# Global Sensitivity Analysis in Stochastic Simulators of Uncertain Reaction Networks

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

[1)]*CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.*

[2)]*LIMSI, CNRS, Université Paris-Saclay, France.*

[3)]*Department of Mechanical Engineering and Materials Science, Duke University, Durham, North Carolina 27708, USA.*

Stochastic models of chemical systems are often subjected to uncertainties in kinetic parameters in addition to the inherent random nature of their dynamics. Uncertainty quantification in such systems is generally achieved by mean of sensitivity analyses in which one characterizes the variability with the uncertain kinetic parameters of the first statistical moments of model predictions. In this work, we propose an original global sensitivity analysis method where the parametric and inherent variability sources are both treated through Sobol's decomposition of the variance into contributions from arbitrary subset of uncertain parameters *and* stochastic reaction channels. The conceptual development only assumes that the inherent and parametric sources are independent, and considers the Poisson processes in the random-time-change representation of the state dynamics as the fundamental objects governing the inherent stochasticity. A sampling algorithm is proposed to perform the global sensitivity analysis, and to estimate the partial variances and sensitivity indices characterizing the importance of the various sources of variability and their interactions. The birth-death and Schlögl models are used to illustrate both the implementation of the algorithm and the richness of proposed analysis method. The output of the proposed sensitivity analysis is also contrasted with a local derivative-based sensitivity analysis method classically used for this type of systems.

## I. INTRODUCTION

Stochastic models are widely used in many scientific fields, including population dynamics, queues and social-network studies, biosciences, and chemical systems. Stochastic modeling is necessary when deterministic models are not able to capture important features of the dynamics, for instance when large-scale effects of unresolved small-scale fluctuations occur, or when systems are subjected to important inherent noise. Regarding chemical modeling, which are the focus of the present work, deterministic models are usually unsuitable for describing the evolution of small systems where reactive species are present in small amounts, such that stochastic fluctuations have great effects on the system response and carry important meaning.

Chemical system modeling requires the selection of a reaction mechanisms (or pathways), that is the set of reactions between the species constituting the system, and the definition of propensity (or intensity) functions that relate the current state of the system to the probability of occurrence of a reaction in the next infinitesimal time interval. Algorithms and stochastic simulators can then be designed for the generation of paths, or trajectories, of the system's stochastic dynamics. Sample sets of paths can be used to compute, *à la* Monte Carlo, the statistical moments of some model output or, more generally, the expected value of some functional of the stochastic state.

Typically, the propensity functions involve kinetic parameters, such as thermodynamic variables and reaction parameters, *e.g.*, the mixture temperature and pressure, the rate constants, the pre-exponential factors and the activation energies of the Arrhenius law. Often, these kinetic parameters are not well known and must be considered as uncertain. In this case, it is critical to assess the impact of the uncertain kinetic parameters on the model predictions.[1] Classically, this is achieved by performing a sensitivity analysis (SA), which, roughly speaking, measures the variability of a model prediction when the uncertain parameters are varied.

A number of SA methods have been proposed for uncertain chemical systems.[2,3] Restricting the discussion to the case of stochastic models, SA methods have been mostly limited to the assessment of parametric variability in the first statistical moments (mean and variance) of the model output due to the uncertain kinetic parameters. Along these lines, one can distinguish between the local and global SA methods. The former essentially consist in the characterization of the variability in the statistical moments, based on their derivatives with respect to the kinetic parameters at some nominal values

---

[a)]Electronic mail: mariaisabel.navarrojimenez@kaust.edu.sa

[b)]Electronic mail: olm@limsi.fr

[c)]Electronic mail: omar.knio@kaust.edu.sa

(see for instance Refs. 4–7).

Global SA (GSA) methods, on the other hand, aim at accounting for the uncertainty effects when the parameters vary over their whole range. The GSA methods are often based on second-order characterizations of the variability (with respect to the parameters) using the so-called Sobol or ANOVA decompositions.[8,9] The Sobol decomposition can be determined efficiently relying on a projection step, where the dependences of the moments are first represented as Polynomial Chaos (PC) expansions,[10,11] before computing the partial variances and so-called sensitivity indices.[12,13]

More recently, we have proposed in Ref. 14 a different approach for GSA, for parametric uncertainties in stochastic systems. Compared to the classical approaches discussed above, the idea is to treat the inherent stochasticity of the system as an additional and independent source of uncertainty in the Sobol decomposition. Within this framework, one proceeds directly with the decomposition of the stochastic model output, and not with the decomposition of its statistical moments. This paradigm shift enables a much more detailed and finer GSA, such the assessment of the partial variances associated to the interactions between parameters and inherent stochasticity, a task that is not possible for a GSA of the moments.

In the GSA method proposed in Refs. 14 and 15, the decomposition of the variance was performed by relying on PC expansions of the individual random trajectories. This was possible owing to the smoothness (differentiability) of the stochastic trajectories with respect to the parameters. Such path-wise smoothness is not to be expected in general for the systems considered in the present work, either because of strong non-linearities and bifurcations, or because of the discrete nature of the stochastic state spaces (typically, the system state is integer valued). Another important challenge in extending the GSA method of Refs. 14 and 15 to stochastic simulators, as explained later, comes from the need to perform conditional re-sampling of the trajectories. Resampling trajectories conditioned on fixed kinetic parameter values is straightforward in the context of stochastic simulators. In contrast, the meaning of resampling trajectories conditioned on fixed "values" of the inherent stochasticity must be clarified and properly defined, prior to be implemented in a stochastic simulator.

This issue was recently addressed in Ref. 16, where it has been proposed to identify a particular realization of the inherent stochasticity with the corresponding realization of the standard Poisson processes appearing the so-called random-time-change representation of the dynamics.[17] Adopting this definition, we were able in Ref. 16 to perform an analysis of the relative influences of different reaction channels on the total variability of some model outputs of a stochastic simulator.

In the present paper, we generalize the methodology in Ref. 16 to propose a GSA method for stochastic simulators having additionally uncertain parameters. In doing so, we aim to preserve the ability to perform a global

variance decomposition, namely one that is not simply based on SA of statistical moments. Achieving this new capability for stochastic simulators requires us to devise effective means to distinguish between the effects of parametric uncertainty and state-dependent stochastic variability, and to quantify their joint contributions. The presently proposed methodology enables us to overcome associated challenges, and in addition offers several key features. First, the method is global and can be applied to models having arbitrary levels of parametric uncertainty as well as inherent variability. Second, the method enables a detailed GSA, through the quantification of the partial variances due to individual parameters or groups of parameters, randomness in individual or groups of the channel dynamics (inherent stochasticity), and finally the interaction between arbitrary groupings of parameters and random channels. Finally, the method is purely based on conditional sampling of the model and is then easily amenable to stochastic simulators relying on a random-time-change formulation, as well as being trivially implemented in parallel.

The outline of the paper is as follows. In Sec. II, we briefly discuss the stochastic simulators and introduce the modified next reaction method, which is subsequently used to perform the conditional sampling of the stochastic model. Sec. III details the proposed GSA method, starting with a discussion of the parametric and inherent variabilities in Sec. III A. The Sobol-Hoeffding variance decomposition and the associated sensitivity indices are introduced in Sec. III B, starting from the simple separation of parametric uncertainty and random channels variability and leading to the computation of general sensitivity indices. A numerical algorithm performing the GSA and its practical implementation are described in Sec. III C. The proposed method is illustrated based on two simple stochastic models, namely the birth-death process, in Sec. IV, and the Schlögl model, in Sec. V. The results of the GSA are contrasted with a local-derivative-based SA in order to highlight the robustness and richness of the proposed method. Concluding remarks are provided in Sec. VI.

## II. STOCHASTIC CHEMICAL SYSTEMS AND SIMULATOR

In this section, we review background materials on the stochastic modeling of chemical systems and relevant simulation methods. In Sec. II B, we introduce the Modified Next Reaction Method (MNRM) which is the basis of subsequent developments.

### A. Stochastic chemical systems

Let $(\Omega, \Sigma, \mu)$ be a probability space, where $\Omega$ is the event space, $\Sigma$ denotes a $\sigma$-algebra on $\Omega$, and $\mu$ is a probability measure. We consider a chemical reaction sys-

tem involving $M_s$ chemical species, labeled $S_1, \ldots, S_{M_s}$, whose evolutions in continuous time are discrete and stochastic. The random state of the system is $\boldsymbol{X}$ : $[0, T] \times \Omega \rightarrow \mathbb{N}_0^{M_s}$, where the $i$-th component $X_i(t)$ is the number of molecules of species $S_i$ at time $t$. The state vector $\boldsymbol{X}$ evolves through $M_r$ reactions channels, $R_1, \ldots, R_{M_r}$, each characterized by a propensity or intensity function $a_j$ and a stoichiometric vector $\boldsymbol{\nu}_j$. The stoichiometric vector $\boldsymbol{\nu}_j \in \mathbb{Z}^{M_s}$ prescribes the changes in the population of the reactant species due to the occurrence of a reaction in channel $R_j$, and are therefore independent of the state and time. The propensity functions, on the contrary, are dependent on the system state and thermodynamical conditions and, therefore, they change in time. For simplicity, we shall restrict ourselves to well-stirred systems at constant temperature and pressure, such that the propensity functions can be expressed as $a_j(\boldsymbol{X}(t); \boldsymbol{c})$ where $\boldsymbol{c} = (c_1, \ldots, c_{N_c}) \in \mathbb{R}^{N_c}$ is a vector of real valued kinetic parameters which are not time dependent. For $\tau \ll 1$, $a_j(\boldsymbol{X}(t); \boldsymbol{c})\tau$ is the probability that, given the current system state $\boldsymbol{X}(t)$, a reaction along channel $R_j$ occurs during the small time interval $(t, t + \tau]$. In other words, the random process $\boldsymbol{X}(t)$ is a continuous-time Markov chain defined by the set of transition probabilities: for $j = 1, \ldots, M_r$,

$$P[\boldsymbol{X}(t + \tau) = \boldsymbol{X} + \boldsymbol{\nu}_j | \boldsymbol{X}(t) = \boldsymbol{X}] = a_j(\boldsymbol{X}; \boldsymbol{c})\tau. \quad (1)$$

Note that the propensity functions and stoichiometric vectors of the reaction channels must satisfy a non-negativity assumption, meaning that the dynamic can not produce a population with negative size.

The Chemical Master Equation[18,19] (CME) describes the evolution of the probability distribution of the stochastic state vector $\boldsymbol{X}(t)$; it is expressed as

$$\frac{\partial P_{t,t_0}[\boldsymbol{x}|\boldsymbol{x}_0]}{\partial t} = \sum_{j=1}^{M_r} (a_j(\boldsymbol{x} - \boldsymbol{\nu}_j; \boldsymbol{c}) P_{t,t_0}[\boldsymbol{x} - \boldsymbol{\nu}_j | \boldsymbol{x}_0] - a_j(\boldsymbol{x}; \boldsymbol{c}) P_{t,t_0}[\boldsymbol{x}|\boldsymbol{x}_0]). \quad (2)$$

In the CME (2), we denoted $P_{t,t_0}[\boldsymbol{x}|\boldsymbol{x}_0]$ the probability of the system to reach a state $\boldsymbol{x}$ at time $t \geq t_0$, knowing that it is in a state $\boldsymbol{x}_0$ at time $t_0$. Equation (2) has no general explicit solution, and its numerical solution is too costly to compute in most cases, especially when $M_s$ exceeds a few units. One has then resort to a simulation method to generate sample trajectories from the distribution that is the solution of Eq. (2) without having to explicitly solve the CME. One classically uses so-called exact simulation algorithms to compute such trajectories and estimate the expected value of some functional, $G$, of the random system state, through

$$\mathbb{E}\left[G(\boldsymbol{X})\right] = \int_\Omega G(\boldsymbol{X}(\omega)) d\mu(\omega) = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^M G(\boldsymbol{X}^{(i)}),$$

where $\omega \in \Omega$ denotes a particular random event and $\boldsymbol{X}^{(i)}$ are sampled trajectories of $\boldsymbol{X}$. Consequently, relying on a simulator or solving the CME is equivalent, although

using stochastic simulators has an inherent sampling error as a finite number $M$ of paths can be generated. Note that there also exist inexact stochastic simulation algorithms, that produce trajectories according to an approximation of the CME solution. Such approximate simulators are not considered in this work.

## B. Modified Next Reaction Method (MNRM)

One of the best-known exact simulation methods is the Stochastic Simulation Algorithm (SSA) which was first introduced by Gillespie.[20,21] Several modifications and variants of the original SSA have been proposed to improve its computational efficiency and adapt the algorithm to specific properties and characteristics of the simulated system (see for instance Refs. 22 and 23). For the purpose of decomposing the variance, as pursued below, a fine distinction and control of the individual channels dynamics is required.[16] The standard SSA is not suitable for purpose of performing the variance decomposition, because the algorithm combines all channels dynamics to determine the next reaction time and the first firing channel. For this reason, we consider the Modified Next Reaction Method[24] (MNRM) (see Algorithm II.1) which is a slightly modified version of the original Next Reaction Method (NRM).[25] The MNRM is an exact simulation algorithm making explicit the respective contributions of the $M_r$ channels in the evolution of the system state. This characteristic is subsequently exploited to perform the variance decomposition, as discussed in Sec. III.

The algorithm uses Kurtz' random time-change (RTC) representation[17] to simulate exact paths, expressing the random state $\boldsymbol{X}(t)$ through:

$$\boldsymbol{X}(t) = \boldsymbol{X}(t_0) + \sum_{j=1}^{M_r} \boldsymbol{\nu}_j Y_j(t_j), \quad (3)$$

where $Y_j$ are independent unit-rate Poisson processes with respective scaled times $t_j$ given by

$$t_j = \int_{t_0}^t a_j(\boldsymbol{X}(s); \boldsymbol{c}) ds, \quad j = 1, \ldots, M_r. \quad (4)$$

Equation (3) shows that the change in the state, from the initial condition $\boldsymbol{X}(t_0)$, is proportional to the number of times ($Y_j$) that each channel has fired, times their respective stoichiometric vectors $\boldsymbol{\nu}_j$. The random number of times a channel has fired is identified as the value of a standard (unit rate) Poisson process, evaluated at a scaled time given by the integral of the propensity (or intensity) function along the trajectory. Hence, in the MNRM, the randomness is encapsulated in the Poisson processes describing the channels' dynamics. This crucial point was exploited in our previous paper,[16] where control of the realizations of the Poisson processes was exploited to perform a variance decomposition and analyze the variability induced by different reaction channels.

An algorithm for the simulation of the stochastic evolution governed by Eq. (3) can be easily set up, remarking that both the system state $\boldsymbol{X}(s)$ and the propensity functions $a_j(\boldsymbol{X}(s); \boldsymbol{c})$ remains constant between successive reactions. Denote $t_j^+ > t_j$ the scaled time at which $Y_j$ jumps from $Y_j(t_j)$ to $Y_j(t_j) + 1$, the main idea of Algorithm II.1 is that, given the current state $\boldsymbol{X}(t)$, scaled times $t_j$ and paths of Poisson processes $Y_j$, we can determine from (4) the (physical) time increments $dt_j = (t_j^+ - t_j)/a_j$ till $Y_j(t_j)$ jumps to $Y_j + 1$. The minimal $dt_j$ hence determines both the time and reaction channel that will first fire. Therefore, it suffices to keep track of the set of $t_j^+$ and $t_j$ to simulate trajectories of the system, as shown in Algorithm II.1.

