Variance Decomposition in Stochastic Simulators

O.P. Le Maître,[1, a)] O.M. Knio,[2, b)] and A. Moraes[3, c)]

[1)] *LIMSI-CNRS, UPR 3251, Orsay, France*

[2)] *Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708*

[3)] *King Adbullah University of Science and Technology, Thuwal, Saudi Arabia*

(Dated: 6 January 2017)

This work aims at the development of a mathematical and computational approach that enables quantification of the inherent sources of stochasticity and of the corresponding sensitivities in stochastic simulations of chemical reaction networks. The approach is based on reformulating the system dynamics as being generated by independent standardized Poisson processes. This reformulation affords a straightforward identification of individual realizations for the stochastic dynamics of each reaction channel, and consequently a quantitative characterization of the inherent sources of stochasticity in the system. By relying on the Sobol-Hoeffding decomposition, the reformulation enables us to perform an orthogonal decomposition of the solution variance. Thus, by judiciously exploiting the inherent stochasticity of the system, one is able to quantify the variance-based sensitivities associated with individual reaction channels, as well as the importance of channel interactions. Implementation of the algorithms is illustrated in light of simulations of simplified systems, including the birth-death, Schlögl, and Michaelis-Menten models.

I. INTRODUCTION

It has long been recognized that stochastic simulators[1–7] offer an attractive computational means for simulating the chemical master equation. These algorithms are based on describing the evolution of the number density of individual chemical species on the basis of propensity functions, which provide a probabilistic characterization of the rate of progress of elementary reactions. In its elementary formulation, the reaction is advanced by sampling the time for next reaction to occur, whereas the system remains frozen in its current state in the mean time. After a reaction has occurred, new samples are drawn, and the procedure is iterated in order to determine successive elements of a Markov process associated with the evolution of the state vector of species number densities. By computing multiple realizations of such chains, the algorithm affords in a straightforward fashion the possibility to determine various statistical moments of the solution.

One of the difficulties associated with stochastic simulators is that it is not generally straightforward to assess the impact of individual reaction channels on the evolution of the solution, nor to quantify the effect of uncertainties on the variability of the solution. This is especially the case for stochastic systems that do not involve parametric uncertainty, which are the focus of the present study. The central question we aim to address is whether it is possible, without perturbing the reaction rates (propensity functions), to exploit the stochasticity inherent in stochastic simulations for the purpose of quantifying the impact of individual channels on the evolution of particular species of interest, and also potentially characterizing the role of channel interactions. As outlined in section II, our approach to this question is based on reformulating classical Stochastic Simulator Algorithms (SSAs, see Refs. 1 and 2) using independent standardized Poisson processes similarly to the Next Reaction Method proposed in Ref. 8. This reformulation affords a straightforward identification of independent random realizations for individual channel dynamics, and consequently enables us to exploit sampling strategies in order to quantify the impact of the stochasticity associated with elementary channels, or groups of channels. As discussed in section III, the Sobol-Hoeffding decomposition is applied for this purpose. In particular, this enables us to adopt well-established Monte-Carlo sampling strategies to obtain estimates of the contributions of individual channels, and of channel interactions, to the variance of specific quantities of interest. Implementation of the resulting methodology is illustrated in section V through applications to well-known systems, including the birth-death, Schlögl, and Michaelis-Menten models. Major conclusions are summarized in section VI.

II. STOCHASTIC SYSTEMS AND SIMULATORS

The present work focuses on the analysis of the respective importance of inherent sources of variance in stochastic simulations of chemical reaction networks. Stochas-

[a)] Visiting Professor, Department Mechanical Engineering and Materials Science, Duke University; Electronic mail: olm@limsi.fr

[b)] Present address: King Adbullah University of Science and Technology, Thuwal, Saudi Arabia; Electronic mail: knio@duke.edu

[c)] Electronic mail: alvaro.moraesgutierrez@kaust.edu.sa

tic effects arise predominantly when the reacting system involves a small number of reactant molecules, due to discrete evolutions. Specifically, we consider a well-stirred chemical system involving M_s molecular species $\{S_1, \ldots, S_{M_s}\}$ and K_r reactions channels $\{R_1, \ldots, R_{K_r}\}$. We denote $\boldsymbol{X}(t) = (X_1(t) \ldots X_{M_s}(t))^T \in \mathbb{N}_0^{M_s}$ the state of the system at time t, where $X_i(t)$ is the instantaneous number of molecules of species S_i. The dynamics of the system are prescribed by the definition of the propensity functions a_j and state-change vectors $\boldsymbol{\nu}_j \in \mathbb{Z}^{M_s}$ associated to the reaction channel R_j. The propensity functions,

$$a_j : \boldsymbol{x} \in \mathbb{N}_0^{M_s} \mapsto a_j(\boldsymbol{x}) \in \mathbb{R}_+, \quad j = 1, \ldots, K_r,$$

are defined such that, for infinitesimal time dt, $a_j(\boldsymbol{x})dt$ is the probability that one j-th reaction occurs in the time interval $[t, t + dt)$ given that $\boldsymbol{X}(t) = \boldsymbol{x}$, while $\boldsymbol{\nu}_j = (\nu_{1,j}, \ldots, \nu_{M_s,j})$ describes the change in the molecular population due to a unique reaction R_j.

The system then obeys the chemical master equation[9] (CME),

$$\frac{\partial P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)}{\partial t} = \sum_{j=1}^{K_r} [a_j(\boldsymbol{x} - \boldsymbol{\nu}_j) P(\boldsymbol{x} - \boldsymbol{\nu}_j, t|\boldsymbol{x}_0, t_0)$$
$$- a_j(\boldsymbol{x}) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)], \quad (1)$$

where $P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)$ is the probability that $\boldsymbol{X}(t) = \boldsymbol{x}$ given that $\boldsymbol{X}(t_0) = \boldsymbol{x}_0$.

Solving (1) is impractical even for systems with small M_s and K_r. The Gillespie stochastic simulation algorithm (SSA) exactly emulates the chemical master equation by generating samples of $\boldsymbol{X}(t)$ with probability law $P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)$. In Section II A we briefly recall the original SSA. In Section II B we introduce a slightly modified version of the SSA, which is suitable for the analysis of the variance.

A. Stochastic Simulation Algorithm

Gillespie's SSA[1,2] consists in the construction of a Markov-chain $\boldsymbol{X}(t)$ having for distribution $P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)$. The first idea of the SSA is that, given $\boldsymbol{X}(t) = \boldsymbol{x}$, the probability of the next reaction (irrespective of the channel) to occur in the infinitesimal time interval $[t, t + dt)$ is

$$a_0(\boldsymbol{x})dt = dt \sum_{j=1}^{K_r} a_j(\boldsymbol{x}).$$

In other words, the time elapsed to the next reaction, τ, follows an exponential distribution with mean $1/a_0(\boldsymbol{x})$. Second, the relative probability of reaction channel R_j to fire first is $p_j(\boldsymbol{x}) = a_j(\boldsymbol{x})/a_0(\boldsymbol{x})$. Gillespie's Algorithm 1 proceeds along these two stages to advance the state of the system: first, drawing at random a time increment to

the next reaction event, and second, selecting at random the firing reaction channel to update the state. In more details, given the current state $\boldsymbol{X}$ of the system, the SSA first evaluates the K_r propensity functions a_j and their sum a_0. Drawing r_1 uniformly in $(0, 1)$ the algorithm sets $\tau = -\log(r_1)/a_0$, and selects the index of the firing channel using the largest l such that $\sum_{j<l} a_j < a_0 r_2$, where r_2 is another random number, independent of r_1, drawn uniformly from $(0, 1)$. Time and system state are then advanced, $t \leftarrow t + \tau$ and $\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l$, and the procedure is repeated until the final time is reached.

It is thus seen that the SSA advances in time using a consolidated next reaction time distribution and relative reactions probabilities, so the method inherently depends on the whole state of the system. In other words, the algorithm fully couples the reaction dynamics. It requires drawing 2 random numbers from the unit interval for every elementary event (occurrence of a reaction). In practice, a pseudo-random number generator RG is used to simulate the sequence of random numbers r_1 and r_2. Classically, the pseudo-random number generator RG provides the exact same sequence of pseudo-random couples (r_1, r_2) when properly seeded (initialized). As further discussed below, this property of pseudo-random generators is attractive for our purpose, because it enables us to generate the same sequence of pseudo-random events multiple times.

However, the SSA has several drawbacks that make it unsuitable for the separation of the variances caused by the different channels. The central drawback is that SSA combines all channels to determine the next reaction time, while the analysis of the variability inherent to specific channels, as proposed in the next section, requires us to distinguish and control the dynamics of individual channels. An alternative way to understand the issue is that for a fixed sequence of pseudo-random numbers, the resulting trajectory of $\boldsymbol{X}(t)$ is not stable to permutation of the reaction indexation. This drawback motivates the use of the Next Reaction Algorithm (NRA) outlined below.

B. Next Reaction Algorithm

Following the Next Reaction Method proposed in Ref. 8, the dynamics of the stochastic state vector $\boldsymbol{X}(t)$ can be expressed through the following random time-change[10]

$$\boldsymbol{X}(t) = \boldsymbol{X}(t_0) + \sum_{j=1}^{K_r} \boldsymbol{\nu}_j N_j(t_j), \quad (2)$$

where $N_j(\tau)$ are *independent Poisson processes* with unit rates, and the scaled time t_j are given by

$$t_j = \int_{t_0}^{t} a_j(\boldsymbol{X}(\tau))d\tau, \quad j = 1, \ldots, K_r. \quad (3)$$

ALGORITHM 1. Stochastic Simulation Algorithm.
Procedure SSA $(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG})$

Require: Initial condition $\boldsymbol{X}_0$, final time T, state-change vectors $\{\boldsymbol{\nu}_j\}$, propensity functions $\{a_j\}$, and seeded pseudo-random number generator $\mathtt{RG}$
1: $t \leftarrow 0$, $\boldsymbol{X} \leftarrow \boldsymbol{X}_0$
2: **loop**
3: Evaluate $a_j(\boldsymbol{X})$, $j = 1, \ldots, \mathrm{K}_r$
4: Set $a_0 = \displaystyle\sum_{j=1}^{\mathrm{K}_r} a_j(\boldsymbol{X})$
5: Get r_1 and r_2 from $\mathtt{RG}$
6: Set $\tau = -\dfrac{1}{a_0} \ln(r_1)$ ▷ time to next reaction
7: **if** $t + \tau > T$ **then**
8:   **Break**   ▷ final time reached
9: **else**
10:   Set largest l s.t. $\displaystyle\sum_{j=1}^{j<l} a_j < (a_0\, r_2)$ ▷ pick the reaction to fire
11:   $\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l$ ▷ update the state vector
12:   $t \leftarrow t + \tau$ ▷ advance time
13: **end if**
14: **end loop**
15: **Return** $\boldsymbol{X}$ ▷ State $\boldsymbol{X}(T)$

Thus, the system state can be expressed as a linear combination of K_r standard Poisson processes, with individual time scaling given by the time-integrals of the corresponding propensity functions from t_0 to t. The expression (2), together with (3), result from the fact that Poisson processes have independent increments distributed according to

$$P(N_j(\tau + \Delta\tau) - N_j(\tau) = k \geq 0) = \frac{e^{-\Delta\tau}\Delta\tau^k}{k!}.$$

Consequently, the probability of R_j to fire once in the infinitesimal time interval $[t, t + dt)$ is

$$P(N_j(t_j + a_j(\boldsymbol{X}(t))dt) - N_j(t_j) = 1) =$$
$$e^{-a_j(\boldsymbol{X}(t))dt} a_j(\boldsymbol{X}(t))dt = a_j(\boldsymbol{X}(t))dt + \mathcal{O}(dt^2), \quad (4)$$

which is precisely the definition of the propensity function a_j.

