Global Sensitivity Analysis in Stochastic Systems

Olivier Le Maître, Omar Knio

LIMSI CNRS, Orsay, France
KAUST, Saudi-Arabia
Duke University, Durham, North Carolina

CEMRACS - CIRM

Stochastic models

Physical systems with

- Complex small scale dynamics (MD, chemical systems, . . .)
- Random forcing and source terms (finance, wind-load, . . .)
- Unresolved scales (turbulence, climate modeling, . . .)

are often tackled by means of stochastic modeling where complex / unknown / unresolved phenomenons are accounted for by the introduction of noisy dynamics.

In addition to the effect of the noise, the model may involve unknown parameters : *e.g.* noise level, physical constants and parameters, initial conditions, . . .

Our general objective is to propagate / assess the impact of parameters uncertainty within such stochastic models while characterizing the effect of inherent noise :

global sensitivity analysis & analysis of the variance

Stochastic ODEs

We consider a simple systems driven by random noise (Ito equation) : for $t \in [0, T] \doteq \mathcal{T}$

$$dX(t) = C(X(t))dt + D(X(t))dW(t), \quad X(t = 0) = X_0,$$

where

- $X(t) \in \mathbb{R}$ is the solution,
- $W(t)$ is the Wiener process,
- $C(\cdot)$ is the drift function,
- and $D(\cdot)$ is the diffusion coefficient.

The solution can be computed through MC simulation, solving *(e.g.)*

$$X_{i+1} = X_i + C(X_i)\Delta t + D(X_i)\Delta W_i, \quad X_i \approx X(i\Delta t),$$

drawing iid random variables $\Delta W_i \sim N(0, \Delta t)$.
Sample estimate expectation, moments, quantiles, probability law, . . ., of the stochastic process $X(t)$:

$$\mathbb{E}\{g(X(i\Delta t))\} \approx \frac{1}{M}\sum_{l=1}^{M} g(X_i^l).$$

Stochastic ODEs with parametric uncertainty

The drift function and diffusion coefficient can involve some uncertain parameters Q :

$$dX(t) = C(X(t); Q)dt + D(X(t); Q)dW(t), \quad X(t = 0) = X_0.$$

We consider that :

- Q random with known probability law,
- Q and W are assumed independent.

The solution can be seen as a functional of $W(t)$ and Q : $X(t) = X(t, W, Q)$. We shall assume, $\forall t \in \mathcal{T}$,

1. $\mathbb{E}\left\{X^2\right\} < \infty$,
2. $\mathbb{E}\left\{\mathbb{E}\left\{X|W\right\}^2\right\} \doteq \mathbb{E}\left\{X_{|W}^2\right\} < \infty$,
3. $\mathbb{E}\left\{\mathbb{E}\left\{X|Q\right\}^2\right\} \doteq \mathbb{E}\left\{X_{|Q}^2\right\} < \infty$.

We want to investigate the respective impact of Q, W on X.

Classical sensitivity analysis

Focusing on the two first moments, global SA for the random parameters Q is based on :

1. approximating the mean and variance of $X_{|Q}$

$$\mathbb{E}\left\{X_{|Q}\right\} = \mu_X(Q), \quad \mathbb{V}\left\{X_{|Q}\right\} = \Sigma_X^2(Q),$$

2. perform a GSA of $\mu_X(Q)$ and $\Sigma_X^2(Q)$ with respect to the input parameters in Q.

In particular, for independent parameters Q, Polynomial Chaos approximations :

$$\mu_X(Q) \approx \sum_\alpha \mu_\alpha \Psi_\alpha(Q), \quad \Sigma_X^2(Q) \approx \sum_\alpha \Sigma_\alpha^2 \Psi_\alpha(Q).$$

PC expansion coefficients can be computed / estimated by means of Non-Intrusive Spectral Projection, Bayesian identification, ...

This approach characterizes the dependence of the first moments with respect to the parameters Q.

Another approach of GSA

Here, we exploit the structure of the model to take an alternative approach, inspired from the hierarchical orthogonal Sobol-Hoeffding decomposition of X :

$$X(W,Q) = \overline{X} + X_W(W) + X_Q(Q) + X_{W,Q}(W,Q), \quad \forall t \in \mathcal{T},$$

where the functionals in the SH decomposition are mutually orthogonal.
In fact, the decomposition is unique and given by

- $\overline{X}(t) \doteq \mathbb{E}\{X(t)\}$,
- $X_W(t,W) \doteq \mathbb{E}\{X(t)|W\} - \mathbb{E}\{X(t)\} = X_{|W}(t) - \overline{X}(t)$,
- $X_Q(t,Q) \doteq \mathbb{E}\{X(t)|Q\} - \mathbb{E}\{X(t)\} = X_{|Q}(t) - \overline{X}(t)$.

Owing to the orthogonality of the SH decomposition, we have

$$\mathbb{V}\{X\} = \mathbb{V}\{X_W\} + \mathbb{V}\{X_Q\} + \mathbb{V}\{X_{W,Q}\},$$

from which follow the definitions of the sensitivity indices

$$\mathrm{S}_W = \frac{\mathbb{V}\{X_W\}}{\mathbb{V}\{X\}}, \quad \mathrm{S}_Q = \frac{\mathbb{V}\{X_Q\}}{\mathbb{V}\{X\}}, \quad \mathrm{S}_{W,Q} = \frac{\mathbb{V}\{X_{W,Q}\}}{\mathbb{V}\{X\}}.$$

Sensitivity indices

The sensitivity indices

$$\mathrm{S}_W = \frac{\mathbb{V}\{X_W\}}{\mathbb{V}\{X\}}, \quad \mathrm{S}_Q = \frac{\mathbb{V}\{X_Q\}}{\mathbb{V}\{X\}}, \quad \mathrm{S}_{W,Q} = \frac{\mathbb{V}\{X_{W,Q}\}}{\mathbb{V}\{X\}},$$

then measure the fraction of the variance due to

- the Wiener noise only, or intrinsic randomness (S_W),
- the parameters only, or parametric randomness (S_Q),
- the combined effect of intrinsic and parametric randomness ($\mathrm{S}_{W,Q}$).

In particular, S_W measure the part of the variance that cannot be reduced through a better knowledge of the parameters.

In addition,

$$\frac{\mathbb{V}_Q\{\mu_X(Q)\}}{\mathbb{V}\{X\}} = \mathrm{S}_Q, \quad \text{but} \quad \frac{\mathbb{E}_Q\{\Sigma^2(Q)\}}{\mathbb{V}\{X\}} = \mathrm{S}_W + \mathrm{S}_{W,Q}.$$

From $\Sigma^2(Q)$, one cannot distinguish the intrinsic and mixed randomness effects.

