A SEPARATED SPLITTING TECHNIQUE FOR DISCONNECTED RARE EVENT SETS

Wander Wadman
Daan Crommelin
Jason Frank

CWI Amsterdam
Science Park 123
Amsterdam, 1098 XG, THE NETHERLANDS

ABSTRACT

A key challenge for an efficient splitting technique is defining the importance function. If the rare event set consists of multiple separated subsets this challenge becomes bigger since the most likely path to the rare event set may be very different from the most likely path to an intermediate level. We propose to mitigate this problem of path deviation by estimating the subset probabilities separately using a modified splitting technique. We compare the proposed separated splitting technique with a standard splitting technique by estimating the probability of entering either of two separated intervals on the real line. The squared relative error of the estimator is shown to be significantly higher when using standard splitting than when using separated splitting. We show that this difference increases if the rare event probability becomes smaller, illustrating the advantage of the separated splitting technique.

1 INTRODUCTION

Efficient estimation of very low probabilities is desired in many application areas like telecommunication networks and power grid reliability (de Boer 2000; Wang et al. 2011). Monte Carlo simulation is a robust and popular technique in case the model does not allow an exact derivation, but Crude Monte Carlo may require prohibitively many samples to be practically feasible. To reduce the computational burden, rare event simulation techniques have been developed (Rubino and Tuffin 2009), and (importance) splitting is a well-known variant. A splitting technique aims to reduce the squared relative error by replicating (splitting) trajectories as soon as the rare event is presumed more likely to occur (L'Ecuyer, Demers, and Tuffin 2006; Glasserman et al. 1999; Villén-Altamirano and Villén-Altamirano 2006).

The efficiency of a splitting technique depends strongly on the choice of the importance function. The optimal importance function strictly increases with the probability of reaching the rare event set from the current state. In this paper we assume no knowledge of this probability except that the underlying Markov process is (the discretization of) a continuous stochastic process. In a simple case with a one-dimensional state space and an interval as the rare event set, any importance function that measures the proximity of the current state to the interval is optimal. In many other cases however, the choice for the importance function is much more difficult.

For example, a challenge arises when the rare event set is disconnected; that is, it consists of multiple nonempty, separated subsets. Such a rare event set is plausible in a reliability assessment of a multiple component system for which various combinations of component behavior may lead to a system failure (Grainger and Stevenson 1994, Billinton and Li 1994). One example involves the reliability estimation of an electrical power grid with nodal power injections modeled as stochastic processes (Wadman, Crommelin, and Frank 2013). Then if at any grid node the injection is very different from that at an adjacent node, the current through the connection in between may exceed its allowed maximum. Hence there could be various combinations of power injections leading to the rare event of a grid fault, and they could involve

opposite extremes of power injections. In case of a disconnected rare event set, an importance function based on the proximity to the complete rare event set may lead to poor results: the algorithm may replicate sample paths near a subset that has only little or even negligible contribution to the rare event probability. Therefore the importance function may be far from optimal as the simulation focuses on areas where the rare event probability is hardly increased or even decreased.

In this paper we assume that we have a partition of the rare event set into subsets that are (or at least might be) separated from each other. We propose to estimate the rare event subset probabilities separately. We use a modified splitting technique to estimate each subset probability, taking care that sample paths hitting multiple sets are not counted multiple times. Standard splitting framework and notation are introduced in Section 2. We discuss some properties of the optimal importance function in Section 3. In Section 3.1 we show a one-dimensional problem where the choice for the importance function is clear, and we describe a technique to calibrate the levels using a pilot run. In Section 3.2, we discuss the challenge of a disconnected rare event set, and we describe how to estimate each subset probability separately. We present the numerical results of an experiment in Section 4, after which conclusions are given in Section 5.