---

**Algorithm II.1** Modified Next Reaction Method

---

Function MNRM($\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \boldsymbol{c}, \mathbf{RG}$)

**Input:** Initial condition $\boldsymbol{X}_0$, final time $T$, stoichiometric vector $\{\boldsymbol{\nu}_j\}$, propensity functions $\{a_j\}$, vector of kinetics parameters $\boldsymbol{c}$, and initialize pseudo-random number generators $\mathbf{RG} = (\mathrm{RG}_1, \ldots, \mathrm{RG}_{M_r})$

**Output:** State $\boldsymbol{X}(T)$

1: **for** $j = 1$ to $M_r$ **do**
2:    Draw $r_j$ from $\mathrm{RG}_j$
3:    $t_j \leftarrow 0$, $t_j^+ \leftarrow -\log(r_j)$     {set next reaction times}
4: **end for**
5: $t \leftarrow 0$, $\boldsymbol{X} \leftarrow \boldsymbol{X}_0$
6: **loop**
7:    **for** $j = 1$ to $M_r$ **do**
8:       Evaluate $a_j(\boldsymbol{X}; \boldsymbol{c})$ and $dt_j = \frac{t_j^+ - t_j}{a_j}$
9:    **end for**
10:   Set $l = \mathrm{argmin}_j dt_j$      {pick next reaction}
11:   **if** $t + dt_l > T$ **then**
12:     **Break**             {final time reached}
13:   **else**
14:     $\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l$    {update the state vector}
15:     $t \leftarrow t + dt_l$           {advance time}
16:     **for** $j = 1$ to $M_r$ **do**
17:       $t_j \leftarrow t_j + a_j dt_l$     {update scaled times}
18:     **end for**
19:     Get $r_l$ from $\mathrm{RG}_l$ and set $t_l^+ \leftarrow t_l^+ - \log(r_l)$   {next reaction time}
20:   **end if**
21: **end loop**

---

We remark that the time-marching method depicted in Algorithm II.1 does not require the storage of all the sequences of jumping times $t_j^+$. Instead, the sequences are determined on the fly using the random number generators (see line 19). In practice, the sequences of jumping times are entirely determined by the seeds initializing the pseudo-random number generators $\mathrm{RG}_j$, and the same sequences of random (unscaled) jumping times can be reproduced by reseeding the pseudo-random number generators accordingly. Another important remark concerning Algorithm II.1 is that it considers $M_r$ distinct random number generators $\mathrm{RG}_j$, one for each reaction channel. This may appear unnecessary, in view of the advancement in time performed in Algorithm II.1 as a single random number generator would be enough. In fact, the use of different random number generators allows for the control of the individual Poisson processes and consequently enable the variance decomposition which requires to reproduce exactly sequences of jumping times, as explained below. To this end, the possibility of identifying the sequence of jumping times for a channel to the random seed of its generator is critical for the computational efficiency because it alleviates the burden of storing sequences of sampled $t_j$.

## III. GLOBAL SENSITIVITY ANALYSIS

In this section, we introduce the variance decomposition method and the related global sensitivity analysis. As discussed in the introduction, our method differs from the classical sensitivity analysis Refs. 4–7 of stochastic systems, that are usually based on the estimation of the local sensitivity of selected statistical moments with respect to uncertain system parameters. Our method, on the contrary, aims at providing a direct second-order (variance-based) characterization of the respective variabilities induced by the uncertain parameters, globally, and quantifying separately the inherent stochasticity due to the channels. To this end, we build of the approach originally proposed in Ref. 16, which focused on the variance decomposition over the reaction channels without considering parametric uncertainty.

We start by discussing parametric uncertainties in the context of stochastic reaction networks in Sec. III A. The decomposition of the variance and the uncertainty characterization using sensitivity indices are detailed in Sec. III B. Finally, in Sec. III C, we describe a Monte Carlo sampling approach for the estimation of the sensitivity indices and we provide a brief outline of the algorithm and implementation used in the present work.

### A. Parametric and stochastic uncertainties

As mentioned before, stochastic systems can involve two sources of randomness. On the one hand, the random dynamics associated to the reaction channels lead to inherent variability that can not be reduced. On the other hand, the definition of the reaction networks involves several kinetic parameters that are usually estimated within a finite confidence level only. To properly apprehend stochastic simulations, it is therefore important to assess the respective impacts of these inherent and parametric uncertainties on the variability of a quantity of interest, $G(X)$, related to the system output.

A typical situation of parametric uncertainty, considered in this work, is the case of propensity functions $a_j$ not perfectly known, but involving some uncertain parameters. Practical examples of uncertainties in the propensity functions will be given in the following sections; we can already mention the case of uncertainties

in the reacting mixture temperature and pressure, in the Arrhenius activation energies, rate constants and pre-exponential factors of the different reaction channels. In these situations, it is relevant to treat the uncertain parameters as random quantities. To this end, we shall denote by $\boldsymbol{C}(\omega) \in \mathbb{R}^{N_c}$ the uncertain vector of random parameters in the propensity functions $a_j$ (recall that $\omega \in \Omega$ is a random event). We shall also denote $p_{\boldsymbol{C}} : \mathbb{R}^{N_c} \mapsto \mathbb{R}_+$ the probability density function of $\boldsymbol{C}$. As a result of the uncertainty in the parameters of the $a_j$, the dynamics given by Eq. (3) are random, because of the stochastic nature of the Poisson processes $Y_j$, and also because of the scaled times in Eq. (4) which involve propensity functions with now random coefficients $\boldsymbol{C}$. Specifically, in the case of random parameters $\boldsymbol{C}$, we have

$$a_j : \mathbb{N}_0^{M_s} \times \Omega \to a_j(\boldsymbol{x}; \boldsymbol{C}(\omega)) \in \mathbb{R}^+ \quad j = 1, \ldots, M_r.$$

In the following, we shall refer collectively to the set of $N_c$ random parameters $\boldsymbol{C}$, as the uncertain kinetic parameters. We shall also assume that the components of $\boldsymbol{C}$ are independent random variables. Consequently, the scaled times $t_j$ are given by:

$$t_j = \int_{t_0}^{t} a_j\left(\boldsymbol{X}(s); \boldsymbol{C}(\omega)\right) ds \quad j = 1, \ldots, M_r.$$

For convenience, we shall write the stochastic state as $\boldsymbol{X}(t, \boldsymbol{Y}, \boldsymbol{C})$ to underline its dependence on the random Poisson processes $\boldsymbol{Y}$ and random parameters $\boldsymbol{C}$. The main difference between the two sources of variability (inherent and parametric), is that while the channel variability cannot be reduced, the parametric uncertainty can eventually be reduced by a better estimation of the propensity function parameters. A fundamental assumption of the proposed method is that $\boldsymbol{Y}$ and $\boldsymbol{C}$ are *independent*. This assumption will allow us to decompose the variance as outlined below; it also implies that the inherent stochastic nature of the system is not dependent on the kinetic parameters. It does not mean, however, that the stochastic dynamics are not dependent on the parameters. In the remainder of the section, we discuss how to separate and quantify the variance associated to the two sources of randomness.

## B. Variance decomposition and sensitivity indices

We begin by introducing the Sobol-Hoeffding (SH) decomposition[8] of a functional of several independent random variables. The decomposition is further exploited to derive an expression for the sensitivity indices, which will provide us with a measure of the impact of the distinct sources of uncertainty on the process.

Recall that we consider $\boldsymbol{Y}$ and $\boldsymbol{C}$ as independent random quantities defined on the probability space $(\Omega, \Sigma, \mu)$. Let $G : \Omega \to G(\boldsymbol{X}(\boldsymbol{Y}; \boldsymbol{C})) \in \mathbb{R}$ be a second-order random functional of the stochastic process, *i.e.* $\mathbb{E}\left[G(\boldsymbol{X}(\boldsymbol{Y}; \boldsymbol{C}))^2\right] < \infty$. With a slight abuse of notation,

we shall denote $G(\boldsymbol{X}(\boldsymbol{Y}; \boldsymbol{C}))$ simply as $G(\boldsymbol{Y}, \boldsymbol{C})$. The SH decomposition separates the function $G(\boldsymbol{Y}, \boldsymbol{C})$ into independent and orthogonal elements, as follows

$$G(\boldsymbol{Y}, \boldsymbol{C}) = \bar{G} + G_{\mathrm{par}}(\boldsymbol{C}) + G_{\mathrm{ch}}(\boldsymbol{Y}) + G_{\mathrm{mix}}(\boldsymbol{Y}, \boldsymbol{C}) \quad (5)$$

where $\bar{G}$ is a constant, $G_{\mathrm{par}}(\boldsymbol{C})$ is a random functional that depends only on the parameters, $G_{\mathrm{ch}}(\boldsymbol{Y})$ depends only on the Poisson processes, and $G_{\mathrm{mix}}(\boldsymbol{Y}, \boldsymbol{C})$ depends on both random inputs. The pairwise orthogonality between the functions on the right-hand-side of Eq. (5) ensures the uniqueness of the decomposition. These functionals are given by $\bar{G} = \mathbb{E}[G]$,

$$G_{\mathrm{par}}(\boldsymbol{C}) = \mathbb{E}[G|\boldsymbol{C}] - \mathbb{E}[G],$$
$$G_{\mathrm{ch}}(\boldsymbol{Y}) = \mathbb{E}[G|\boldsymbol{Y}] - \mathbb{E}[G],$$
$$G_{\mathrm{mix}}(\boldsymbol{Y}, \boldsymbol{C}) = G(\boldsymbol{Y}, \boldsymbol{C}) + \mathbb{E}[G] - \mathbb{E}[G|\boldsymbol{Y}] - \mathbb{E}[G|\boldsymbol{C}].$$

In the previous expressions, the conditional expectations are defined by

$$\mathbb{E}[G(\boldsymbol{Y}, \boldsymbol{C})|\boldsymbol{C} = \boldsymbol{c}] = \int_{\Omega} G(\boldsymbol{X}(\boldsymbol{Y}, \boldsymbol{C} = \boldsymbol{c})) d\mu(\omega),$$
$$\mathbb{E}[G(\boldsymbol{Y}, \boldsymbol{C})|\boldsymbol{Y} = \boldsymbol{y}] = \int_{\Omega} G(\boldsymbol{X}(\boldsymbol{Y} = \boldsymbol{y}, \boldsymbol{C})) d\mu(\omega),$$

where $\boldsymbol{c}$ and $\boldsymbol{y}$ are deterministic realizations of the parameters and the Poisson processes respectively. Because the decomposition in Eq. (5) is orthogonal, the variance of $G(\boldsymbol{Y}, \boldsymbol{C})$ can be split according to

$$\mathbb{V}[G] = V_{\mathrm{par}} + V_{\mathrm{ch}} + V_{\mathrm{mix}}, \quad (6)$$

where $V_{\mathrm{par}} = \mathbb{V}[G_{\mathrm{par}}]$, $V_{\mathrm{ch}} = \mathbb{V}[G_{\mathrm{ch}}]$ and $V_{\mathrm{mix}} = \mathbb{V}[G_{\mathrm{mix}}]$ are called the partial variances. These partial variances account for the variability due to parameters, the reaction channels and the interactions between them, respectively. Finally, the Sobol sensitivity indices are obtained by normalizing partial variances:

$$S_{\mathrm{par}} = \frac{V_{\mathrm{par}}}{\mathbb{V}[G]}, \quad S_{\mathrm{ch}} = \frac{V_{\mathrm{ch}}}{\mathbb{V}[G]} \quad S_{\mathrm{mix}} = \frac{V_{\mathrm{mix}}}{\mathbb{V}[G]}.$$

The indices $S_{\mathrm{par}}$ and $S_{\mathrm{ch}}$ are the fractions of the variance in $G$ explained by the parametric uncertainty and stochastic channel dynamics, respectively. The index $S_{\mathrm{mix}}$ accounts for the fraction of the variance due to the interactions between the parametric and channel variabilities. To measure the total impacts of parametric uncertainty and channel stochasticity, we introduce the total indices

$$T_{\mathrm{par}} = \frac{V_{\mathrm{par}} + V_{\mathrm{mix}}}{\mathbb{V}[G]}, \quad T_{\mathrm{ch}} = \frac{V_{\mathrm{ch}} + V_{\mathrm{mix}}}{\mathbb{V}[G]}.$$

The decomposition of the variance of $G(\boldsymbol{Y}, \boldsymbol{C})$ above can be generalized to perform finer sensitivity analyses and measure, for instance, the impact of a subset of reaction channels, of a single parameter, or some joint effects of channel(s) and parameter(s), provided that the components of $\boldsymbol{C}$ are mutually independent (recall

that the Poisson processes $Y_j$ of different channels are independent). In this case, the random vector $\boldsymbol{U} = (Y_1, \ldots, Y_{M_r}, C_1, \ldots, C_{N_c})$ has components $U_i(\omega)$ which are all mutually independent. We shall set $N_U = M_r + N_c$ the total number of variability sources, defined as the sum of the number of reaction channels and uncertain kinetic parameters. With a slight abuse of notation, we shall now denote $G(\boldsymbol{X}(\boldsymbol{Y}; \boldsymbol{C}))$ as $G(\boldsymbol{U})$. The general SH decomposition can immediately be generalized by considering the variability sources in the random vector $\boldsymbol{U}$. Let us denote the index set $\mathrm{U} = \{1, \ldots, N_U\}$ and $\mathscr{P}_{\mathrm{U}}$ its power set. For $\mathrm{u} \subseteq \mathrm{U}$, let $\mathrm{u}_\sim = \mathrm{U} \setminus \mathrm{u}$ be its complement in U. With these notations, the generalized SH decomposition of $G$ is

$$G(\boldsymbol{U}) = \sum_{\mathrm{u} \in \mathscr{P}_{\mathrm{U}}} G_{\mathrm{u}}(\boldsymbol{U}_{\mathrm{u}}), \qquad (7)$$

where the $G_{\mathrm{u}}(\boldsymbol{U}_{\mathrm{u}})$ are mutually orthogonal random functionals, each depending only on the subvector of $\boldsymbol{U}$ having components $\boldsymbol{U}_i$ with $i \in \mathrm{u}$. These functionals are recursively defined by:

$$G_{\mathrm{u}}(\boldsymbol{U}_{\mathrm{u}}) = \mathbb{E}\left[G | \boldsymbol{U}_{\mathrm{u}}\right] - \left(\sum_{\mathrm{v} \subset \mathrm{u}} G_{\mathrm{v}}(\boldsymbol{U}_{\mathrm{v}})\right),$$

with the conditional expectations given by

$$\mathbb{E}\left[G(\boldsymbol{U}) | \boldsymbol{U}_{\mathrm{u}} = \boldsymbol{u}_{\mathrm{u}}\right] = \int_\Omega G(\boldsymbol{u}_{\mathrm{u}}, \boldsymbol{U}_{\mathrm{u}\sim}(\omega)) d\mu(\omega). \quad (8)$$

Due to the orthogonality of SH decomposition in Eq. (7), and observing that $G_\emptyset = \mathbb{E}[G]$, the variance of $G(U)$ can be expressed as

$$\mathbb{V}[G] = \sum_{\mathrm{u} \in \mathscr{P}_{\mathrm{U}} \setminus \emptyset} V_{\mathrm{u}}, \quad V_{\mathrm{u}} = \mathbb{V}[G_{\mathrm{u}}(\boldsymbol{U}_{\mathrm{u}})]. \qquad (9)$$

Similarly, "first-order" and total sensitivity indices associated to variability sources in u can be defined by extending the previous expressions according to

$$S_{\mathrm{u}} = \frac{V_{\mathrm{u}}}{\mathbb{V}[G]}, \quad T_{\mathrm{u}} = \sum_{\mathrm{v} \subseteq \mathscr{P}_{\mathrm{U}}}^{\mathrm{v} \cap \mathrm{u} \neq \emptyset} S_{\mathrm{v}} = 1 - S_{\mathrm{u}\sim}. \qquad (10)$$

### C. Monte Carlo algorithm

We now show how to estimate the sensitivity indices by means of Monte Carlo sampling and we briefly discuss the practical implementation of the sampling method.