Polynomial Chaos expansion

We express the dependence of X on Q as a PC expansion

$$X(t, W, Q) = \sum_{\alpha} X_{\alpha}(t, W)\Psi_{\alpha}(Q),$$

where

- $\{\Psi_{\alpha}\}$ is a CONS of $L^2(Q, p_Q)$,
- the expansion coefficients X_{α} are random processes.

The random processes $X_{\alpha}(t)$ are the solutions of the coupled system of SODEs

$$dX_{\beta}(t) = \left\langle F\left(\sum_{\alpha} X_{\alpha}(t)\Psi_{\alpha}; Q\right), \Psi_{\beta}\right\rangle dt + \left\langle G\left(\sum_{\alpha} X_{\alpha}(t)\Psi_{\alpha}; Q\right), \Psi_{\beta}\right\rangle dW,$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L^2(Q, p_Q)$.

This system can be solved by MC simulation (upon truncation).

PC expansion

Assuming $\Psi_0 = 1$, it comes

$$\mathbb{E}\{X\} = \mathbb{E}\{X_0\}, \quad X_Q(Q) = \sum_{\alpha \neq 0} \mathbb{E}\{X_\alpha\} \Psi_\alpha(Q), \quad X_W(W) = X_0(W) - \mathbb{E}\{X_0\},$$

and

$$X_{W,Q}(W,Q) = \sum_{\alpha \neq 0} (X_\alpha(W) - \mathbb{E}\{X_\alpha\}) \Psi_\alpha(Q).$$

Finally, the partial variances have for expression :

$$\mathbb{V}\{X_Q\} = \sum_{\alpha \neq 0} \mathbb{E}\{X_\alpha\}^2, \quad \mathbb{V}\{X_W\} = \mathbb{V}\{X_0\}, \quad \mathbb{V}\{X_{W,Q}\} = \sum_{\alpha \neq 0} \mathbb{V}\{X_\alpha\}.$$

Observe :

1. $X_Q(Q) + \mathbb{E}\{X\} = \mu_X(Q)$,
2. $\sum_\alpha \mathbb{V}\{X_\alpha\} = \sum_\alpha \mathbb{E}\{X_\alpha^2\} - \mathbb{E}\{X_\alpha\}^2 = \mathbb{E}_Q\{\Sigma_X^2\}$.

Linear additive system

- Consider SODE with drift and diffusion terms given by :

$$C(X, Q) = Q_1 - X \qquad D(X, Q) = (\nu X + 1)Q_2$$

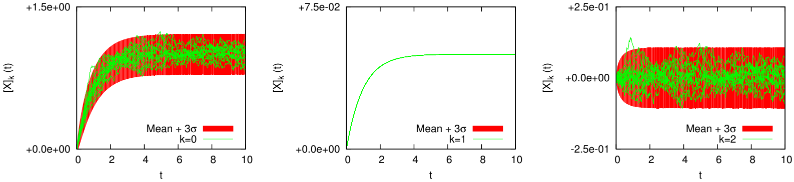
  where Q_1 and Q_2 are independent, uniformly-distributed, random variables with mean $\mu_{1,2}$ and standard deviation $\sigma_{1,2}$.
- The orthonormal PC basis consists of tensorized Legendre polynomials.
- We use for initial condition $X(t = 0) = 0$ almost surely.

Additive Noise

Additive noise model ($\nu = 0$) with $\mu_1 = 1$, $\mu_2 = 0.1$, $\sigma_1 = \sigma_2 = 0.05$:

$$dX(Q) = (Q_1 - X(Q))dt + Q_2 dW,$$

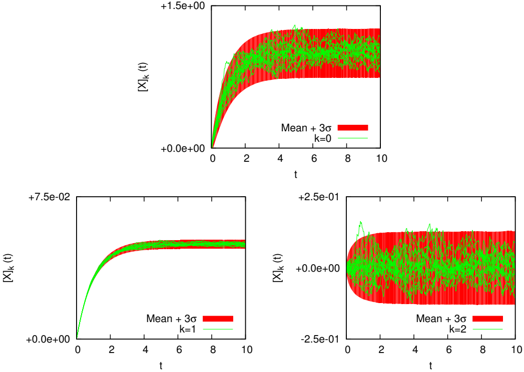
a first-order expansion suffices to exactly represent $X(Q)$.



Selected trajectories and variability ranges for $[X_k](t, W)$. The plots correspond to $k = 0$, 1 and 2, arranged from left to right.

Multiplicative Noise – I

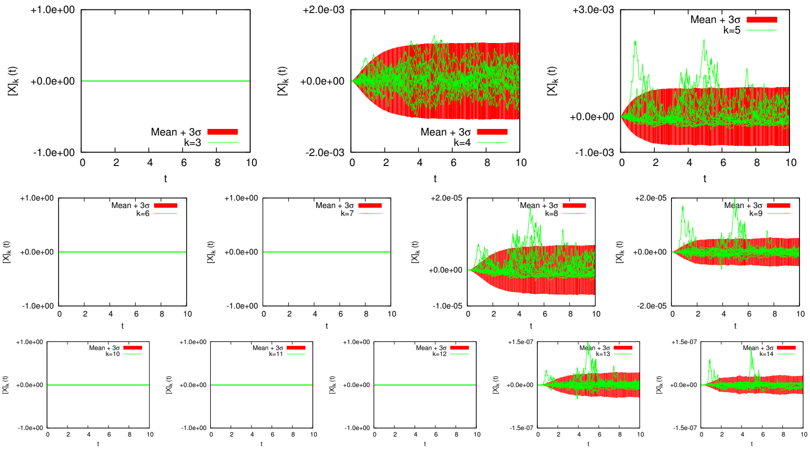
Multiplicative noise : $Q_1 \sim \mathcal{U}[1, 0.05]$, $Q_2 \sim \mathcal{U}[0.1, 0.05]$, $\nu = 0.2$



Sample trajectories of $[X_k]$, $0 \leq k \leq 2$. Top row : order 0, bottom row : order 1 with and decreasing order in Q_1 from left to right.

Multiplicative Noise – II

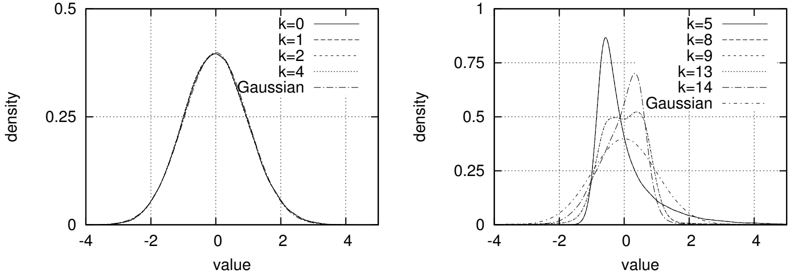
Multiplicative noise : $Q_1 \sim \mathcal{U}[1, 0.05]$, $Q_2 \sim \mathcal{U}[0.1, 0.05]$, $\nu = 0.2$



Sample trajectories of $[X_k]$, $3 \leq k \leq 14$. The total order ranges from 2 (top row) to 4 (bottom row), with and decreasing order in Q_1 from left to right.