The representation of $\boldsymbol{X}(t)$ through (2) and (3) suggests an alternative version of the original SSA, hereafter referred as the Next Reaction Algorithm (NRA), which is a slightly modified version of the Next Reaction Method proposed in Ref. 8. Algorithm 2 introduces two additional vectors, $(\tau_1, \ldots, \tau_{\mathrm{K}_r})$ and $(\tau_1^+, \ldots, \tau_{\mathrm{K}_r}^+)$ respectively, to store the unscaled time of the Poisson processes and their next reaction times. These two vectors are used to keep track of the trajectories of the K_r independent Poisson processes. Given the current state $\boldsymbol{X}$ of the system, normalized times $(\tau_1, \ldots, \tau_{\mathrm{K}_r})$ and next reaction times $(\tau_1^+, \ldots, \tau_{\mathrm{K}_r}^+)$, one can deduce which channel will fire first from the scaled times advancement rules (3). Specifically, if no other channel fires before it does, the j-th channel will react after a time lapse $dt_j = (\tau_j^+ - \tau_j)/a_j$.

Letting $l = \arg\min_j dt_j$ be the first channel to fire, the procedure selects at once both the firing channel and the time lapse (dt_l) to the next reaction, unlike the SSA. Having found l, time is advanced $(t \leftarrow t + dt_l)$, system state $\boldsymbol{X}$ is updated $(\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l)$, the unscaled times vector is updated $(\tau_j \leftarrow \tau_j + a_j dt_l)$ and the unscaled next reaction time τ_l^+ is drawn at random $(\tau_l^+ \leftarrow \tau_l^+ - \log(r_l)$ where r_l is drawn at random from $(0, 1))$. The procedure is then repeated until t reaches the final time. The NRA 2 involves drawing a unique random variable per time-step (whereas SSA draws two) but requires the initialization, computation and storage of the scaled-times and next jump-times associated to the K_r Poisson processes associated to the reaction channels. However, both algorithms result in states $\boldsymbol{X}$ that are equal in distribution, that is

$$X^{SSA}(t) \stackrel{d}{=} X^{NRA}(t).$$

Our interest in considering the NRA lies in the possibility of maintaining constant realizations of selected standardized Poisson processes N_j while varying others. It amounts to keeping invariant the sequences $\{\tau_j^+\}$ of unscaled jump-times. This possibility is enabled by the introduction of the K_r individual pseudo-random number generators, one for each reaction channel, when in fact a unique random number generator would have been sufficient in NRA 2. The use of K_r independent pseudo-random number generators distinguishes NRA 2 from the Next Reaction Method in 8. As for SSA, the storage of these K_r independent sequences is not required as they are generated on the fly using the pseudo-random number generators $\mathtt{RG}_j$. Doing so, the realizations of the Poisson processes are fully determined by imposing the seeds of the K_r pseudo-random number generators. As a result, we now have a *clear understanding of the inherent source of stochasticity* which is kept invariant when using the same set of seeds, namely the underlying Poisson processes, and a fine control over this source of stochasticity. This control over the inherent stochasticity enables a decomposition of the variance of the state vector, and the analysis of the variability induced by distinct channels, as discussed in the following section.

III. VARIANCE DECOMPOSITION

In parametric sensitivity analysis, one is interested in characterization of the variability of the deterministic output (or quantity of interest) of a model when some parameters are changed. Local sensitivity analyses characterize the dependence of the deterministic output about a particular value of the parameters, usually by computing the first derivatives of the output with respect to the parameters. Contrary to local methods, global sensitivity analyses aim to elucidate the global contribution of a particular input on the total output variability, assuming a certain distribution of the parameters. In the ANOVA (analysis of the variance) or Sobol's

ALGORITHM 2. Next Reaction Algorithm.
Procedure NRA$(X_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG}_1, \ldots, \mathtt{RG}_{\mathrm{K}_r})$

Require: Initial condition $\boldsymbol{X}_0$, final time T, state-change vectors $\{\boldsymbol{\nu}_j\}$, propensity functions $\{a_j\}$, and seeded pseudo-random number generators $\mathtt{RG}_{j=1,\ldots\mathrm{K}_r}$

1: **for** $j = 1, \ldots, \mathrm{K}_r$ **do**
2:    Draw r_j from $\mathtt{RG}_j$
3:    $\tau_j \leftarrow 0$, $\tau_j^+ \leftarrow -\log r_j$ ▷ set next reaction times
4: **end for**
5: $t \leftarrow 0, \boldsymbol{X} \leftarrow \boldsymbol{X}_0$
6: **loop**
7:    **for** $j = 1, \ldots, \mathrm{K}_r$ **do**
8:       Evaluate $a_j(\boldsymbol{X})$ and $dt_j = \frac{\tau_j^+ - \tau_j}{a_j}$
9:    **end for**
10:   Set $l = \arg\min_j dt_j$ ▷ pick next reaction
11:   **if** $t + dt_l > T$ **then**
12:     break   ▷ Final time reached
13:   **else**
14:     $t \leftarrow t + dt_l$ ▷ update time
15:     $\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l$ ▷ update the state vector
16:     **for** $j = 1, \ldots, \mathrm{K}_r$ **do**
17:       $\tau_j \leftarrow \tau_j + a_j \, dt_l$ ▷ update unscaled times
18:     **end for**
19:     Get r_l from $\mathtt{RG}_l$
20:     $\tau_l^+ \leftarrow \tau_l^+ - \log r_l$ ▷ next reaction time
21:   **end if**
22: **end loop**
23: **Return** $\boldsymbol{X}$ ▷ State $\boldsymbol{X}(T)$

---

variance decomposition, the sensitivity indices are associated to the input parameters *via* a decompositions of the output variance.[11,12] When the input parameters vary independently, the Sobol's decomposition is orthogonal and the definition of the sensitivity indices is immediate from the so-called partial variances (see below). The case of dependent inputs requires appropriate definition and interpretation of the sensitivity indices (see for instance Refs. 13 and 14). The Sobol's decomposition is also related to High Dimensional Model Representation (HDMR), which corresponds to a truncated SH decomposition; such decompositions have in particular been used for the construction of surrogate models (see for instance Refs. 13, 15–17). In the context of reaction networks presenting an inherent stochastic dynamics, analyses have been restricted to averaged functionals of the stochastic model solution: the analyses characterizes the sensitivity of such averages with respect to parameters, for instance coefficients in the propensity functions defining the stochastic network.[18–22]

Parametric sensitivity analyses have to be contrasted with the use of the Sobol decomposition we are proposing in the present work. Here, we assume no variability in any model parameters. Instead, we seek to quantify the respective contributions of different reaction channels to the variance of a given functional of the stochastic model solution.

In this section, we summarize relevant aspects of the Sobol-Hoeffding decomposition, provide a brief outline of how these concepts can be applied to define useful sensitivity indices, and detail a Monte-Carlo sampling procedure for their estimation.

A. Sobol-Hoeffding decomposition

Consider a vector $\boldsymbol{N} = (N_1, \cdots, N_{\mathrm{D}})$ of D independent random quantities defined on an abstract probability space $\mathcal{P} = (\Omega, \Sigma, \mu)$. Let $F : \boldsymbol{N} \mapsto F(\boldsymbol{N}) \in \mathbb{R}$ be a second-order random functional in $\boldsymbol{N}$, that is

$$F \in L_2(\Omega, \mu) \Leftrightarrow \mathbb{E}\left\{F(\boldsymbol{N})^2\right\} < \infty. \tag{5}$$

Let $\mathcal{D}$ be the power set of $\{1, \cdots, \mathrm{D}\}$. For $\mathfrak{u} \in \mathcal{D}$ denote $|\mathfrak{u}| = \mathrm{Card}(\mathfrak{u})$ and $\mathfrak{u}_\sim = \mathcal{D} \setminus \mathfrak{u}$, such that $\mathfrak{u} \cup \mathfrak{u}_\sim = \mathcal{D}$ and $\mathfrak{u} \cap \mathfrak{u}_\sim = \emptyset$. Given $\mathfrak{u} \in \mathcal{D}$ we denote $\boldsymbol{N}_\mathfrak{u}$ the sub-vector of $\boldsymbol{N}$ with components $(N_{\mathfrak{u}_1}, \cdots, N_{\mathfrak{u}_{|\mathfrak{u}|}})$, so $\boldsymbol{N} = (\boldsymbol{N}_\mathfrak{u}, \boldsymbol{N}_{\mathfrak{u}_\sim})$.

