Adaptive importance sampling for reliability assessment of an industrial system modeled by a Piecewise Deterministic Markov Process

Guillaume Chennetier [1,2] · Anne Dutfoy[1] · Hassane Chraibi [1] · Josselin Garnier[2]

[1] EDF Lab Paris-Saclay, Boulevard Gaspard Monge, 91120 Palaiseau, France.
[2] CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France.

Content

We wish to estimate the probability of failure of hybrid dynamic industrial systems represented by piecewise deterministic Markov processes (PDMP). Crude Monte Carlo methods (CMC) are not suitable for this purpose because the typical failure probabilities are very low. We propose instead an adaptive importance sampling method with cross entropy procedure that achieves tremendous variance reduction.

The success of our method relies on the ability to approximate the committor function of the PDMP. Our main contribution is to use the reliability concept of minimal path sets of the system to build a good approximation of the committor function.

Bibliography

[1] M. H. A. Davis (1984), Piecewise-Deterministic Markov Processes : A general class of non-diffusion stochastic models. *J. R. Stat. Soc. B 46 :353–388.*

[2] H. Chraibi, A. Dutfoy, T. Galtier et J. Garnier (2019), On the optimal importance process for piecewise deterministic Markov process. *ESAIM : PS. 23 :893–921.*

[3] R. Rubinstein (1999), The Cross-Entropy method for combinatorial and continuous optimization. *Methodology and Computing in Applied Probability 1 :127–190.*

[4] B. Delyon et F. Portier (2018), Asymptotic optimality of adaptive importance sampling. *NIPS 2018.*

1. PDMP – Piecewise Deterministic Markov Processes

Hybrid process : $Z_t = (X_t, M_t) \in E$
- position X_t is continuous,
- mode M_t is discrete.

Deterministic process between two jumps :
- mode remains constant $M_{s+t} = M_s = m$,
- position follows the **flow** Φ,

$$(X_{s+t}, m) = \Phi_{X_s,m}(t).$$

Jumps at the boundaries of E :

$$t_z^\partial = \inf\{t > 0 : \Phi_z(t) \in \partial E\}.$$

Jumps at random times according to **jump intensity** λ. Let T_z be the waiting time from z,

$$\mathbb{P}(T_z > t) = \mathbb{1}_{t<t_z^\partial} e^{-\int_0^t \lambda(\Phi_z(u))du}.$$

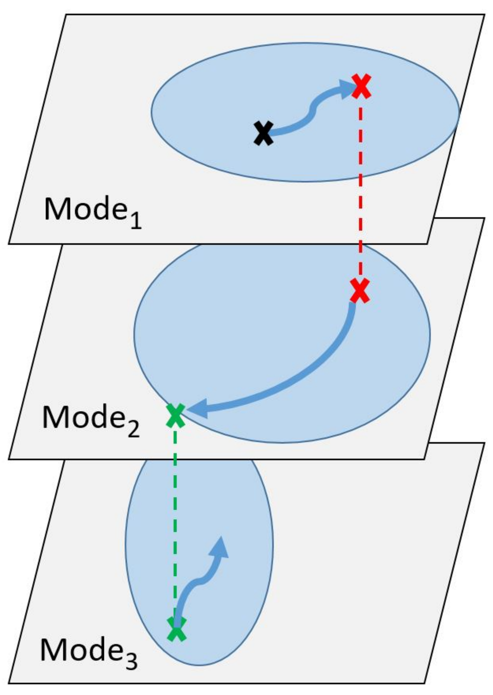
The state of the process after a jump is randomly selected by **jump kernel** K. Jumping from z^- :

$$\text{for } B \subset E, \quad \mathbb{P}_{Z^-=z^-}(Z \in B) = \int_B K(z^-, dz).$$

PDMP trajectories of duration $t_{\max} > 0$.

$$\mathcal{Z} := (Z_t)_{t\in[0,t_{\max}]} \sim \pi_{\lambda,K}$$

where $\pi_{\lambda,K}$ is the distribution of the PDMP characterized by jump intensity λ and kernel K.



What are we trying to do ?

Let $\pi_0 \equiv \pi_{\lambda_0,K_0}$ be the distribution of the PDMP $\mathcal{Z}$ and $\mathcal{D}$ a subset of the possible trajectories on E.

Goal : estimating $P := \mathbb{P}_{\pi_0}(\mathcal{Z} \in \mathcal{D})$ when P is too small to be estimated by a crude Monte-Carlo method.

Application case : the PDMP models an industrial system. $\mathcal{D}$ is the set of trajectories that encounter system failure and the probability of failure P is about 10^{-5}.