Distribution functions

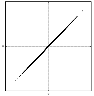
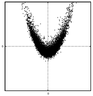
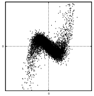
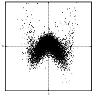
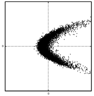
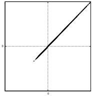
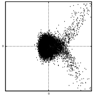
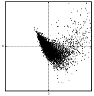
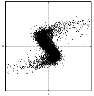
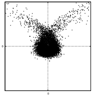
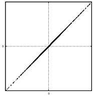
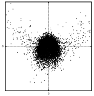
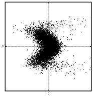
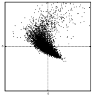
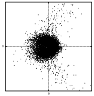
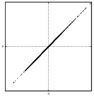
Multiplicative noise : $Q_1 \sim \mathcal{U}[1, 0.05]$, $Q_2 \sim \mathcal{U}[0.1, 0.05]$, $\nu = 0.2$



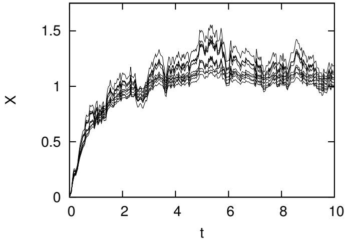
Probability density functions of the modes $[X_k]$ at $t = 10$. The modes have been centered and normalized to facilitate the comparison ; the standard Gaussian distribution is also reported for reference.

Mode correlations

Projections in the planes $([X_k], [X_{k'}])$ of realizations of the centered and normalized solution vector $\boldsymbol{X}$ at time $t = 10$, for selected indices

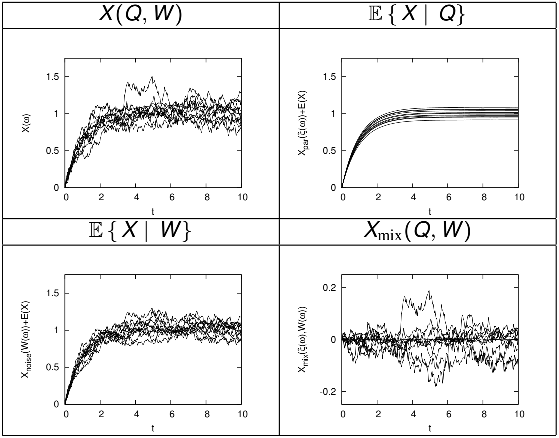
$[X_k]$ vs $[X_{k'}]$	$k' = 0$	$k' = 5$	$k' = 9$	$k' = 14$
$k = 0$				
$k = 5$				
$k = 9$				
$k = 14$				

Conditional trajectories



Left : trajectories for samples of Q and a *fixed* realization of W
Right : trajectories for samples of W at a *fixed* value of the parameters.

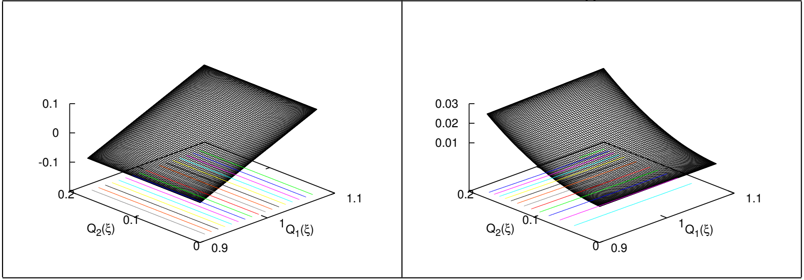
SH functions



Selected trajectories of X and its SH functions.

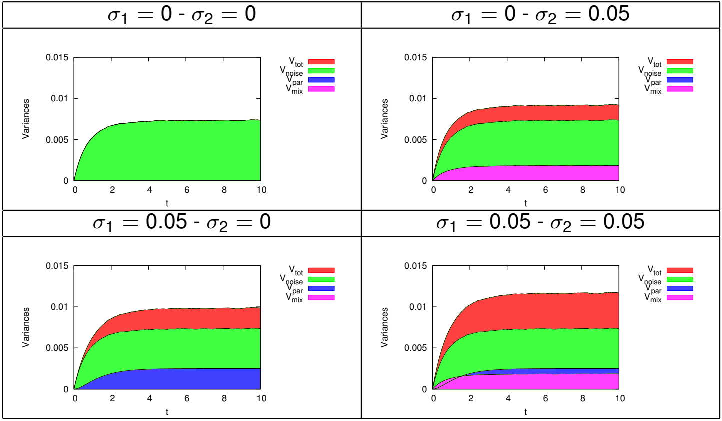
Parametric sensitivity

$\mu_X(Q) - \mathbb{E}\{X\}$	$\Sigma^2_X(Q)$



Effect of Q_1 and Q_2 of the (centered) conditional mean $\mu_X(Q) = \mathbb{E}\{X \mid Q\} - \mathbb{E}\{X\}$ and variance $\Sigma^2_X(Q) = \mathbb{V}\{X \mid Q\}$ at time $t = 10$

ANOVA



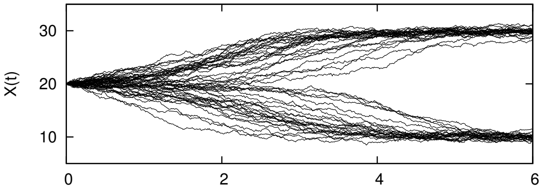
Evolution of the components of the total variance. Shown are variance decompositions obtained for different values of σ_1 and σ_2

Non-linear system

Consider a system with additive noise and non-linear drift

$$dX = F(X)dt + \delta dW = -\gamma(X-a)(X-b)(X-c)dt + \delta dW$$

where $\delta > 0$ is an additional parameter controlling the noise level, and as before W is a Wiener process. Again the IC is $X_0 = X(t=0)$.



Sample trajectories with $a = 10$, $b = 20$, $c = 30$, $\gamma = 0.01$, and $\delta = 1$. In all cases, the initial condition coincides with $x^0 = b$.

