State-dependent Importance Sampling for Estimating Expectations of Functionals of Sums of Independent Random Variabless

Eya Ben Amar,
Dr. Nadhir Ben Rached, Prof. Raúl Tempone, and Dr. Abdul-Lateef Haji-Ali.

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

14 International Conference on Monte Carlo Methods and Applications, Sorbonne University, Paris.

June 28, 2023

- Rare events are events that happen with very low probabilities but their occurrences can lead to critical consequences.
- In the context of communication systems, a rare event can be the event that the wireless system is in an outage and hence fail to operate properly.
- For sophisticated networks such as ultra-reliable 5G or 6G systems, one can encounter the problem of estimating failure probabilities of the order of 10^{-9} (Ben Rached et al. 2020).

Objective

Let $\mathrm{X} = (X_1, X_2, \cdots, X_N)^t$ a random vector composed of independent positive components with joint PDF $f(\boldsymbol{x}) = \prod_{n=1}^{N} f_{X_n}(x_n)$. Let $S_n = \sum_{i=1}^{n} X_i$ and $g : \mathbb{R}_+ \to \mathbb{R}$ a given function. The aim is to develop a generic state-dependent Importance sampling (IS) algorithm via a novel Stochastic Optimal Control (SOC) formulation to estimate rare event problems that could be written in the following form

$$\alpha = \mathbb{E}\left[g\left(S_N\right)\right]. \tag{1}$$

Right Tail: $g(x) = 1_{(x \geq \gamma)}$, $\alpha = \mathbb{P}(S_N \geq \gamma)$, where $S_N = \sum_{i=1}^{N} X_i$

Example: Ruin probability of an insurance company. S_N represents the total sum of claims, and γ is the initial reserve (Asmussen and Glynn 2007).

Left Tail: $g(x) = 1_{(x \leq \gamma)}$, $\alpha = \mathbb{P}(S_N \leq \gamma)$,

Example: The outage probability (OP): probability that the signal-to-noise ratio (SNR) at EGC and MRC diversity receivers falls below a given threshold γ.

- $SNR = \frac{E_s}{N_0\sqrt{N^{1-p+q}}} \left(\sum_{i=1}^{N} R_i^p\right)^q$, where N is the number of diversity branches, $\frac{E_s}{N_0}$ is the SNR per symbol, R_i, $i = 1, 2, ..., N$, is the fading channel envelope, $(p, q) = (1, 2)$ for EGC or $(p, q) = (2, 1)$ for MRC (Ben Rached et al. 2016).
- $OP = \mathbb{P}\left(SNR \leq \gamma_{th}\right) = \mathbb{P}\left(\sum_{i=1}^{N} X_i \leq \gamma\right)$, with $X_i = R_i^2$ and $\gamma = \gamma_{th} N_0/E_s$ (for MRC) or $X_i = R_i$ and $\gamma = \sqrt{\gamma_{th} N_0 N/E_s}$ (for EGC), $i = 1, 2, \cdots, N$.

CDF of the ratio of independent RVs:

- For SISO systems and in the presence of interferences and noise, the OP is expressed as (Ben Rached et al. 2017)

$$P_{out} = \mathbb{P}(\mathrm{SINR} \leq \gamma) = \mathbb{P}\left(\frac{X_0}{\sum_{n=1}^{N} X_n + \eta} \leq \gamma\right),$$

where X_0 is the useful signal power, $X_1, \cdots, X_N$ are the received powers of the N interfering signals and η is the variance of the noise. We assume that $X_0, \cdots, X_N$ are independent.

- By conditioning on $X_1, X_2, \cdots, X_N$ and using the law of total expectation, we write

$$\mathbb{E}\left[F_{X_0}\left(\gamma\left(\sum_{n=1}^{N} X_n + \eta\right)\right)\right], \quad (2)$$

where $F_{X_0}(\cdot)$ is the CDF of the RV X_0.