2. Application case – Spent fuel pool system

Spent nuclear fuel is stored in a cold water pool. If the system does not cool the pool, the nuclear fuel evaporates the water then damages the structure and contaminates the outside.

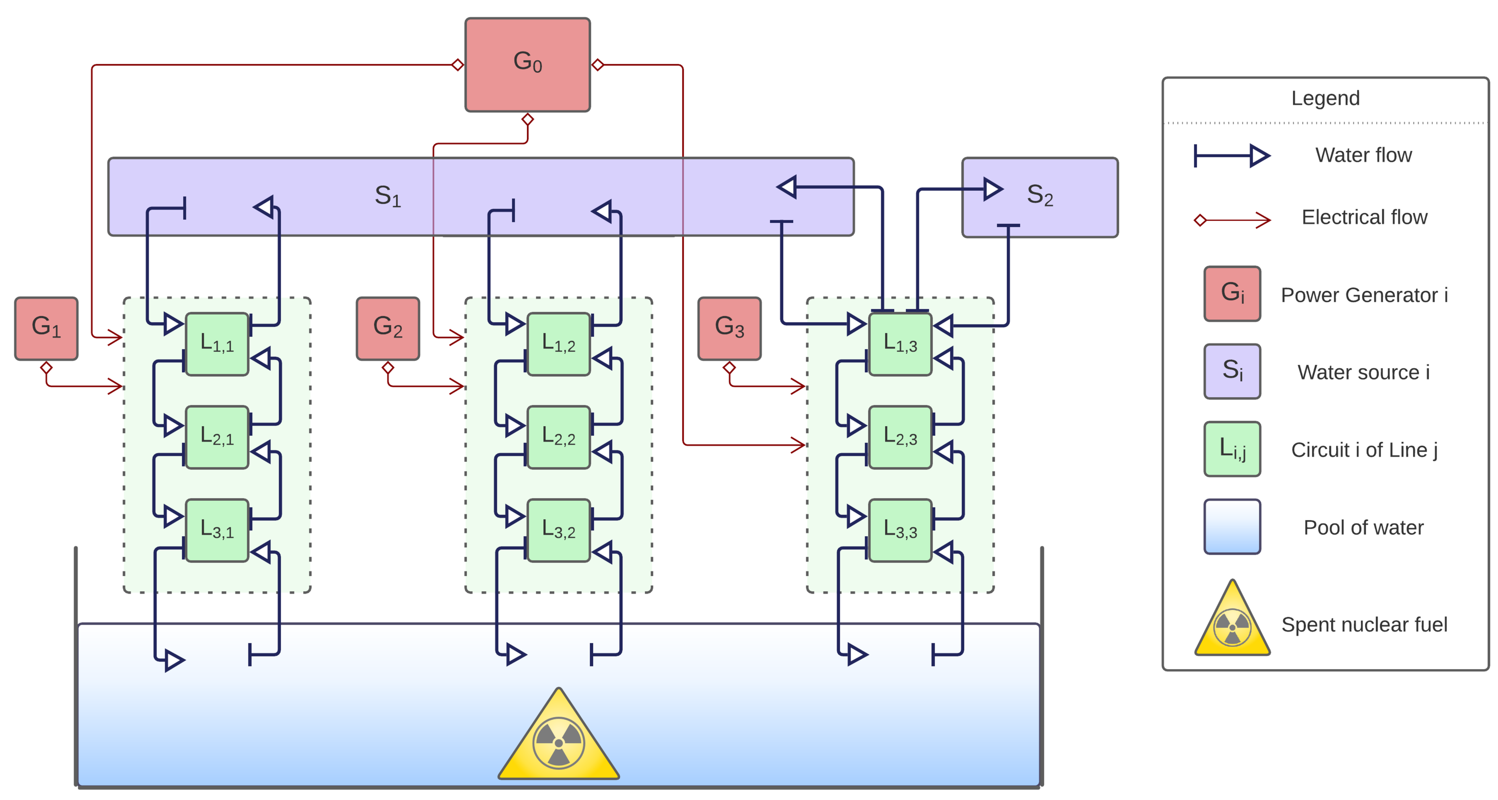


Figure 1 – **Representation of the spent fuel pool.** The temperature of an outside water source S_1 is transfered to the pool through three sealed circuits connected by heat exchangers $L_{1,1}$, $L_{2,1}$ and $L_{3,1}$ forming a line L_1. The system has a general power supply G_0. In the event of a problem with one of these components, the system is equipped with two other lines L_2 and L_3 identical to L_1, an emergency diesel generator for each line G_1, G_2 and G_3, and a second outside water source S_2 accessible only to the third line L_3.