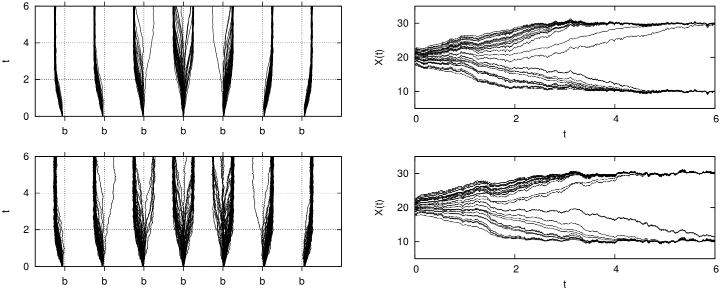
Parametric uncertainty

- Consider an uncertain initial condition, $Q_1 \sim \mathcal{R}[17.5, 22.5]$, and forcing amplitude, $Q_2 \sim \mathcal{R}[0.5, 1.5]$.

$$dX = F(X)dt + Q_1 dW \quad X_0 = Q_2.$$

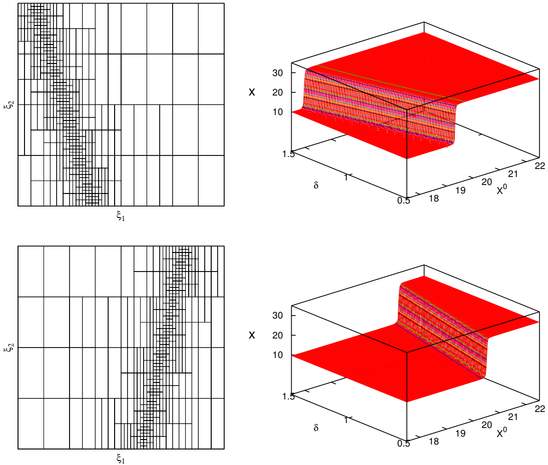
- Q_1 and Q_2 independent.
- The PC representation is based on an adaptive multiwavelet basis expansion, which enables us to accommodate for bifurcation(s).
- The use of a non polynomial basis complicates the sensitivity analysis, but the framework is essentially unaltered.

Sample trajectories



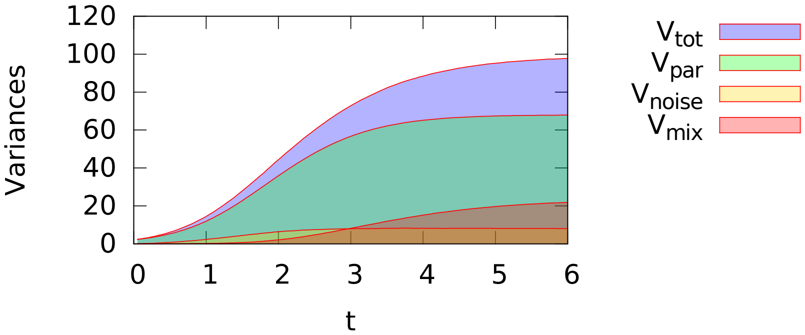
Left plots : sample set of realizations of W, the trajectories of X (time running up) for different initial conditions and two noise levels $Q_2 = 0.65$ (top plot) and $Q_2 = 1.35$ (bottom).
Right plots show for two realizations of W (top and bottom), the trajectories of X for a random sample set of values of $Q_1 0$ and Q_2.

MW expansion



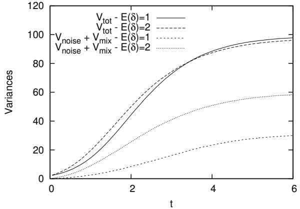
Left : partitions of the parametric domain and surface plots for $X(t = 6, W)$ as a function of Q.

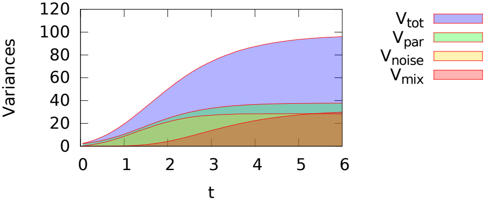
Variance decomposition



Partial variances of $X(t)$

Effect of Noise amplitude





Left : comparison of the total variances $\mathbb{V}\{X\}$ and total noise contributions $V_{\text{noise}} + V_{\text{mix}}$ to the variance, for two expected values of $\mathbb{E}\{\delta\} = 1$ and 2.
Right : partial variances of the stochastic process $X(t)$ for the case $\mathbb{E}\{\delta\} = 2$.

Extension to Non-Intrusive Projection

- The PC expansion of $X(t, W, Q)$ can be estimated non-intrusively, *e.g.* :

$$X(t, W^{(i)}, Q) \approx \sum_{\alpha} X_\alpha(t, W^{(i)})\Psi_\alpha(Q), \quad X_\alpha(t, W^{(i)}) \approx \sum_{\beta=1}^{N_Q} \Pi_{\alpha,\beta} X(t, W^{(i)}, Q^{(\beta)})$$

- For instance sparse grid pseudo spectral projection operator [Π] [Conrad, Marzouk] & [Constantine et al]
- Provides accurate Q-statistics for each path $W^{(i)}$ from only N_Q simulations
- Yields complexity reduction when Q-variance is dominant
- Applied to non-smooth QoI $g(X)$, such as exit time. [Navarro, OLM, Knio, JUQ 2016]

Stochastic Systems

Stochastic Simulator

Work with Omar Knio, Alvaro Moraes and Maria Navarro (KAUST)

Stochastic Systems
Stochastic systems

governed by probabilistic evolution rules expresses by the master equation

$$\frac{\partial P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0)}{\partial t} = \sum_{j=1}^{\mathrm{K}_r} \left[a_j(\boldsymbol{x} - \boldsymbol{\nu}_j) P(\boldsymbol{x} - \boldsymbol{\nu}_j, t|\boldsymbol{x}_0, t_0) - a_j(\boldsymbol{x}) P(\boldsymbol{x}, t|\boldsymbol{x}_0, t_0) \right],$$

- $\boldsymbol{x}(t) \in \mathbb{Z}^{\mathrm{M}_s}$: state of the system at time t,
- K_r reactions channels,
- propensity functions a_j and state-change vectors $\boldsymbol{\nu}_j \in \mathbb{Z}^{\mathrm{M}_s}$,
- $P(\boldsymbol{x}, t|\boldsymbol{x}_o, t_o)$: probability of $\boldsymbol{X} = \boldsymbol{x}$ at time t, given $\boldsymbol{X} = \boldsymbol{x}_0$ at time t_0,
- Markov process.

Examples includes Reactive Networks (chemistry, biology), social networks, ...

- Direct resolution of the master equation is usually not an option,
- Simulate trajectories $\boldsymbol{X}(t) \sim P(\boldsymbol{x}, t|\boldsymbol{x}_o, t_o)$, using a **stochastic simulator**.

Gillespie's Algorithm

Given $\boldsymbol{X}(t) = \boldsymbol{x}$, the probability of the next reaction to occur in the $[t, t + dt)$ is

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

The time to the next reaction, τ, follows an exponential distribution with mean $1/a_0(\boldsymbol{x})$.

