger |n\rangle &= |s_1 \dots (s_i + 1) \dots s_k\rangle \\ \alpha_i |n\rangle &= s_i |s_1 \dots (s_i - 1) \dots s_k\rangle, \end{aligned} \quad (5)$$

with linearity determining how  $\alpha_i^\dagger$ , and  $\alpha_i$  operate on a general element of the Fock space. It is straightforward to confirm that the commutation rule  $[\alpha_i, \alpha_j^\dagger] = \delta_{ij}$  is satisfied, whence the master equation can be recast as a Schrödinger equation, with  $i\hbar = 1$ ,

$$\frac{\partial |\Psi(t)\rangle}{\partial t} = -H(\alpha_1^\dagger, \alpha_1, \dots, \alpha_k^\dagger, \alpha_k) |\Psi(t)\rangle. \quad (6)$$

The resulting solution is explicitly given in terms of the quasi-Hamiltonian  $H$ ,

$$|\Psi(t)\rangle = \exp(-H(\alpha_1^\dagger, \alpha_1, \dots, \alpha_k^\dagger, \alpha_k)t) |\Psi(0)\rangle. \quad (7)$$

With  $\pi_i^\dagger, \rho_i^\dagger, \pi_i, \rho_i$  respectively denoting the creation and annihilation operators for the species  $P_i$  and  $R_i$  for the individual reaction specified by relation (1), the contribution to the quasi-Hamiltonian,  $H$ , from the forward reaction can be written as [3,5,26]

$$-k_+ \left[ \prod_{j=1}^M (\pi_j^\dagger)^{p_j} - \prod_{i=1}^N (\rho_i^\dagger)^{r_i} \right] \prod_{i=1}^N \rho_i^{r_i}, \quad (8)$$

with analogous additive contributions from the back-reaction and all other reactions. In particular, the first term above corresponds to an annihilation operator for each reactant and a creation operator for each product in the reaction, while the second one is an annihilation and creation operator for each reactant. A fundamental part of this formalism relies on the calculation of a matrix representation of  $H$  using a suitable representation for the basis from the associated Fock space. In general, this is implemented by setting an upper bound  $N_{\max}(s_i)$  for each species and, via symbolic computation, generating all possible  $\prod_{i=1}^k [N_{\max}(s_i) + 1]$  pure Fock states  $|i\rangle = |s_1 \dots s_k\rangle$ , with  $s_i \in \{0, \dots, N_{\max}(s_i)\}$ , so a matrix element of  $H$  is given by  $H_{ij} = \langle i | H | j \rangle$ .

This procedure is general, as well as computationally feasible, for small systems. Moreover, as detailed below, one can take advantage of intrinsic symmetries, such as conservation of number of particles in some species, to reduce the dimensionality of the Fock space using a Noether theorem for Markov processes [26]. Once a matrix representation of  $H$  is found, the calculation of  $\exp(-Ht)$  can be achieved using a variety of available methods [27], thus allowing the calculation of the Fock-space state, and thus the state of the system, at any particular time given the initial state of the system. In summary, the above framework is general and provides solutions of any law of mass action stochastic system via three simple steps: construction of the quasi-Hamiltonian using the Fock space representation, followed by diagonalization, and, finally, projection of initial data to later-time solutions.

In particular, by taking the Laplace transform, defined via  $\mathcal{L}[f(t)] := \int_0^\infty f(t) e^{-pt} dt$ , we obtain

$$\mathcal{L}[|\Psi(t)\rangle] = (p I_n - H)^{-1} |\Psi(0)\rangle. \quad (9)$$

Assuming a representation of the Fock space, and thus the quasi-Hamiltonian  $H$ , in terms of matrices,  $|\Psi(t)\rangle$  can also be written in terms of the inverse Laplace transform

$$|\Psi(t)\rangle = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} dp e^{pt} \frac{\text{adj}(p I_n - H) |\Psi(0)\rangle}{\det(p I_n - H)}. \quad (10)$$

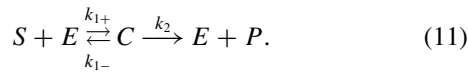
where  $\text{adj}(M)$  denotes the adjugate of the matrix  $M$  and this integral can be evaluated by a sum of residues indexed by the eigenvalues of  $H$ . Alternatively, the Jordan normal form of the quasi-Hamiltonian can often be computed instead,  $J_H = Q^{-1} H Q$ , so  $|\Psi(t)\rangle = Q \exp(-J_H t) Q^{-1} |\Psi(0)\rangle$ , or one, more generally, can numerically compute the exponential  $\exp(-Ht)$  at required time points. Furthermore, the sparsity of the matrix representing  $H$  in typical cases entails that the numerical linear

algebra described above in solving these systems does not require as many operations, or as much memory, as for a general densely populated matrix of the same size.