#### 1. Monte Carlo estimation of the sensitivity indices

In this work, we rely on the method of Homma and Saltelli[26] (see also Ref. 9) to estimate the sensitivity indices. For simplicity, we restrict the exposition to the

separation of the effects associated to the channels, on the one hand, and kinetic parameters, on the other hand. The task is then to estimate $S_{\mathrm{par}}$, $S_{\mathrm{ch}}$ and $S_{\mathrm{mix}}$.

The Monte Carlo estimation makes use of independent sample sets for both the Poisson processes and the kinetic parameters. The estimation of the partial variances is derived from evaluations of the functional $G(\boldsymbol{Y}, \boldsymbol{C})$ at selected sample points resulting from the appropriate combination of the sample sets.

Let $\mathscr{Y}$ and $\mathscr{C}$ denote two sample sets of $M$ independent realizations $\boldsymbol{Y}^{(i)}$ and $\boldsymbol{C}^{(i)}$ of the Poisson processes and kinetics parameters, respectively, and where the superscript $1 \leq i \leq M$ refers to the sample index. As discussed at the end of Section II B, the sampling of the Poisson processes in $\boldsymbol{Y}$ simply amounts to sampling (independently) the seeds of the random number generators that yields the respective sequences of jumping times in $\boldsymbol{Y}^{(i)}$. On the other hand, the sampling of $\boldsymbol{C}$ amounts to drawing random vectors $\boldsymbol{C}^{(i)}$ from the probability density function $p_{\boldsymbol{C}}$. Standard methods, such as probabilistic transformations, can be used to generate the $\boldsymbol{C}^{(i)}$. From these sample sets, the Monte Carlo estimators of the expectation and variance of $G$ are

$$\widehat{E_G} \doteq \frac{1}{M} \sum_{i=1}^M G(\boldsymbol{Y}^{(i)}, \boldsymbol{C}^{(i)}),$$

$$\widehat{V_G} \doteq \frac{1}{M-1} \sum_{i=1}^M G(\boldsymbol{Y}^{(i)}, \boldsymbol{C}^{(i)})^2 - \widehat{E_G}^2.$$

Consider now independent replicas of the sample sets $\mathscr{Y}$ and $\mathscr{C}$, denoted respectively $\widetilde{\mathscr{Y}}$ and $\widetilde{\mathscr{C}}$, with elements $\widetilde{\boldsymbol{Y}}^{(i)}$ and $\widetilde{\boldsymbol{C}}^{(i)}$ respectively defined as discussed previously. The original sample sets and their replicas can be appropriately combined to obtain Monte Carlo estimates $\widehat{V}_{\mathrm{par}}$ and $\widehat{V}_{\mathrm{ch}}$ of the partial variances of $G$; we use:

$$\widehat{V_{\mathrm{par}}} = \frac{1}{M-1} \sum_{i=1}^M G(\boldsymbol{Y}^{(i)}, \boldsymbol{C}^{(i)}) G(\widetilde{\boldsymbol{Y}}^{(i)}, \boldsymbol{C}^{(i)}) - \widehat{E_G}^2,$$

$$\widehat{V_{\mathrm{ch}}} = \frac{1}{M-1} \sum_{i=1}^M G(\boldsymbol{Y}^{(i)}, \boldsymbol{C}^{(i)}) G(\boldsymbol{Y}^{(i)}, \widetilde{\boldsymbol{C}}^{(i)}) - \widehat{E_G}^2.$$

The sensitivity indices are finally estimated as follows

$$\widehat{S_{\mathrm{par}}} = \frac{\widehat{V_{\mathrm{par}}}}{\widehat{V_G}}, \quad \widehat{S_{\mathrm{ch}}} = \frac{\widehat{V_{\mathrm{ch}}}}{\widehat{V_G}}, \quad \widehat{S_{\mathrm{mix}}} = 1 - \widehat{S_{\mathrm{par}}} - \widehat{S_{\mathrm{ch}}}.$$

The MC estimators above can be extended to compute sensitivity indices $S_{\mathrm{u}}$ associated with different groupings u the input variables, for instance the individual effects of a variable, $\boldsymbol{U}_{\mathrm{u}} = (Y_j)$ or $\boldsymbol{U}_{\mathrm{u}} = (C_j)$. The estimate $\widehat{S_{\mathrm{u}}}$ can be immediately achieved using the appropriate combination of the original sample set and replicas. For instance, the estimation of the partial variance $V_{\mathrm{u}}$ associated to $C_j$ involves the estimator

$$V_{\mathrm{u}} = \frac{1}{M} \sum_{i=1}^M G(\boldsymbol{Y}^{(i)}, \boldsymbol{C}^{(i)}) G(\widetilde{Y}^{(i)}, \widetilde{\widetilde{\boldsymbol{C}}}^{(i)}),$$

where $\widetilde{\widetilde{\boldsymbol{C}}}_k^{(i)} = \widetilde{\boldsymbol{C}}_k^{(i)}$ for $k \neq j$ and $\widetilde{\widetilde{\boldsymbol{C}}}_j^{(i)} = \boldsymbol{C}_j^{(i)}$. We estimate the total sensitivity indices $T_\mathrm{u}$ using the relation $T_\mathrm{u} = 1 - S_\mathrm{u}$ and the Jansen formula[27] which, while also based on combining two replicas of sample sets, was shown to be more accurate.[28]

### 2. Implementation details

We have just seen that the Monte Carlo estimation of the sensitivity indices involves multiple evaluations of the quantity of interest for suitable combinations of samples set of the Poisson and kinetic parameters. Algorithm III.1 provides a sketch of a practical implementation in the case of the simple channel and parametric separation. The algorithm can be easily adapted to handle the estimation of more general sensitivity indices $S_\mathrm{u}$ and $T_\mathrm{u}$. It is adapted from the original algorithm proposed in Ref. 16 for the separation of the variance due to different channels (without parametric uncertainty). Specifically, instead of combining different samples sets for the individual Poisson processes $Y_j$, Algorithm III.1 combines different samples sets of the vector of Poisson processes $\boldsymbol{Y}$ and parameter $\boldsymbol{C}$.

---

**Algorithm III.1** GSA for stochastic chemical reactions

Function MNRM-GSA$(M, \boldsymbol{X}_0, T, \{a_j\}, \{\boldsymbol{\nu}_j\}, \boldsymbol{C})$

**Input:** Number $M$ of samples, initial condition $\boldsymbol{X}_0$, final time $T$, propensity functions $\{a_j\}$ and stoichiometric vector $\{\boldsymbol{\nu}_j\}$, random kinetic parameters $\boldsymbol{C}$.

**Output:** $\widehat{S}_\mathrm{par}$, $\widehat{S}_\mathrm{ch}$ and $\widehat{S}_\mathrm{mix}$

1: $\widehat{E_G} \leftarrow 0$, $\widehat{V_G} \leftarrow 0$,  {init mean and variance}
2: $\widehat{V}_\mathrm{par} \leftarrow 0$, $\widehat{V}_\mathrm{ch} \leftarrow 0$ and $\widehat{V}_\mathrm{mix} \leftarrow 0$  {init partial var}
3: **for** $i = 1$ to $M$ **do**
4:  Draw seeds $\boldsymbol{S}$ and replicas $\widetilde{\boldsymbol{S}}$ for the random number generators
5:  Draw $\boldsymbol{c}$ and replicas $\widetilde{\boldsymbol{c}}$ from density $p_{\boldsymbol{C}}$
6:  $\boldsymbol{X} \leftarrow \mathrm{MNRM}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \boldsymbol{c}, \mathrm{RG}(\boldsymbol{S}))$
7:  $\boldsymbol{X}_\mathrm{ch} \leftarrow \mathrm{MNRM}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \widetilde{\boldsymbol{c}}, \mathrm{RG}(\boldsymbol{S}))$
8:  $\boldsymbol{X}_\mathrm{par} \leftarrow \mathrm{MNRM}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \boldsymbol{c}, \mathrm{RG}(\widetilde{\boldsymbol{S}}))$
9:  $\widehat{E_G} \leftarrow \widehat{E_G} + G(\boldsymbol{X})$
10:  $\widehat{V_G} \leftarrow \widehat{V_G} + G(\boldsymbol{X})^2$
11:  $\widehat{V}_\mathrm{ch} \leftarrow \widehat{V}_\mathrm{ch} + G(\boldsymbol{X})G(\boldsymbol{X}_\mathrm{ch})$
12:  $\widehat{V}_\mathrm{par} \leftarrow \widehat{V}_\mathrm{par} + G(\boldsymbol{X})G(\boldsymbol{X}_\mathrm{par})$
13: **end for**
14: $\widehat{E_G} \leftarrow \widehat{E_G}/M$
15: $\widehat{V_G} \leftarrow \widehat{V_G}/(M-1) - \widehat{E_G}^2$
16: $\widehat{V}_\mathrm{par} \leftarrow \widehat{V}_\mathrm{par}/(M-1) - \widehat{E_G}^2$, $\widehat{V}_\mathrm{ch} \leftarrow \widehat{V}_\mathrm{ch}/(M-1) - \widehat{E_G}^2$
17: $\widehat{S}_\mathrm{par} = \widehat{V}_\mathrm{par}/\widehat{V_G}$, $\widehat{S}_\mathrm{ch} = \widehat{V}_\mathrm{ch}/\widehat{V_G}$ and $\widehat{S}_\mathrm{mix} = 1 - \widehat{S}_\mathrm{par} - \widehat{S}_\mathrm{ch}$

---

To this end, within the main loop of the MC sampler (loop over $i$), two vectors of seeds are first independently drawn ($\boldsymbol{S}$ and its replica $\widehat{\boldsymbol{S}}$, see line 4). These seeds are subsequently used to initialize the random number generators of the stochastic simulator, so as to appropriately set the sequence of jumping times of the Poisson processes and individual reaction channels. Second, two in-

dependent realizations of the parameters are drawn from the density $p_{\boldsymbol{C}}$ ($\boldsymbol{c}$ and $\widetilde{\boldsymbol{c}}$, see line 5). In the present work, we shall consider independent $C_j$ with uniform distributions, such their sampling is straightforward.

The trajectories $X(\boldsymbol{Y}, \boldsymbol{c})$, $X(\boldsymbol{Y}, \widetilde{\boldsymbol{c}})$ and $X(\widetilde{\boldsymbol{Y}}, \widetilde{\boldsymbol{c}})$ are then determined using the MNRM Algorithm II.1 (see lines 6-8); the statistic and correlations of the functional $G(X)$ between these trajectories are subsequently accumulated, before proceeding to the next sample. Finally, the sensitivity indices of $G$ are estimated when $M$ samples have been computed.

We remark that the procedure depicted in Algorithm III.1 aims at clarifying the sampling procedure and is not designed in view of optimizing the computational efficiency. In particular, it is possible the parallelize the main loop of the algorithm.

## IV. BIRTH-DEATH EXAMPLE

We illustrate the variance decomposition and global sensitivity analysis presented in Sec. III on one of the simplest stochastic reaction network: the birth-death (BD) process. A more complex reaction network will be considered in Sec. V.

### A. Birth-death model

The BD system models the dynamic of a single species $S$ ($M_s = 1$), evolving through $M_r = 2$ reaction channels, simply indexed as $R_b$ and $R_d$, according to

$$\oslash \xrightarrow{B} S, \quad S \xrightarrow{D} \oslash. \tag{11}$$
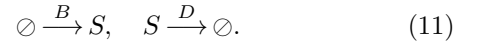
The first reaction channel describes the spontaneous creation (birth) of individuals, at a time-rate $B$; the second channel describes the disappearance (death) of individual at a rate proportional to the population size times the time-rate $D$. In the following, we shall consider the birth and death rates as uncertain parameters, that is $C_1 = B$ and $C_2 = D$, and proceed with their modeling as two independent random variables having uniform distributions. For simplicity, we shall directly denote $B$ and $D$ the two uncertain parameters in the rest of this section. We denote $\mu_B$ and $\sigma_B$ (resp. $\mu_D$ and $\sigma_D$) the mean and standard deviation of the birth rate (resp. death rate). The variability ranges of the two rates are characterized by their coefficients of variation, $\mathrm{CoV}_B = \sigma_B/\mu_B$ and $\mathrm{CoV}_D = \sigma_D/\mu_D$. For simplicity the two coefficients of variation will be set equal to $\mathrm{CoV}_B = \mathrm{CoV}_D = \mathrm{CoV}$, such that the two channels have comparable amounts of parametric variability. Finally, the propensity functions are given by

$$a_b(x; B) = B(\omega), \quad a_d(x; D) = D(\omega)x.$$

By Eqs. (3) and (4), and from the stoichiometric coefficients associated to the BD model, the state can be

rewritten as the difference between two Poisson processes,

$$X(t,\omega) = X(t_0) + Y_b(t_b) - Y_d(t_d) \qquad (12)$$

where the scaled times are given by

$$t_b(\omega) = \int_{t_0}^{t} B(\omega)ds, \quad t_d(\omega) = \int_{t_0}^{t} D(\omega)X(s)ds. \quad (13)$$

Unless specified otherwise, we set $\mu_B = 200$, $\mu_D = 1$ and CoV $= 0.1$. The initial population size at $t_0 = 0$ is set to zero.

### 1. Channel variability

Figure 1 illustrates the effects of channel variability. Specifically, in the three plots of the figure, we condition the trajectories to a fixed value of the kinetic parameters $B(\omega) = \mu_B$ and $D(\omega) = \mu_D$. In Figure 1(a), we report different trajectories obtained using the MNRM algorithm, and generating different realizations of the Poisson processes $Y_b$ and $Y_d$. The trajectories of $X(t)$ quickly grow from the initial condition $X = 0$ and converge to an invariant distribution as $t$ increases. The invariant distribution corresponds to a Poisson distribution with mean population size $\mu_B/\mu_D$ (see below). For different realizations of the Poisson processes, all trajectories have qualitatively the same behavior and asymptotically fluctuate around this terminal expected value.