2 SETTINGS AND NOTATION

2.1 Rare Event Simulation

We consider the class of rare events where a discrete-time Markov chain $\{X_t, t \geq 0\}$ with state space $\mathcal{X}$ enters the rare event set $B \subset \mathcal{X}$ before stopping time ζ. We define

$$\tau_B = \inf\{t > 0 : X_t \in B\}$$

as the first entrance time in B and $x_0 \notin B$ as the initial state. We want to estimate the rare event probability

$$\gamma := \mathbb{P}(\tau_B < \zeta | X_0 = x_0).$$

Two typical choices for ζ are a fixed finite time $\zeta = T$ and stopping time $\zeta = \tau_A := \inf\{t > 0 : \exists\, 0 \leq s < t \text{ s.t. } X_s \notin A \text{ and } X_t \in A\}$, the first time the chain reenters A, where $A \subset \mathcal{X}$ is some set that contains x_0. The Crude Monte Carlo (CMC) estimator

$$\tilde{\gamma}_n := \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\tau_B < \zeta \text{ in sample } i\}}$$

for γ is unbiased but its squared relative error

$$\text{SRE}(\tilde{\gamma}_n) := \frac{\text{Var}\, \tilde{\gamma}_n}{\gamma^2} = \frac{\gamma(1-\gamma)}{\gamma^2 n} = \frac{1-\gamma}{\gamma n} \tag{1}$$

diverges to infinity as $O(1/\gamma)$ when $\gamma \to 0$. Therefore, to estimate a very small probability using CMC simulation, one will need a number of samples that may be computationally too intensive in practice.

2.2 Importance Splitting

Any splitting technique starts by defining an importance function $h : \mathcal{X} \mapsto \mathbb{R}$ that assigns an importance value to each chain state. We assume that A and B are determined in terms of this importance function:

$$A = \{x \in \mathcal{X} : h(x) \leq 0\} \text{ and } B = \{x \in \mathcal{X} : h(x) \geq 1\}.$$

The interval $[0,1]$ is divided into m subintervals with thresholds $0 = l_0 < l_1 < \cdots < l_m = 1$. We define $T_k = \inf\{t > 0 : h(X_t) \geq l_k\}$ as the first time of hitting the k-th level and $D_k = \{T_k < \zeta\}$ as the event that

the k-th level is hit during $[0, \zeta]$. Obviously, $\mathbb{P}(D_0) = 1$, and $\mathbb{P}(D_m) = \gamma$ is the value of interest. As $D_m \subset D_{m-1} \subset \cdots \subset D_0$, we decompose

$$\gamma = \prod_{k=1}^{m} \mathbb{P}(D_k | D_{k-1})$$

in m conditional probabilities $p_k := \mathbb{P}(D_k | D_{k-1})$, which we will estimate separately. Generating independent sample paths from the conditional distribution of the entrance state $(T_{k-1}, X_{T_{k-1}})$ given D_{k-1} would give us an estimate for p_k. However, as we do not know this distribution for levels $k > 1$, we use its empirical distribution, obtained from samples of the previous level. In this way, we proceed recursively, and at each level k we estimate p_k by the proportion $\hat{p}_k$ of sample paths for which D_k occurs (see Figure 1). Then the product

$$\hat{\gamma} := \prod_{k=1}^{m} \hat{p}_k$$

is an unbiased estimator for γ for several variants of the splitting technique, for example Fixed Splitting, Fixed Effort and Fixed Number of Successes. Proofs that the estimator is unbiased for these three variants can be found in Dean and Dupuis (2009), Del Moral and Garnier (2005) and Amrein and Künsch (2011), respectively. Fixed Splitting determines the number of splittings per level hit in advance. A disadvantage of this technique is that, especially when the number of levels and the thresholds are far from optimally chosen, paths may hit the next level many times or never. The former situation results in a high computational burden, the latter situation in a useless estimate $\hat{\gamma} = 0$. Fixed Effort attempts to overcome these issues by fixing the total number of samples per level. Unfortunately, as the realized number of hits may differ from the expected number of hits, the same issues of path explosion and extinction, although to a lesser extent, may persist. Fixed Number of Successes (FNS) fixes the number of hits r_k for all levels k, repeating the simulation at one level until r_k hits are observed. This technique avoids the issues of path explosion and extinction by adjusting the computational effort. We use the FNS splitting technique in the rest of this paper.
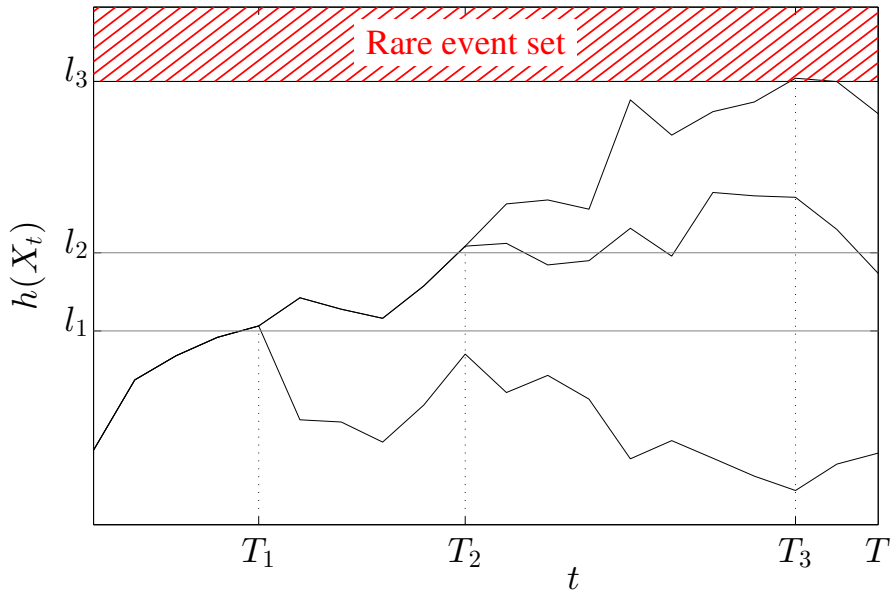