In the following section, we apply the general framework developed here to a biochemical Michaelis-Menten enzymatic system. First, we will analytically diagonalize the system for one enzyme and an arbitrary number of substrates. Further, for an arbitrary number of enzymes and substrates, we will find a matrix representation of the system which, consequently, gives us all its information apart from computing  $\exp(-Ht)$ . Furthermore, the necessary MAPLE code to diagonalize the system and compute the matrix exponential is available in the Supplemental Material [17]. In particular, we will explore stiffness in the low-copy-number Michaelis-Menten system, as well as highlighting an exploration of parameter space, all of which are considered without ensembles.

### III. APPLICATION TO MICHAELIS-MENTEN ENZYME KINETICS

The Michaelis-Menten model consists of an enzyme,  $E$ , binding to a substrate  $S$  to form a complex,  $C$ , which is then converted into a product,  $P$ , while returning the enzyme for further catalysis, as summarized by the reactions:



Furthermore, the most common initial conditions for Michaelis-Menten systems are deterministic with no product or complex and with the initial number of substrates,  $N_S$ , and enzymes,  $N_E$ , specified with unit probability [28].

The quasi-Hamiltonian, written in terms of creation and annihilation operators as in Eq. (5), is given by

$$H = -k_{1+}(e^\dagger s^\dagger - c^\dagger)es - k_{1-}(s^\dagger e^\dagger - c^\dagger)c - k_2(e^\dagger p^\dagger - c^\dagger)c. \quad (12)$$

Furthermore, the ket  $|s_i e_i c_i p_i\rangle$  represents a state with  $s_i$  substrates,  $e_i$  enzymes,  $c_i$  complexes, and  $p_i$  products. Given that the stochastic process is considered in the context of a second-quantized notation, that is, in terms of annihilation and creation operators as specified above, a Fock space basis can be symbolically generated, as exemplified by the MAPLE code included in the Supplemental Material [17]. In particular, this simply requires looping through all possible kets of the system, while observing any symmetries, i.e., conserved quantities. This basis subsequently enables a matrix representation of operators on the Fock space associated with the reaction scheme, in particular the quasi-Hamiltonian,  $H$ .

In detail, for the Michaelis-Menten system, a basis that spans the reaction scheme Fock space  $\mathcal{S} \otimes \mathcal{E} \otimes \mathcal{C} \otimes \mathcal{P}$  is constituted by all kets  $|s_i e_i c_i p_i\rangle$  with  $s_i \in \mathcal{S}, e_i \in \mathcal{E}, c_i \in \mathcal{C}$ , and  $p_i \in \mathcal{P}$ , respectively, such that  $N_E = e_i + c_i$  and  $N_S = s_i + c_i + p_i$  are fixed, representing the conservation of the total number of substrates and enzymes. A general state of the system can be written as

$$|\Psi(t)\rangle = \sum_{s_i, e_i, c_i, p_i} P(s_i, e_i, c_i, p_i; t) |s_i e_i c_i p_i\rangle, \quad (13)$$

where  $P(s_i, e_i, c_i, p_i; t)$  is the probability of the system be found in a state with  $s_i$  substrates,  $e_i$  enzymes,  $c_i$  complexes, and  $p_i$  products at time  $t$ , respectively, while the sum runs over all vectors in the basis set. Equation (13) corresponds to the solution of the Schrodinger's equation

$$|\Psi(t)\rangle = \exp(-Ht)|\Psi(0)\rangle, \quad (14)$$

where  $|\Psi(0)\rangle = |N_S N_E 0 0\rangle$  is the initial condition of the system.

#### A. Single enzyme-substrate exact solutions

Before studying the general solution for an arbitrary, but relatively small, number of enzymes and substrates, it is instructive to analyze the simplest case with only one enzyme and one substrate, which allows direct analytical progress and pedagogical understanding of the proposed methodology. In this case, applying the conservation constraint on the number of substrates and enzymes, the basis of the reaction scheme Fock space is simply given by

$$|1\rangle = |0 0 1 0\rangle, \quad |2\rangle = |0 1 0 1\rangle, \quad |3\rangle = |1 1 0 0\rangle.$$

Note that with this generated basis order, the last ket always corresponds to the usual deterministic initial condition for Michaelis-Menten kinetics.