In Figure 1(b), not only we sample trajectories conditionally on the parameters mean values, but we also condition the dynamics on a particular realization of the Poisson process associated to channel $R_d$. Therefore, the trajectories plotted reflect the variability induced by the birth channel $R_b$ only. It can be observed that, compared to trajectories in Figure 1(a), the set of trajectories in Figure 1(b) is more correlated. A similar behavior is reported in Figure 1(c), where instead of conditioning the system on the mean parameters and a particular realization of $Y_d$, it is now conditioned on a realization $Y_b(\omega)$ of the birth channel. Note that the conditioning of the channels in the last two plots is based on the realizations of $Y_b$ and $Y_d$ corresponding to the trajectory reported with a thick magenta line all plots.

### 2. Parametric variability

Similarly to the previous exercise, we now condition the trajectories of $X(t)$ on the same realizations used previously to condition on $Y_b$ and $Y_d$; the MNRM algorithm is then applied for different realizations of the kinetic parameters $B$ and $D$. We also report in Figure 2 the previous trajectory corresponding to the expected rate parameters $\mu_B$ and $\mu_D$, shown again with a thick magenta line.

Figure 2(a), which should be contrasted with Figure 1(a), illustrates the effect of sampling the kinetic parameters only. It is seen that this results in a variability of the trajectories that is significantly higher than the variability obtained sampling only the Poisson processes. Moreover, about the trajectories in Figure 2(a) highlight the fact that the parametric variability induces trajectories that are significantly more correlated than for the channels variability. The parametric uncertainty is also seen to have a greater impact on the population size, with fluctuations of the trajectories around a time-averaged that changes from a trajectory to another. The variability among trajectories of the asymptotic time-average contrasts with the previous case of the channels variability, where the trajectories fluctuate around the same value.

To gain further insight into the effects of the uncertain kinetic parameters, we subsequently also condition the trajectories on $D = \mu_D$ or $B = \mu_D$, see Figure 2(b) and 2(c) respectively. It is seen that the correlation structure between the sampled trajectories are significantly different depending on the sampled parameter, $B$ or $D$. In particular, it is seen that considering variability in $B$ only (Figure 2(b)) induces trajectories that remain very correlated in time, while for a variability in $D$ only (Figure 2(c)) a time shift between the trajectories is also visible. These different effects can be understood from the structure of the solution as expressed in (12) and (13). From the plots, the variability of the trajectories reported in Figure 2(c) appears larger than in Figure 2(b), suggesting a more important impact of the uncertainty in $D$, than in $B$, on the variability of $X$. However, the experiments reported in this section are not sufficient to confidently draw such a conclusion: the complete variance decomposition must be performed to properly assess the relative importance of each of the variability sources.

### B. Variance decomposition

The variance decomposition is performed using Algorithm III.1, and its generalization. We take advantage of the BD model simplicity to accurately estimate the partial variances, using large sample sets with $M = 5 \cdot 10^7$.

We start by performing the simple parametric / channel variance decomposition, on the quantity of interest consisting of the system state, $G = X(t)$, at different times $t$. Figure 3 reports the evolution with $t$ of the total variance of $X$ and its decomposition in parametric ($V_{\text{par}}$), channels ($V_{\text{ch}}$) and mixed ($V_{\text{mix}}$) contributions. The results in Figure 3(a) focus on the early transient stage for $t < 10$. We see that the variability in $X(t)$ is clearly dominated by the parametric uncertainty, except at the early stages, $t < 0.5$, where the channels variability is larger than parametric variability. The results also indicate that for $t > 5$ the total and parametric variances are essentially constant, the latter accounting for roughly 80% of the former.
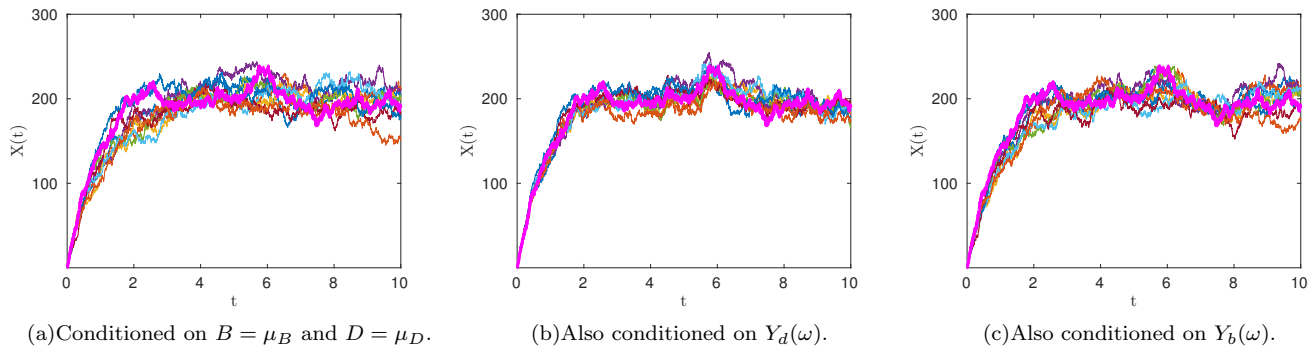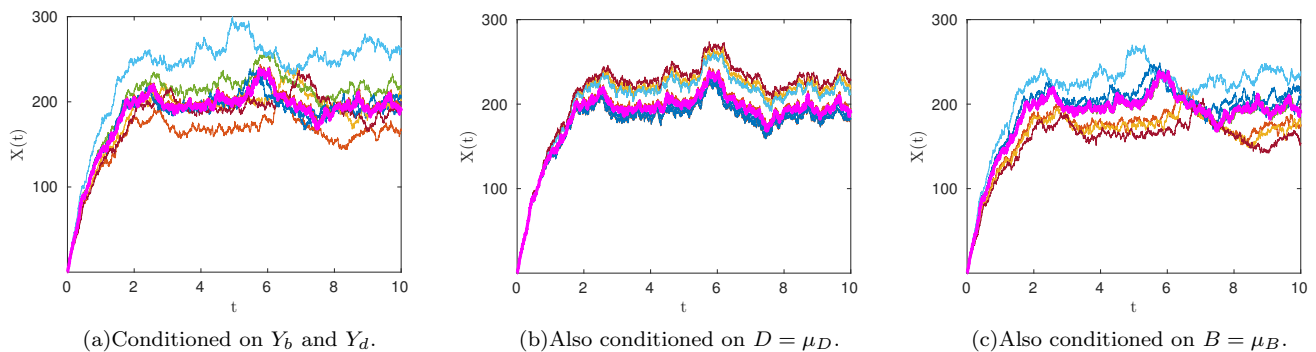
(a)Conditioned on $B = \mu_B$ and $D = \mu_D$.　　　　(b)Also conditioned on $Y_d(\omega)$.　　　　(c)Also conditioned on $Y_b(\omega)$.

FIG. 1. Trajectories of the DB system conditioned on $B = \mu_B$ and $D = \mu_D$.



(a)Conditioned on $Y_b$ and $Y_d$.　　　　(b)Also conditioned on $D = \mu_D$.　　　　(c)Also conditioned on $B = \mu_B$.

FIG. 2. Trajectories of the DB system conditioned on channels realization $\boldsymbol{Y}(\omega) = \boldsymbol{y}$.

In contrast, the channels' partial variance $V_{\text{ch}}$ is seen to reach a maximum for $t \approx 2.5$ and then progressively decreases while the variance associated to mixed effects monotonically increases. Inspecting Figure 3(b), which reports the evolution of the partial variances over an extended time range, we observe that $V_{\text{ch}}$ continues to decrease and eventually becomes negligible at $t \approx 100$, while the variability of the mixed effects levels to 20% of the total variance. These evolutions can be understood from Eq. (13): as time advances, the impact of the uncertain parameters on the scaled times becomes more and more pronounced such that $V_{\text{mix}}$ increases and $V_{\text{ch}}$ decreases.

To gain additional insight into individual effects of the parameters and the channels, we provide in Figure 4 the complete separation of the partial variances $V_{\text{par}}$, $V_{\text{ch}}$ and $V_{\text{mix}}$. Figure 4(a) shows the decomposition of $V_{\text{par}}$ into pure contributions from $B$ and $D$, and mixed effects of $B$ and $D$. It is seen that the partial variances $V_B$ and $V_D$ become very close for $t > 6$, while the interaction between the two parameters remains negligible.

Regarding the decomposition of $V_{\text{ch}}$ into contributions of the two channels and their interaction, we observe in Figure 4(b) that the two variances $V_{Y_b}$ and $V_{Y_d}$ decay at the same rate and remain sensibly equal, while the mixed contribution $V_{Y_b,Y_d} \approx 0$ denotes a negligible interaction between the two channels. It should be remarked that

this finding is seemingly in contradiction with the finding reported in our previous work,[16] where the interaction between the two channels was increasing over time with asymptotically $V_{Y_b} \approx V_{Y_b,Y_d}$ and $V_{Y_d} \approx 0$. However, no parametric uncertainty was considered in Ref. 16.

Figure 4(c) indicates that mixed effects are still existing in the present system, but they essentially consist in interactions between parameter $B$ and the two channels: $V_{\text{mix}} \approx V_{B,Y_b} + V_{B,Y_d}$. Note also that third order variances, e.g. $V_{B,Y_b,Y_d}$, are negligibly small and are thus not reported.

Figure 5 reports the evolution of the total sensitivity indices $T_{B,D,Y_d,Y_b}$ scaled by the total variance. The results in Figure 5(a) correspond to CoV = 0, with all the variability due to the channels and a larger effect of $Y_b$ compared to $Y_d$. In Figure 5(b), corresponding to CoV = 0.015, a low contribution of the parameters is visible with a more noticeable impact of $B$ compared to $D$, but the variability due to the two channels remains predominant. The total variance is also seen to be slightly higher than for CoV = 0.0.

Increasing CoV to 0.05, in Figure 5(c), we see that the birth rate parameter $B$ becomes the dominant source of variability for $t > 2.5$, while parameter $D$ and the channels have comparable impact. Note also that the variance is larger than previously observed. Finally, Figure 5(d) corresponding to CoV = 0.15 shows that in this case
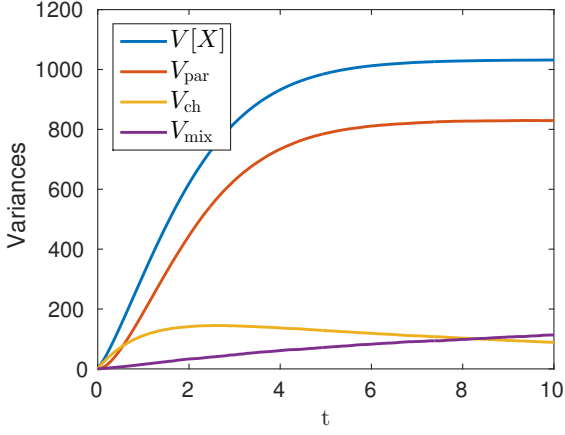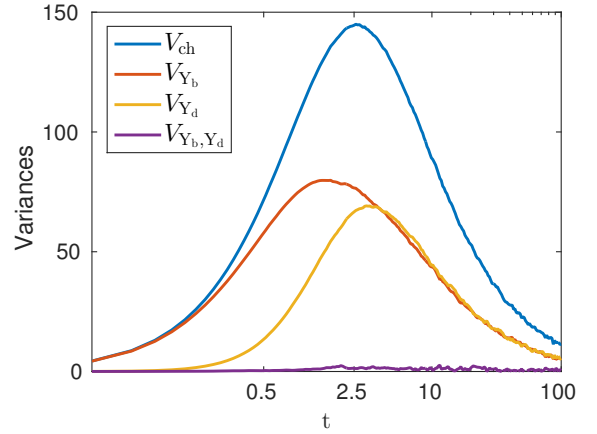
(a)For $t < 10$.
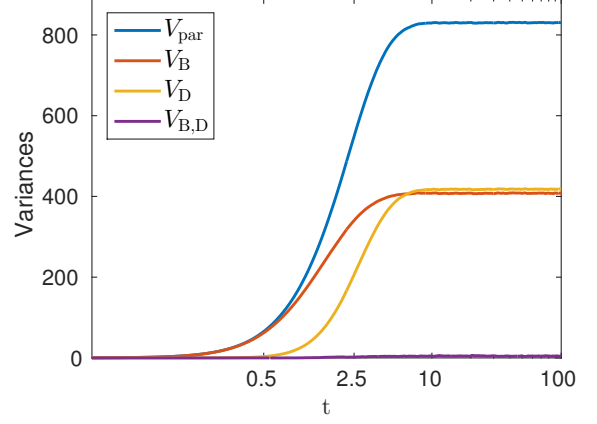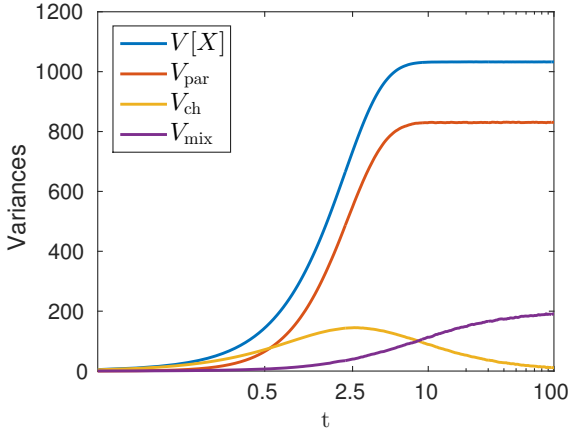


(b)Long time behavior.

FIG. 3. Evolution of the components of the total variance of the birth-dead model for $t \in [0, 10]$.

most of the variability is due to the kinetic parameters, which contribute roughly equally at $t \approx 10$. The total contribution of the channels is small and dominated by interactions with the parameters. We also observe that the total variance is up to 10 times larger for CoV $= 0.15$ compared to the case of CoV $= 0$.

We have just seen that the different variabilities in $X(t)$ increase with CoV, making difficult the quantitative comparison of their relative impacts. We provide in Figure 6 the dependence on CoV of the first order (Figure 6(a)) and total (Figure 6(b)) sensitivity indices of $X$ at a fixed time $t = 10$. The figures show that as CoV increases, the first order sensitivity indices associated to the channels, $S_{Y_b}$ and $S_{Y_d}$ decrease in favor of $S_B$ and $S_D$ which become dominant for CoV $\gtrsim 0.05$. A similar trend is reported for the total sensitivity indices, shown in Figure 6(b), though $T_{Y_b}$ and $T_{Y_d}$ decrease with a slower rate than $S_{Y_b}$ and $S_{Y_d}$, denoting the emergence of stronger interactions between the parametric and channel variabilities and the emergence of non-additive effects.
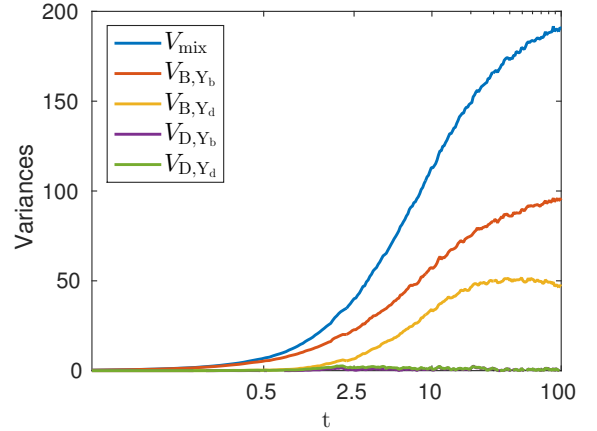


(a)Separation of $V_{\text{par}}$



(b)Separation of $V_{\text{ch}}$



(c)Separation of $V_{\text{mix}}$

FIG. 4. Decomposition of the partial variances $V_{\text{par}}$ (a), $V_{\text{ch}}$ (b) and $V_{\text{mix}}$ (c), into contributions from the uncertain kinetic parameters $B$ and $D$, and channel $Y_b$ and $Y_d$.