Gillespie's Algorithm : [Gillespie, 1970's]

1. Set $t = t_0$, $\boldsymbol{X} = \boldsymbol{x}_0$.
2. Repeat until $t > T$
   - Draw $\tau \sim \exp a_0(\boldsymbol{X})$
   - Pick randomly $k \in \{1 \ldots \mathrm{K}_r\}$ with relative probability $p_k(a_k)$
   - update $t \leftarrow t + \tau$, $\boldsymbol{X} \leftarrow \boldsymbol{X} + \nu_k$
3. Return $X(T) \sim P(\boldsymbol{x}, t|\boldsymbol{x}_o, t_o)$.

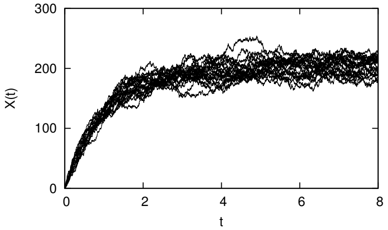
From a sample set of trajectory, estimate expectation of functionals $\mathbb{E}\left\{g(X)\right\}$.

Sobol Analysis of the variance

From the stochastic state $\boldsymbol{X}(t)$ and a given functional g, we would like :

assess the contributions of different reaction channels (or group of) on the variability of $g(\boldsymbol{X})$

For instance : which channel(s) is (are) responsible for most of the variance in $g(\boldsymbol{X})$?



This is **not to be confused with parametric sensitivity analyses** where one wants to estimate the sensitivity of $\mathbb{E}\{g(\boldsymbol{X})\}$ with respect to some parameters $\boldsymbol{q}$ in the definition of the dynamics (*e.g.* propensity functions).

Sobol Analysis of the variance

- $\boldsymbol{N}(\omega) = (N_1, \cdots, N_{\mathrm{D}})$ a set of D independent random inputs N_i,
- $F(\boldsymbol{N})$ a (second-order) random functional in $\boldsymbol{N}$,

$F(\boldsymbol{N})$ has a unique orthogonal decomposition [Sobol, 2002 ; Homma & Saltelli, 1996]

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

where $\mathcal{D}$ is the power set of $\{1, \cdots, \mathrm{D}\}$ and $\boldsymbol{N}_{\mathfrak{u}} = (N_{\mathrm{u}_1}, \cdots, N_{\mathrm{u}_{|\mathfrak{u}|}})$. The orthogonality condition reads

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

so

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

where $\mathbb{V}\{F_{\mathfrak{u}}\}$ are the partial variances.

Sobol Analysis of the variance (cont...)

From the variance decomposition,

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

- First order sensitivity indices $\mathrm{S}_{\mathfrak{u}}$: fraction of the variance caused by the random inputs $\boldsymbol{N}_{\mathfrak{u}}$ only

$$\mathbb{V}\{F\}\,\mathrm{S}_{\mathfrak{u}} = \sum_{\mathfrak{s} \supseteq \mathfrak{u}}^{\mathfrak{s} \neq \emptyset} \mathbb{V}\{F_{\mathfrak{u}}\}$$

- Total order sensitivity indices $\mathrm{T}_{\mathfrak{u}}$: fraction of the variance caused by the random inputs $\boldsymbol{N}_{\mathfrak{u}}$ and interaction

$$\mathbb{V}\{F\}\,\mathrm{T}_{\mathfrak{u}} \sum_{\mathfrak{s} \in \mathcal{D}}^{(\mathfrak{s} \cap \mathfrak{u}) \neq \emptyset} \mathbb{V}\{F_{\mathfrak{s}}\}$$

The partial variances $\mathbb{V}\{F_{\mathfrak{u}}\}$ can be expressed as conditional variances : [Homma & Saltelli, 1996]

$$\mathbb{V}\{F_{\mathfrak{u}}\} = \mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathfrak{u}}\}\} - \sum_{\substack{\mathfrak{s} \in \mathcal{D} \setminus \emptyset \\ \mathfrak{s} \subsetneq \mathfrak{u}}} \mathbb{V}\{F_{\mathfrak{s}}\},$$

or

$$\mathbb{V}\{F\}\,\mathrm{S}_{\mathfrak{u}} = \mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathfrak{u}}\}\}, \quad \mathbb{V}\{F\}\,\mathrm{T}_{\mathfrak{u}} = \mathbb{V}\{F\} - \mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathfrak{u}_\sim}\}\} = \mathbb{V}\{F\}(1 - \mathrm{S}_{\mathfrak{u}_\sim}),$$

where $\mathfrak{u}_\sim = \{1, \dots, \mathrm{D}\} \setminus \mathfrak{u}$.

Decomposition of the Variance = Estimation of conditional variances

Monte-Carlo estimation of the sensitivity indices

Consider two independent sample sets $\mathcal{N}^{①}$ and $\mathcal{N}^{②}$ of M realizations of $\boldsymbol{N}$.
The conditional variance $\mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathfrak{u}}\}\}$ can be estimated as [Sobol, 2001]

$$\mathbb{V}\{\mathbb{E}\{F \mid \boldsymbol{N}_{\mathfrak{u}}\}\} + \mathbb{E}\{F\}^2 = \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} F(\boldsymbol{N}_{\mathfrak{u}}^{①,(i)}, \boldsymbol{N}_{\mathfrak{u}\sim}^{①,(i)})F(\boldsymbol{N}_{\mathfrak{u}}^{①,(i)}, \boldsymbol{N}_{\mathfrak{u}\sim}^{②,(i)}),$$

such that

$$\widehat{\mathrm{S}_{\mathfrak{u}}} = \frac{\frac{1}{M}\sum_{i=1}^{M} F(\boldsymbol{N}^{①,(i)})F(\boldsymbol{N}_{\mathfrak{u}}^{①,(i)}, \boldsymbol{N}_{\mathfrak{u}\sim}^{②,(i)}) - \widehat{\mathbb{E}\{F\}}^2}{\widehat{\mathbb{V}\{F\}}},$$

and

$$\widehat{\mathrm{T}_{\mathfrak{u}}} = 1 - \frac{\frac{1}{M}\sum_{i=1}^{M} F(\boldsymbol{N}^{①,(i)})F(\boldsymbol{N}_{\mathfrak{u}}^{②,(i)}, \boldsymbol{N}_{\mathfrak{u}\sim}^{①,(i)}) - \widehat{\mathbb{E}\{F\}}^2}{\widehat{\mathbb{V}\{F\}}}$$

where $\widehat{\mathbb{E}\{F\}}$ and $\widehat{\mathbb{V}\{F\}}$ are the classical MC estimators for the mean and variance.
The computational complexity scales linearly with the number of indices to be computed

Application to Stochastic Simulators

To assess the respective impacts of different reaction channels through Sobol's decomposition of $\mathbb{V}\{g(\boldsymbol{X})\}$, when $\boldsymbol{X}$ is the output of a stochastic simulator, we need to condition $\boldsymbol{X}$ on the channels dynamics :

What is a particular realization of a channel dynamics ?