For the assumptions considered, the function $F(\boldsymbol{N})$ has a *unique* orthogonal decomposition of the form[23]

$$F(\boldsymbol{N}) = \sum_{\mathfrak{u} \in \mathcal{D}} F_\mathfrak{u}(\boldsymbol{N}_\mathfrak{u}), \tag{6}$$

Equation (6) is called the Sobol-Hoeffding (SH) decomposition of F. For instance, in the case D = 3 the decomposition is expressed as:

$$\begin{aligned}
F(N_1, N_2, N_3) = &F_\emptyset + F_{\{1\}}(N_1) + F_{\{2\}}(N_2) + F_{\{3\}}(N_3) \\
&+ F_{\{1,2\}}(N_1, N_2) + F_{\{1,3\}}(N_1, N_3) \\
&+ F_{\{2,3\}}(N_2, N_3) + F_{\{1,2,3\}}(N_1, N_2, N_3).
\end{aligned}$$

The orthogonality condition for the functions implies that $\forall \mathfrak{u} \neq \mathfrak{s}$

$$\mathbb{E}\{F_\mathfrak{u}F_\mathfrak{s}\} = \int_\Omega F_\mathfrak{u}(\boldsymbol{N}_\mathfrak{u}(\omega))F_\mathfrak{s}(\boldsymbol{N}_\mathfrak{s}(\omega))d\mu(\omega) = 0,$$

and the SH functions $F_\mathfrak{u}$ are recursively defined according to:[11]

$$F_\mathfrak{u}(\boldsymbol{N}_\mathfrak{u}) = \mathbb{E}\{F \mid \boldsymbol{N}_\mathfrak{u}\} - \sum_{\substack{\mathfrak{s} \in \mathcal{D} \\ \mathfrak{s} \subsetneq \mathfrak{u}}} F_\mathfrak{s}(\boldsymbol{N}_\mathfrak{s}), \tag{7}$$

where $\mathbb{E}\{F \mid \boldsymbol{n}_\mathfrak{u}\}$ is the conditional expectation of $F(\boldsymbol{N})$ given $\boldsymbol{N}_\mathfrak{u} = \boldsymbol{n}_\mathfrak{u}$, namely

$$\mathbb{E}\{F \mid \boldsymbol{n}_\mathfrak{u}\} = \int_\Omega F(\boldsymbol{n}_\mathfrak{u}, \boldsymbol{N}_{\mathfrak{u}_\sim}(\omega))d\mu(\omega).$$

For instance, in the previous example with D = 3, we have

$$\begin{aligned}
F_{\{2,3\}}(N_2, N_3) = &\mathbb{E}\{F \mid N_2, N_3\} - F_\emptyset - F_{\{2\}}(N_2) \\
&- F_{\{3\}}(N_3).
\end{aligned}$$

Clearly, $F_\emptyset = \mathbb{E}\{F\}$, and the orthogonality of the decomposition implies $\mathbb{E}\{F_\mathfrak{u}\} = 0$ for $\mathfrak{u} \neq \emptyset$. The decomposition (6) being orthogonal, the variance $\mathbb{V}\{F\}$ is decomposed into the sum of partial variances corresponding to the variances of the SH functions:

$$\mathbb{V}\{F\} = \sum_{\mathfrak{u} \in \mathcal{D} \setminus \emptyset} \mathbb{V}\{F_\mathfrak{u}\}. \tag{8}$$

Recalling that $\mathbb{E}\left\{F_{\mathbf{u}\neq\emptyset}\right\} = 0$, and using the definition (7), the partial variances $\mathbb{V}\{F_{\mathbf{u}}\}$ can be expressed as:

$$\mathbb{V}\{F_{\mathbf{u}}\} = \mathbb{E}\left\{\left(\mathbb{E}\{F \mid \boldsymbol{N_u}\} - \sum_{\substack{\mathbf{s}\in\mathcal{D} \\ \mathbf{s}\subsetneq\mathbf{u}}} F_{\mathbf{s}}(\boldsymbol{N_s})\right)^2\right\}$$

$$= \mathbb{E}\left\{(\mathbb{E}\{F \mid \boldsymbol{N_u}\} - F_\emptyset)^2\right\}$$

$$-2\sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\subsetneq\mathbf{u}}}\mathbb{E}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}F_{\mathbf{s}}(\boldsymbol{N_s})\} + \sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\subsetneq\mathbf{u}}}\mathbb{V}\{F_{\mathbf{s}}\}$$

$$= \mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}\} - \sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\subsetneq\mathbf{u}}}\mathbb{V}\{F_{\mathbf{s}}\}, \quad (9)$$

where we have used $\mathbb{E}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}\} = \mathbb{E}\{F\} = F_\emptyset$ and the orthogonality of the SH functions.

B. Sensitivity indices

The partial variance $\mathbb{V}\{F_{\mathbf{u}}\}$ measures the contribution to $\mathbb{V}\{F\}$ of the interactions between the variables $\boldsymbol{N_u}$. Since there are 2^{D} such partial variances, the sensitivity analysis is usually reduced to a simpler characterization, based on first and total-order sensitivity indices associated to individual variables $N_i = \boldsymbol{N}_{\{i\}}$ or group of variables $\boldsymbol{N_u}$.

The *first-order sensitivity index* $\mathtt{S}_{\mathbf{u}}$, defined as[12]

$$\mathtt{S}_{\mathbf{u}} \doteq \frac{1}{\mathbb{V}\{F\}}\sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\subseteq\mathbf{u}}}\mathbb{V}\{F_{\mathbf{s}}\} = \frac{\mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}\}}{\mathbb{V}\{F\}}, \quad (10)$$

represents the fraction of the variance $\mathbb{V}\{F\}$ due to the variable $\boldsymbol{N_u}$ only, *i.e.* without any interaction with variables in $\boldsymbol{N}_{\mathbf{u}_\sim}$. For instance,

$$\mathtt{S}_{\{1,2\}} = \frac{\mathbb{V}\{F_{\{1\}}\} + \mathbb{V}\{F_{\{2\}}\} + \mathbb{V}\{F_{\{1,2\}}\}}{\mathbb{V}\{F\}}.$$

The *total-order sensitivity index* $\mathtt{T}_{\mathbf{u}}$, defined as[12]

$$\mathtt{T}_{\mathbf{u}} \doteq \frac{1}{\mathbb{V}\{F\}}\sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\cap\mathbf{u}\neq\emptyset}}\mathbb{V}\{F_{\mathbf{s}}\}$$

$$= \frac{1}{\mathbb{V}\{F\}}\left(\sum_{\mathbf{s}\in\mathcal{D}\backslash\emptyset}\mathbb{V}\{F_{\boldsymbol{\nu}}\} - \sum_{\substack{\mathbf{s}\in\mathcal{D}\backslash\emptyset \\ \mathbf{s}\subseteq\mathbf{u}_\sim}}\mathbb{V}\{F_{\mathbf{s}}\}\right)$$

$$= \frac{\mathbb{V}\{F\} - \mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathbf{u}_\sim}\}\}}{\mathbb{V}\{F\}} = 1 - \mathtt{S}_{\mathbf{u}_\sim}, \quad (11)$$

accounts in addition for the interactions between variables $\boldsymbol{N_u}$ and $\boldsymbol{N}_{\mathbf{u}_\sim}$. For our example with $\mathrm{D} = 3$, we have

$$\mathbb{V}\{F\}\mathtt{T}_{\{1,2\}} = \mathbb{V}\{F_{\{1\}}\} + \mathbb{V}\{F_{\{2\}}\} + \mathbb{V}\{F_{\{1,2\}}\}$$

$$+\mathbb{V}\{F_{\{1,3\}}\} + \mathbb{V}\{F_{\{2,3\}}\} + \mathbb{V}\{F_{\{1,2,3\}}\}$$

$$= \mathbb{V}\{F\} - \mathbb{V}\{F_{\{3\}}\} = \mathbb{V}\{F\}(1 - \mathtt{S}_{\{3\}}).$$

In fact, it is immediate to show from the definition of the total-order sensitivity index that $\mathtt{S}_{\mathbf{u}} \leq \mathtt{T}_{\mathbf{u}} \leq 1$. In addition, the following inequalities hold for the case of sensitivity indices associated to a single variable:

$$\sum_{i=1}^{\mathrm{D}}\mathtt{S}_{\{i\}} \leq 1, \quad \sum_{i=1}^{\mathrm{D}}\mathtt{T}_{\{i\}} \geq 1,$$

while

$$\sum_{i=1}^{\mathrm{D}}\mathtt{S}_{\{i\}} = 1 \Rightarrow \mathtt{T}_{\{i\}} = \mathtt{S}_{\{i\}} \text{ for } i = 1,\ldots,\mathrm{D}.$$

In the latter case, F is said to be additive since there is no interaction between the N_i's.

C. Monte Carlo estimation of the sensitivity indices

From (10) and (11) it is seen that the computation of the first and total order sensitivity indices essentially amounts to the determination of the variance of conditional expectations. For this purpose, we shall rely on Monte Carlo method proposed in Ref. 24.

Consider two *independent* random sample sets $\mathcal{N}^I$ and $\mathcal{N}^{II}$ of M realizations of $\boldsymbol{N}$. The conditional variance $\mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}\}$ can be estimated using the following average[24]

$$\mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N_u}\}\} + \mathbb{E}\{F\}^2 =$$

$$\lim_{M\to\infty}\frac{1}{M}\sum_{i=1}^{M}F(\boldsymbol{N}_{\mathbf{u}}^{I,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{I,(i)})F(\boldsymbol{N}_{\mathbf{u}}^{I,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{II,(i)}). \quad (12)$$

Here we have denoted $\boldsymbol{N}^{I,(i)}$ and $\boldsymbol{N}^{II,(i)}$ the i-th elements of $\mathcal{N}^I$ and $\mathcal{N}^{II}$ respectively. Introducing the classical ($\mathcal{N}^I$-)sample estimator for the mean

$$\widehat{\mathbb{E}\{F\}} \approx \frac{1}{M}\sum_{i=1}^{M}F(\boldsymbol{N}^{I,(i)}),$$

and variance of F,

$$\widehat{\mathbb{V}\{F\}} = \frac{1}{M-1}\sum_{i=1}^{M}\left(F(\boldsymbol{N}^{I,(i)}) - \widehat{\mathbb{E}\{F\}}\right)^2,$$

the Monte Carlo estimator of the first-order sensitivity index is

$$\widehat{\mathtt{S}_{\mathbf{u}}} = \frac{\frac{1}{M}\sum_{i=1}^{M}F(\boldsymbol{N}^{I,(i)})F(\boldsymbol{N}_{\mathbf{u}}^{I,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{II,(i)}) - \widehat{\mathbb{E}\{F\}}^2}{\widehat{\mathbb{V}\{F\}}}.$$

$$(13)$$

A similar expression can be derived for the estimation of the total-order sensitivity index $\mathtt{T_u}$ in (11),

$$\widehat{\mathtt{T_u}} = 1 - \frac{\frac{1}{M}\sum_{i=1}^{M} F(\boldsymbol{N}^{I,(i)}) F(\boldsymbol{N}_{\mathbf{u}}^{II,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{I,(i)}) - \widehat{\mathbb{E}\{F\}}^2}{\widehat{\mathbb{V}\{F\}}}.$$
(14)

It is seen that the MC evaluation of the sensitivity indices requires M evaluations of F (for the elements of $\mathcal{N}^I$), and M new function evaluations for each $\mathtt{S_u}$ or $\mathtt{T_u}$. In particular, the MC estimation of the first and total-order sensitivity indices $\mathtt{S}_{\{i\}}$ and $\mathtt{T}_{\{i\}}$ for the D input variables requires a total of $(2D+1)M$ model evaluations.