- This corresponds to the form in (1) with $g(x) = F_{X_0}(\gamma(x+\eta))$.

Inefficiency of naive Monte Carlo

- The naive MC estimator of our quantity of interest $\alpha = \mathbb{E}\left[g(S_N)\right]$ is given by: $\hat{\alpha}_{mc} = \frac{1}{M}\sum_{k=1}^{M} g(S_N^{(k)})$, where M is the number of simulation runs and $\{S_N^{(k)}\}_{k=1}^{M}$ represent independent realizations of the RV S_N.
- Statistical Error:

$$|\alpha - \hat{\alpha}_{MC}| \approx C\sqrt{\frac{var[g(S_N)]}{M}},$$

- Relative Error: (for $g(x) = 1_{(x \in A)}$)

$$\epsilon_{MC} = \frac{|\alpha - \hat{\alpha}_{MC}|}{\alpha} \approx C\frac{\sqrt{\alpha(1-\alpha)}}{\sqrt{M\alpha}} \approx \frac{C}{\sqrt{M\alpha}}$$

.

- To ensure a relative error equal to TOL, $M = \frac{C^2}{TOL^2\alpha}$ samples are required.
- When α is small (rare event probabilities), MC method is computationally expensive $\rightarrow$ Use appropriate variance reduction techniques, such as IS.

Idea: Introduce a new probability measure $\tilde{f}(\cdot)$, which reduces $\mathrm{Var}\left[g(S_N)\right]$ but keeps $\mathbb{E}\left[g(S_N)\right]$ unchanged (Kroese et al. 2011).

$$\begin{aligned}\alpha &= \int_{\mathbb{R}^N} g(S_N) f(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^N} \underbrace{g(S_N)\frac{f(\mathbf{x})}{\tilde{f}(\mathbf{x})}}_{\tilde{g}(\mathbf{x})} \tilde{f}(\mathbf{x}) d\mathbf{x} \\ &= \mathbb{E}_{\tilde{f}}\left[\tilde{g}\left(X_1, \cdots, X_N\right)\right],\end{aligned} \tag{3}$$

where $\mathbb{E}_{\tilde{f}}[\cdot]$ is the expectation under which the vector $(X_1, X_2, \cdots, X_N)^T$ has the PDF $\tilde{f}(\cdot)$. The ratio $\frac{f(\mathbf{x})}{\tilde{f}(\mathbf{x})}$ is called the **likelihood ratio**.

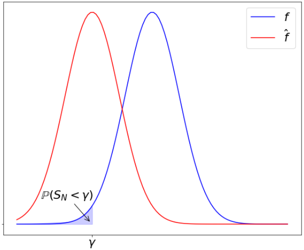


Figure 1: Importance Sampling.

- The IS estimator is $\hat{\alpha}_{IS} = \frac{1}{M}\sum_{k=1}^{M} \tilde{g}\left(X_1^{(k)}, \cdots, X_N^{(k)}\right)$, where $\left\{\left[X_1^{(k)}, \cdots, X_N^{(k)}\right]^T\right\}_{k=1}^{M}$ represent independent realizations of $[X_1, \cdots, X_N]^T$ sampled according to $\tilde{f}(\cdot)$.
- The aim is then to propose a sub-optimal change of measure that leads to a substantial amount of variance reduction.

Literature

- The literature is rich of variance reduction techniques dealing with the right-tail ($g(x) = 1_{(x \geq a)}$) and the left-tail ($g(x) = 1_{(x \leq a)}$) problems.
- In particular, various state-independent IS techniques have been proposed (Rajhaa and Juneja 2021, Juneja 2007, Murthy et al. 2015).
- State-independent change of measure for estimating certain rare events involving sums of heavy-tailed RVs are not efficient (Bassamboo et al. 2007).
- Complex state-dependent IS have been proposed in the literature over the last few years (Blanchet and Lam 2012, Blanchet and Li 2011, Blanchet Liu 2006, Dupuis and Wang. 2004, Dupuis et al. 2007).
- By state-dependent IS, we mean that the IS parameter is dynamically chosen as a function of the current step and current state of the dynamical system.
- In the i.i.d case and for distributions with finite MGFs, an approach has been developed based on connecting IS with SOC (Dupuis and Wang 2004).