Gillespie's algorithm is not suited, and we have to recast the stochastic algorithm in terms of

independent processes associated to each channel.

Next Reaction Formulation. [Ethier &Kurtz, 2005, Gibson & Bruck, 2000]

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

where the $N_j(t)$ are independent standard (unit rate) Poisson processes, and the scaled times t_j are given by

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

Then, $g(\boldsymbol{X})$ can be seen as

$$g(\boldsymbol{X}) = F(N_1, \ldots, N_{\mathrm{K}_r}).$$

Application to Stochastic Simulators (cont.)

The random functional $g(\boldsymbol{X}) = F(N_1, \dots, N_{\mathrm{K}_r})$ can then be decomposed *à la* Sobol.

A particular realization of a channel dynamic is identified with a realization of the underlying standard Poisson processes.

For instance, the conditional variance writes

$$\mathbb{E}\left\{ g(\boldsymbol{X}) \mid \boldsymbol{N}_{\mathbf{u}} = \boldsymbol{n}_{\mathbf{u}} \right\} = \mathbb{E}\left\{ g\left(\boldsymbol{X}(t_0) + \sum_{j \in \mathbf{u}} \boldsymbol{\nu}_j n_j(t_j) + \sum_{j \in \mathbf{u}_\sim} \boldsymbol{\nu}_j N_j(t_j) \right) \right\},$$

with $t_j = \int_{t_0}^{t} a_j(\boldsymbol{X}(\tau)) d\tau$.

Note that

- in general, all firing times t_j remain random for given $\boldsymbol{n}_{\mathbf{u}}(t)$, as they depend on $\boldsymbol{N}_{\mathbf{u}_\sim}$
- in practice, the standard Poisson processes N_j are entirely specified by their random seeds and pseudo-random number generator :

  the Poisson processes don't have to be stored but are computed on the fly

The birth-death (BD) process

Single species S ($\mathrm{M}_s = 1$) and $\mathrm{K}_r = 2$ reaction channels :

$$\emptyset \xrightarrow{b} S, \quad S \xrightarrow{d} \emptyset,$$

with propensity functions

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

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

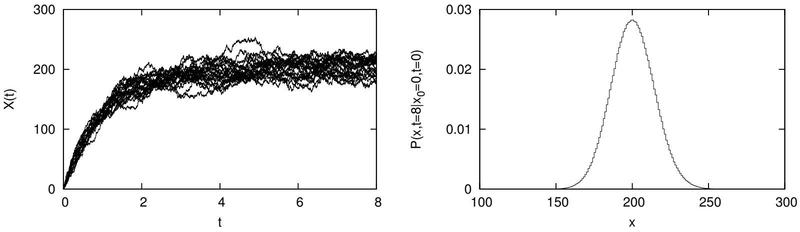
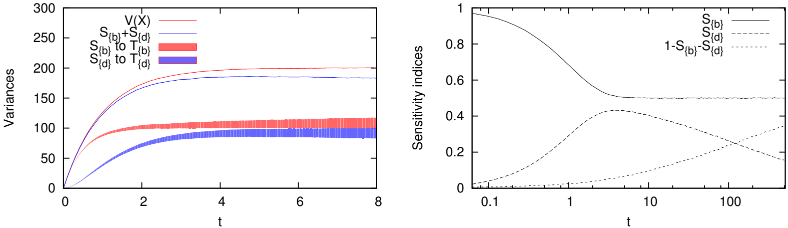


FIGURE: Left : Selected trajectories of $X(t)$ generated using Next Reaction Algorithm. Right : histogram of $X(t = 8)$.

B-D process. Variance decomposition of $g(X) = X(t)$



Left : scaled first-order and total sensitivity indices (scaled by the variance) of the birth-death model and $t \in [0, 8]$. Right : long-time evolution of the first-order sensitivity indices, and of the mixed interaction term.

- Variance in X is predominantly caused by the birth channel stochasticity for early time $t < 1$
- For $1 \leq t \leq 4$, the variability induced by R_d only continues to grow with the population size (first order reaction), while mixed effects develops
- Eventually, effect of R_b stabilize (zero-order reaction, rate independent of X) while effect of R_d only slowly decays to benefit the mixed term (stochasticity of N_b affects more and more the death process).

Schlögl model

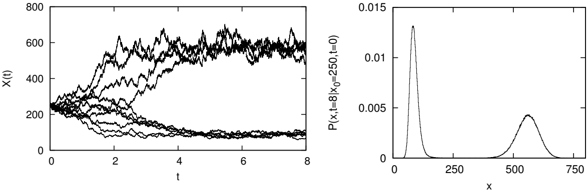
System with $\mathrm{K}_r = 4$ reaction channels :

$$B_1 + 2S \underset{c_2}{\overset{c_1}{\rightleftharpoons}} 3S, \quad B_2 \underset{c_4}{\overset{c_3}{\rightleftharpoons}} S,$$

with B_1 and B_2 in large excess and constant population over time, $X_{B_1} = X_{B_2}/2 = 10^5$ and a single evolving species S with $\mathrm{M}_s = 1$. The propensity functions are given by

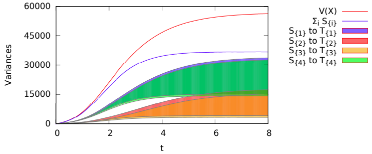
$$a_1(x) = \frac{c_1}{2} X_{B_1} x(x-1), \quad a_2(x) = \frac{c_2}{6} x(x-1)(x-2), \quad a_3(x) = c_3 X_{B_2}, \quad a_4(x) = c_4 x.$$

We set $c_1 = 3 \times 10^{-7}$, $c_2 = 10^{-4}$, $c_3 = 10^{-3}$, $c_4 = 3.5$ and deterministic initial condition $X(t=0) = 250$. Results in a bi-modal dynamic

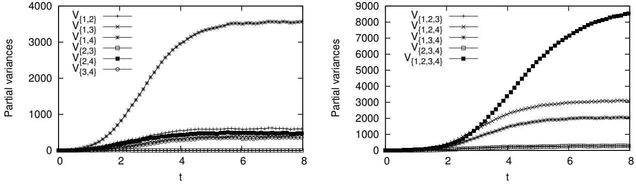


Left : selected trajectories of $X(t)$ showing the bifurcation in the stochastic dynamics. Right : histogram of $X(t=8)$.

Schlögl model - Variance decomposition of $g(X) = X(t)$



First and total order partial variances.



Higher order partial variances.