The matrix element of the quasi-Hamiltonian is given by  $H_{ij} = \langle i | H | j \rangle$ , where  $H$  is specified in Eq. (12), allowing its matrix representation to be written as

$$H = \begin{bmatrix} (-k_2 - k_{1-}) & 0 & k_{1+} \\ k_2 & 0 & 0 \\ k_{1-} & 0 & -k_{1+} \end{bmatrix}. \quad (15)$$

The general solution is  $|\Psi(t)\rangle = \exp(-Ht)|\Psi(0)\rangle$ , while the characteristic polynomial of  $H$  is

$$p(\lambda) = \lambda(k_{1+}k_2 + \lambda(k_2 + k_{1-} + k_{1+}) + \lambda^2), \quad (16)$$

with the associated eigenvalues

$$\lambda_1 = 0, \quad \lambda_{\pm} = -\frac{k_2 + k_{1-} + k_{1+} \pm \sqrt{\Delta}}{2}, \quad (17)$$

where  $\Delta = (k_2 + k_{1-} + k_{1+})^2 - 4k_2k_{1+}$ . In order to evaluate  $\exp(-Ht)$ , we compute the Jordan normal form of  $H$ ,  $J_H$ , and write  $H = QJ_HQ^{-1}$  to find  $|\Psi(t)\rangle = Q \exp(-J_H t) Q^{-1} |\Psi(0)\rangle$ , which can be further simplified on noting that

$$\exp(J_H) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{\lambda_-} & 0 \\ 0 & 0 & e^{\lambda_+} \end{bmatrix}, \quad (18)$$

with

$$Q = \begin{bmatrix} 0 & \frac{1}{2} \frac{\sqrt{\Delta} + k_2 + k_{1-} - k_{1+}}{\sqrt{\Delta}} & \frac{1}{2} \frac{\sqrt{\Delta} - k_2 - k_{1-} + k_{1+}}{\sqrt{\Delta}} \\ 1 & -\frac{1}{2} \frac{\sqrt{\Delta} + k_2 - k_{1-} - k_{1+}}{\sqrt{\Delta}} & -\frac{1}{2} \frac{\sqrt{\Delta} - k_2 + k_{1-} + k_{1+}}{\sqrt{\Delta}} \\ 0 & -\frac{k_{1-}}{\Delta} & \frac{k_{1-}}{\Delta} \end{bmatrix}. \quad (19)$$

With the information embedded in the above matrices, an analytical expression for  $|\Psi(t)\rangle$  can be readily determined

$$|\Psi(t)\rangle = P(0,0,1,0;t)|1\rangle + P(0,1,0,1;t)|2\rangle + P(1,1,0,0;t)|3\rangle, \quad (20)$$



where  $P(0,0,1,0;t)$ ,  $P(0,1,0,1;t)$ , and  $P(1,1,0,0;t)$  are given by:

$$P(0,0,1,0;t) = \frac{2k_{1+}e^{-1/2(k_2+k_{1+}+k_{1-})t} \sinh\left(\frac{\sqrt{\Delta}t}{2}\right)}{\sqrt{\Delta}} \quad (21)$$

$$P(0,1,0,1;t) = 1 - 4 \frac{e^{-1/2(k_{1+}+k_2+k_{1-})t} \left[ \cosh\left(\frac{\sqrt{\Delta}t}{2}\right) + \frac{(k_{1+}+k_2+k_{1-})}{\sqrt{\Delta}} \sinh\left(\frac{\sqrt{\Delta}t}{2}\right) \right] k_2 k_{1+}}{(k_2 + k_{1+} + k_{1-} + \sqrt{\Delta})(k_2 + k_{1+} + k_{1-} - \sqrt{\Delta})} \quad (22)$$

$$P(1,1,0,0;t) = e^{-\frac{(k_2+k_{1-}+k_{1+})}{2}t} \left[ \sinh\left(\frac{\sqrt{\Delta}t}{2}\right) \frac{(k_{1-} + k_2 - k_{1+})}{\sqrt{\Delta}} + \cosh\left(\frac{\sqrt{\Delta}t}{2}\right) \right], \quad (23)$$

respectively. The latter further provides a straightforward manner to evaluate all the moments for the system via

$$\langle \alpha^l \rangle = \sum_{s_i, e_i, c_i, p_i} \alpha_i^l P(s_i, e_i, c_i, p_i; t), \quad (24)$$

with  $l > 0$ ,  $\alpha \in \{S, E, C, P\}$  and  $\alpha_i \in \{s_i, e_i, c_i, p_i\}$ . In particular, for this simple case of one substrate and one enzyme,  $\langle \alpha_i^l \rangle = \langle \alpha_i \rangle$ ,  $\forall \alpha_i, l > 0$ , and thus, for example,