- Embed the static problem with the evolution of a Markov chain with the following dynamics: $S_0 = 0$ and

$$S_{n+1} = S_n + X_{n+1}, \quad n = 0, 1, \cdots, N-1 \tag{4}$$

- Perform a change of measure such that, given S_n, X_{n+1} is distributed according to $\tilde{f}_{X_{n+1}}(\cdot; \mu_{n+1}(S_n))$, where μ_{n+1} is a function of S_n.
- The new joint PDF can be written as

$$\tilde{f}(\mathrm{x}) = \prod_{n=1}^{N} \tilde{f}_{X_n}\left(x_n; \mu_n\left(s_{n-1}\right)\right), \tag{5}$$

where $s_{n-1} = \sum_{i=1}^{n-1} x_i$.

Objective

Find the optimal controls $\mu_n : \mathbb{R}_+ \to A \subset \mathbb{R}$, $n = 1, 2, \cdots, N$, that minimizes the second moment of the IS estimator.

- The cost function for $\mu_{n+1}, \cdots, \mu_N \in \mathcal{A}^{N-n},\ n = 0, \cdots, N-1$, is

$$C_{n,s}(\mu_{n+1}, \cdots, \mu_N) = \mathbb{E}_{\tilde{f}}\left[\left(g\left(S_N\right)\right)^2 \times \prod_{i=n+1}^{N} \left(\frac{f_{X_i}\left(X_i\right)}{\tilde{f}_{X_i}\left(X_i; \mu_i(S_{i-1})\right)} \right)^2 \mid S_n = s \right], \tag{6}$$

where $\mathcal{A}$ is the set of admissible Markov controls.

- Define also the value function as follows

$$u(n, s) = \inf_{\mu_{n+1}, \cdots, \mu_N \in \mathcal{A}^{N-n}} C_{n,s}(\mu_{n+1}, \cdots, \mu_N). \tag{7}$$

Proposition (Ben Amar et al. 2023)

For all $n \in \{0, 1, \cdots, N-1\}$ and $s \geq 0$, we have

$$u(n,s) = \inf_{\mu \in A} \mathbb{E}_{\tilde{f}} \left[\left(\frac{f_{X_{n+1}}(X_{n+1})}{\tilde{f}_{X_{n+1}}(X_{n+1};\mu)} \right)^2 u(n+1, S_{n+1}) \mid S_n = s \right], \qquad (8)$$

and if the minimum is attained, we have

$$\mu_{n+1}(s) = \arg\min_{\mu \in A} \mathbb{E}_{\tilde{f}} \left[\left(\frac{f_{X_{n+1}}(X_{n+1})}{\tilde{f}_{X_{n+1}}(X_{n+1};\mu)} \right)^2 u(n+1, S_{n+1}) \mid S_n = s \right], \qquad (9)$$

with $u(N,x) = (g(x))^2$, $S_{n+1} = s + X_{n+1}$ and X_{n+1} is distributed according to $\tilde{f}_{X_{n+1}}(\cdot;\mu)$.

- Originally developed in (Juneja and Shahabuddin 2002, Ben Rached et al. 2018).
- Define the hazard rate $\lambda_{X_i}(\cdot)$ associated to the RV X_i as

$$\lambda_{X_i}(x) = \frac{f_{X_i}(x)}{1 - F_{X_i}(x)}, \quad x > 0, \tag{10}$$

where $F_{X_i}(x) = \mathbb{P}(X_i \leq x)$ is the CDF of X_i , $i = 1, \cdots, N$.