IV. APPLICATION TO STOCHASTIC SIMULATORS

Our objective is now to propose a decomposition of the variance for a functional $g(\boldsymbol{X}(t))$, where $\boldsymbol{X}(t)$ is the stochastic output of a stochastic simulator. In the examples of the next section, the functional g is simply a component of $\boldsymbol{X}(t)$ at some time $t = T$, but the proposed methodology can be readily extended to more complex functionals including path integrals and exit times, provided that $g(\boldsymbol{X}(t))$ is a second-order random variable.

To decompose $\mathbb{V}\{g(\boldsymbol{X}(t))\}$ into contributions arising from the different K_r reaction channels and their interaction, and accordingly define the first and total-order sensitivity indices, associated with the variability of an individual channel (or group of reaction channels), an appropriate definition of the conditional expectations is required. Specifically, we need to provide a clear meaning to the expectation of $g(\boldsymbol{X}(t))$ conditioned on a realization of a particular (or group of) reaction channel(s).

A. Conditional expectations

The central idea proposed in the present work is to identify a realization of the reaction channels with the realization of the underlying vector of independent standard Poisson processes

$$\boldsymbol{N}(\omega) \doteq (N_1(\omega), \ldots, N_{\mathrm{K}_r}(\omega)).$$

To this end, let us denote $\boldsymbol{X}(t, \boldsymbol{N}(\omega))$ the solution of the Poisson process formulation of the stochastic dynamics given by (2), that is

$$\boldsymbol{X}(t, \boldsymbol{N}(\omega)) = \boldsymbol{X}(t_0) + \sum_{j=1}^{\mathrm{K}_r} \boldsymbol{\nu}_j N_j(t_j, \omega),$$

where

$$t_j = \int_{t_0}^{t} a_j(\boldsymbol{X}(\tau, \boldsymbol{N}(\omega))) d\tau, \quad j = 1, \ldots, K.$$

Then, defining $F(\boldsymbol{N}(\omega)) \doteq g(\boldsymbol{X}(t, \boldsymbol{N}(\omega)))$, with $\boldsymbol{X}(t, \boldsymbol{N}(\omega))$ given by (2), and assuming F has finite second order moment, the variance decomposition with respect to the Poisson processes in $\boldsymbol{N}$ can proceed as discussed previously in Section III, because the K_r channels are associated with *independent* standard Poisson processes $N_k(\omega)$. To this end, let $\mathcal{K}$ be the power set of $\{1, \ldots, \mathrm{K}_r\}$; for $\mathbf{u} \in \mathcal{K}$, the conditional expectation of $g(\boldsymbol{X}(t, \boldsymbol{N}))$ given $\boldsymbol{N_u} = \boldsymbol{n_u}$ becomes

$$\mathbb{E}\left\{ g(\boldsymbol{X}(t, \boldsymbol{N})) \mid \boldsymbol{N_u} = \boldsymbol{n_u} \right\} =$$

$$\mathbb{E}\left\{ g\left(\boldsymbol{X}(t_0) + \sum_{j\in\mathbf{u}} \boldsymbol{\nu}_j n_j(t_i) + \sum_{j\in\mathbf{u}_\sim} \boldsymbol{\nu}_j N_j(t_j, \omega) \right) \right\}.$$
(15)

Observe that given $\boldsymbol{n}_u$, the physical times and sequence of firing reaction channels remains in general stochastic, that is dependent on $\boldsymbol{N_{u_\sim}}$, even for the channels R_j with index $j \in \mathbf{u}$, owing to the definition of the scaled time t_j: the conditioning fixes the sequence of firing channels only in the unscaled time. However, $\mathbb{V}\{\mathbb{E}\{g(\boldsymbol{X}(t, \boldsymbol{N})) \mid \boldsymbol{N_u}\}\}$ does measure the variability in $g(\boldsymbol{X}(t))$ induced by the underlying stochastic processes $N_j(\omega)$ with $j \in \mathbf{u}$, and so characterizes the variance caused by the inherent stochasticity in channels R_j, $j \in \mathbf{u}$. As a result, the sensitivity indices defined above can be used to quantify and characterize the impact of individual or group of reaction channels on the variability of $g(\boldsymbol{X}(t))$. Even though the indices are not characterizing a sensitivity with respect to some changes in the definition of the propensity functions associated to the channels (case of parametric studies), but rather characterize a contribution to the inherent stochasticity of the system, we shall continue to refer to them as sensitivity indices.

B. Implementation details

We complete the section by discussing the practical implementation of the Monte Carlo procedure for the estimation of the sensitivity indices of $g(\boldsymbol{X}(t, \boldsymbol{N}))$, and drop the time dependence of $\boldsymbol{X}$ for simplicity.

In addition to the classical mean and variance estimations for $g(\boldsymbol{X}(\boldsymbol{N}))$, the Monte Carlo estimation of sensitivity indices involves correlations of the form (see (13) and (14))

$$C_{\mathbf{u}} = \frac{1}{M}\sum_{i=1}^{M} g(\boldsymbol{X}(\boldsymbol{N}_{\mathbf{u}}^{I,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{I,(i)})) g(\boldsymbol{X}(\boldsymbol{N}_{\mathbf{u}}^{I,(i)}, \boldsymbol{N}_{\mathbf{u}_\sim}^{II,(i)})),$$
(16)

where $\boldsymbol{N}^{I,(i)}$ and $\boldsymbol{N}^{II,(i)}$ are elements of two independent sample sets of $\boldsymbol{N}(\omega)$.

Direct sampling of the vector of K_r independent standard Poisson processes is generally not practical, as it would require the storage of a prohibitively large number of firing times for each channel. Instead, it is more

convenient to construct the Poisson processes "on the fly", as shown in Algorithm 2, where the next unscaled firing time τ_j^+ for channel j is determined only when it actually fires (that is when the unscaled time τ_j reaches τ_j^+).

However, for the correct MC estimation of the correlation in (16), it is crucial to ensure that the sequence of (unscaled) firing times τ_j^+ remains the same for the channels $j \in \mathfrak{u}$ when computing $\boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{I,(i)})$ and $\boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{II,(i)})$. In other words, the same sequence of time increments between the successive firing events of channel $j \in \mathfrak{u}$ must be repeated; because the time increment between two successive firing events for channel j is simulated by means of the pseudo-random number generators $\texttt{RG}_j$, as shown in line 20 of Algorithm 2, this implies that the generator $\texttt{RG}_j$ must reproduce the same sequence of pseudo-random numbers r_j when computing $g(\boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{I,(i)}))$ and $g(\boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{II,(i)}))$, whenever $j \in \mathfrak{u}$. On the contrary, for a channel index $j \notin \mathfrak{u}$, two independent sequences of random numbers r_j have to be used. The control of sequence of pseudo-random numbers delivered by $\texttt{RG}_j$ can be classically enforced through the "seed" of the generator. In this context, the Poisson process $N_j(\tau)$ is totally determined by $\texttt{RG}_j$ and its seed, s_j. For such generators, consistent solutions

$$\boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{I,(i)}) \text{ and } \boldsymbol{X}(\boldsymbol{N}_{\mathfrak{u}}^{I,(i)}, \boldsymbol{N}_{\mathfrak{u}_\sim}^{II,(i)})$$

are obtained by calling Algorithm 2 twice, with the same initialization (seeding) $\texttt{RG}_j(s_j)$ if $j \in \mathfrak{u}$, and two random and independent realizations $\texttt{RG}_j(s_j^I)$ and $\texttt{RG}_j(s_j^{II})$ for $j \in \mathfrak{u}_\sim$.

The whole procedure for the Monte Carlo estimation of the whole set of K_r first and total-order sensitivity indices $\texttt{S}_{\{i\}}$ and $\texttt{T}_{\{i\}}$ is schematically illustrated in Algorithm 3, in the case of $g(\boldsymbol{X}) = g(\boldsymbol{X}(t = T))$. Note that Algorithm 3 is not optimized; in particular the computationally intensive part of the algorithm, namely the $(\text{K}_r + 1)$ calls to $\texttt{NRA}$ can be carried out in parallel. Nonetheless, the present formulation provides a general framework that can be easily adapted to the estimation of generic sensitivity indices associated to $\mathfrak{u} \in \mathcal{K}$, and to other functionals $g(\boldsymbol{X})$.

V. EXAMPLES

In this section, we apply and illustrate the proposed sensitivity analysis and variance decomposition methods to well-known stochastic models.

A. Birth-Death model

The birth-death (BD) process[25] involves a single species S ($\text{M}_s = 1$) and $\text{K}_r = 2$ reaction channels:

$$\emptyset \xrightarrow{b} S, \quad S \xrightarrow{d} \emptyset,$$
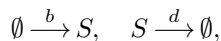
ALGORITHM 3. Computation of the first and total-order sensitivity indices $\texttt{S}_{\{j\}}$ and $\texttt{T}_{\{j\}}$ of $g(\boldsymbol{X}(T))$.

Procedure $\texttt{Compute\_SI}(M, \boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, g)$

Require: Sample set dimension M, initial condition $\boldsymbol{X}_0$, final time T, state-change vectors $\{\boldsymbol{\nu}_j\}$, propensity functions $\{a_j\}$ and functional g