$$\begin{aligned} \langle E \rangle &= \langle E^2 \rangle = P(0,1,0,1,0;t) + P(1,1,0,0;t), \\ \langle C \rangle &= \langle C^2 \rangle = P(0,0,1,0;t), \\ \langle P \rangle &= \langle P^2 \rangle = P(0,1,0,1;t), \\ \langle S \rangle &= \langle S^2 \rangle = P(1,1,0,0;t), \end{aligned} \quad (25)$$

which is exact and immediately allows a calculation of the variances of the reactants' numbers.

The time evolution of the above moments are depicted in Fig. 1 for  $k_{1+} = 100$ ,  $k_{1-} = 50$ , and  $k_2 = 1$ . In the inset, stiffness in the early stages of the dynamics is observed, and despite the inherent complexity of multiple time scales in stochastic systems, this stiffness can be easily analyzed via the expressions in Eqs. (21)–(23). For instance, from Eq. (23) one can see that  $\langle S \rangle$  is characterized by two time scales in the current parameter regime. In particular, given

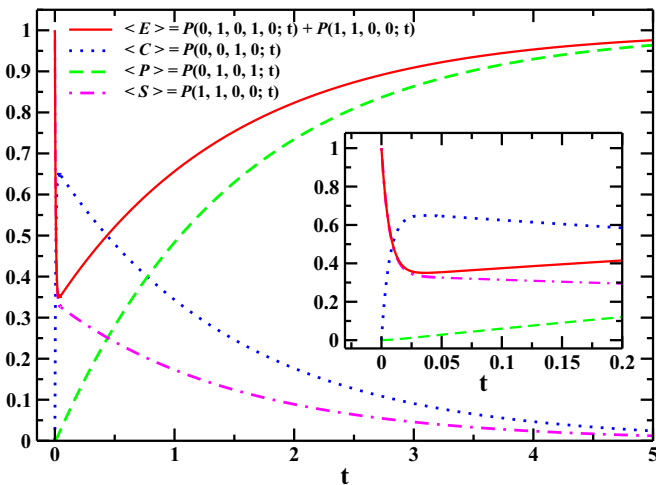


FIG. 1. (Color online) All units are SI units. Michaelis-Menten enzyme kinetics: Exact solution for one substrate and one enzyme for  $k_{1+} = 100$ ,  $k_{1-} = 50$ , and  $k_2 = 1$  from Eqs. (21)–(23). The inset depicts the stiffness of this enzymatic system at the onset of the dynamics, with the two characteristic time scales.

$k_{1+}, k_{1-} \gg k_2$ , expanding Eq. (23) around  $t = 0$  reveals that  $\langle S \rangle \approx \exp(-t/\tau_1)$ , with  $\tau_1 = 1/k_{1+}$ ; in contrast, for sufficiently large  $t$ ,  $\langle S \rangle \approx \langle S(\tau) \rangle \exp(-t/\tau_2)$ , where  $\tau_2 = 1/\lambda_+$ , and  $\tau = 2\tau_1\tau_2/(\tau_1 + \tau_2)$ .

## B. Solutions for a finite number of enzymes and substrates

We now focus on the general solution of an enzymatic system composed by a fixed number of enzymes and substrates. The general element  $|i\rangle$  of the basis set, under the substrate and enzyme conservation constraint, may be written as  $|i\rangle = |s_i e_i c_i p_i\rangle$ , with  $s_i \in \{0, \dots, N_S\}$ ,  $e_i \in \{0, \dots, N_E\}$ ,  $c_i \in \{0, \dots, N_E - e_i\}$ ,  $p_i \in \{0, \dots, N_S - s_i - N_E + e_i\}$ .

After symbolic manipulation, the quasi-Hamiltonian, Eq. (12), can be generated from the following nonzero contributions

$$\begin{aligned} \mathcal{H}_{ij} = & \delta_{p_i, p_j} \left[ -\delta_{c_i, c_j} \delta_{e_i, e_j} ((k_{1-} + k_2) c_j + k_{1+} s_j e_j) \delta_{s_i, s_j} \right. \\ & + k_{1+} s_j e_j \delta_{s_i, s_j-1} \delta_{e_i, e_j-1} \delta_{c_i, c_j+1} \\ & + k_{1-} c_j \delta_{s_i, s_j+1} \delta_{e_i, e_j+1} \delta_{c_i, c_j-1} \left. \right] \\ & + k_2 c_j \delta_{s_i, s_j} \delta_{e_i, e_j+1} \delta_{c_i, c_j-1} \delta_{p_i, p_j+1}. \end{aligned} \quad (26)$$

Furthermore, this can be converted to a matrix representation for an arbitrary, fixed, number of substrates and enzymes via the physics package of MAPLE, as illustrated within in the Supplemental Material [17].