- Define the hazard function as

$$\Lambda_{X_i}(x) = -\log\left(1 - F_{X_i}(x)\right), \quad x > 0. \tag{11}$$

- From (10) and (11), the PDF of X_i can be expressed as

$$f_{X_i}(x) = \lambda_{X_i}(x) \exp\left(-\Lambda_{X_i}(x)\right), \quad x > 0. \tag{12}$$

- The HRT change of measure is obtained by twisting the hazard rate of each component X_i by a quantity $\mu_i \in A =]-\infty, 1[$

$$\begin{aligned} \tilde{f}_{X_i}(x; \mu_i) &= (1 - \mu_i)\lambda_{X_i}(x) \exp\left(-(1 - \mu_i)\Lambda_{X_i}(x)\right) \\ &= (1 - \mu_i) f_{X_i}(x) \exp\left(\mu_i \Lambda_{X_i}(x)\right), \quad x > 0. \end{aligned} \tag{13}$$

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of Science and Technology

Aim: Compute $u(n, s_k)$ and μ_n for all $n = 0, 1, \cdots, N-1$ and for all s_k, $k = 0, 1, \cdots, K$.

We truncate the space $\mathbb{R}_+$ and work in the interval $[0, S]$, where S is a large number in $\mathbb{R}_+$. Let us consider a mesh in the one dimensional s-space: $0 = s_0 < s_1, \cdots < s_K = S$.

1. For each s_k in the mesh, we solve

$$\begin{aligned} u(N-1, s_k) &= \min_{\mu \in A} \mathbb{E}_{\tilde{f}} \left[\left(\frac{f_{X_N}(X_N)}{\tilde{f}_{X_N}(X_N; \mu)} \right)^2 (g(s_k + X_N))^2 \right] \\ &= \min_{\mu \in A} \int_0^{+\infty} \frac{(f_{X_N}(t))^2}{\tilde{f}_{X_N}(t; \mu)} (g(s_k + t))^2 \, dt, \end{aligned} \tag{14}$$

and

$$\mu_N(s_k) = \arg\min_{\mu \in A} \int_0^{+\infty} \frac{(f_{X_N}(t))^2}{\tilde{f}_{X_N}(t; \mu)} (g(s_k + t))^2 \, dt. \tag{15}$$

2. Having obtained $u\left(N-1, s_k\right)$ for all s_k in the grid, the next step corresponds to

$$u\left(N-2, s_k\right)=\min_{\mu \in A} \int_0^{+\infty} \frac{\left(f_{X_{N-1}}(t)\right)^2}{\tilde{f}_{X_{N-1}}(t ; \mu)} u\left(N-1, s_k+t\right) d t. \tag{16}$$

Interpolating $u\left(N-1, s_k\right)$, $k=0,1, \cdots, K$, in the s-space is used. Also linear extrapolation is used for $s>S$ when needed.

3. Having $u\left(n, s_k\right)$ for all $n \in\{0,1, \cdots, N-1\}$ and for all s_k in the grid $k=0,1,2, \cdots, K$, the following step is to solve for $\mu_n, n=1,2, \cdots, N$, by going forward in time. Let $S_0=0$ and sample from $\tilde{f}_{X_1}\left(\cdot, \mu_1\right)$ to get S_1. Note that μ_1 has been already computed in the resolution of the backward problem. Then compute μ_2 as

$$\mu_2\left(\tilde{s}_1\right)=\underset{\mu \in A}{\arg \min } \int_0^{\infty} \frac{\left(f_{X_2}(t)\right)^2}{\tilde{f}_{X_2}(t ; \mu)} u\left(2, \tilde{s}_1+t\right) d t. \tag{17}$$

Having computed μ_2, we simulate S_2 as $S_2 = \tilde{s}_1 + X_2$, with X_2 sampled from $\tilde{f}_{X_2}\left(.; \mu_2\right)$. We keep repeating this procedure until we get μ_N and then we sample X_N.