Figure 1: A minimal example of a splitting simulation with $\hat{p}_1 = 1$, $\hat{p}_2 = \hat{p}_3 = 1/2$.

3 FINDING A SUITABLE IMPORTANCE FUNCTION

Finding a suitable importance function is crucial to reduce the variance of the splitting estimator (Garvels 2000, L'Ecuyer, Demers, and Tuffin 2006). The importance function should 'reward good behavior' by splitting trajectories that are more likely to hit the rare event set. Glasserman et al. (1998) show that the levels should be chosen in a way consistent with the most likely path to the rare event set. Ideally, this path should coincide with the most likely path to any intermediate level. In case $\zeta = \tau_A$ is a stopping time as defined in Section 2.1, Garvels, Van Ommeren, and Kroese (2002) propose to use the importance function

$$g(h^*(x)),\tag{2}$$

with $g : [0,1] \mapsto \mathbb{R}$ some strictly increasing function and

$$h^*(x) := \mathbb{P}(\tau_B < \zeta | X_0 = x).\tag{3}$$

As knowing $h^*(x)$ would defeat the point of using simulation, the lesson is to find an importance function that is close to (2). Some results on the efficiency of splitting can be derived under the assumption that the probability of hitting the next level does not depend on the current entrance state

$$\mathbb{P}\left(D_k | D_{k-1}, (T_{k-1}, X_{T_{k-1}})\right) = p_k \quad \forall \ (T_{k-1}, X_{T_{k-1}}).\tag{4}$$

Under this assumption and using the optimal choice for m and p_k Amrein and Künsch (2011) show that the squared relative error of the FNS estimator is proportional to

$$\mathrm{SRE}(\hat{\gamma}) \propto \frac{(\log \gamma)^2}{n}.$$

Hence the $\mathrm{SRE}(\hat{\gamma})$ diverges as $O((\log \gamma)^2)$ when $\gamma \to 0$. This squared logarithmic divergence rate is slower than the divergence rate of the CMC squared relative error in (1). Splitting estimators outperfom the CMC estimator in this sense. Furthermore, under assumption (4) the authors show that $\hat{p}_1, \ldots, \hat{p}_m$ are independent and derive the bound

$$\mathrm{SRE}(\hat{\gamma}) \leq -1 + \prod_{k=1}^{m} \frac{r_k - 1}{r_k - 2}\tag{5}$$

for the squared relative error of the FNS splitting estimator. That is, the accuracy of the estimator can be controlled by choosing the number of hits r_k sufficiently large for all k. In Section 4 we will check if the simulation results of an experiment fulfill bound (5). As the validity of this bound is a necessary condition for (4) to hold, this check will serve as a test for the quality of the chosen importance function.