One can diagonalize the associated matrix numerically, or sometimes symbolically, and thus find a solution for the system. For a single enzyme,  $N_E = 1$ , we will symbolically find an exact solution for the system with a finite number of substrate  $N_S$ . We stress that the solution found in the form of summations of marginal generating function methods by Aranyi [29] was explicitly illustrated only for  $N_E = N_S = 1$ .

In fact, using the above quasi-Hamiltonian, we can symbolically write its Fock space matrix representation for an arbitrary deterministic initial condition, with  $N_S$  substrates and  $N_E$  enzymes and no complexes or products, in terms of a block tridiagonal matrix. For one initial enzyme we have

$$H_{N_S} = \begin{bmatrix} \mathbf{D}_0 & \mathbf{U}_0 & 0 & 0 \\ \mathbf{L}_0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{U}_{N_S-1} \\ 0 & & \mathbf{L}_{N_S-1} & \mathbf{D}_{N_S} \end{bmatrix}, \quad (27)$$

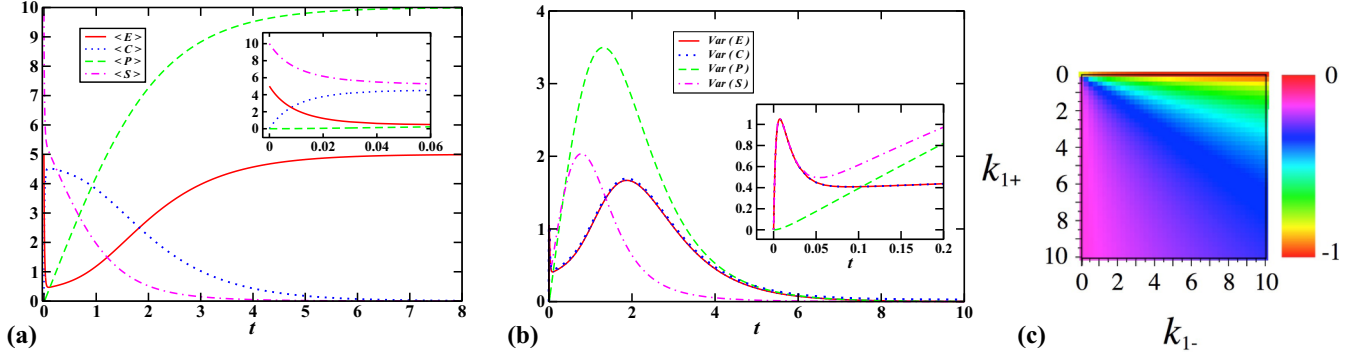


FIG. 2. (Color online) All units are SI units. Michaelis-Menten kinetics for  $N_S = 10$ ,  $N_E = 5$ ,  $k_{1+} = 10$ ,  $k_{1-} = 5$ , and  $K - 2 = 1$ : (a) Mean values and (b) variances. Short-time behavior can be seen in the insets. (c) The decay rate of the substrate as a function of  $k_{1+}$  and  $k_{1-}$  given  $k_2 = 1$ .

where

$$\mathbf{D}_{s_i} = \begin{bmatrix} -(k_2 + k_{1-}) & 0 \\ k_2 & -s_i k_{1+} \end{bmatrix}, \quad (28)$$

$$\mathbf{U}_{s_i} = \begin{bmatrix} 0 & (s_i + 1)k_{1+} \\ 0 & 0 \end{bmatrix}, \quad \mathbf{L}_{s_i} = \begin{bmatrix} k_{1-} & 0 \\ 0 & 0 \end{bmatrix}. \quad (29)$$

Furthermore, when there is more than one initial enzyme,  $N_E > 1$ , the resultant tridiagonal matrices are similar in structure to the above, though there is an increase in size of the tridiagonal blocks, that depends on the number of enzymes. The MAPLE code in the Supplemental Material can be used to immediately generate these matrix representations [17].

