Coupling of Stochastic Processes

This section briefly revisits the coupling method proposed by Anderson [1] for the estimation of gradients in well-mixed reaction network systems. We recall that the sensitivity index of an observable of the stochastic process, $X$, with respect to the parameter $\theta_k$ is defined as

$$ S_{k,\ell} = \frac{\partial}{\partial \theta_k} \mathbb{E}_{Q_{[0,\tau]}}[F_\ell((X_t)_{t=0}^T)], \quad k = 1, \ldots, K \quad \& \quad \ell = 1, \ldots, L. $$

The partial derivative can be approximated by a second-order finite-difference scheme as

$$ S_{k,\ell} \approx \tilde{S}_{k,\ell} = \frac{1}{2\epsilon_0} \left( \mathbb{E}_{x_0}[F_\ell(X_{t}^{\theta_k+\epsilon_0})] - \mathbb{E}_{x_0}[F_\ell(X_{t}^{\theta_k-\epsilon_0})] \right), $$

where $\epsilon_0 \in \mathbb{R}$ and $\epsilon_0 \ll 1$. The expected value of the observable function $u(t, x_0) := \mathbb{E}_{x_0} F(X_t)$ satisfies the initial value problem

$$ \frac{d}{dt} u(t, x_0) = \mathcal{L}^\theta u(t, x_0), \quad u(0, x_0) = F(x_0) $$

where $\mathcal{L}^\theta$ is the generator operator defined by

$$ \mathcal{L}^\theta F(x) := \sum_{j=1}^J a_j^\theta(x) [F(x + \nu_j) - F(x)], $$

and $a_j^\theta(\cdot)$, $j = 1, \ldots, J$, are the propensity functions.

In order to reduce the variance of the statistical estimator for $\tilde{S}_{k,\ell}$, we have to correlate the process $X_{t}^{\theta_k+\epsilon_0}$ with the process $X_{t}^{\theta_k-\epsilon_0}$. It has been shown in [1, 2] that coupling indeed reduces the variance of the estimator, even in high-dimensional systems. In the coupling method, a new process is considered which is the pair of stochastic processes $Z_t = (X_{t}^{\theta_k+\epsilon_0}, X_{t}^{\theta_k-\epsilon_0})$. We define the vector $e_k$ which has zeros in all positions and 1 in the $k$-th position and set $\epsilon = \epsilon_0 e_k$. Then $Z_t$ is a Markov process with infinitesimal generator,

$$ \mathcal{L} G(z) = \sum_{j=1}^J \min\{a_j^{\theta+\epsilon}(x), a_j^{\theta-\epsilon}(y)\} \left[ G(x + \nu_j, y + \nu_j) - G(x, y) \right] $$

$$ + \sum_{j=1}^J \left[ a_j^{\theta+\epsilon}(x) - \min\{a_j^{\theta+\epsilon}(x), a_j^{\theta-\epsilon}(y)\} \right] \left[ G(x + \nu_j, y) - G(x, y) \right] $$

$$ + \sum_{j=1}^J \left[ a_j^{\theta-\epsilon}(y) - \min\{a_j^{\theta+\epsilon}(x), a_j^{\theta-\epsilon}(y)\} \right] \left[ G(x, y + \nu_j) - G(x, y) \right], $$

where $z = (x, y)$ is the state of the $Z_t$ process and $G$ an observable function on $Z_t$. The transitions of the coupled generator are shown in Figure 1. More specifically, the intuition behind this coupling is that both processes move to the same updating direction with the minimum rate $c_j(x, y) = \min\{a_j^{\theta+\epsilon}(x), a_j^{\theta-\epsilon}(y)\}$.
for $j = 1, \ldots, J$, leading to trajectories that most of the time track each other. On the other hand, every process has to do some independent jumps in order to keep the correct dynamics. Indeed, it can be proved, [3], that the coupled generator (3) satisfies

$$
\mathbb{E}_{z_0}\left[G(X_{t_{j}}^{\theta+\epsilon} , X_{t_{j}}^{\theta-\epsilon})\right] = \mathbb{E}_{x_0}\left[F(X_{t_{j}}^{\theta+\epsilon})\right] \text{ when } G(x, y) = F(x) \text{ and }
\mathbb{E}_{z_0}\left[G(X_{t_{j}}^{\theta+\epsilon} , X_{t_{j}}^{\theta-\epsilon})\right] = \mathbb{E}_{y_0}\left[F(X_{t_{j}}^{\theta-\epsilon})\right] \text{ when } G(x, y) = F(y),
$$

where $\mathbb{E}_{z_0}\left[G(X_{t_{j}}^{\theta+\epsilon} , X_{t_{j}}^{\theta-\epsilon})\right]$ denotes the mean value of the observable function $G$ of the state variables $(X_{t_{j}}^{\theta+\epsilon} , X_{t_{j}}^{\theta-\epsilon})$ with respect to the law imposed by the initial data $z_0 = (x_0, y_0)$. These relations imply that all averages of the two coupled processes, and thus all observables $F$, coincide with the averaged observables generated by the uncoupled generators $\mathcal{L}^{\theta+\epsilon}, \mathcal{L}^{\theta-\epsilon}$.

References