The system fails when the water level drops below a critical level. This is only possible when specific combinations of components are broken.

Minimal path sets (MPS) of an industrial system

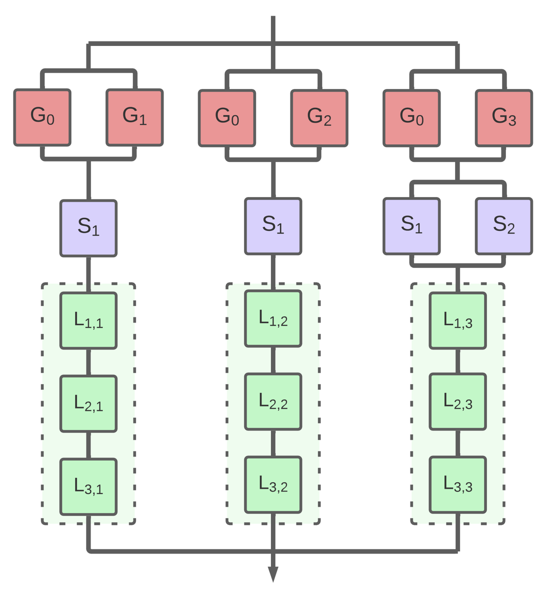


Figure 2 – **8 MPS of the SFP system.**
$(G_0, S_1, \mathbf{L_1})$, $(G_1, S_1, \mathbf{L_1})$, $(G_0, S_1, \mathbf{L_2})$, $(G_2, S_1, \mathbf{L_2})$, $(G_0, S_1, \mathbf{L_3})$, $(G_3, S_1, \mathbf{L_3})$, $(G_0, S_2, \mathbf{L_3})$, $(G_3, S_2, \mathbf{L_3})$.

The path sets of a system are the sets of components such that :
- keeping all components of any path set intact prevents system failure.
- keeping one component broken in each path set ensures system failure.

A **M**inimal **P**ath **S**et is a path set that does not contain any other path set.

- d_{MPS} is the number of MPS of the system,
- β_z is the number of MPS with at least one broken component in state $z \in E$.

3. Importance sampling for PDMP

Importance sampling for rare events : we generate trajectories from an auxiliary distribution $\widetilde{\pi}$ which produces more trajectories in $\mathcal{D}$ than π_0 then we fix the bias with the proper likelihood ratio.

$$\widehat{P}_{\text{IS}} := \frac{1}{N}\sum_{k=1}^{N} \mathbb{1}_{\mathcal{Z}_k\in\mathcal{D}} \frac{\pi_0(\mathcal{Z}_k)}{\widetilde{\pi}(\mathcal{Z}_k)} \xrightarrow[N\to\infty]{\text{a.s.}} \mathbb{E}_{\widetilde{\pi}}\left[\mathbb{1}_{\mathcal{Z}\in\mathcal{D}}\frac{\pi_0(\mathcal{Z})}{\widetilde{\pi}(\mathcal{Z})}\right] = \mathbb{E}_{\pi_0}[\mathbb{1}_{\mathcal{Z}\in\mathcal{D}}] = P.$$

Variance reduction : strongly depends on the choice of $\widetilde{\pi}$. Poor choices lead to a very high variance estimator but optimal choice $\pi_{\text{opt}}(\mathcal{Z}) := \frac{1}{P}\mathbb{1}_{\mathcal{Z}\in\mathcal{D}}\,\pi_0(\mathcal{Z})$ leads to a zero variance estimator.

Optimal importance distribution for PDMP

Distribution π_{opt} : same state space E and same flow Φ as for π_0 but optimal jump intensity λ_{opt} and optimal jump kernel K_{opt} depend on U_{opt} the committor function of the process.

$$\lambda_{\text{opt}}(\Phi_{z^-}(t)\,;\,s) = \lambda_0(\Phi_{z^-}(t)) \times \frac{U^-_{\text{opt}}(\Phi_{z^-}(t), s+t)}{U_{\text{opt}}(\Phi_{z^-}(t), s+t)}, \tag{1}$$

$$K_{\text{opt}}(z^-, z\,;\,s) = K_0(z^-, z) \times \frac{U_{\text{opt}}(z, s)}{U^-_{\text{opt}}(z^-, s)}, \tag{2}$$

$$\text{with} \quad U_{\text{opt}}(z, s) = \mathbb{P}_{\pi_0}(\mathcal{Z} \in \mathcal{D} \,|\, Z_s = z) \quad \text{and} \quad U^-_{\text{opt}}(z^-, s) = \int_E U_{\text{opt}}(z, s)\, K_0(z^-, dz).$$

What does that mean ?