3.1 A Simple Case with a Connected Rare Event Set

Consider the one-dimensional case $\mathcal{X} = \mathbb{R}$, X_t the discretization of a continuous stochastic process and $B = [b, \infty)$ an interval for some $b > x_0$. Then $\gamma = \mathbb{P}(\tau_B < \zeta | X_0 = x_0)$ is a barrier crossing probability and it is fair to assume that $\mathbb{P}(\tau_B < \zeta | X_0 = x)$ increases in x as long as $x < b$, since for the continuous counterpart of X_t this is certainly true. Therefore, we may assume that the normalized proximity function

$$h_p(x) := 1 - (b - x)/(b - x_0)\tag{6}$$

is of the form (2) even though g and h^* are not explicitly known. Stated in terms of Glasserman et al. (1998), the most likely path to the rare event set will automatically coincide with the most likely path to any intermediate level in this simple case.

The remaining challenge is to find a suitable partition of the interval $[0,1]$ into level thresholds. Amrein and Künsch (2011) propose a pilot run to obtain a rough idea of the optimal number and height of the thresholds for a subsequent final run. That is, they suggest a splitting simulation with a large number of levels and a moderate number of level hits such that the computational intensity is small compared to that of the final run. The authors showed that the optimal value for p_k in terms of variance reduction is $p_{\text{opt}} := 0.2032$. Therefore, a pilot estimate of p_k close to one suggests to merge level k with a adjacent level, whereas a pilot estimate close to zero suggests to divide the level into multiple levels.

Along these lines we now describe such a calibration procedure in more detail. That is, we search for m, and for all $k = 1, \ldots, m-1$ we search for the threshold l_k such that

$$p(l_k) = p_{\text{opt}}^{m-k},$$

where $p(l)$ is the conditional probability of the rare event given that the importance function has exceeded l. Assuming no knowledge of h^*, we perform a pilot run using many levels, say 20, and equidistant thresholds $l_k^{\text{pilot}} = k/20$. As this will yield information on $p(l)$ only at the pilot threshold values $l = l_i^{\text{pilot}}$, we assume a power law relationship for $p(l)$ between each pair of subsequent pilot thresholds:

$$p(\alpha \Delta l) = p(\Delta l)^{\alpha}, \tag{7}$$

for $\alpha > 0$. For example, if the top level probability $\hat{p}_m^{\text{pilot}} = p_{\text{opt}}^2$, then (7) suggests to set the penultimate threshold l_{m-1} twice as close to the top as the penultimate pilot threshold l_{19}^{pilot}. Hence, we approximate $p(l)$ by the log-linear interpolation $p^*(l)$ of the estimators

$$\prod_{i=k+1}^{20} \hat{p}_i^{\text{pilot}}$$

for $p\left(l_k^{\text{pilot}}\right)$, as shown in Figure 2. Then we simply solve the obtained piecewise linear equation

$$\log p^*(l_k) = (m-k) \log p_{\text{opt}},$$

for l_k. The number of levels is automatically found by computing $l_{m-1}, l_{m-2}, \ldots, l_i$ where i is the last index such that $l_i > 0$. After re indexing we have found the levels $0 = l_0 < \cdots < l_m = 1$ that we use for the final run.
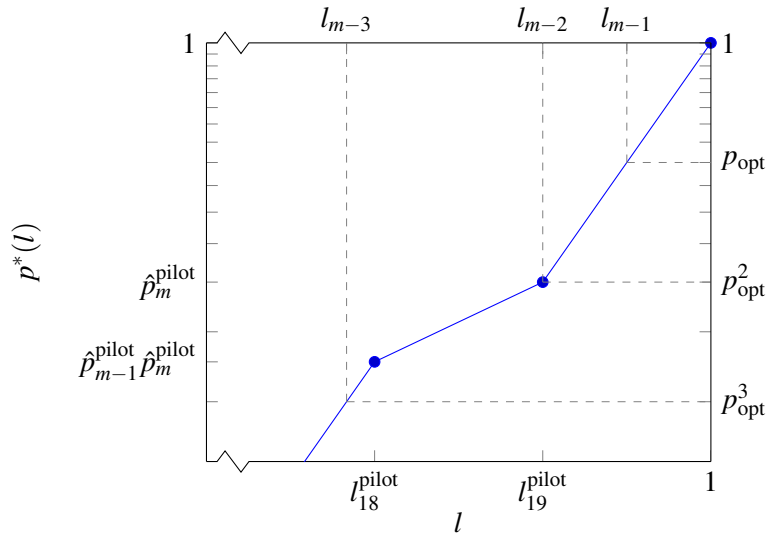

Figure 2: An updated set of thresholds $l_0, \ldots, l_m$ is found by a log-linear interpolation of the estimates of reaching the rare event set from the pilot thresholds.