## C. Comparison with local sensitivity estimates

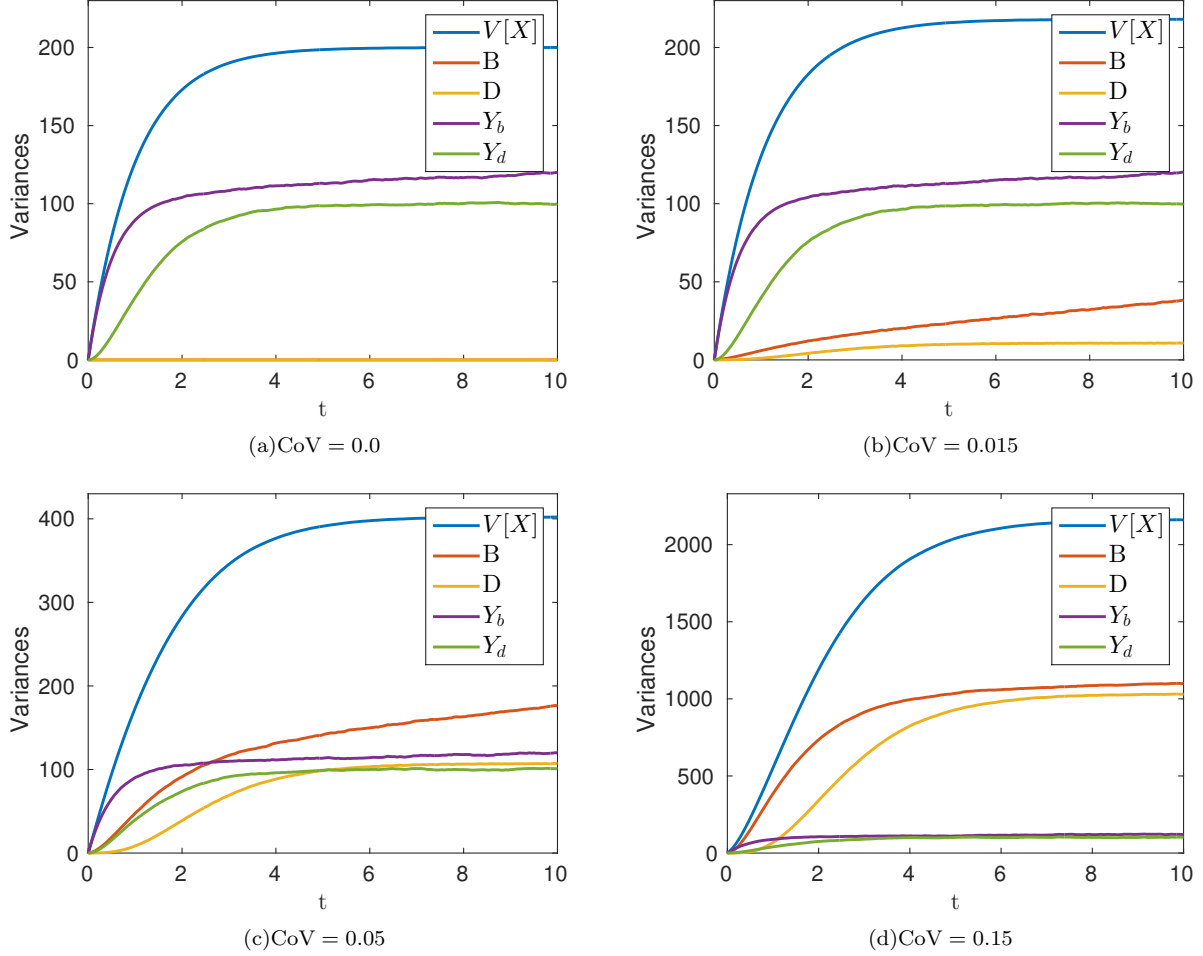The proposed method is global, in the sense that

FIG. 5. Total sensitivity indices of $X(t)$ scaled by $\mathbb{V}[X(t)]$, for different values of the coefficient of variation CoV of the rate parameters $B$ and $D$ as indicated.

it measures the variabilities induced by the considered sources, either parametric or inherent, over the whole events set. It also provides a natural and objective way to compare the respective influences of different variability sources. These characteristics are contrasted with the classical approaches of local sensitivity measures, involving derivatives at nominal points, and which can be limited to low variability levels. In addition, comparing local sensitivity associated to different sources may be difficult, as it usually involves derivatives with respect to different quantities, sometimes hardly comparable (temperature, pre-exponential factor, ...).

Our objective in this section is to demonstrate that our approach is consistent, and agrees with the local derivative-based parametric sensitivity estimates when the coefficient of variation of the kinetic parameters goes to zero. Specifically, considering again for quantity of interest $G(X) = X$, we introduce the local sensitivity

index,[4] associated to the birth rate parameter, as follows

$$L_B(b,d) = \frac{\partial \mathbb{E}[X|b,d]}{\partial b} = \lim_{h \to 0} \frac{\mathbb{E}[X|b+h,d] - \mathbb{E}[X|b,d]}{h}.$$
(14)

A similar expression can be written for $L_D(b,d)$. In fact, for the birth-death process, the exact expression for the conditional expectation in Eq. (14) is available:

$$\mathbb{E}[X(t)|b,d] = x_0 e^{-dt} + \frac{b}{d}(1 - e^{-dt}).$$

Accounting for the initial condition, $x_0 = 0$, we see that the local sensitivity indices at the nominal parameter values are given by

$$L_B(\mu_B, \mu_D) = (1 - e^{-\mu_D t})/\mu_D,$$

and

$$L_D(\mu_B, \mu_D) = \frac{\mu_B}{\mu_D^2}(e^{-\mu_D t}(1 + \mu_D t) - 1).$$

Thus, for large enough times, one gets in our case $L_B(\mu_B, \mu_D) \approx \mu_D^{-1} = 1$ and $L_D(\mu_B, \mu_D) \approx \mu_B/\mu_D^2 = 200$.
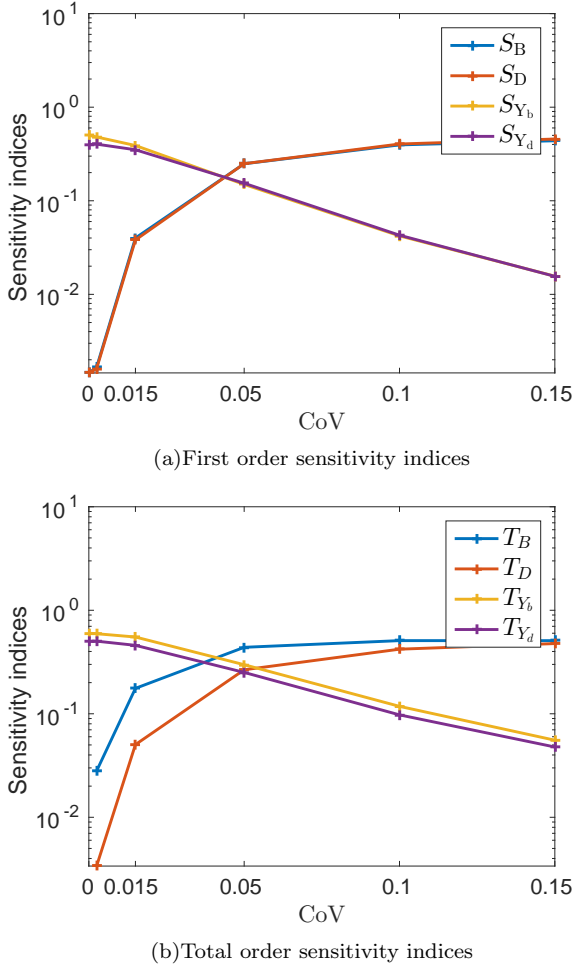
(a)First order sensitivity indices



(b)Total order sensitivity indices

FIG. 6. First and total order sensitivity indices of $X$ at time $t = 10$. Note the log scale.

In order to show how global and local sensitivity indices relate to each other, we consider the case of the birth rate, and assume a sufficient regularity of $\mathbb{E}\left[X|b,d\right]$ with respect to $b$; we then have the following Taylor series,

$$\mathbb{E}\left[X|b+h,d\right] = \mathbb{E}\left[X|b,d\right] + h\frac{\partial\mathbb{E}\left[X|b,d\right]}{\partial b} + \mathscr{O}(h^2).$$

From this expansion, one can easily establish the expression of the variance of the conditional expectation,

$$\mathbb{V}\left[\mathbb{E}\left[X|B,\mu_D\right]\right] = L_B^2(\mu_B,\mu_D)\mathbb{E}\left[(B-\mu_B)^2\right] + hot,$$

where $hot$ stands for higher order terms involving higher order moments in $(B - \mu_B)$, and that are assumed to vanish faster than $\sigma_B^2$ when $\mathrm{CoV}_B \to 0$. Therefore, it follows that

$$\lim_{\mathrm{CoV}\to 0} \frac{V_B}{\sigma_B^2} = L_B^2(\mu_B,\mu_D), \qquad (15)$$

with a similar expression in the case of the uncertain death rate $D$. Note that these expressions hold in the

limit of a vanishing variability of all kinetic parameters. A bias term, e.g. $\mathbb{E}\left[X|\mu_B\right] - \mathbb{E}\left[X|\mu_B,\mu_D\right]$, would be otherwise introduced.

In Figure 7 we check the validity of relation in Eq. (15) in our computations, reporting as a function of the number of MC samples $M$ the ratios of the partial variances, $V_B$ and $V_D$ for $X$ at time $t = 10$, with the respective variances of the parameters, $\sigma_B^2$ and $\sigma_D^2$. Different values of CoV are reported as indicated. The plots show that the MC estimates of $V_B/\sigma_B^2$ and $V_D/\sigma_D^2$ are converging with increasing number $M$ of samples. However, the magnitude of the fluctuations are increasing as CoV decreases, denoting an increasing sampling error. This trend can be explained by the large number of samples needed to accurately estimate the partial variances $V_B$ and $V_D$ when the global variance becomes more and more dominated by the channels' partial variances.

Conversely, the sampling error is seen to converge quickly with $M$ when CoV is larger, but we should not expect the ratios to necessarily converge to the exact squared local sensitivity index, $L_B^2$ and $L_D^2$ which are shown with dashed lines in the plots, because of the contributions of higher order moments. This is a classical trade-off problem, between bias and sampling errors, that is also present in the numerical estimation of the local sensitivity indices. In this context, one has to select an appropriate value for $h$ that leads to a small enough bias in the approximation of the derivative by the difference formula, and a small enough sampling error in the MC estimate of the difference which demands more samples as $h$ decreases.

Different methods (see for instance Refs. 4, 6, and 7) have been proposed recently to reduce and balance the bias and sampling errors in the estimation of the local sensitivity indices. In particular, correlations between the estimates of the two expectations in the difference formula (14) and multilevel Monte Carlo constructions have been proposed to obtain accurate estimators of the local sensitivity indices at a reduced cost; we plan to adapt these techniques to improve the computational efficiency of the MC estimation of the global sensitivity indices in future works.

Anyway, we emphasize that, in the case of finite CoV, a difference between the ratios and the local sensitivities is not a limitation of our method. It underlines the importance of accounting *globally* for the parametric uncertainty effects, and reveals that a local characterization is not appropriate. Further, even if a local sensitivity coefficient (say $L_B$) agrees with its counterpart ($V_B/\sigma_B^2$), the presently developed global sensitivity analysis indicates that the local index may fail to account for some of the variability actually induced by the parameter ($B$). Indeed, the parameter may by responsible for an additional variability through its interactions with the channels (and the other parameters). In this case, the *total* sensitivity index is the appropriate measure of induced variability and restricting the analysis to local indices may be misleading.
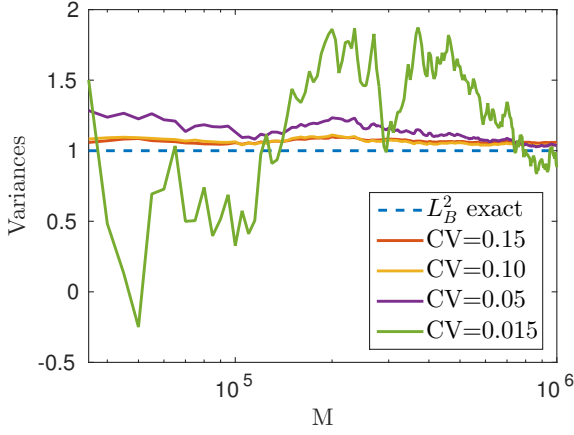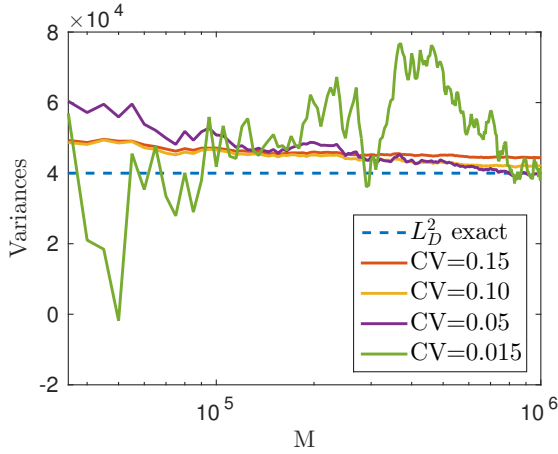
(a)Case of birth rate $B$.



(b)Case of death rate $D$.

FIG. 7. Convergence with the samples set size, $M$, of the ratios of partial to parametric variances $V_B/\sigma_B^2$ and $V_D/\sigma_D^2$, and different coefficient of variations as indicated. Also shown are the exact theoretical values for CoV $\to 0$, denoted $L_B^2$ and $L_D^2$ respectively.