For only one initial enzyme and  $N_S$  initial substrates the characteristic polynomial is given by

$$\lambda_H(p; N_S) = p \prod_{s_i=1}^{N_S} [s_i k_{1+} k_2 + p(k_2 + k_{1-} + s_i k_{1+}) + p^2] \quad (30)$$

and associated eigenvalues by

$$\Lambda(N_S) = \{0, \lambda_{\pm}(s_i)\}, \quad s_i \in \{1 \dots, N_S\}, \quad (31)$$

with

$$\lambda_{\pm} = \frac{-(k_2 + k_{1-} + s_i k_{1+}) \pm \sqrt{\Delta(s_i)}}{2},$$

$$\Delta(s_i) = (k_2 + k_{1-} + s_i k_{1+})^2 - 4s_i k_2 k_{1+}.$$

Furthermore, the residues associated with the inverse Laplace transform of Eq. (10) are readily computed, and, consequently,  $|\Psi(t)\rangle$  can be determined. The basis is chosen, without loss of generality, so the last vector  $|k\rangle$  corresponds to the usual deterministic initial condition of a Michaelis-Menten system, i.e.,  $s_i = N_S$ ,  $e_i = N_E$ , and  $|k\rangle = |N_S N_E 0 0\rangle$ . This convenient choice for the basis ensures that  $|\Psi(t)\rangle$  corresponds to the last column of  $\exp(-Ht)$ , and thus we have the simplification

$$\begin{aligned} |\Psi(t)\rangle_l &= \frac{1}{2\pi i} \int_{\gamma} dp e^{pt} \frac{\text{adj}(p I_n - H)_{lk}}{\det(p I_n - H)} \\ &= \sum_{\text{residues}} e^{pt} \frac{\text{adj}(p I_n - H)_{lk}}{\det(p I_n - H)}, \end{aligned} \quad (32)$$

with, once more, a summation over all the residues, as indexed by the eigenvalues for the Fock space matrix representation of  $H$ , which can be readily evaluated. In the particular example of one enzyme, the eigenvalues are available analytically, so the analytical solution is given via the use of symbolic computation to evaluate the sum of residues in Eq. (32).

More generally, note that if the eigenvalues have to be determined numerically, which is the case if the characteristic polynomial does not factorise simply, one needs to only evaluate the eigenvalues and matrix adjugates once to determine the behavior of the system, including all of its moments, at any desired time. Further note this applies even when the eigenvalues have multiplicities greater than unity. Analogous comments also hold if the solution is obtained via the Jordan normal form and one also has the option of just evaluating  $\exp(-Ht)$  at desired time points if only early, or late, time behavior is required, which is also the methodology used for generating the exemplifying results of Fig. 2. Furthermore, any and all of these procedures grant access to the time evolution of the system and associated moments with minimal effort, providing the system size is small enough to ensure the linear algebra is tractable.

To illustrate the results of the methodology with multiple enzymes, where the characteristic polynomial does not factorise simply, the time evolution of all mean values and variances are depicted in Fig. 2 for a Michaelis-Menten system composed of 10 substrates and five enzymes with  $k_{1+} = 10$ ,  $k_{1-} = 5$ , and  $k_2 = 1$ . Interestingly, the larger number of molecules is not observed to alter the striking stiffness of the mean behaviors predicted earlier for the single enzyme-substrate system in Fig. 1 though, not surprisingly, higher moments and thus variances are altered.

These results more generally demonstrate the numerous opportunities to probe the sensitivity of small law-of-mass-action stochastic systems for an increasing number of interacting molecules, with the additional possibility of exact solutions for the simplest systems. In particular, fast searches in the parameter space can be effortlessly performed without ensemble numerical averaging using, as demonstrated in Fig. 2(c), where the decay ratios of  $\langle S \rangle$ , for  $k_{1\pm} \in [0, 10]$  and  $k_2 = 1$ , are portrayed. This is in contrast to the traditionally employed stochastic simulation algorithms for which explorations of the parameter space are computationally demanding, especially

for stiff systems, and have limited potential for analytical insights.

#### IV. DISCUSSION

We have implemented statistical and quantum mechanical tools for the study of law-of-mass-action stochastic systems, where randomness is due to low copy number rather than quantum fluctuations, as has been recognized and pursued previously, though only at an abstract level in the context of this study. Deducing a matrix representation of the quasi-Hamiltonian,  $H$ , associated with the time evolution of the chemical master equation within Fock space has enabled the exploitation of symbolic computation and numerical linear algebra to determine the solution of the chemical master equation via the evolution operator  $\exp(-Ht)$ . This resulting solution can be determined in the absence of ensemble averaging via numerous techniques, ranging from matrix exponentiation to the residue summations associated with inverse Laplace transforms, as explicitly demonstrated in the context of the low-reactant-number Michaelis-Menten system. In particular, we have reproduced simpler forms of previous analytical results derived via marginal generating functions and deduced conditions for the presence of stiffness at low reactant numbers, even though a large ratio of substrate to enzyme features in the standard asymptotic scalings to deduce stiffness in the large number continuum model [23].