Reaction channels R_1 and R_4 are the dominant sources of variance

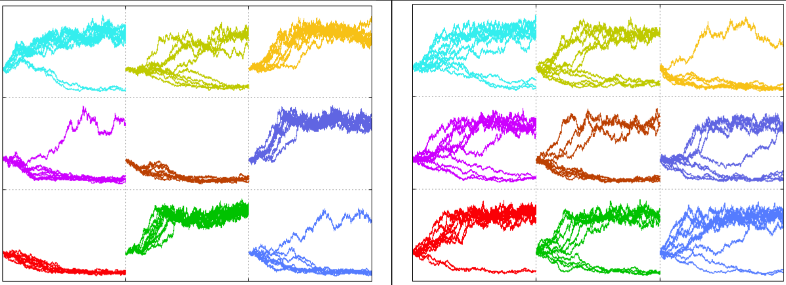
Dynamic essentially additive up to $t \sim 2$

Schlögl model - Variance decomposition of $g(X) = X(t)$

Analysis of the partial variance revealed that R_1 and R_4 are the main sources of stochasticity.

It suggests a dominant role in selecting the bifurcation branch, as illustrated below

(a) Conditioned on N_1 and N_4	(b) Conditioned on N_2 and N_3



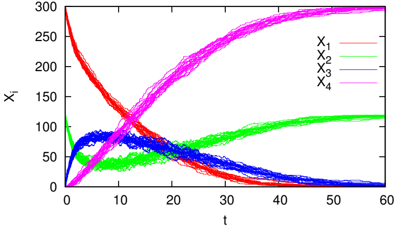
Trajectories of $X(t)$ conditioned on (a) $N_1(\omega) = n_1$ and $N_4(\omega) = n_4$, and (b) $N_2(\omega) = n_2$ and $N_3(\omega) = n_3$. Each sub-plot shows 10 conditionally random trajectories for fixed realizations n_1 and n_4 in (a), and n_2 and n_3 in (b).

Michaelis-Menten system

$\mathrm{M}_s = 4$ species and $\mathrm{K}_r = 3$ reaction channels :

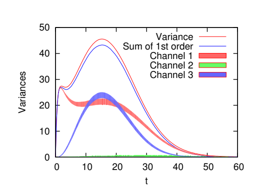
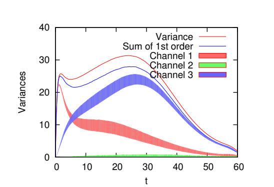
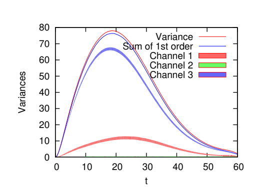
$$S_1 + S_2 \underset{c_2}{\overset{c_1}{\rightleftharpoons}} S_3, \quad S_3 \overset{c_3}{\rightarrow} S_4 + S_2$$

with $a_1(\boldsymbol{x}) = c_1 x_1 x_2$, $a_2(\boldsymbol{x}) = c_2 x_3$, and $c_3(\boldsymbol{x}) = c_3 x_3$,.
We set $c_1 = 0.0017$, $c_2 = 10^{-3}$ and $c_3 = 0.125$, and initial conditions $X_1(t=0) = 300$, $X_2(t=0) = 120$ and $X_3(t=0) = X_4(t=0) = 0$



Michaelis-Menten system - Variance decomposition of $g(\boldsymbol{X}) = X_i(t)$

Note : $X_2 + X_3 =$ const, the sensitivity indices for S_2 and S_3 are equal

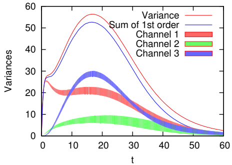
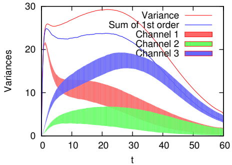
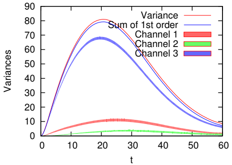
(a) S_1	(b) S_2	(c) S_4
		

Michaelis-Menten model : First-order and total sensitivity indices $\mathtt{S}_{\{j\}}$ and $\mathtt{T}_{\{j\}}$ for $j = 1, \ldots, 4$. Plots are a generated for (a) X_1, (b) X_2 and (c) X_4

- Relative importance of R_1 and R_3 changes in time for S_1 and S_2
- Stochastic dynamic of S_4 is essentially additive and dominated by R_3
- Channel R_2 induces nearly no variance in $X(t)$: **here** the dissociation reaction R_2 can be simply disregarded without affecting significantly the dynamics.

Michaelis-Menten system - Variance decomposition of $g(\boldsymbol{X}) = X_i(t)$

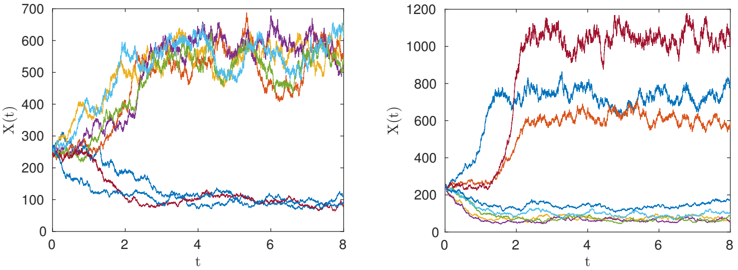
On the contrary, increasing c_2 by an order of magnitude, the effect of R_2 on the variances becomes apparent :

(a) S_1	(b) S_2	(c) S_4
		

Schlögl model - Effect of parameters $g(X) = X(t)$

Consider parametric uncertainty on the propensity function :

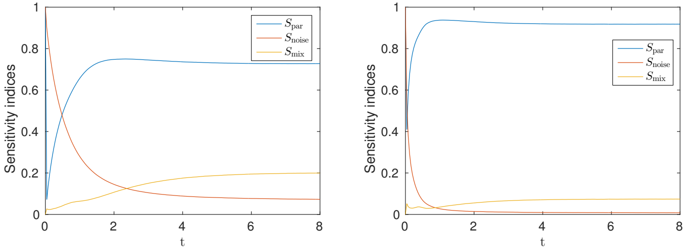
$$a_k(X; Q).$$



Trajectories of $X(t)$ for different Poisson processes and fixed propensity functions (left) or different propensity functions and fixed Poisson processes (right).

Schlögl model - Effect of parameters $g(X) = X(t)$

Consider parametric uncertainty on the propensity function (ind. all with same CV)



(a) $CV = 0.05$ (b) $CV = 0.15$

Sensitivity indices associated to the propensity function parameters S_{par}, inherent stochastic dynamic S_{noise} and their interaction S_{mix}.