Committor function : probability of reaching the rare event knowing the current state of the process.

Equation (1) : if the probability of reaching $\mathcal{D}$ is k times higher by jumping from a specific state than by not jumping, then the jump intensity on that state must be multiplied by k.

Equation (2) : if the probability of reaching $\mathcal{D}$ is k times higher by jumping to a specific state than by jumping randomly according to K, then the probability of jumping to that state must be multiplied by k.

If you know the committor function, you can build the optimal IS estimator !

4. Approximating the committor function

Idea : build a near-optimal importance distribution π_α by using an approximation U_α instead of the unknown function committor U_{opt} in equations (1) and (2).

$$U_\alpha(z) = e^{\left(\sum_{i=1}^{\beta_z}\alpha_i\right)^2}, \quad \alpha \in \mathbb{A} \subset \mathbb{R}^{d_{\text{MPS}}}. \tag{3}$$

The closer β_z is to d_{MPS}, the closer the process is to $\mathcal{D}$. U_α is therefore an increasing function in β_z.

Cross entropy procedure

Sequential algorithm : we jointly tune α and estimate P.

$$\underset{\alpha\in\mathbb{A}}{\arg\min}\, \mathcal{D}_{\text{KL}}(\pi_{\text{opt}} \| \pi_\alpha) = \underset{\alpha\in\mathbb{A}}{\arg\min}\, \{-\mathbb{E}_{\pi_0}[\mathbb{1}_{\mathcal{Z}\in\mathcal{D}} \log(\pi_\alpha(\mathcal{Z}))]\}$$

At iteration $q = 1, \ldots, Q$, we minimize an estimate of the KL divergence using all the trajectories drawn :
- **Simulation phase.** Generate a new sample of n_q trajectories $\mathcal{Z}_1^{(q)}, \ldots, \mathcal{Z}_{n_q}^{(q)} \overset{\text{i.i.d.}}{\sim} \pi_{\alpha^{(q)}}$.
- **Optimization phase.** Update the parameter α with the q last samples $(\mathcal{Z}_k^{(1)})_{k=1}^{n_1}, \ldots, (\mathcal{Z}_k^{(q)})_{k=1}^{n_q}$.

$$\alpha^{(q+1)} = \underset{\alpha\in\mathbb{A}}{\arg\min}\left\{-\sum_{r=1}^{q}\sum_{k=1}^{n_r} \mathbb{1}_{\mathcal{Z}_k^{(r)}\in\mathcal{D}} \frac{\pi_0(\mathcal{Z}_k^{(r)})}{\pi_{\alpha^{(r)}}(\mathcal{Z}_k^{(r)})} \log\left[\pi_\alpha(\mathcal{Z}_k^{(r)})\right]\right\}.$$

Estimation phase at the final iteration Q (with $N_Q = \sum_{q=1}^{Q} n_q$), we reuse all past samples to estimate P :

$$\widehat{P}_{N_Q} = \frac{1}{N_Q}\sum_{q=1}^{Q}\sum_{k=1}^{n_q} \mathbb{1}_{\mathcal{Z}_k^{(q)}\in\mathcal{D}} \frac{\pi_0(\mathcal{Z}_k^{(q)})}{\pi_{\alpha^{(q)}}(\mathcal{Z}_k^{(q)})}.$$

5. Numerical results

Method	Sample size N	Estimated probability $\widehat{P}$	Coefficient of variation	95% confidence interval
CMC	10^5	2×10^{-5}	223.60	$[0\,;\, 4.77 \times 10^{-5}]$
	10^6	1.3×10^{-5}	277.35	$[5.93 \times 10^{-6}\,;\, 2.01 \times 10^{-5}]$
	10^7	1.77×10^{-5}	237.68	$[1.51 \times 10^{-5}\,;\, 2.03 \times 10^{-5}]$
IS	10^2	2.18×10^{-5}	4.69	$[1.76 \times 10^{-5}\,;\, 4.18 \times 10^{-5}]$
	10^3	2.19×10^{-5}	3.01	$[1.78 \times 10^{-5}\,;\, 2.60 \times 10^{-5}]$
	10^4	1.99×10^{-5}	1.01	$[1.96 \times 10^{-5}\,;\, 2.03 \times 10^{-5}]$

Table 1 – **Comparison between crude Monte-Carlo (CMC) and our adaptive importance sampling method (IS).**

Our method reduces the variance of the estimation by a factor greater than 10^4 compared to a CMC method.