## V. SCHLÖGL MODEL

In this section we consider the Schlögl model,[29] which has $M_r = 4$ reaction channels, $\{R_i, i = 1, \ldots, 4\}$, corresponding to the reaction mechanism

$$B_1 + 2S \underset{C_2}{\overset{C_1}{\rightleftharpoons}} 3S, \quad B_2 \underset{C_4}{\overset{C_3}{\rightleftharpoons}} S.$$
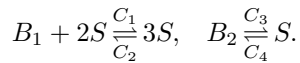
Species $B_1$ and $B_2$ are assumed in large excess with constant population size, set respectively to $x_1 = x_2/2 = 10^5$. The stochastic state then reduces to the single evolving species $S$ with $M_s = 1$. The uncertain propensity functions are

$$a_1(x; \boldsymbol{C}) = C_1 x_1 x(x-1), \quad a_2(x; \boldsymbol{C}) = C_2 x(x-1)(x-2),$$

$$a_3(x; \boldsymbol{C}) = C_3 x_2, \quad \text{and} \quad a_4(x; \boldsymbol{C}) = C_4 x.$$

The parameters $C_1$ to $C_4$ in the definition of the propensity functions are considered as independent random variables with uniform distributions. As previously, we define the uniform distributions by their means $\mu_{1 \leq j \leq M_r}$ and coefficients of variation CoV; again the latter is taken to be the same for all the parameters. Regarding the expected values of the $C_i$, we set $\mu_1 = 0.015, \mu_2 = 10^{-4}/6, \mu_3 = 200$, and $\mu_4 = 3.5$. Finally the initial condition is $X(t = 0) = 250$.

### A. Qualitative analysis

To gain an understanding of the qualitative impact of the different sources of variability we select a reference trajectory, corresponding to the mean value of the kinetic parameters and a specific realization of the Poisson processes. This particular trajectory is plotted using a thick magenta line in the following plots. We draw samples of trajectories conditioned on the same parameter values or Poisson processes.

Figure 8(a) shows several trajectories of the state conditioned on fixed values of the kinetic parameters $C_j(\omega) = \mu_j$, and sampling of the Poisson processes $\boldsymbol{Y}$. Theses trajectories illustrate the random dynamics of the system, that exhibit a bifurcation: starting from the initial condition, the system selects in an initial stage one or the other branch, called lower and upper branches in the following, and later fluctuates around the selected branch. It is seen that the inherent (channels) variability acts on both the branch-selection process and the subsequent fluctuations.

Figure 8(b) depicts trajectories of the system for sampled values of the parameters, but conditioned on the reference realization of the Poisson processes. The sampling of the kinetic parameters uses CoV = 0.1. The plot shows that although the trajectories are conditioned on the same realization of the channel dynamics, the branch selected changes from one sample of the kinetic parameters to another. This behavior highlights the sensitivity of the trajectories and of the branch selection process with regards to both sources of variability.

In addition, qualitative differences are reported between the two sets of sampled trajectories in Figure 8. First, it is seen that the variability level is much larger when sampling the kinetic parameters, in Figure 8(b), than when resampling the Poisson processes in Figure 8(a). More precisely, the fluctuation amplitudes around the selected branches appear to be essentially the same for the inherent and parametric resampling, but the latter induces an additional variability by displacing the location of the attracting branches.

The analysis of the inherent and parametric variability can be further refined by distinguishing the elementary effects of each channel and parameter. This is done by resampling the Poisson processes, $Y_j$, and kinetic parameters, $C_j$, one at a time. Our previous work[16] focused on the resampling of the Poisson processes $Y_j$ to investigate
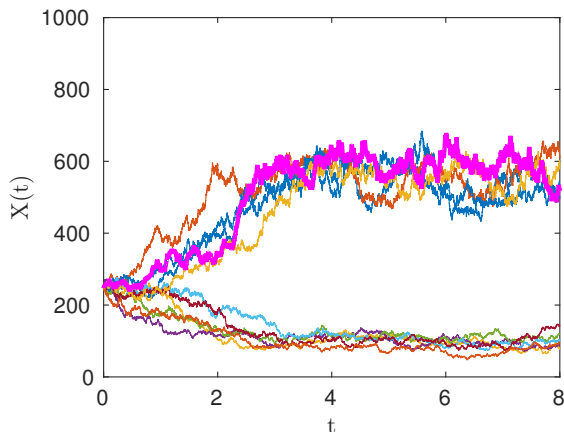
(a)Conditioned on the parameter mean values
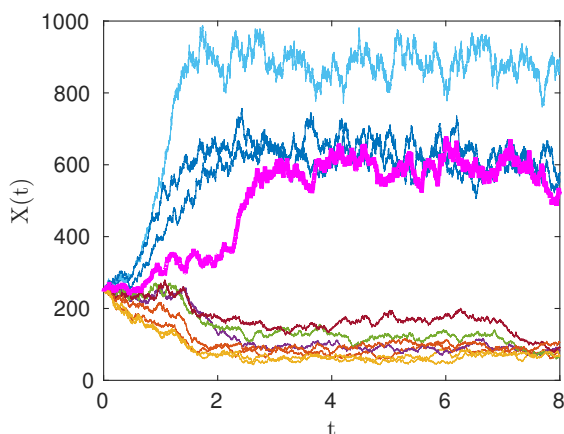


(b)Conditioned on fixed Poisson processes

FIG. 8. Sampled trajectories of the Schlögl model: (a) conditioned on the mean parameter values and, (b) conditioned on a particular realization of the Poisson processes $\boldsymbol{Y}(\omega)$.

the relative variability induced by the different channels in absence of parametric uncertainty. The conclusion of this analysis is briefly summarized in Figure 9, which shows that channels $R_1$ and $R_4$ are the principal sources of inherent variability in the state $X$, while $R_2$ and $R_3$ are less important. This is explained by the selected branch which depends more heavily on the processes $Y_1$ and $Y_4$ than on $Y_2$ and $Y_3$. Indeed, one can observe that for all the resampled processes $Y_2$ and $Y_3$ the system has selected the same (upper) branch.

Figure 10 shows the individual effect of each kinetic parameter on the reference trajectory. The resampling of $C_j$ uses increasing coefficients of variation: CoV = 0.05 (top row), 0.10 (center row) and 0.15 (bottom row). Focusing first on the case of the largest parametric variability CoV = 0.15, we observe that the resampling of the reference trajectory leads to very different effects depending on the considered parameter. Specifically, the resampling of $C_1$ and $C_2$ induces a large variability of the resulting trajectories. In fact, both the location of the attracting

branch and the time scale of the selection process are evidently affected when $C_1$ and $C_2$ are resampled. On the contrary, resampling $C_3$ and $C_4$ appears to essentially affect the selected branch but not its location. Also, $C_4$ seems to have a greater impact on the time scale to reach the selected branch, compared to $C_3$. Finally, we observe that reducing the coefficient of variation CoV reduces the variability in the trajectories obtained by resampling $C_1$ and $C_2$. This is due to the decreasing uncertainty in the location of the attracting branches. A similar effect can be seen on the time scale for reaching an attracting branch and the exchange of attracting branch when resampling $C_2$ and $C_3$ with a lower CoV.

### B.   Global sensitivity analysis

The previous exercises with different resampling of the channel processes and kinetic parameters highlight the complex structure governing the effects and interactions between inherent and parametric variability sources. This complexity stresses the need for quantitative and generic methodologies enabling the quantification and importance ranking of individual sources, as performed below, in order to fully understand the variability in the model output. It also underlines the importance of the global sensitivity approach, as the stochastic model response can substantially depend on the input uncertainty level (CoV).

#### 1.   Variance decomposition

As for the birth-death example, we start by performing the parametric / channel variance decomposition, for the system state $G(X) = X(t)$; unless otherwise noted CoV = 0.1. The MC estimators of the partial variances reported in this section use sample sets with size $M = 10^6$.

Figure 11 depicts the time evolution of decomposition of $\mathbb{V}[X(t)]$, into parametric, channels and mixed contributions, for $t \in [0,8]$. The results show that, for this value of CoV, the variability in $X(t)$ is predominantly caused by the uncertainty in the kinetic parameters, which accounts for up to 80% of the total variability. The remaining fraction of the variance is essentially due to the mixed effects, particularly for $t > 2$ where the channels contribution $V_{\mathrm{ch}}$ is negligible. The dominance of $V_{\mathrm{par}}$ is explained by the parametric impact on the attracting branch locations; changing the kinetic parameters immediately affects the asymptotic time-averages and even possibly the selected branch, causing significant variability. The reason for the low contribution of $V_{\mathrm{ch}}$ to $\mathbb{V}[X]$ is less obvious and requires additional analysis, as further discussed below.

To have a better perception of the structure of the total variability in the model state, we present in Figure 12 the time evolution of the first-order partial vari-

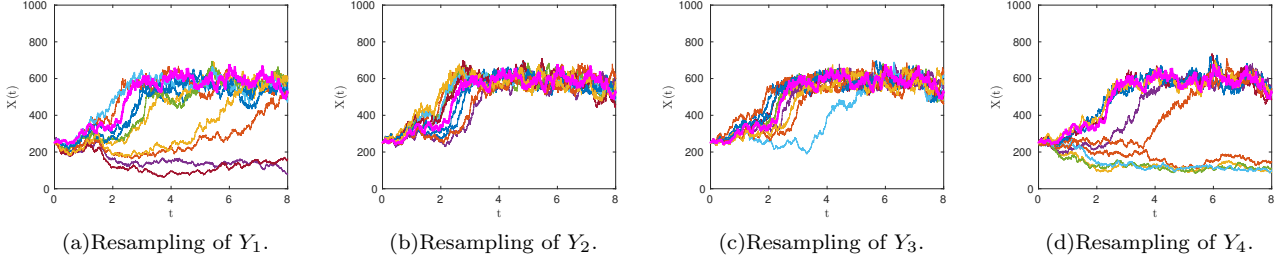(a)Resampling of $Y_1$.     (b)Resampling of $Y_2$.     (c)Resampling of $Y_3$.     (d)Resampling of $Y_4$.

FIG. 9. Trajectories of the Schlögl model in which the kinetic parameters fixed to their mean values, and the only source of variability is due to the Poisson process specified in the caption. The reference trajectory is shown with a thick magenta line.
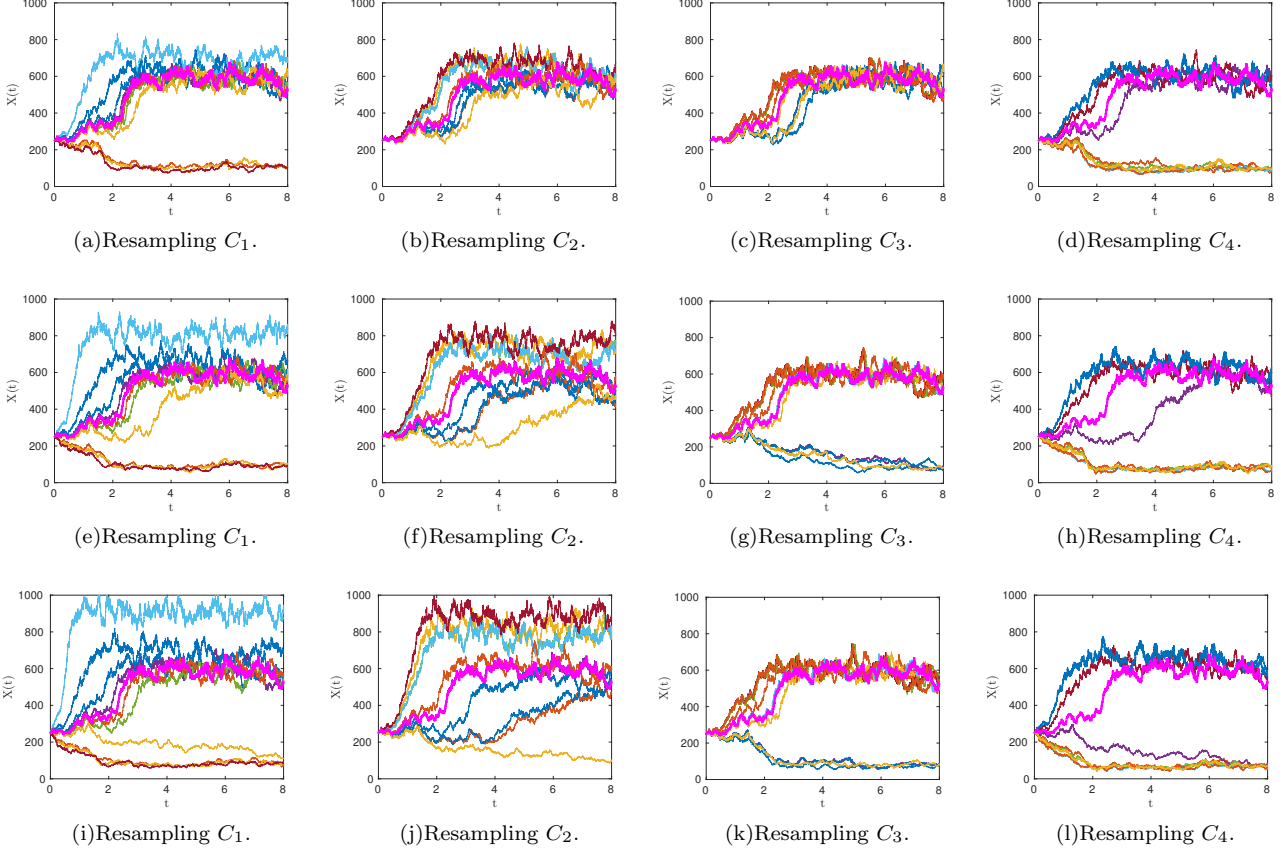


(a)Resampling $C_1$.     (b)Resampling $C_2$.     (c)Resampling $C_3$.     (d)Resampling $C_4$.

(e)Resampling $C_1$.     (f)Resampling $C_2$.     (g)Resampling $C_3$.     (h)Resampling $C_4$.

(i)Resampling $C_1$.     (j)Resampling $C_2$.     (k)Resampling $C_3$.     (l)Resampling $C_4$.

FIG. 10. Trajectories of the Schlögl model in which the Poisson processes are fixed and the only source of variability is due to the parameter specified in the caption. First row corresponds to CoV = 0.05, second row to CoV = 0.10 and third row to CoV = 0.15. The reference trajectory is shown with a thick magenta line.

ances. Starting with the first-order partial variances associated with the kinetic parameters, $V_{C_1}$ to $V_{C_4}$, shown in Figure 12(a), we observe that the parameter having the largest effect on $X(t)$ is $C_1$, accounting for roughly 50% of the parametric variance, followed by $C_4$ ($\approx 25\%$ of $V_{\mathrm{par}}$) and $C_2$ ($\approx 10\%$ of $V_{\mathrm{par}}$), while the single effect of $C_3$ is relatively small. The two dominant parameters are in fact associated with the propensity functions of the two channels that were shown in Ref. 16 to be responsible for most of the state variability. The figure also reports the sum of the first order partial variances $V_{C_i}$, which is seen

to amount to roughly 85% of $V_{\mathrm{par}}$, indicating the presence of some interactions between the parameter variabilities, independently of the interactions with the channels. In words, the effects of the parameters are not additive.