The absence of handling ensembles in such studies, and an analytical emphasis, entails that this methodology readily allows further analytical insights as well as the prospect of simpler and faster parameter inference from empirical data, for instance, via standard curve fitting or Bayesian techniques, especially in higher-dimensional parameter spaces where computation can still be a significant issue. There are analogous advantages for stiffness, sensitivity, and robustness analyses and also in the analysis of objective function optimization, which is likely to find application in disciplines such as synthetic biology, which address the question of how to design function into biomimetic systems, with the immediate generalization of improving and optimizing design.

Furthermore, the presented Fock space methodology can be readily implemented, especially by taking advantage of tools developed for, and by, the physics community, as illustrated by the extensive use of the physics package for MAPLE within the code provided in the Supplemental Material [17]. Parameter space surveys are also easily pursued, as highlighted by the presented results for the time scales in the small-reactant-number Michaelis-Menten system with an initial total of 10 substrate and five enzyme molecules. In

addition, there is often the possibility of generating an exact analytical solution, as highlighted when only a single enzyme molecule is initially present. In addition, the specific choice of Michaelis-Menten kinetics emphasizes that the Fock space approach can readily accommodate stiff systems, whereas stiffness typically induces computational inefficiencies in stochastic simulation algorithms.

Although the framework developed here is formally valid for an arbitrary finite number of particles, personal computing can only be utilized for sufficiently small systems, though with an increased efficiency when conservation laws and inherent symmetries are exploited to reduce the configuration space, as illustrated earlier. The ease of application of these techniques for small systems can be observed via use of the MAPLE code in the Supplemental Material [17], which can also be readily adapted to analyze other chemical reactions by changing the corresponding quasi-Hamiltonian.

However, for sufficiently large reactant numbers, and no configuration space reduction, computational memory is the limiting resource for the symbolic construction of the quasi-Hamiltonian. Even in this case, the size of the quasi-Hamiltonian matrix only grows polynomially with the total number of reactants, while its inherent structure ensures very sparse matrices, as one proceeds through the Fock space one reaction at a time, as illustrated by Eq. (12). Hence there is also a wide range of fast algorithms that could be employed to exploit the sparse nature of the quasi-Hamiltonian for the numerical evaluation of the system solution.

Finally, the link with the Fock space of quantum and statistical physics not only entails access to numerous insights and computational techniques for generating and studying the matrix representation of the quasi-Hamiltonian,  $H$ , and the time evolution operator  $\exp(-Ht)$  but also provides the prospect of hijacking further tools and insights from 80 years of physics research for the purpose of studying and understanding the chemical master equation. For instance, relating the chemical master equation to Feynman path integrals allows novel perturbation methods to be employed, while mapping the chemical master equation onto exactly solvable statistical mechanical models [8], such as the Ising model or Potts model, may generate further novel analytical solutions. Finally, the fact reaction-diffusion processes can be seen as physical realizations of Hecke algebras [30] implies there are future prospects of applying analogous approaches to the study of stochastic reaction diffusion systems via Hecke algebra-based techniques.

#### ACKNOWLEDGMENTS

This work was partially supported by CNPq and FACEPE (Brazilian agencies).

- 
- [1] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer-Verlag, Berlin, 1997).
  - [2] M. Doi, Second quantization representation for classical many-particle system, *J. Phys. A* **9**, 1465 (1976); see also: M. Doi, Stochastic theory of diffusion-controlled reaction, *ibid.* **9**, 1479 (1976).
  - [3] P. Grassberger and M. Scheunert, Fock-space methods for identical classical objects, *Fortschr. Phys.* **28**, 547 (1980).
  - [4] L. Peliti, Path integral approach to birth-death processes in a lattice, *J. Phys.* **46**, 1469 (1985).
  - [5] For a review see: D. C. Mattis and M. Lawrence Glasser, The uses of quantum field theory in diffusion-limited reactions,