4. The forward problem is repeated M times. The proposed IS estimator is then given as

$$\hat{\alpha}_{\mathrm{IS}} = \frac{1}{M} \sum_{k=1}^{M} g\left(S_N^{(k)}\right) \prod_{i=1}^{N} \frac{f_{X_i}\left(X_i^{(k)}\right)}{\tilde{f}_{X_i}\left(X_i^{(k)}, \mu_i(S_{i-1}^{(k)})\right)}. \tag{18}$$

Remark

In the case of smooth controls, the optimization problem (17) can be avoided by using instead an interpolation between the controls, obtained in the backward step, on the grid $s_1, \cdots, s_K$.

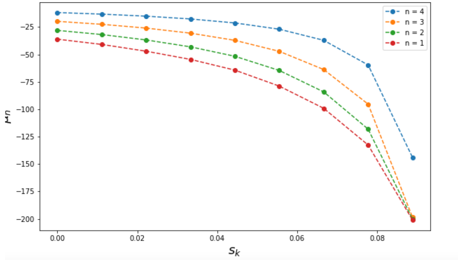


Figure 2: Optimal control in the case of the left-tail i.i.d standard Log-normal RVs with $N = 4,\ K = 10,\ m = 0$ dB, $\sigma = 8.5$ dB, and $\gamma_{th} = -10$ dB.

- Denote our estimator by $\mathrm{HRT\text{-}SOC}$ and HRT to be the estimator without SOC, i.e. the control is constant, independent of the state and time (Ben Rached et al. 2016).
- Let M_{HRT} and $M_{\mathrm{HRT\text{-}SOC}}$ be the number of required simulation runs to ensure a relative error equal to TOL.
- The total costs have the following expressions

$$\mathrm{W}_{\mathrm{HRT\text{-}SOC}} = \underbrace{N \times K \times T_b}_{\text{Backward cost}} + \underbrace{M_{\mathrm{HRT\text{-}SOC}} \times T_f}_{\text{Forward cost}}, \tag{19}$$

$$\mathrm{W}_{\mathrm{HRT}} = \underbrace{M_{\mathrm{HRT}} \times T_f}_{\text{Forward cost}}, \tag{20}$$

where T_b is the time required in the backward algorithm to calculate a single control and T_f the cost per sample in the forward step (it is approximately the same for both approaches).

Idea: divide the sum S_N into B blocks and compute the controls for each block rather than for each $X_i, i = 1, \cdots, N$
$\rightarrow$ reduce the backward cost from $N \times K \times T_b$ to $B \times K \times T_b$. We call this method aggregate method.

- Choose B blocks, such that $B \in \{1, 2, \cdots, N\}$ and consider the following dynamics

$$S_{n_m + b_{m+1}} = S_{n_m} + \sum_{i=n_m+1}^{n_m+b_{m+1}} X_i, \quad m = 0, 1, \cdots, B-1, \tag{21}$$

  where $n_m = \sum_{j=1}^{m} b_j$, and $b_m, \quad m = 1, 2, \cdots, B$, are chosen such that $n_B = \sum_{j=1}^{B} b_j = N$.
- The idea is to have the same control $\mu_m(S_{n_m})$ for each X_i from $i = n_m + 1$ to $i = n_{m+1}$.

Remark

With this proposed approach, we decrease the cost of the backward step with the price of increasing the variance.

OP in a Log-Normal environment without co-channel interference

$$P_{out} = \mathbb{P}\left(\sum_{n=1}^{N} X_i < \gamma_{th}\right) = \mathbb{E}\left[g\left(\sum_{n=1}^{N} X_i\right)\right], \tag{22}$$

where $g(x) = 1_{(x<\gamma_{th})}$.

Comparison with the HRT (Ben Rached et al. 2016) and to the exponential twisting estimator (Asmussen et al. 2016).