Figure 12(b) depicts the evolution of the first-order partial variances, $V_{Y_i}$, associated with the reaction channels. Again, channels $Y_1$ and $Y_4$ are seen to dominate the two other channels in terms of first-order partial variances. It is remarked that the partial variances $V_{Y_i}$ are lower than their parametric counterparts, $V_{C_i}$, and that the interactions between channels appear to be less pro-
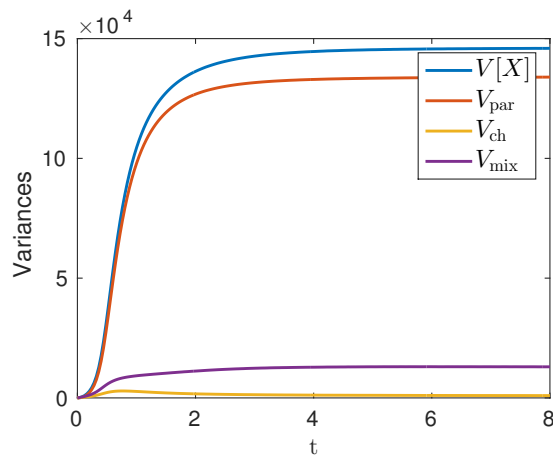
FIG. 11. Decomposition of the variance $\mathbb{V}[X(t)]$, for $t \in [0, 8]$, in parametric, channels and mixed contributions. Case of the Schlögl model with CoV = 0.1.

nounced than the interactions between parameters, as $\sum_i V_{Y_i}$ is close to $V_{ch}$. Regarding the time behavior, the partial variances associated with the channels are observed to initially grow, reaching a maximum for $t \approx 1$, then decrease progressively to roughly half of their maximum values and become constant. The initial growth phase corresponds to the stage where the system evolves from the initial condition to the randomly selected attracting branch, with an increasing effect of the inherent stochasticity during the selection stage. Shortly after the branch has been selected, not only the interactions between the channels develop, as in the case of certain parameters,[16] but the interactions with the parameters becomes more and more pronounced, with a resulting significant reduction of the first-order partial variances, $V_{Y_i}$.

Interactions between the channels and the parameters can be assessed from the plots of Figure 13. In Figure 13(a) we report the *total* effects associated with each parameter $C_i$ and its interactions with the channels, that is the total contribution of $C_i$ to $V_{mix}$. It is seen that once more the parameters $C_1$ and $C_4$ have the highest contributions to $V_{mix}$ while $C_2$ and $C_3$ have contributions roughly one half smaller. Further, the sum of these total contributions largely exceeds $V_{mix}$ indicating the importance of not only interactions between the $C_i$ and the channels stochasticity, but also of the joint interactions between the set of parameters and channels. In fact, the partial variances due to the interaction between single parameter $C_i$, only, and all Poisson processes $\boldsymbol{Y}$ are all very low compared to $V_{mix}$ (not shown). Similarly, one can compute the total contribution of individual channels $Y_i$ to $V_{mix}$. Because $V_{ch}$ is small we directly report in Figure 13(b) the time evolution of the total contributions of the $Y_i$ to the total variance. Compared to the corresponding first-order partial variances $V_{Y_i}$, shown in Figure 12(b), we see that the total channel contributions
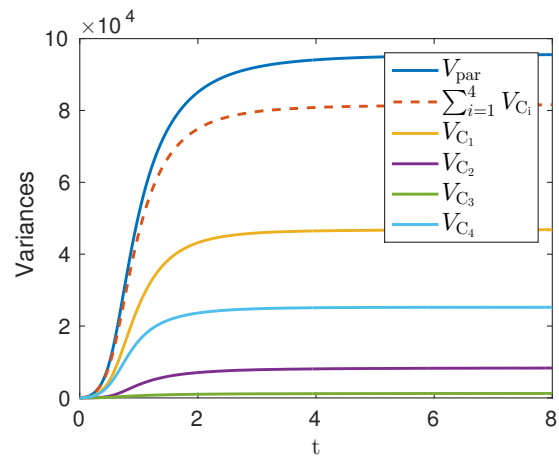


(a)Parametric first-order partial variances.
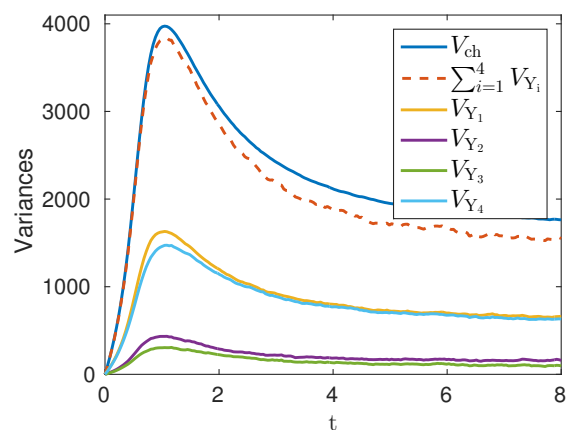


(b)Channels' first-order partial variances.

FIG. 12. Time evolution of the first-order partial variances. The top plot corresponds to partial variances associated with the parameters, $C_i$, whereas the bottom plot depicts partial variances associated with the reaction channels' Poisson processes, $Y_i$.

are significantly higher and exhibit a different behavior in time, monotonically increasing to asymptotically constant values. Again, channels $R_1$ and $R_4$ are found to contribute roughly equally and twice as much as channels $R_2$ and $R_3$.

### 2. Total sensitivity indices

Provided in Figure 14 are the total sensitivity indices of $X(t)$, associated with the kinetic parameters $C_i$ and channels Poisson processes $Y_i$. The total sensitivity indices are scaled by the total variance $\mathbb{V}[X(t)]$, and are reported for different coefficients of variation, CoV = 0, 0.05, 0.1 and 0.15.

Figure 14(a), where CoV = 0.0, corresponds to the case of a model having no parametric uncertainty as studied in Ref. 16. The total sensitivity indices $T_{C_i}$ of
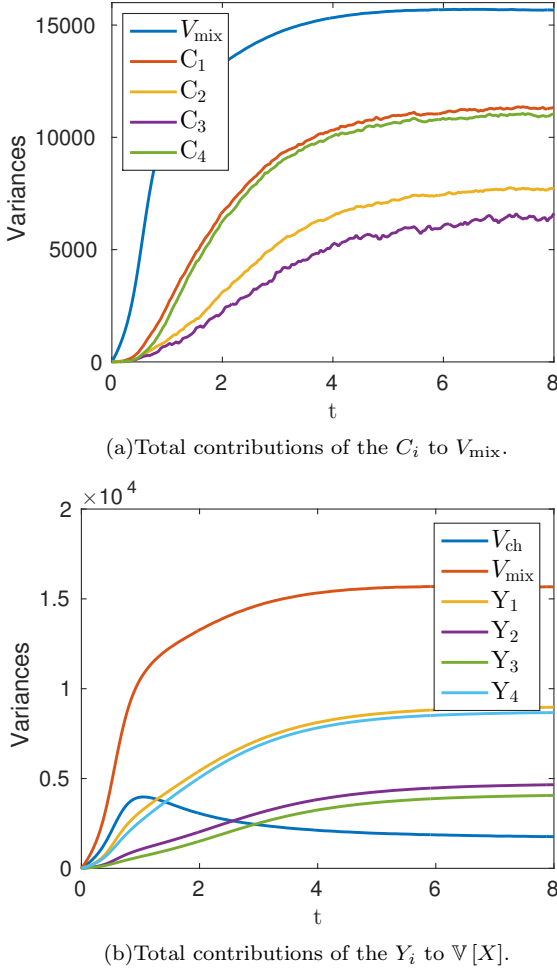
(a)Total contributions of the $C_i$ to $V_{\text{mix}}$.



(b)Total contributions of the $Y_i$ to $\mathbb{V}[X]$.

FIG. 13. Time evolution of the total contributions of each $C_i$ to the partial variance $V_{\text{mix}}$ (top plot) and total contribution to the variance of the channels $Y_i$ (bottom plot).

the parameters are then 0, and the variability is principally due to channels $R_1$ and $R_4$ roughly equally, and to a lesser exteny to channels $R_2$ and $R_3$. Increasing the coefficient of variation of the kinetic parameters, we observe successively in Figure 14(b)-14(d) that the total sensitivity indices of the parameters, $T_{C_i}$, become progressively dominant compared to the total sensitivity indices of the channels, $T_{Y_i}$. It is remarked that the total sensitivity indices of the parameters do not increase proportionally with the squared coefficient of variation, as one would have expected for a linear parametric dependence. Also, the total variances due to the channels are seen to slightly diminish with increasing CoV. Regarding the parametric total indices $T_{C_i}$ the plots indicate that $T_{C_1} > T_{C_4} > T_{C_2} > T_{C_3}$ for the whole range of CoV values presented, and that $T_{C_1}$ becomes more dominant as CoV increases.

## C.   Local sensitivity analysis

We finally contrast the outputs of the global sensitivity analysis above with those of a local sensitivity characterization. It should be clear that a local characterization of the Schlögl model is prone to be inappropriate, because of the complicated non-linear behavior of the system, which involves third-order reactions, and the non-trivial dependencies on the parameters. Although trajectories of the model have complex dependencies with respect to the uncertain kinetic parameters, the statistical moments of the quantity of interest may behave well and smoothly as the parameters changes; conducting derivative-based local approaches is then quite legitimate.

However, we show below that using local information does not necessarily provide a useful characterization of the variability in the quantity of interest. In particular, following the local approach previously introduced in Sec. IV C for the birth-death model, we wish to estimate by $\tilde{V}_{C_i}$ the first-order variances $V_{C_i}$ associated with the uncertain kinetic parameters, using (see Eq. (15))

$$\tilde{V}_{C_i} = L_{C_i}^2 \sigma_i^2, \qquad (16)$$

where $L_{C_i}$ is the derivative, with respect to the $i$-th kinetic parameter, of the averaged quantity of interest at the parameter mean values $C_i = \mu_i$. In our case, $L_{C_i}$ is then the derivative of the mean population size, since $G(X) = X$. It should be clear that relying on moments to characterize the influence of parameters can be of limited relevance, because of the branching nature of the Schlögl model. Specifically, the average trajectory and its sensitivity to parameters is not necessarily representative of the system response, especially after the system has selected one or the other branch.

To support these claims, we provide in Figure 15 a comparison of the first-order and total-order variances associated with the four uncertain kinetic parameters $C_i$, and the local estimate $\tilde{V}_{C_i}$ as defined above. In practice, we estimate the derivatives $L_{C_i}$ of $\mathbb{E}[G(X)]$ with a large Monte Carlo sample set (typically $M = 10^6$ samples) using different values for $h$ in Eq. (14) to optimize the tradeoff between bias and sampling error. The errors in $\tilde{V}_{C_i}$ reported below were estimated to be roughly 5%. Figure 15(a) corresponds to the case of CoV = 0.1 at an early time $t = 0.1$, where the system has not yet clearly selected the random branch. For such a short time, the distribution of $G(X)$ for the nominal kinetic parameters is still uni-modal and the local approach is meaningful. Indeed, the estimates $\tilde{V}_{C_i}$ are not far from the first-order partial variances $V_{C_i}$, as one would expect. However, it is already observed that the variability due to the interactions of $C_i$ with other parameters and inherent stochasticity is not accounted by $\tilde{V}_{C_i}$, although these interactions amounts to roughly half of the induced total variability. Figure 15(b) provides the same comparison but at a later time, $t = 4$. At this time, the system has, with high probability, already selected its attracting branch: the distribution of $G(X)$ for the nominal parameters is
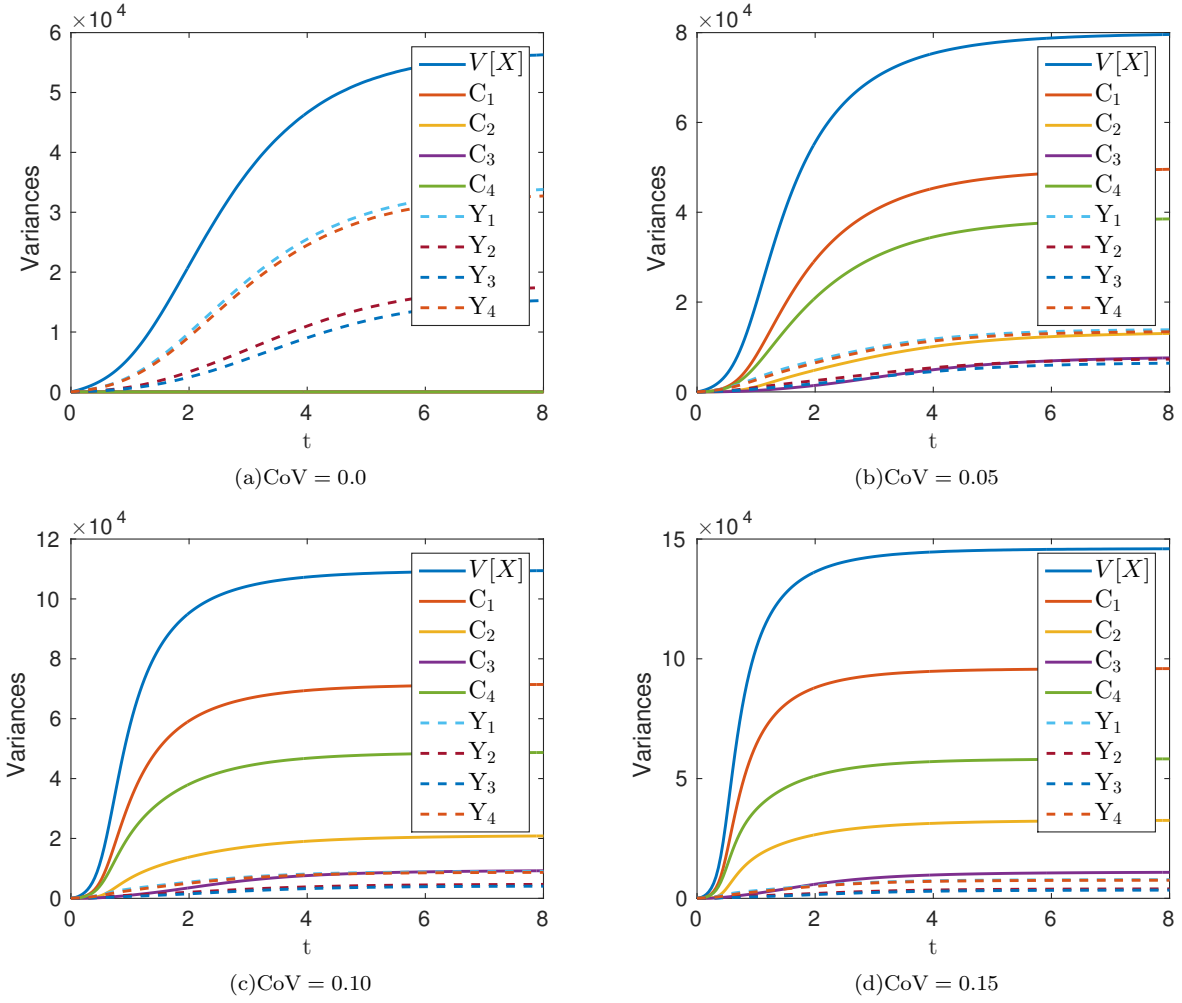
FIG. 14. Total sensitivity indices $T_{C_i}$ and $T_{Y_i}$, scaled by the total variance, versus time. Plotted are results for different values of CoV as indicated.
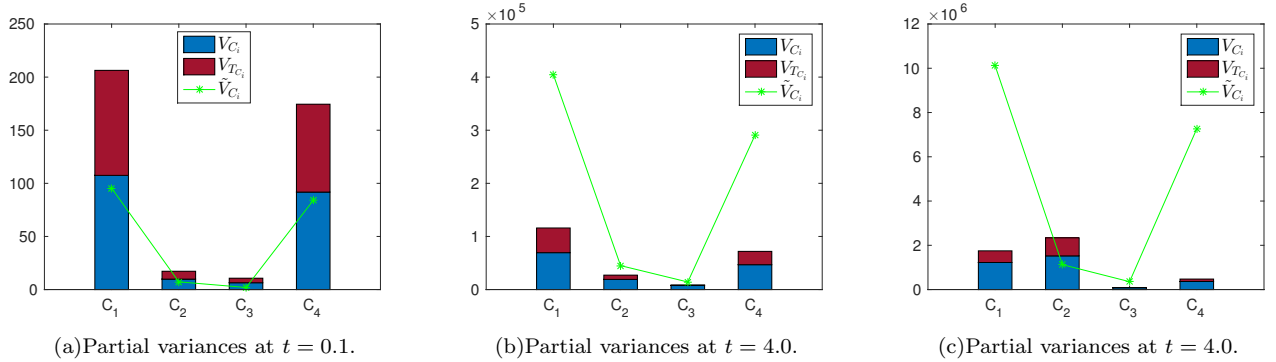