1: $\mu \leftarrow 0$, $\sigma^2 \leftarrow 0$ ⊳ Init. Mean and Variance
2: **for** $j = 1$ to K_r **do**
3: $\quad \texttt{S}(j) \leftarrow 0$, $\texttt{T}(j) \leftarrow 0$ ⊳ Init. first and total-order SIs
4: **end for**
5: **for** $m = 1$ to $m = M$ **do**
6: $\quad$ Draw two independent set of seeds $\boldsymbol{s}^I$ and $\boldsymbol{s}^{II}$
7: $\quad \boldsymbol{X} \leftarrow \texttt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \texttt{RG}_1(s_1^I), \ldots, \texttt{RG}_{\text{K}_r}(s_{\text{K}_r}^I))$
8: $\quad \mu \leftarrow \mu + g(\boldsymbol{X})$, $\sigma^2 \leftarrow \sigma^2 + g(\boldsymbol{X})^2$ ⊳ Acc. mean and variance
9: $\quad$ **for** $j = 1$ to K_r **do**
10: $\quad\quad \boldsymbol{X}_\texttt{S} \leftarrow \texttt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \texttt{RG}_1(s_1^{II}), \ldots,$
11: $\quad\quad\quad\quad \ldots, \texttt{RG}_j(s_j^I), \ldots, \texttt{RG}_{\text{K}_r}(s_{\text{K}_r}^{II}))$
12: $\quad\quad \boldsymbol{X}_\texttt{T} \leftarrow \texttt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \texttt{RG}_1(s_1^I), \ldots,$
13: $\quad\quad\quad\quad \ldots, \texttt{RG}_j(s_j^{II}), \ldots, \texttt{RG}_{\text{K}_r}(s_{\text{K}_r}^I))$
14: $\quad\quad \texttt{S}(j) \leftarrow \texttt{S}(j) + g(\boldsymbol{X}) \times g(\boldsymbol{X}_\texttt{S})$ ⊳ Acc. 1-st order
15: $\quad\quad \texttt{T}(j) \leftarrow \texttt{T}(j) + g(\boldsymbol{X}) \times g(\boldsymbol{X}_\texttt{T})$ ⊳ Acc. total order
16: $\quad$ **end for** ⊳ Next channel
17: **end for** ⊳ Next sample
18: $\mu \leftarrow \mu/M$, $\sigma^2 \leftarrow \sigma^2/(M-1) - \mu^2$
19: **for** $j = 1$ to K_r **do**
20: $\quad \texttt{S}(j) \leftarrow \frac{\texttt{S}(j)}{(M-1)\sigma^2} - \frac{\mu^2}{\sigma^2}$ ⊳ Estim. 1-st order
21: $\quad \texttt{T}(j) \leftarrow 1 - \frac{\texttt{T}(j)}{(M-1)\sigma^2} + \frac{\mu^2}{\sigma^2}$ ⊳ Estim. total order
22: **end for**
23: **Return** $\texttt{S}(j)$ and $\texttt{T}(j)$, $j = 1, \ldots, \text{K}_r$ ⊳ First and total-order sensitivity indices $\texttt{S}_{\{j\}}$ and $\texttt{T}_{\{j\}}$ of $g(\boldsymbol{X}(T))$

---

with propensity functions

$$a_1(x) = b, \quad a_2(x) = d \times x.$$

We set $b = 200$, $d = 1$, and use $M = 1,000,000$ Monte Carlo samples to compute the estimates presented in this section.

Figure 1 reports typical trajectories of the system for the initial condition $X(t = 0) = 0$. The plot shows the transient growth of X toward its asymptotic distribution, which is the Poisson distribution with mean $b/d = 200$. As depicted in Figure 2, the histogram of $X(t = 8)$ is very close to the asymptotic distribution.

For this simple model with only two reaction channels, R_b and R_d, the analysis of the variance for $g(X(t)) = X(t)$ yields only three terms corresponding to the two first-order sensitivity indices $\texttt{S}_{\{b\}}$ and $\texttt{S}_{\{d\}}$, and the mixed contribution $\texttt{S}'_{\{b,d\}} \doteq \mathbb{V}\{X_{\{b,d\}}\}/\mathbb{V}\{X\}$; we have

$$\texttt{T}_{\{b\}} = \texttt{S}_{\{b\}} + \texttt{S}'_{\{b,d\}}, \quad \texttt{T}_{\{d\}} = \texttt{S}_{\{d\}} + \texttt{S}'_{\{b,d\}},$$

and $\texttt{S}_{\{b\}} + \texttt{S}_{\{d\}} + \texttt{S}'_{\{b,d\}} = 1$.

Figure 3 shows the evolution of the first-order and total sensitivity indices of X for $t \in [0, 8]$. To appreciate the transient dynamic, the sensitivity indices have been scaled by the variance $\mathbb{V}\{X\}$ whose evolution is
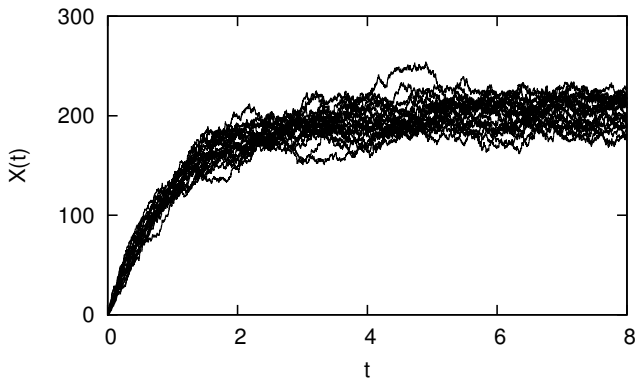
FIG. 1. Birth-Death model: Selected trajectories of $X(t)$ generated using Algorithm 2.
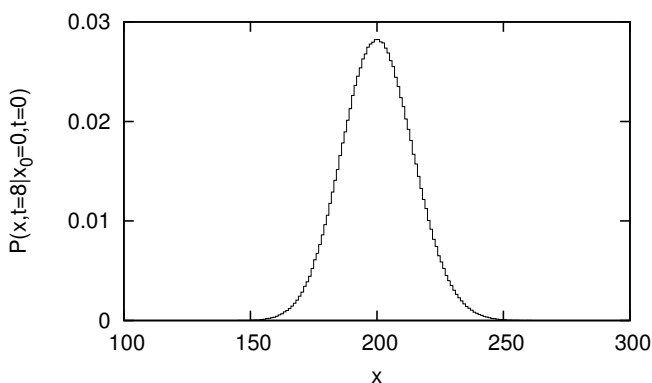

FIG. 2. Birth-Death model: Histogram of $X(t = 8)$. Results are based on 1 million realizations of the Poisson Process, Algorithm 2.

also provided. For the each reaction channel, the scaled first-order and total sensitivity indices are reported as a colored strip (red for R_b and blue for R_d) spanning vertically the interval $[\mathbb{V}\{X\}\,\mathtt{S}, \mathbb{V}\{X\}\,\mathtt{T}]$, so the extent of the strips is equal to $\mathbb{V}\{X_{\{b,d\}}\}$. The results show that at early times during the transient ($t < 1$) the variability in X is predominantly caused by the birth channel stochasticity, whereas the death channel induces a variability that grows in time at a much lower rate. This can be explained by the fact that R_b is a zero-order reaction with constant rate, whereas R_d is a first-order reaction and thus requires a sufficiently large population X to fire at a significant rate. Further, during this transient stage, the dynamics are essentially additive since $\mathtt{S}_{\{b\}} + \mathtt{S}_{\{d\}} \approx 1$. For $1 \leq t \leq 4$, the variability induced by R_d only, measured by $\mathtt{S}_{\{d\}}$, continues to grow with the population size (recall that $\mathbb{E}\{X\} = \mathbb{V}\{X\}$ for this model) and becomes eventually close to, but lower than $\mathtt{S}_{\{b\}}$. In the mean time, the mixed effects also increase, as reflected by the growing extent of the strips. Finally, for $t > 4$, interval for which the variance has essentially achieved its asymptotic value ($\mathbb{V}\{X\} \sim 200$), the first order sensitivity index $\mathtt{S}_{\{b\}}$ becomes constant while the

mixed effect $\mathtt{S}_{\{d\}}$ starts to decrease with a very low rate. The slow decay of $\mathtt{S}_{\{d\}}$ underlines the slow emergence of an increasingly important mixed contribution of the two channels on the variability of X.
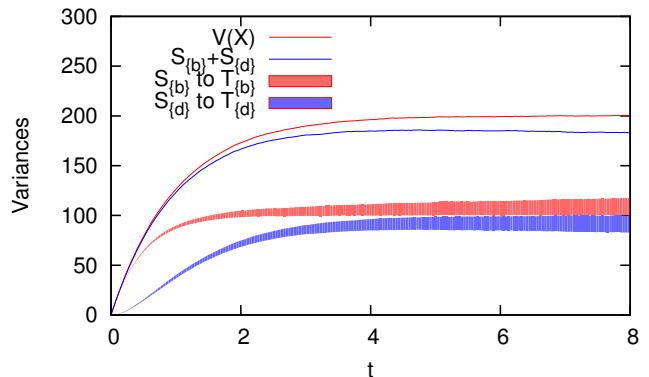

FIG. 3. Birth-death model. First-order and total sensitivity indices of the birth-death model and $t \in [0, 8]$. For clarity, the sensitivity indices have been scaled by $\mathbb{V}\{X(t)\}$.

To illustrate the channels interaction, we provide in Figure 4 the time evolution up to $t = 500$ of the first-order sensitivity indices $\mathtt{S}_{\{b\}}$ and $\mathtt{S}_{\{d\}}$, together with the interaction term $\mathtt{S}'_{\{b,d\}} = 1 - \mathtt{S}_{\{b\}} - \mathtt{S}_{\{d\}}$. The figure shows that while the variability induced by R_b remains constant for large t and amounts to 50% of the overall variance, the sole impact of channel R_d monotonically decays to the benefit of the mixed term. Again, this behavior is explained by the different natures of the two channels. First, because channel R_b is a zero-order reaction, the stochasticity that it induces is independent of the state X and therefore becomes asymptotically constant, accounting for half the variance. Second, equations (2) and (3) indicate that the contribution of channel R_d (a first-order reaction) to the dynamics of X, involves not only the intrinsic stochasticity of the standardized Poisson process N_d, but also the variability of the scaled-time t_d which involves the path-integral of $X(t)$. As a result, the variance due to $N_d(t_d)$ involves also the variability brought by channel R_b, through its effect on $X(t)$, and our computation quantifies how the sole effect of the variability of N_d monotonically decays while the interaction of the two channels grows.

B. Schlögl model

The Schlögl model[26] has $\mathrm{K}_r = 4$ reaction channels, expressed according to:

$$B_1 + 2S \underset{c_2}{\overset{c_1}{\rightleftharpoons}} 3S, \quad B_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} S.$$
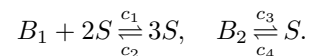
Species B_1 and B_2 are assumed in large excess with constant population over time: $X_{B_1} = X_{B_2}/2 = 10^5$. The stochastic state then reduces to the single evolving
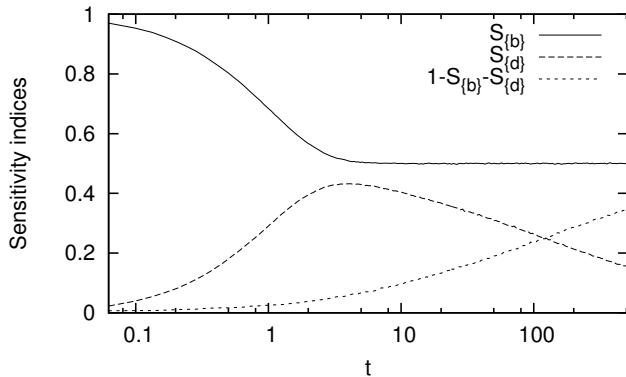
FIG. 4. Birth-death model. Long-time evolution of the first-order sensitivity indices, and of the mixed interaction term. Note the log-scale in time.