- Rev. Mod. Phys. **70**, 979 (1998); U. C. Tauber, M. Howard, and B. P. Vollmayr-Lee, Applications of field-theoretic renormalization group methods to reaction-diffusion problems, *J. Phys. A* **38**, R79 (2005); J. Baez and J. Biamonte, A course on quantum techniques for stochastic mechanics, [arXiv:1209.3632](https://arxiv.org/abs/1209.3632).
- [6] F. C. Alcaraz, M. Droz, M. Henkel, and V. Rittenberg, Reaction-diffusion processes, critical dynamics, and quantum chains, *Ann. Phys. (NY)* **230**, 250 (1994).
- [7] M. Henkel, E. Orlandini, and J. Santos, Reaction-diffusion processes from equivalent integrable quantum chains, *Ann. Phys. (NY)* **259**, 163 (1997).
- [8] R. Baxter, *Exactly Solved Model in Statistical Mechanics* (Dover, London, 2005).
- [9] J. L. England and J. Cardy, Morphogen Gradient from a Noisy Source, *Phys. Rev. Lett.* **94**, 078101 (2005).
- [10] M. Sasai and P. G. Wolynes, Stochastic gene expression as a many-body problem, *Proc. Natl. Acad. Sci. USA* **111**, 2374 (2003).
- [11] P. J. Dodd and N. M. Ferguson, A many-body field theory approach to stochastic models in population biology, *PLoS ONE* **4**, e6855 (2009).
- [12] B. Zhang and P. G. Wolynes, Stem cell differentiation as a many-body problem, *Proc. Natl. Acad. Sci. USA* **111**, 10185 (2014).
- [13] G. L. Eyink, Action principle in nonequilibrium statistical dynamics, *Phys. Rev. E* **54**, 3419 (1996).
- [14] Y. Lan, P. G. Wolynes, and G. A. Papoian, A variational approach to stochastic aspects of cellular signal transduction, *J. Chem. Phys.* **125**, 124106 (2006).
- [15] A. Mugler, A. M. Walczak, and C. H. Wiggins, Spectral solutions to stochastic models of gene expression with bursts and regulation, *Phys. Rev. E* **80**, 041921 (2009).
- [16] B. Munsky and M. Khammash, The finite state projection algorithm for the solution of the chemical master equation, *J. Chem. Phys.* **124**, 044104 (2006).
- [17] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.92.062714> for MAPLE codes.
- [18] A. Gupta and M. Khammash, Methods and algorithms for scientific computing unbiased estimation of parameter sensitivities for stochastic chemical reaction networks, *SIAM. J. Sci. Comput.* **35**, A2598 (2013).
- [19] R. Grima, N. G. Walter, and S. Schnell, Single-molecule enzymology a la Michaelis-Menten, *FEBS J.* **281**, 518 (2014).
- [20] S. C. Kou, B. J. Cherayil, W. Min, B. P. English, and X. S. Xie, Single-molecule Michaelis-Menten equations, *J. Phys. Chem. B* **109**, 19068 (2005).
- [21] P. S. Swain, M. B. Elowitz, and E. D. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, *Proc. Natl. Acad. Sci. USA* **99**, 12795 (2002).
- [22] A. Raj and A. van Oudenaarden, Single-molecule approaches to stochastic gene expression, *Annu. Rev. Biochem.* **38**, 255 (2009).
- [23] J. D. Murray, *Mathematical Biology* (Springer-Verlag, Berlin, 1993).
- [24] E. Weinan, D. Liu, and E. Vanden-Eijnden, Nested stochastic simulation algorithm for chemical kinetic systems with disparate rates, *J. Chem. Phys.* **123**, 194107 (2005).
- [25] D. A. McQuarrie, Stochastic approach to chemical kinetics, *J. Appl. Prob.* **4**, 413 (1967).
- [26] J. C. Baez and B. Fong, A Noether theorem for Markov processes, *J. Math. Phys.* **54**, 013301 (2013).
- [27] C. Moler and C. Van Loan, Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, *SIAM Rev.* **45**, 3 (2003).
- [28] Other arbitrary initial conditions can be chosen. In particular because total enzyme and substrate are conserved given enzyme and substrate within the complex and product are also counted, stochasticity in the initial total numbers of enzyme and substrate can be considered explicitly via the techniques presented using a simple decomposition of the Fock space into subspaces where, within each subspace, the total enzyme and substrate are fixed. Alternatively, this is equivalent to conditioning on the initial number of substrates,  $N_S$ , and enzymes,  $N_E$ .
- [29] P. Aranyi and J. Toth, A full stochastic description of the Michaelis-Menten reaction for small systems, *Acta Biochem. Biophys. Acad. Sci. Hung.* **12**, 375 (1977).
- [30] F. C. Alcaraz, V. Rittenberg, Reaction-diffusion processes as physical realizations of Hecke algebras, *Phys. Lett. B* **314**, 377 (1993).