3.2 A Disconnected Rare Event Set

Now we consider a possibly multidimensional state space $\mathcal{X}$, and a rare event set B that can be partitioned into two nonempty, separated subsets B_1 and B_2. By an induction argument the rest of this section can be generalized to a partition into more than two subsets. As an extension of (6), an intuitive choice for the importance function is the maximum

$$h_{12}(x) := \max_{i=1,2} h_i(x) \tag{8}$$

of the normalized proximities of x to each subset:

$$h_i(x) := 1 - \frac{d(x,B_i)}{d(x_0,B_i)}. \tag{9}$$

Here $d(v,V)$ denotes the distance between point $v \in \mathcal{X}$ and set $V \subset \mathcal{X}$. However, as we assume no prior knowledge of $h^*(x)$, it could be that the most likely path to an intermediate level approaches B_2, whereas the most likely path to B does not and instead goes directly to B_1. If X_t is the discrete counterpart of a continuous stochastic process, we may expect that this deviation between the most likely paths to different level sets becomes smaller as the distance between B_1 and B_2 goes to zero.

Therefore, we propose not to estimate the rare event probability $\gamma = \mathbb{P}(\tau_B < \zeta | X_0 = x_0)$ using h_{12} in (8) for disconnected B. Instead, we write $\gamma = \gamma_1 + \gamma_2$ with

$$\begin{aligned} \gamma_1 &:= \mathbb{P}(\tau_{B_1} < \zeta | X_0 = x_0), \\ \gamma_2 &:= \mathbb{P}(\tau_{B_2} < \zeta \cup \tau_{B_1} \geq \zeta | X_0 = x_0). \end{aligned}$$

We can interpret this as the rare event being partitioned into a part where B_1 is hit in due time (and maybe B_2 as well), and a part where B_2 is hit in due time while B_1 is never hit during $[0,\zeta]$. We propose to separately estimate γ_1 and γ_2 using two independent splitting simulations. To obtain an unbiased estimator for γ_1 we simply choose h_1 in (9) as importance function (see Figure 3). To obtain an unbiased estimator for γ_2 we choose h_2 in (9) as importance function and a modified splitting framework rewarding sample paths approaching B_2 while killing those that hit B_1 (see Figure 4). We call these two runs together the *separated splitting technique*, as opposed to the single *standard splitting technique* using importance function h_{12} in (8).

More precisely, to estimate γ_2 we start from a standard splitting setup for $\mathbb{P}(\tau_{B_2} < \zeta | X_0 = x_0)$ as described in subsection 2.2. We set l_k, T_k and D_k for all k as before, but we modify the levels by defining

$$\begin{aligned} \bar{D}_k &:= D_k \cap \{\tau_{B_1} \geq T_k\} \quad \forall\, k = 0, \dots, m, \\ \bar{D}_{m+1} &:= D_m \cap \{\tau_{B_1} \geq \zeta\}. \end{aligned}$$

So to enter level k, we additionally require the chain not to have entered B_1 in the mean time. Further, the occurrence of the additional level $\bar{D}_{m+1}$ requires that after entering B_2 in due time, the chain should not have entered B_1 during the remaining time domain $[T_m, \zeta]$. Note that $\bar{D}_{m+1}$ is exactly the rare event, and since again $\bar{D}_0 \supset \bar{D}_1 \supset \cdots \supset \bar{D}_m \supset \bar{D}_{m+1}$ we can write

$$\gamma_2 = \mathbb{P}(\bar{D}_{m+1}) = \prod_{k=1}^{m+1} \mathbb{P}(\bar{D}_k | \bar{D}_{k-1}).$$