FIG. 5. Schlögl model. Selected trajectories of $X(t)$ showing the bifurcation in the stochastic dynamics.

species S with $M_s = 1$. The propensity functions are given by

$$a_1(x) = \frac{c_1}{2} X_{B_1} x(x-1), \quad a_2(x) = \frac{c_2}{6} x(x-1)(x-2),$$

and

$$a_3(x) = c_3 X_{B_2}, \quad a_4(x) = c_4 x.$$

We set $c_1 = 3 \times 10^{-7}$, $c_2 = 10^{-4}$, $c_3 = 10^{-3}$, and $c_4 = 3.5$, so that the dynamics involve a zero-order reaction (channel R_3), a first-order reaction (channel R_4), a second-order reaction (channel R_1) and a third-order reaction (channel R_2). In addition, we use the deterministic initial condition $X(t=0) = 250$.

For the present settings, the system exhibits a bifurcation with two attracting branches. This is illustrated in Figures 5 and 6, which depict selected trajectories of $X(t)$ and the empirical histogram of $X(t = 8)$ respectively. The bi-modal character of the solution at $t = 8$ is clearly evidenced, with a first peak at $x \sim 100$ that is well separated from a second peak occurring at $x \sim 600$. Also note the different spreads of the distributions, as the second peak is significantly broader than the first.

We first examine sensitivity indices of the K_r reaction channels on $g(X(t)) = X(t)$. For each channel, Figure 7 shows the ranges $[S_{\{j\}}(t), T_{\{j\}}(t)]$ scaled using the state variance $\mathbb{V}\{X(t)\}$ as scaling factor. Also plotted is the sum of the first-order contributions. Focusing first on the first-order indices, we observe that for $t \le 8$, we have

$$S_{\{3\}} \approx S_{\{2\}} < S_{\{4\}} \approx S_{\{1\}}.$$

Since most of the variance arises from the bi-modality of the solution, this suggests that channels R_1 and R_4 are the primary contributors to the solution branch selection. Figure 7 also indicates that up to $t \approx 2$ the system dynamics are essentially additive, while at later times mixed interactions between channels become increasingly significant. In fact, it is seen that the variance induced by the
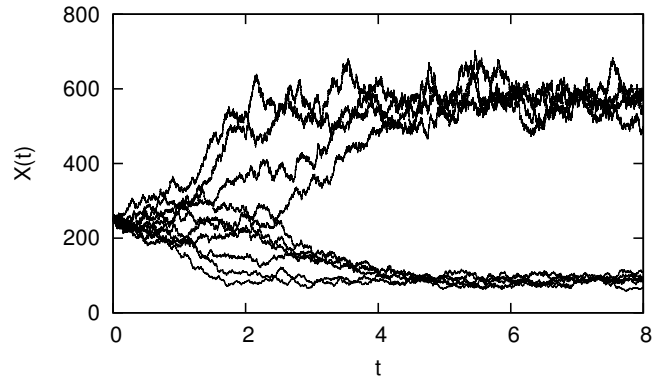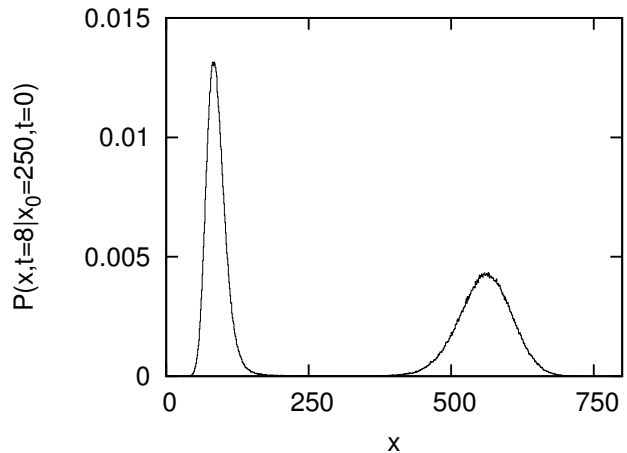

FIG. 6. Schlögl model. Histogram of $X(t = 8)$ based on one million realizations of the Poisson Process Algorithm 2, highlighting the bimodal nature of the stochastic solution.

first-order effects, $\mathbb{V}\{X\} \times \sum_{j=1}^{K_r} S_{\{j\}}$, i.e. by the individual channels without interaction, is essentially constant for $t > 4$ while the total variance continues to grow as interaction contributions develop. At $t = 8$, X is close to its asymptotic distribution, and the importance of the mixed effects in all the channels, $T_{\{j\}} - S_{\{j\}}$, can be appreciated from the extents of the colored strips in the plots. Again, mixed effects for channels R_1 and R_4 appear to be comparable, but significantly larger than they are in channels R_2 and R_3.

To gain further insight into the structure of the variability, the characterization of high-order partial variances is needed. For the present model with $K = 4$ we can afford to estimate all first-order sensitivity indices $S_{\mathbf{u}}$, $\mathbf{u} \in \mathcal{K}$, from which the partial variances $\mathbb{V}\{X_{\mathbf{u}}\}$ can be retrieved. We provide in Figure 8 the second-order partial variances $\mathbb{V}\{X_{\mathbf{u}}\}$ for $|\mathbf{u}| = 2$. It is seen that they all assume relatively low values ($\lesssim 500$) except for $\mathbb{V}\{X_{\{1,4\}}\}$ which asymptotically peaks at approximately $3,500$, confirming the preponderance of these two channels. Note
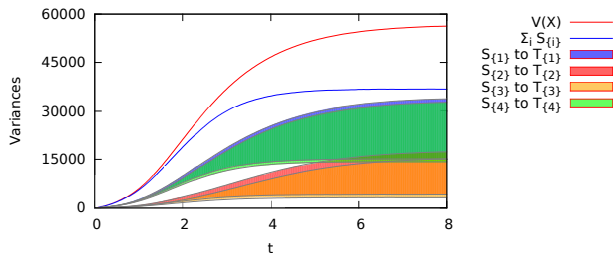
FIG. 7. Schlögl model. Sensitivity indices scaled by the local variance as well as the total variance and sum of first-order variance. The estimates are based on sample sets of one million trajectories of the NRA.

also that $\mathbb{V}\{X_{\{2,3\}}\}$ is estimated to be zero, so not only these two channels have lower first-order variance, but also have no second-order interaction.

Regarding the higher-order partial variances for $|\mathbf{u}| > 2$, reported in Figure 9, we observe again that $\mathbb{V}\{X_\mathbf{u}\}$ have significant values only for $\{1,4\} \in \mathbf{u}$. The partial variance $\mathbb{V}\{X_{\{1,2,3,4\}}\}$, which quantifies the variability due to the interaction between all the channels, becomes asymptotically dominant over all other partial variances except $\mathbb{V}\{X_{\{1,2,4\}}\}$ and $\mathbb{V}\{X_{\{1,3,4\}}\}$, underscoring the importance of complex high-order interactions between the channels. Further, unlike the third-order partial variances, $\mathbb{V}\{X_{\{1,2,3,4\}}\}$ has not yet converged at $t = 8$, highlighting the slower time scales at which full interactions develop.
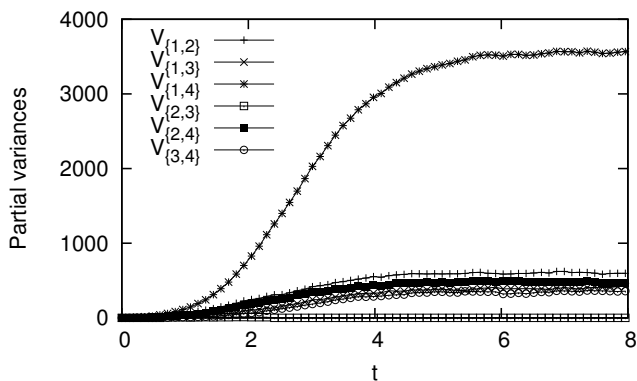
FIG. 8. Schlögl model. Evolution of second-order partial variances. The estimates are based on sample sets of one million trajectories of the NRA.

To support our conclusion that channels R_1 and R_4 are the main sources of stochasticity, through their role as the dominant mechanism for selecting the bifurcation branch, we provide in Figure 10 sample sets of trajectories of X for $t \in [0,8]$ conditioned on the Poisson processes (N_1, N_4) and (N_2, N_3) respectively. In Figure 10(a) we show, for 9 realizations (n_1, n_4) of $(N_1(\omega), N_4(\omega))$, 10 trajectories of X. Similarly, Figure 10(b) depicts, for 9 realizations (n_2, n_3) of
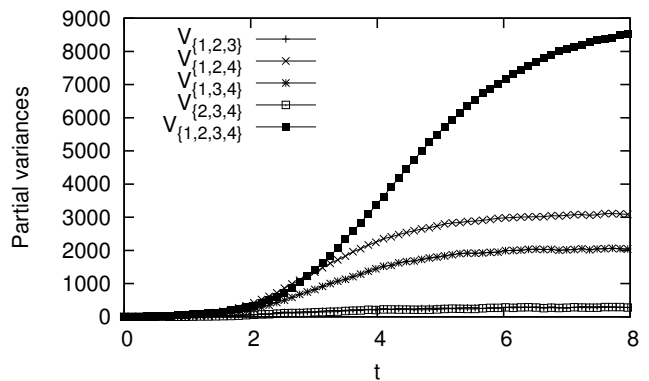

FIG. 9. Schlögl model. Evolution of third-order and fourth-order partial variances. The estimates are based on sample sets of one million trajectories of the NRA.

$(N_2(\omega), N_3(\omega))$, 10 trajectories of X. The results indicate that given the Poisson processes of channels R_1 and R_4 the stochastic system tends to select more systematically one of the two attracting branches of the dynamics. In contrast, the branch selection appears to be less influenced by channels R_2 and R_3.

C. Michaelis-Menten model

The Michaelis-Menten system[27] has $M_s = 4$ species and $K_r = 3$ reaction channels, and is expressed according to:

$$S_1 + S_2 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} S_3, \quad S_3 \overset{c_3}{\rightarrow} S_4 + S_2.$$
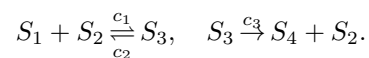
The system models the creation of a product S_4 through the binding of the enzyme S_2 with the substrate S_1; S_3 is an intermediate species which can dissociate back into enzyme and substrate (channel 2) or decompose to product S_4 and enzyme S_2 (channel 3). Note that $X_2 + X_3$ is a conserved quantity. The propensity functions of the model are given by $a_1(\boldsymbol{x}) = c_1 x_1 x_2$, $a_2(\boldsymbol{x}) = c_2 x_3$, and $c_3(\boldsymbol{x}) = c_3 x_3$. We set $c_1 = 0.0017$, $c_2 = 10^{-3}$ and $c_3 = 0.125$, and use as initial conditions $X_1(t = 0) = 300$, $X_2(t = 0) = 120$ and $X_3(t = 0) = X_4(t = 0) = 0$.