OP in the presence of co-channel interference in a Log-Normal environment for SISO systems

$$P_{out} = \mathbb{P}\left(\mathrm{SINR} < \gamma_{th}\right) = \mathbb{P}\left(\frac{X_0}{\sum_{n=1}^{N} X_n + \eta} \leq \gamma_{th}\right) = \mathbb{E}\left[g\left(\sum_{n=1}^{N} X_i\right)\right], \tag{23}$$

where $g(x) = F_{X_0}(\gamma_{th}(x+\eta))$.

Comparison with the estimator based on a covariance matrix scaling (CS) technique (Ben Rached et al .2017) the exponentially tilted (ET) estimator (Botev and L'Ecuyer 2017).

1. We fix $N = 10$ and $TOL = 0.05$ and we vary the threshold γ_{th} (range of OP: $[2 \times 10^{-12}, 6 \times 10^{-6}]$).

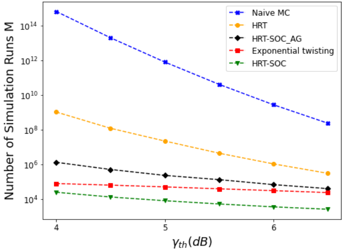


Figure 3: Number of required simulation runs for 5% relative error. For the aggregate method, we choose a constant parameter b, i.e. $b_m = 2$ for all $m = 1, \cdots, B$ with $B = \frac{N}{2}$.

- The HRT-SOC approach requires the smallest number of simulation runs.
- The number of simulations is reduced by about 41775 times for a small threshold ($4\ dB$) which corresponds to an OP value of 2×10^{-12}.

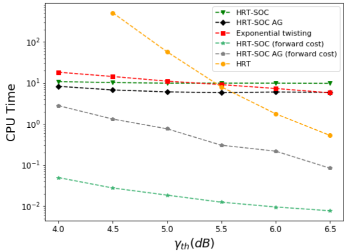
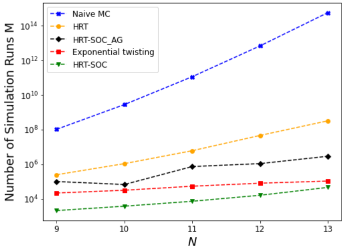


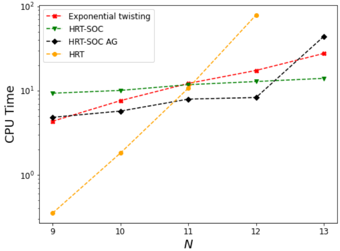
Figure 4: CPU time required for 5% relative error. For the aggregate method, we choose a constant parameter b, i.e. $b_m = 2$ for all $m = 1, \cdots, B$ with $B = \frac{N}{2}$.

- As the event becomes rarer, the time gap between the proposed approach and other IS techniques increases significantly.
- The efficiency of the aggregate method in terms of time reduction exceeds the loss in terms of variance.
- The HRT-SOC-AG reduces the CPU time by about 1.7 times compared to the HRT-SOC approach for $\gamma_{\text{th}} \geq 5$ dB.

2. We fix $\gamma_{th} = 10$ and $TOL = 0.05$ and we vary N. For HRT-SOC-AG, $b_m = 2$, $m = 1, \cdots, \frac{N}{2}$ for N even, and $b_m = 2, m = 1, \cdots, \frac{N-3}{2}$, $b_{\frac{N-1}{2}} = 3$ for N odd. (range of OP: $[2.5 \times 10^{-12}, 6 \times 10^{-5}]$).



(a) Number of required simulation runs.



(b) CPU time required.

- For $N = 13$, HRT-SOC requires 7455 times less simulation runs than the HRT technique to meet the same accuracy requirement.