We may now resume the standard splitting technique as described in subsection 2.2 using $D_k = \bar{D}_k$ for all k. Therefore, an unbiased estimator for γ_2 is given by
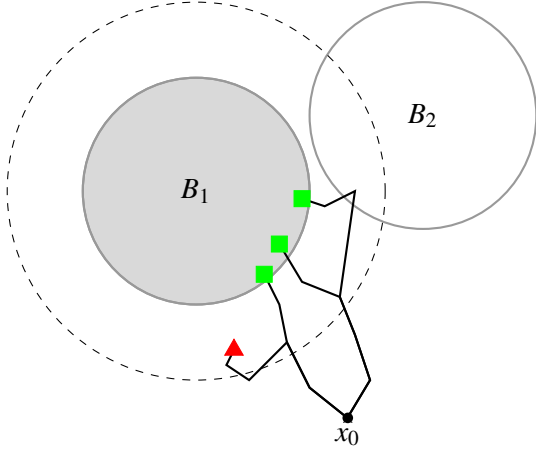
$$\hat{\gamma}_2 := \prod_{k=1}^{m+1} \hat{p}_k,$$

Figure 3: Illustration of splitting sample paths to estimate $\gamma_1 = \mathbb{P}(\tau_{B_1} < \zeta | X_0 = x_0)$, ignoring B_2. Squares correspond to subpaths for which the rare event does occur, triangles correspond to subpaths for which it does not.
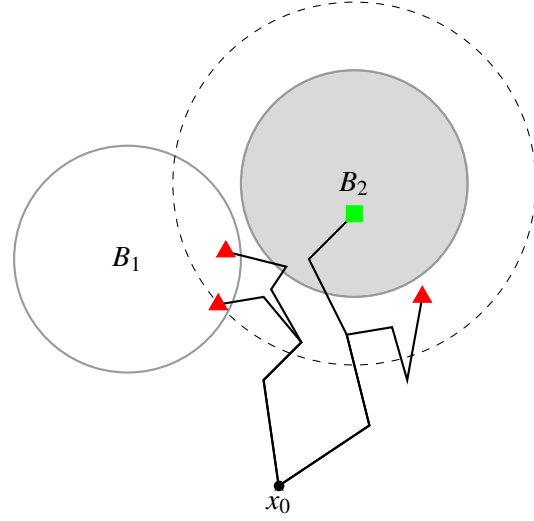
Figure 4: Illustration of splitting sample paths to estimate $\gamma_2 = \mathbb{P}(\tau_{B_2} < \zeta \cup \tau_{B_1} \geq \zeta | X_0 = x_0)$, killing paths that enter B_1. Squares correspond to subpaths for which the rare event does occur, triangles correspond to subpaths for which it does not.

where for all levels k $\hat{p}_k$ is the proportion of sample paths for which $\bar{D}_k$ occurs. We conclude that $\hat{\gamma}_1 + \hat{\gamma}_2$ is an unbiased estimator for γ. Furthermore, if assumption (4) holds we can compute bounds for the squared relative errors $\mathrm{SRE}(\hat{\gamma}_1)$ and $\mathrm{SRE}(\hat{\gamma}_2)$ using (5). Then a bound for $\hat{\gamma}_1 + \hat{\gamma}_2$ follows immediately, since

$$\mathrm{SRE}(\hat{\gamma}_1 + \hat{\gamma}_2) = \frac{\mathrm{Var}\,\hat{\gamma}_1 + \mathrm{Var}\,\hat{\gamma}_2}{(\gamma_1 + \gamma_2)^2} \leq \frac{\mathrm{Var}\,\hat{\gamma}_1}{\gamma_1^2} + \frac{\mathrm{Var}\,\hat{\gamma}_2}{\gamma_2^2} = \mathrm{SRE}(\hat{\gamma}_1) + \mathrm{SRE}(\hat{\gamma}_2). \tag{10}$$

We used independence of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ in the first equality.

4 EXPERIMENT

We compare the proposed *separated splitting technique* using importance functions h_1 and h_2 in (9) with the *standard splitting technique* using importance function h_{12} in (8). The state space $\mathcal{X} = \mathbb{R}$ is one-dimensional and the rare event set $B = B_1 \cup B_2$ consists of two separate intervals $B_1 = [b_1, \infty)$ and $B_2 = (-\infty, b_2]$, $b_1 > b_2$. We assume that $\{X_t, t \geq 0\}$ is a modification of an Ornstein-Uhlenbeck process, where we change the sign of the long-term mean whenever the process passes zero:

$$X(t) = \theta(\mu \operatorname{sgn}(X_t) - X_t)dt + \sigma dW(t). \tag{11}$$

The drift of this stochastic differential equation corresponds to the potential

$$V(x) = -\int \theta(\mu \operatorname{sgn}(x) - x)dx = \frac{\theta}{2}(x - \mu \operatorname{sgn}(x))^2. \tag{12}$$