Figure 11 shows selected stochastic trajectories of the state vector. Qualitatively, the trajectories initially show a fast increase in the population of the intermediate species, S_3, with a decay in substrate and enzyme concentrations, respectively X_1 and X_2. When X_3 is large enough, S_3 starts to decompose producing S_4, whose population grows monotonically. As S_4 is produced, the substrate and intermediate species populations are progressively exhausted. As $t \to \infty$, (i) all the substrate will eventually by consumed into product, i.e. $X_4 \to X_1(t = 0)$; and (ii) the intermediate species population X_3 will vanish thus releasing the enzyme, thus $X_2 \to X_2(t = 0)$.

Figure 12 shows the time evolution of the first-order and total sensitivity indices of the three reaction chan-

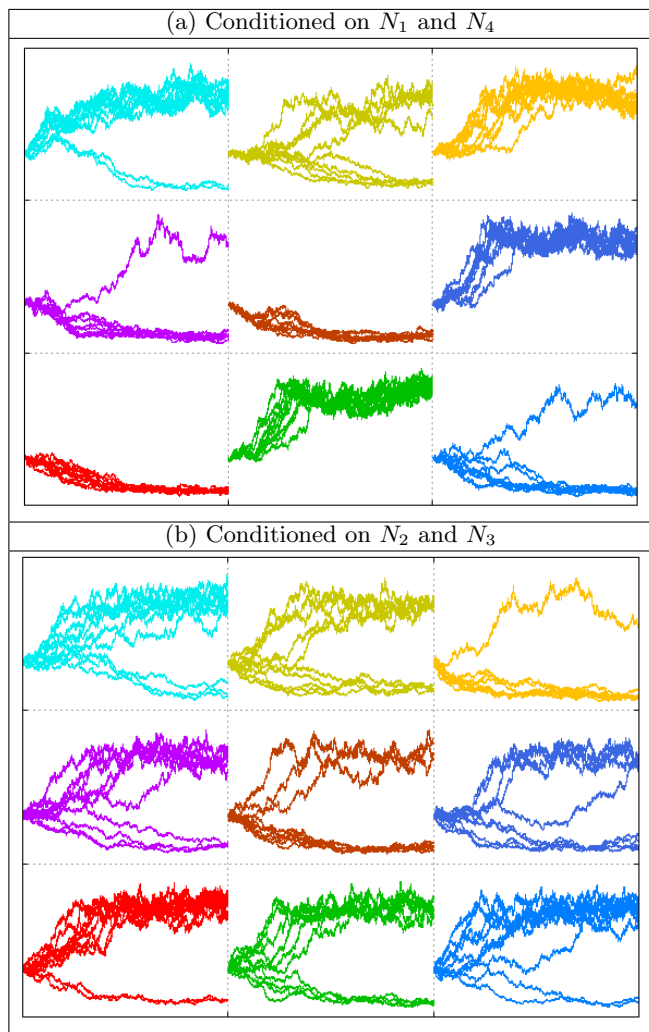FIG. 10. Schlögl model. Trajectories of $X(t)$ conditioned on (a) $N_1(\omega) = n_1$ and $N_4(\omega) = n_4$, and (b) $N_2(\omega) = n_2$ and $N_3(\omega) = n_3$. Each sub-plot shows 10 conditionally random trajectories for fixed realizations n_1 and n_4 in (a), and n_2 and n_3 in (b).
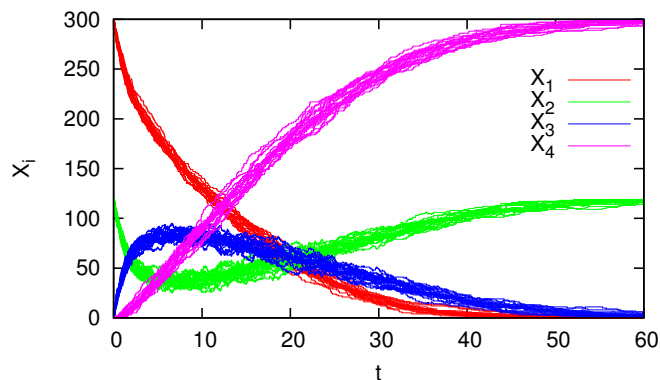

FIG. 11. Michaelis-Menten model. Selected trajectories of the state vector $\boldsymbol{X}$ for $t \in [0, 60]$.

nels, for species (a) S_1, (b) S_2 and (c) S_4 using decompositions of $g(\boldsymbol{X}) = X_i$ with $i = 1, 2$ and 4 respectively. Also plotted for each species is the scaled sum of first-order indices. Because of the conservation rule $X_2 + X_3 = $ const, the sensitivity indices for S_2 and S_3 are equal, and so the results for S_3 are not shown. Starting with S_1 (Fig. 12a), we see that the binding process R_1 is initially the main source of stochasticity; the variability due to the decomposition process R_3 then catches-up at later time, and even becomes dominant for $t \in [10, 20]$, before decaying to zero, while the effect of R_1 decays at a slower rate. At the time of maximum of variance in X_1, the two channels R_1 and R_3 contribute roughly equally to the variability. In contrast, the variability induced by channel R_2 (dissociation of S_3) is always negligible. It is also interesting to note that the effects of the channels are essentially additive.

Regarding S_2 (Fig. 12b), we first remark that, again, channel R_2 induces a negligible fraction of the variance. Second, as for S_1, channel R_1 is initially the dominant source of uncertainty, but R_3 catches up sooner, and then remains the main factor in $\mathbb{V}\{X_2\}$. Further, interaction effects between channel R_1 and R_3 appears to have a more significant impact on the variability of S_2 than for S_1.

Finally, the sensitivity analysis for the product species S_4 (Fig. 12c) highlights the dominance of channel R_3 on the variability at all time. In addition, the variability in X_4 caused by channel R_2 is negligible, and so are the interaction effects.

Overall, it can be concluded that channel R_2 (dissociation of S_3) has essentially no impact on the variability of $\boldsymbol{X}$. In fact, repeating the simulation disregarding channel R_2 (i.e. setting $c_2 = 0$) yields essentially similar variances (not shown), implying that for the present setting the dissociating process can be neglected, and the system accordingly simplified. However, this is not necessarily the case in general, e.g. when system parameters are changed. To demonstrate this claim, we provide in Figure 13 the same analysis as above, but now using $c_2 = 25 \times 10^{-3}$. Globally, the increase in c_2 leads to a significant contribution of channel R_2 to the variance of X_1, X_2, and X_3, and to a lesser extent to the variance of X_4. The increase in the dissociation rate also results in a noticeable increase in the effects of interaction terms on the variance of X_1 and X_2.

VI. DISCUSSION AND CONCLUSIONS

This work focused on the development of methods and algorithms that enable an orthogonal decomposition of the variance of in the functional of the output of a stochastic simulators. The key to this development consists in formulating the stochastic dynamics as being generated by independent standardized Poisson processes. At the manageable expense of representing the individual Poisson processes, this enables us to identify individual
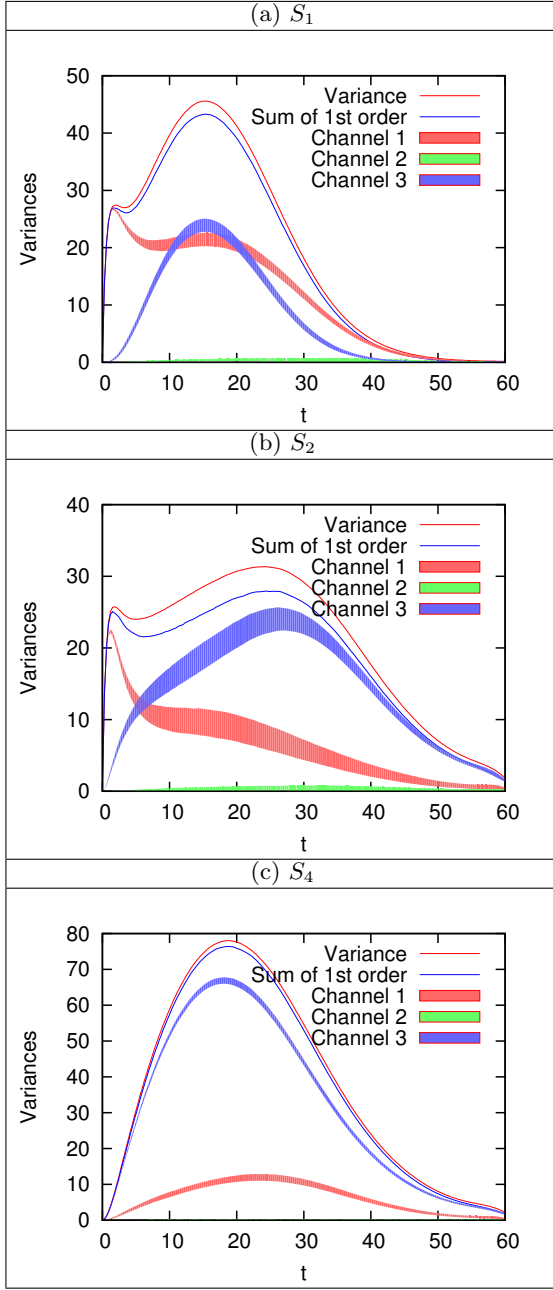
FIG. 12. Michaelis-Menten model: First-order and total sensitivity indices $S_{\{j\}}$ and $T_{\{j\}}$ for $j = 1, \ldots, 4$. Plots are generated for (a) X_1, (b) X_2 and (c) X_4. The estimates are based on a sample set of one million trajectories of the NRA.