Remarks

- For small γ and large N, constant block b is less efficient in terms of CPU time than the HRT-SOC approach.
- In these cases, it is more efficient to reduce the variance rather than to reduce the cost of the backward step.
- $b_m, \ m = 1, \cdots, B$ should be adaptively chosen to give better results

$$\min_{b,M,K} \quad B \times K \times T_b + M \times T_f,$$

$$\text{subject to } C^2 \frac{\mathrm{Var}\left[T_{\text{HRT-SOC-AG}}(b, K)\right]}{M\alpha^2} \leq \mathrm{TOL}^2.$$

- An optimal choice of b_m in the case of a very rare event is $b_m = 1, \ m = 1, \cdots, B$ with $B = N$.
- When the event becomes less rare, an optimal choice of B is to take a single block; i.e. $b_1 = N$. By doing this, the HRT-SOC-AG technique reduces to the HRT technique since in this case the controls are state-independent.

OP in a Log-Normal environment without co-channel interference

$$P_{out} = \mathbb{P}\left(\sum_{n=1}^{N} X_i < \gamma_{th}\right) = \mathbb{E}\left[g\left(\sum_{n=1}^{N} X_i\right)\right], \tag{24}$$

where $g(x) = 1_{(x<\gamma_{th})}$.

Comparison with the HRT (Ben Rached et al. 2016) and to the exponential twisting estimator (Asmussen et al. 2016).

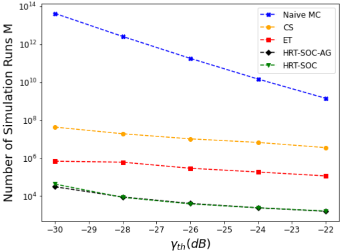
OP in the presence of co-channel interference in a Log-Normal environment for SISO systems

$$P_{out} = \mathbb{P}\left(\mathrm{SINR} < \gamma_{th}\right) = \mathbb{P}\left(\frac{X_0}{\sum_{n=1}^{N} X_n + \eta} \leq \gamma_{th}\right) = \mathbb{E}\left[g\left(\sum_{n=1}^{N} X_i\right)\right], \tag{25}$$

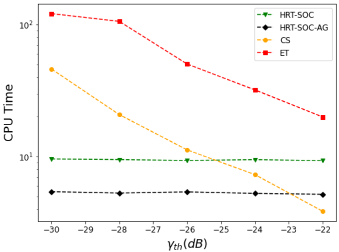
where $g(x) = F_{X_0}(\gamma_{th}(x+\eta))$.

Comparison with the estimator based on a covariance matrix scaling (CS) technique (Ben Rached et al .2017) the exponentially tilted (ET) estimator (Botev and L'Ecuyer 2017).

1. We fix $N = 10$ and $TOL = 0.05$ and we vary γ (range of OP: $[2 \times 10^{-10}, 5 \times 10^{-7}]$).



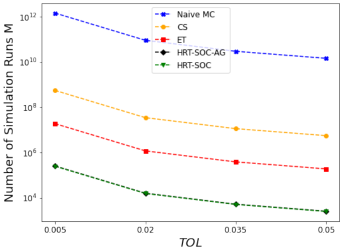
(a) Number of required simulation runs.



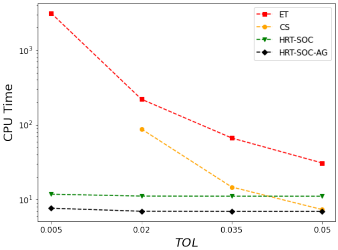
(b) CPU time required.

- The CS technique requires approximately 2000 times as many simulations as needed by the HRT-SOC scheme.
- When $\gamma_{\text{th}} = -30$ dB, HRT-SOC is 13 times more efficient than the ET scheme.
- The HRT-SOC-AG technique requires less time than the HRT-SOC technique.

2. We fix $\gamma_{th} = -24\ dB$ and $N = 10$ $(\alpha = 10^{-7})$ and we vary TOL.



(a) Number of required simulation runs.



(b) CPU time required.

- Our approaches are 2000 times (respectively 65 times) more efficient than the CS (respectively the ET) approaches for all values of TOL.
- The required time for the proposed methods compared to the other algorithms remains unchanged for the considered range of TOL.