Conclusions and Future Work

We have proposed

- A variance decomposition for parametric SA in stochastic systems
- PC expansion when parametric dependence is pathwise smooth
- Development of methods and algorithms to enable variance decomposition in stochastic simulators
- Identify the channels dynamics with their associated standard Poisson processes
- Assessment of the relative importance of different reaction channels

Current works

- Application to complex non-smooth functional $g(\boldsymbol{X})$: exit-time, path integrals, ...
- Account for parametric uncertainty in the definition of the propensity functions
- Improve stochastic simulators for computational complexity reduction, *e.g.* Tau-Leaping method and variance reduction methods.

Acknowledgement : This work was supported by the US Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research, under Award Number DE-SC0008789. Support from the Research Center on Uncertainty Quantification of the King Abdullah University of Science and Technology is also acknowledged.

References

Thank you for your attention

- O.P. Le Maître, O. Knio and A. Moraes, Variance decomposition in stochastic simulators, *J. Chemical Physics*, **142**, pp. 244115, (2015).
- M. Navarro Jimenez, O.P. Le Maître and O.M. Knio, Global sensitivity analysis in stochastic simulators of uncertain reaction networks, *J. Chem. Phys.*, **145**, pp. 244106, (2016).
- O. Le Maître and O. Knio, PC analysis of stochastic differential equations driven by Wiener noise, *Reliability Engineering and System Safety*, **135**, pp. 107-124 (2015).
- M. Navarro Jimenez, O.P. Le Maître and O.M. Knio, Non-intrusive Polynomial Chaos expansions for sensitivity analysis in stochastic differential Equations, *SIAM/ASA J. Uncertainty Quantification*, **5** :1, pp. 378-402, (2017).

Acknowledgement : This work was supported by the US Department of Energy (DOE), Office of Science, Office of Advanced Scientific Computing Research, under Award Number DE-SC0008789. Support from the Research Center on Uncertainty Quantification of the King Abdullah University of Science and Technology is also acknowledged.

Algorithm

ALGORITHM 3. Computation of the first and total-order sensitivity indices $\mathtt{S}_{\{j\}}$ and $\mathtt{T}_{\{j\}}$ of $g(\boldsymbol{X}(T))$.

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

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

$\mu \leftarrow 0$, $\sigma^2 \leftarrow 0$ ▷ Init. Mean and Variance
for $j = 1$ to K_r **do**
  $\mathtt{S}(j) \leftarrow 0$, $\mathtt{T}(j) \leftarrow 0$ ▷ Init. first and total-order SIs
end for
for $m = 1$ to $m = M$ **do**
  Draw two independent set of seeds $\boldsymbol{s}^I$ and $\boldsymbol{s}^{II}$
  $\boldsymbol{X} \leftarrow \mathtt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG}_1(s_1^I), \ldots, \mathtt{RG}_{\mathrm{K}_r}(s_{\mathrm{K}_r}^I))$
  $\mu \leftarrow \mu + g(\boldsymbol{X})$, $\sigma^2 \leftarrow \sigma^2 + g(\boldsymbol{X})^2$ ▷ Acc. mean and variance
  for $j = 1$ to K_r **do**
    $\boldsymbol{X}_{\mathtt{S}} \leftarrow \mathtt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG}_1(s_1^{II}), \ldots,$
    $\ldots, \mathtt{RG}_j(s_j^I), \ldots, \mathtt{RG}_{\mathrm{K}_r}(s_{\mathrm{K}_r}^{II}))$
    $\boldsymbol{X}_{\mathtt{T}} \leftarrow \mathtt{NRA}(\boldsymbol{X}_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG}_1(s_1^I), \ldots,$
    $\ldots, \mathtt{RG}_j(s_j^{II}), \ldots, \mathtt{RG}_{\mathrm{K}_r}(s_{\mathrm{K}_r}^I))$
    $\mathtt{S}(j) \leftarrow \mathtt{S}(j) + g(\boldsymbol{X}) \times g(\boldsymbol{X}_{\mathtt{S}})$ ▷ Acc. 1-st order
    $\mathtt{T}(j) \leftarrow \mathtt{T}(j) + g(\boldsymbol{X}) \times g(\boldsymbol{X}_{\mathtt{T}})$ ▷ Acc. total order
  end for ▷ Next channel
end for ▷ Next sample
$\mu \leftarrow \mu/M$, $\sigma^2 \leftarrow \sigma^2/(M-1) - \mu^2$
for $j = 1$ to K_r **do**
  $\mathtt{S}(j) \leftarrow \frac{\mathtt{S}(j)}{(M-1)\sigma^2} - \frac{\mu^2}{\sigma^2}$ ▷ Estim. 1-st order
  $\mathtt{T}(j) \leftarrow 1 - \frac{\mathtt{T}(j)}{(M-1)\sigma^2} + \frac{\mu^2}{\sigma^2}$ ▷ Estim. total order
end for
Return $\mathtt{S}(j)$ and $\mathtt{T}(j)$, $j = 1, \ldots, \mathrm{K}_r$ ▷ First and total-order sensitivity indices $\mathtt{S}_{\{j\}}$ and $\mathtt{T}_{\{j\}}$ of $g(\boldsymbol{X}(T))$

- Procedure NRA implement the Next Reaction Algorithm
- Poisson processes defined by two independent sets of seeds and RNG
- Obvious parallelization

ALGORITHM 2. Next Reaction Algorithm.

Procedure $\mathtt{NRA}(X_0, T, \{\boldsymbol{\nu}_j\}, \{a_j\}, \mathtt{RG}_1, \ldots, \mathtt{RG}_{\mathrm{K}_r})$

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

for $j = 1, \ldots, \mathrm{K}_r$ **do**
  Draw r_j from $\mathtt{RG}_j$
  $\tau_j \leftarrow 0$, $\tau_j^+ \leftarrow -\log r_j$ ▷ set next reaction times
end for
$t \leftarrow 0, \boldsymbol{X} \leftarrow \boldsymbol{X}_0$
loop
  for $j = 1, \ldots, \mathrm{K}_r$ **do**
    Evaluate $a_j(\boldsymbol{X})$ and $dt_j = \frac{\tau_j^+ - \tau_j}{a_j}$
  end for
  Set $l = \arg\min_j dt_j$ ▷ pick next reaction
  if $t + dt_l > T$ **then**
    break ▷ Final time reached
  else
    $t \leftarrow t + dt_l$ ▷ update time
    $\boldsymbol{X} \leftarrow \boldsymbol{X} + \boldsymbol{\nu}_l$ ▷ update the state vector
    for $j = 1, \ldots, \mathrm{K}_r$ **do**
      $\tau_j \leftarrow \tau_j + a_j\, dt_l$ ▷ update unscaled times
    end for
    Get r_l from $\mathtt{RG}_l$
    $\tau_l^+ \leftarrow \tau_l^+ - \log r_l$ ▷ next reaction time
  end if
end loop
Return $\boldsymbol{X}$ ▷ State $\boldsymbol{X}(T)$