FIG. 13. Michaelis-Menten model: same as in Figure 12 but with c_2 increased 25 folds.

realizations of the channels dynamics, and subsequently apply a Sobol decomposition with respect to the channels of the functional variance. A Monte-Carlo sampling approach was adopted for this purpose. This leads to an algorithm having a complexity that is linear in the number of conditional variances (sensitivity indices) that one wishes to compute. In addition, each conditional variance can be computed in parallel (lines 9–16 in Algorithm 3). Implementation of the algorithms resting on this foundation was illustrated through application to simple
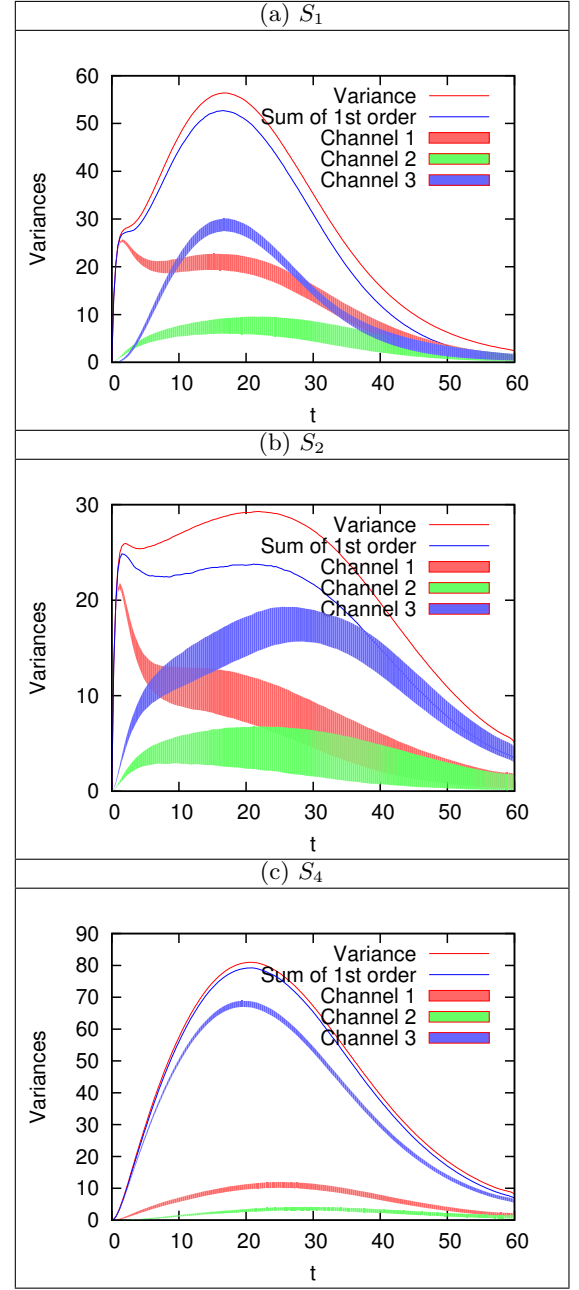
stochastic systems, namely the birth-death, Schlögl, and Michaelis-Menten models. The computations were used to demonstrate the possibility of quantifying contributions of individual Poisson processes associated with elementary reactions to the variance of species concentrations, of characterizing the role of mixed interactions, and consequently assessing the variability of the functional to reaction pathways.

While the scope of the computations performed was restricted to simple demonstrations, the methods developed in this work can be readily applied to investigate broader questions. An interesting area to explore would

consist in investigating more complex functionals of the solution. Examples would include variance decomposition of arrival, exit, means hitting times, reaction time scales, progress variables and full path functionals.

In addition, a number of possible algorithmic improvements and fundamental extensions can be conceived from the methodology and algorithms presented in the paper. In particular, complex systems involving large numbers of channels and species would benefit from the incorporation of more elaborate time-integrators; for instance, if an unaffordable number of elementary reactions fires before reaching the final time, tau-leap methods[4–6,28,29] should be considered to accelerate the simulations. Extending our methodology to tau-leap methods will require consistent time-approximations for the individual realizations of the Poisson processes associated with each channel. It would also be beneficial to rely on improved sampling strategies and variance reduction methods, such as multi-level Monte Carlo,[30–32] to reduce the number of simulations needed to estimate the sensitivity indices. On the fundamental side, interesting generalizations include the potential of leveraging the present approach for the purpose of model reduction,[33–36] and for quantifying the impact of uncertainties that may affect system parameters.[22,37–43] In particular, such extensions would offer the promise of accounting for and distinguishing between the impacts of parametric sensitivity and irreducible noise,[43] as well as providing estimates where reduced model approximations are valid. These topics are the subject of ongoing work.

ACKNOWLEDGMENTS

[1] D.T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403 – 434, 1976.

[2] D.T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[3] D.T. Gillespie. A rigorous derivation of the chemical master equation. *Physica A: Statistical Mechanics and its Applications*, 188(13):404 – 425, 1992.

[4] D.T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.

[5] D.T. Gillespie and L.R. Petzold. Improved leap-size selection for accelerated stochastic simulation. *The Journal of Chemical Physics*, 119(16):8229–8234, 2003.

[6] M. Rathinam, L.R. Petzold, Y. Cao, and D.T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *The Journal of Chemical Physics*, 119(24):12784–12794, 2003.

[7] W. E, D. Liu, and E. Vanden-Eijnden. Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales. *Journal of Computational Physics*, 221(1):158 – 180, 2007.

[8] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, 104:1876–1889, 2000.

[9] C. Gardiner. *Stochastic Methods*, volume 13 of *Springer Series in Synergetics*. Springer, 2009.

[10] S.N. Ethier and T.G. Kurtz. *Markov Processes: characterization and convergence*. Wiley Series in Probability and Statistics. Wiley, 2005.

[11] I.M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.*, 1:407–414, 1993.

[12] T. Homma and A. Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, 1996.

[13] G. Li, H. Rabitz, P.E. Yelvington, O.O. Oluwole, F. Bacon, C.E. Kolb, and J. Schoendorf. Global sensitivity analysis for systems with independent and/or correlated inputs. *The Journal of Physical Chemistry A*, 114(19):6022–6032, 2010.

[14] G. Chastaing, F. Gamboa, and C. Prieur. Generalized Sobol sensitivity indices for dependent variables: numerical methods. *Journal of Statistical Computation and Simulation*, 2014. to appear.

[15] G. Li, S.-W. Wang, H. Rabitz, S. Wang, and P. Jaffé. Global uncertainty assessments by high dimensional model representations (HDMR). *Chemical Engineering Sciences*, 57:4445–4460, 2002.

[16] T. Crestaux, O. Le Maître, and J.M. Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172, 2009.

[17] MM Zuniga, S Kucherenko, and N Shah. Metamodelling with independent and dependent inputs. *Computer Physics Communications*, 184:1570–1580, 2013.

[18] M. Rathinam, P.W. Sheppard, and M. Khammash. Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *The Journal of Chemical Physics*, 132:–, 2011.

[19] P.W. Sheppard, M. Rathinam, and M. Khammash. A pathwise derivative approach to the computation of parameter sensitivities in discrete stochastic chemical systems. *The Journal of Chemical Physics*, 136(3):–, 2012.

[20] E.S. Wolf and D.F. Anderson. A finite difference method for estimating second order parameter sensitivities of discrete stochastic chemical reaction networks. *The Journal of Chemical Physics*, 137(22):–, 2012.

[21] E.S. Wolf and D.F. Anderson. Hybrid pathwise sensitivity methods for discrete stochastic models of chemical reaction systems. *The Journal of Chemical Physics*, 142(3):–, 2015.

[22] A. Alexanderian, F. Rizzi, M. Rathinam, O.P. Le Maître, and O.M. Knio. Preconditioned Bayesian Regression for Stochastic Chemical Kinetics. *Journal of Scientific Computing*, 58:592–626, 2014.

[23] W. Hoeffding. A class of statistics with asymptotically normal distribution. *The annals of Mathematical Statistics*, 19:293–325, 1948.

[24] I.M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math. Comput. in Simul.*, 55:271–280, 2001.

[25] M.A. Nowak. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, 2006.

[26] F. Schlögl. Chemical reaction models fornon-equilibrium phase transitions. *Z. Phys.*, 253:247, 1972.

[27] D.J. Wilkinson. *Stochastic modeling for system biology*. Mathematical and Computational series. Chapman and Hall/CRC, 2006.

[28] D.F. Anderson. A modified next reaction method for simulating chemical systems with time dependent propensities and delays. *J. Chem. Phys.*, 127(21):214107, 2007.

[29] A. Moraes, R. Tempone, and P. Vilanova. Hybrid Chernoff Tau-Leap. *Multiscale Modeling & Simulation*, 12:581–615, 2014.

[30] M.B. Giles. Multi-level Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[31] D.F. Anderson and D.J. Higham. Multi-level Monte Carlo for continuous time Markov chains, with applications in biochemical kinetics. *Multiscale Model. Simul.*, 10(1):146–179, 2012.

[32] A. Moraes, R. Tempone, and P. Vilanova. Multilevel hybrid Chernoff Tau-Leap. *BIT Numer. Math.*, submitted.

[33] X. Han and H.N. Najm. Dynamical structures in stochastic chemical reaction systems. *SIAM J. Applied Dynamical Systems*, 13:1328–1351, 2014.

[34] M. Bruna, J.S. Chapman, and M.J. Smith. Model reduction for slow-fast stochastic systems with metastable behaviour. *J. Chem. Phys.*, 140:174107, 2014.

[35] M. Salloum, A. Alexanderian, O.P. Le Maître, H.N. Najm, and O.M. Knio. Simplified CSP Analysis of a Stiff Stochastic ODE System. *Computer Methods in Applied Mechanics and Engineering*, 217-220:121–138, 2012.

[36] C.A. Gómez-Uribe, G.C. Verghese, and A.R. Tzafriri. Enhanced identification and exploitation of time scales for model reduction in stochastic chemical kinetics. *J. Chem. Phys.*, 129:244112, 2008.

[37] K. Sargsyan, B. Debusschere, H.N. Najm, and Y. Marzouk. Bayesian inference of spectral expansions for predictability assessment in stochastic reaction networks. *Journal of Computational and Theoretical Nanoscience*, 6(10):2283–2297, 2009.

[38] K. Sargsyan, C. Safta, B. Debusschere, and H. Najm. Multi-parameter spectral representation of noise-induced competence in bacillus subtilis. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2012.

[39] F. Rizzi, H.N. Najm, B.J. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson, and O.M. Knio. Uncertainty quantification in MD simulations. Part I: Forward propagation. *Multiscale Model. Simul.*, 10:14281459, 2012.

[40] F. Rizzi, H.N. Najm, B.J. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson, and O.M. Knio. Uncertainty quantification in MD simulations. Part II: Bayesian inference of force-field parameters. *Multiscale Model. Simul.*, 10:14601492, 2012.

[41] F. Rizzi, R. Jones, B.J. Debusschere, and O.M. Knio. Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part I: sensitivity to physical parameters of the pore. *J. Chem. Phys.*, 138:194104, 2013.

[42] F. Rizzi, R. Jones, B.J. Debusschere, and O.M. Knio. Uncertainty quantification in MD simulations of concentration driven ionic flow through a silica nanopore. Part II: uncertain potential parameters. *J. Chem. Phys.*, 138:194105, 2013.

[43] O.P. Le Maître and O.M. Knio. PC analysis of stochastic differential equations driven by Wiener noise. *Reliability Engineering and System Safety*, 135:107–124, 2015.