Summary

- We developed a generic state dependent IS algorithm in order to efficiently estimate rare events quantities that could be written in a form of an expectation of some functional of sums of independent RVs.
- Within a pre-selected class of change of measures, the optimal IS parameters are determined via a connection to a SOC formulation.
- Showed a substantial amount of variance reduction compared to other well-known estimators.
- Proposed an aggregate method to further improve the efficiency in terms of computational time.

Future directions: further optimize the aggregate method, and extend the approach to the multivariate case.

Further details in : Ben Amar, E., Ben Rached, N., Haji-Ali, A. L., Tempone, R. (2023). State-dependent importance sampling for estimating expectations of functionals of sums of independent random variables. Statistics and Computing, 33(2), 40.

References

- N. Ben Rached, A. Kammoun, M.-S. Alouini, and R. Tempone. An accurate sample rejection estimator of the outage probability with equal gain combining. IEEE Open Journal of the Communications Society, 2020.
- S. Asmussen and P. W. Glynn. Stochastic simulation: Algorithms and Analysis. Stochastic modelling and applied probability. Springer, New York, 2007.
- N. Ben Rached, A. Kammoun, M.-S. Alouini, and R. Tempone. Unified importance sampling schemes for efficient simulation of outage capacity over generalized fading channels. IEEE Journal of Selected Topics in Signal Processing, Mar. 2016.
- N. Ben Rached, A. Kammoun, M.-S Alouini, and R. Tempone. On the efficient simulation of outage probability in a Log-normal fading environment. IEEE Transactions on Communications, 2017.
- D. P. Kroese, T. Taimre, and Z.I. Botev. Handbook of Monte Carlo methods. Wiley, N.J, 2011.
- K. Rajhaa AM and S. Juneja. State-independent importance sampling for estimating large deviation probabilities in heavy-tailed random walks. In proc. of the IEEE 6th International ICST Conference on Performance Evaluation Methodologies and Tools, 2012.

- A. Bassamboo, S. Juneja, and A. Zeevi. On the inefficiency of state-independent importance sampling in the presence of heavy tails. Operations research letters, 2007.
- K. RA Murthy, S. Juneja, and J. Blanchet. State-independent importance sampling for random walks with regularly varying increments. Stochastic Systems, 2015.
- A. Bassamboo, S. Juneja, and A. Zeevi. On the inefficiency of state-independent importance sampling in the presence of heavy tails. Operations research letters, 2007.
- J. Blanchet and H. Lam. State-dependent importance sampling for rare-event simulation: An overview and recent advances. Surveys in Operations Research and Management Science, 2012.
- J. Blanchet and C. Li. Efficient rare event simulation for heavy-tailed compound sums. ACM Transactions on Modeling and Computer Simulation, 2011.
- J. Blanchet and J. Liu. Efficient simulation for large deviation probabilities of sums of heavy-tailed increments. In proc. of the IEEE Winter Simulation Conference, 2006.

References (cont.)

- P. Dupuis, K. Leder, and H. Wang. Importance sampling for sums of random variables with regularly varying tails. ACM Transactions on Modeling and Computer Simulation, 2007.
- P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. Stochastics: An International Journal of Probability and Stochastic Processes, 2004.
- S. Juneja and P. Shahabuddin. Simulating heavy tailed processes using delayed hazard rate twisting. ACM Transactions on Modeling and Computer Simulation, 2002.
- N. Ben Rached, Z. I. Botev, A. Kammoun, M.-S. Alouini, and R. Tempone. On the sum of order statistics and applications to wireless communication systems performances. IEEE Transactions on Wireless Communications, Nov. 2018.
- S. Asmussen, J. L. Jensen, and L. Rojas-Nandayapa. Exponential family techniques for lognormal left tail, Scandinavian Journal of Statistics, Sep. 2016.
- Z. Botev and P. l'Ecuyer. Accurate computation of the right tail of the sum of dependent Log-normal variates. In IEEE Winter Simulation Conference (WSC), 2017.