FIG. 15. Comparison of the first order partial variances $V_{C_i}$, total order partial variances $V_{T_{C_i}} = \mathbb{V}[X] T_{C_i}$ and local estimates $\tilde{V}_{C_i}$ using (16). Plots corresponds to CoV = 0.1 for $t = 0.1$ (left) and $t = 4$ (center), and CoV = 0.5 for $t = 4$ (right).

bimodal, and the derivative of its expected value is not representative of the effects of the parameters. Specifically, we see that the estimates $\tilde{V}_{C_i}$ completely overestimate not only the first-order partial variances $V_{C_i}$, but

also the total variances induced by parameter $C_i$, namely the $\mathbb{V}[X] T_{C_i}$. It could be argued that, though the local estimates of the variances are far off the actual partial variances associated with the parameters, the local ap-

proach is still capable of ranking the relative importance of the kinetic parameters. This is, however, not true in general, as can be seen in Figure 15(c) which reports the previous variances at $t = 4$, in the case of a coefficient of variation CoV $= 0.5$: the global sensitivity analysis indicates that the most important source of variability is $C_2$ (because of non-linear saturation effects), whereas with the local approach the ranking is obviously the same as for CoV $= 0.1$, because in the local approach it is independent of CoV.

## VI.  CONCLUSIONS AND FUTURE WORK

In this work, we have extended the variance decomposition approach proposed in Ref. 16 to perform sensitivity analysis in stochastic simulators having uncertain parameters. The parametric uncertainty is treated as an additional source of variability, assumed to be independent of the inherent stochasticity of the reaction channels. This assumption permits the decomposition of the variance of second-order quantities of interest derived from the system state, distinguishing between different variability sources, in particular the channel randomness and the parametric contributions. The approach differs from the classical parametric sensitivity analyses based on the parametric dependences of some statistical moments of the quantities of interest. It offers practical means to fully characterize the complex interplays between inherent sources of noise and individual uncertain parameters effects, through the evaluation of global sensitivity indices associated with an arbitrary set of parameters and random channels.

A Monte Carlo algorithm has been proposed and implemented to illustrate the computation of these sensitivity indices. It has been applied to simple stochastic systems, namely the birth-death and Schlögl models. Through these examples, we have shown that inherent and parametric variability sources influence the model output in a different fashion. In particular, the impact of uncertain parameters becomes more pronounced when their coefficients of variation are increased. In contrast, the inherent variability due to the reaction channels is less affected by the amount of parametric uncertainty, demonstrating the irreducibility of the channels stochasticity.

Numerical examples have also highlighted the importance of a global sensitivity approach, as opposed to local sensitivity methods using derivatives with respect to the uncertain parameters of statistical moments. In particular, our computations showed that first-order derivatives of the first moment was not able to properly account for the interactions between parameters, and is not suitable to assess the global variability induced by a given parameter, even for moderate uncertainty levels. To some extent, these limitations can be remedied considering high-order and mixed derivatives, a global-derivative-based approach[30] or more advanced representations of the para-

metric dependencies of the moments.[31–33] We advocate for the simplicity and generality of the global variance decomposition method, as other sensitivity approaches may miss the characterization of the variance incurred due to the stochasticity of the channels, which is essential for a complete understanding of the stochastic dynamics. The advantages afforded by global sensitivity approach are thus regarded as highly beneficial in a range of applications. Specifically, besides providing an effective means for avoiding the limitations of local sensitivity analyses and moment based approaches, which were highlighted above based on simple model simulations, the presently approach also offers the capability of quantifying the impact of parametric uncertainties, the inherent stochasticity of individual channels and of groups of channels, as well mixed effects involving reducible and irreducible sources of uncertainty. Examples where these features would be of special importance include the analysis of system robustness, model reduction, and parameter calibration.[4,34–38]

The application of the proposed method to complex stochastic systems remains challenging because of the sampling effort. Even though the method is trivial to implement in parallel, future works should focus on the improvement of the methods to lower its computational complexity. Specifically, the results presented in this work relied on exact simulators, such that the errors in the MC estimation of the sensitivity indices are directly related to the sampling error. As for any standard MC method, a rate in $1/\sqrt{M}$ is expected for the convergence of the sensitivity indices, and a large enough samples set must be used to properly explore both the parametric and Poisson processes domains and reduce the variance of the estimators. Note that a different number of samples $M$ could be used depending on the considered sensitivity indices, as their estimators may have fairly different variances. In any case, the convergence with the number of samples is slow and we plan to reduce the computational complexity of Algorithm III.1 in the future, by relying on Multi-Level Monte Carlo strategies.[39–41] Specifically, we envision to alleviate the computational cost of estimating the correlations, between quantities of interest at different parameter values and Poisson processes, using multi-level estimators. The main difficulty in designing such a Multilevel Monte Carlo approach for the estimation of the sensitivity indices will be the enforcement of consistent realizations for the channels Poisson processes in tau-leaping approximations.[23,42–44] Indeed, one would need to sample the same realization of a Poisson process for different time step *and* parameter values, and not just to generate highly correlated but distinct realizations of the Poisson process. This can be achieved *via* a conditional sampling of the Poisson process increments. This would necessitate devising a multi-level procedure capable of balancing the sampling and the bias errors of a set of estimators for the sensitivity coefficients. Progress along this direction will be reported elsewhere.

## ACKNOWLEDGMENTS

## Appendix A: Glossary

| | |
|---|---|
| $\Omega, \omega \in \Omega$ | Event space, random event |
| $\mu, \Sigma$ | Probability measure, $\sigma-$algebra |
| $M_s$ | Numbers of chemical species |
| $S_1, \ldots, S_{M_s}$ | Chemical species |
| $X_i$ | Random number of $S_i$ molecules |
| $\boldsymbol{X} = (X_1 \ldots X_{M_s})$ | Random state vector |
| $M_r$ | Number of reaction channels |
| $R_1, \ldots, R_{M_r}$ | Reaction channels |
| $Y_j$ | Poisson process of channel $R_j$ |
| $\boldsymbol{Y} = (Y_1 \ldots Y_{M_r})$ | Vector of Poisson processes |
| $a_j$ | Propensity functions of channel $R_j$ |
| $\boldsymbol{\nu}_j$ | Stoichiometric vector of channel $R_j$ |
| $t, t_j$ | Time and scaled time of channel $R_j$ |
| $\mathrm{RG}_j$ | Random number generator for $Y_j$ |
| $N_c$ | Number of kinetic parameters |
| $\boldsymbol{c}, \boldsymbol{C}$ | Deterministic and uncertain vectors of kinetic parameters |
| $p_{\boldsymbol{C}}(\boldsymbol{c})$ | Probability density function of $\boldsymbol{C}$ |
| $N_U = M_r + N_c$ | Total number of stochastic sources |
| $\boldsymbol{U} = (\boldsymbol{C}, \boldsymbol{Y})$ | Vector of all stochasticity sources |
| $G(\boldsymbol{X})$ | Second order random functional |
| $\mathbb{E}[\cdot], \mathbb{V}[\cdot]$ | Expectation and variance operators |
| $V_{\mathrm{par}}$ | Variance of $G$ due to parameters ($\boldsymbol{C}$) |
| $V_{\mathrm{ch}}$ | Variance of $G$ due to the channels ($\boldsymbol{Y}$) |
| $V_{\mathrm{mix}}$ | Variance of $G$ due to joint effects |
| $S_{\mathrm{par,ch,mix}}$ | First order sensitivity indices |
| $T_{\mathrm{par,ch,mix}}$ | Total order sensitivity indices |
| $\mathrm{u} \in \{1, \ldots, N_U\}$ | Subset of stochastic sources |
| $S_{\mathrm{u}}, T_{\mathrm{u}}$ | First and Total order sensitivity indices associated to u |
| $V_{S_{\mathrm{u}}}, V_{T_{\mathrm{u}}}$ | Scaled version of the sensitivity indices |
| $L_{C_i}$ | Local sensitivity index |
| $\widehat{V_\bullet}, \widehat{T_\bullet}, \widehat{S_\bullet}, \ldots$ | Monte Carlo estimates |
| $\tilde{V}_{C_i}$ | Estimate of $V_{T_{C_i}}$ based on $L_{C_i}$ |
| CoV | Coefficient of Variation |
| $B, D$ | Uncertain birth and death rates |

[1] H. Rabitz, M. Kramer, and D. Dacol, Annual Review of Physical Chemistry **34**, 419–461 (1983).
[2] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo, Chemical Reviews **105**, 2811–2828 (2005).
[3] M. Reagan, H. Najm, B. Debusschere, O. Le Maître, O. Knio, and R. Ghanem, Combustion Theory and Modelling **8**, 607 (2004), http://www.tandfonline.com/doi/pdf/10.1088/1364-7830/8/3/010.
[4] M. Rathinam, P. W. Sheppard, and M. Khammash, J Chem Phys **132**, 034103 (2010).
[5] P. W. Sheppard, M. Rathinam, and M. M. Khammash, The Journal of Chemical Physics **136**, 034115 (2012).
[6] E. S. Wolf and D. F. Anderson, The Journal of Chemical Physics **137**, 224112 (2012).
[7] E. S. Wolf and D. F. Anderson, The Journal of Chemical Physics **142**, 034103 (2015).
[8] I. Sobol', Math. Modeling & Comput. Exp. **1**, 407 (1993).
[9] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity analysis in practice: a guide to assessing scientific models* (John Wiley & Sons, 2004).
[10] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach* (Springer Verlag, 1991).
[11] O. Le Maître and O. Knio, *Spectral Methods for Uncertainty Quantification* (Springer, New York, NY, 2010).
[12] T. Crestaux, O. Le Maître, and J. M. Martinez, Reliability Engineering & System Safety **94**, 1161 (2009).
[13] A. Alexanderian, F. Rizzi, M. Rathinam, O. P. Le Maître, and O. M. Knio, Journal of Scientific Computing **58**, 592 (2013).
[14] O. Le Maître and O. Knio, Reliability Engineering & System Safety **135**, 107 (2015).
[15] M. Navarro Jimenez, O. Le Maître, and O. Knio, SIAM J. Uncertainty Quantification (submitted).
[16] O. P. Le Maître, O. M. Knio, and A. Moraes, The Journal of Chemical Physics **142**, 244115 (2015).
[17] T. G. Kurtz, *Approximation of Population Processes (CBMS-NSF Regional Conference Series in Applied Mathematics)* (SIAM, 1981).
[18] P. Érdi and G. Lente, *Stochastic Chemical Kinetics: Theory and (Mostly) Systems Biological Applications* (Springer Series in Synergetics, Springer, 2009).
[19] C. W. Gardiner, S. Chaturvedi, and D. F. Walls, Annals of the New York Academy of Sciences **316**, 453 (1979).
[20] D. T. Gillespie, Journal of Computational Physics **22**, 403 (1976).
[21] D. T. Gillespie, Journal of Physical Chemistry **81**, 2340 (1977).
[22] Y. Cao, D. Gillespie, and L. Petzold, The Journal of Chemical Physics **104**, 044109 (2006).
[23] A. Chatterjee, D. Vlachos, and M. Katsoulakis, The Journal of Chemical Physics **122**, 024112 (2005).
[24] D. F. Anderson, The Journal of Chemical Physics **127**, 214107 (2007).
[25] M. A. Gibson and J. Bruck, Journal of Physical Chemistry **104**, 1876 (2000).
[26] A. Saltelli, Computer Physics Communications **145**, 280 (2002).
[27] M. J. W. Jansen, Computer Physics Communications **117**, 35 (1999).
[28] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola, Computer Physics Communications **181**, 259 (2010).
[29] F. Schlögl, Z. Physik **253**, 147 (1972).
[30] I. Sobol and S. Kucherenko, Mathematics and Computers in Cimulation **79**, 3009 (2009).
[31] F. Rizzi, R. Jones, B. Debusschere, and O. Knio, J. Chem. Phys. **138**, 194104 (2013).
[32] F. Rizzi, R. Jones, B. Debusschere, and O. Knio, J. Chem. Phys. **138**, 194105 (2013).
[33] F. Rizzi, H. Najm, B. Debusschere, K. Sargsyan, M. Salloum, H. Adalsteinsson, and O. Knio., Multiscale Model. Simul. **10**, 14281459 (2012).
[34] D. Gillespie, Y. Cao, K. Sanft, and L. Petzold, The Journal of Chemical Physics **130**, 064103 (2009).

[35] J. Wu, B. Vidakovic2, and E. Voit, BMC Systems Biology **5**, 187 (2011).

[36] P. Thomas, A. Straube, and R. Grima, BMC Systems Biology **6**, 39 (2012).

[37] S. Wu, J. Fu, H. Li, and L. Petzold, The Journal of Chemical Physics **137**, 034106 (2012).

[38] M. Barrio, A. Leier, and T. Marquez-Lago, The Journal of Chemical Physics **138**, 104114 (2013).

[39] M. Giles, Operations Research **56**, 607 (2008).

[40] D. F. Anderson and D. J. Higham, Multiscale Modeling & Simulation **10**, 146 (2012), http://dx.doi.org/10.1137/110840546.

[41] C. Lester, C. A. Yates, M. B. Giles, and R. E. Baker, The Journal of Chemical Physics **142**, 024113 (2015), http://dx.doi.org/10.1063/1.4904980.

[42] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie, The Journal of Chemical Physics **119**, 12784 (2003).

[43] A. Moraes, R. Tempone, and P. Vilanova, Multiscale Model. Simul. **12**, 581 (2014).

[44] D. F. Anderson, A. Ganguly, and T. G. Kurtz, Ann. Appl. Probab. **21**, 2226 (2011).