Figure 5 shows that the double well potential (12) is just the original Ornstein-Uhlenbeck potential, mirrored in the y-axis. We substitute $\mu \operatorname{sgn}(X_t)$ for μ in the exact solution of the Ornstein-Uhlenbeck process (Gillespie 1996) to obtain a discretization of (11):

$$X_{i+1} = X_i e^{-\theta \Delta} + \mu \operatorname{sgn}(X_i) \left(1 - e^{-\theta \Delta}\right) + \sigma \sqrt{\frac{1 - e^{-2\theta \Delta}}{2\theta}} Z_i,$$
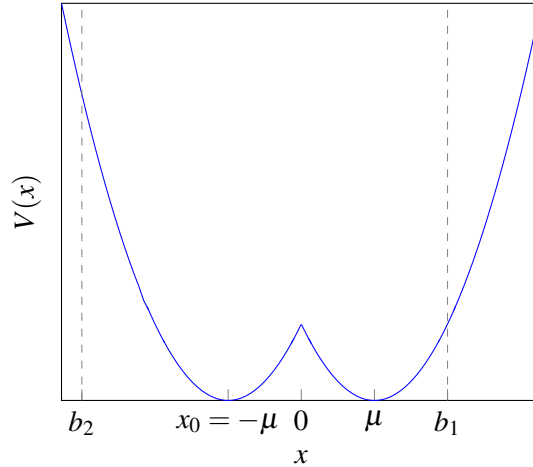
Figure 5: The double well potential (12) of the modified OU process in (11).

with $X_i = X(i\Delta)$ with time step size Δ, and $Z_i \sim N(0,1)$ i.i.d. standard normal for all i. We choose parameter values $\theta = \sigma = 1$, $\mu = -x_0 = 2$, $b_1 = 4$, $b_2 = -6$, $\Delta = 0.001$ and a deterministic stopping time $\zeta = T = 10$. We will estimate the rare event probability $\gamma := \mathbb{P}(\tau_B < T | X_0 = x_0)$. Obviously $\gamma = \gamma_1 + \gamma_2$, with $\gamma_1 := \mathbb{P}(\tau_{B_1} < T | X_0 = x_0)$ and $\gamma_2 = \mathbb{P}(\tau_{B_2} < T \cup \tau_{B_1} \geq T | X_0 = x_0)$. Figure 5 shows that the process starts in the left well and initially has equal probability of going right or left. To arrive at B_1 the process should overcome the potential barrier $V(0) = 2$ at $x = 0$. In spite of this barrier, CMC estimates in Table 1 show that $\gamma_1 \gg \gamma_2$, relating to the fact that $V(b_1) < V(b_2)$. We regard these CMC estimates as reference values for splitting estimates.

We use both the standard splitting technique and the separated splitting technique to estimate γ. We choose the levels as described in Section 3.1 and use $r_k = r = 100$ hits per level. The standard splitting estimator for γ is denoted $\hat{\gamma}_{12}$, whereas the separated splitting technique computes estimators $\hat{\gamma}_1$ for γ_1 and $\hat{\gamma}_2$ for γ_2, yielding $\hat{\gamma}_1 + \hat{\gamma}_2$ for γ. We investigate both techniques by comparing their estimates with the CMC estimates. After repeating both simulations 1000 times we check if the squared relative errors $\text{SRE}(\hat{\gamma}_{12})$, $\text{SRE}(\hat{\gamma}_1)$ and $\text{SRE}(\hat{\gamma}_2)$ are indeed bounded by (5), which should hold under assumption (4).

Table 2 shows that the squared relative error $\text{SRE}(\hat{\gamma}_{12})$ of the standard splitting estimator is more than twice as high as the bound suggested by (5). This suggests that condition (4) under which this bound is derived does not hold, that is, hitting the next level does depend on the entrance state. Table 3 confirms this: more than two third of the first level hits occurred at the left side, whereas hits at the third and fourth level only occurred by exceeding the right side threshold. So hitting the next level clearly depends on the side at which paths crossed the previous threshold, i.e. on the entrance state, whereby (4) cannot hold when using the standard splitting technique.

Separated splitting is not affected by this problem, and indeed the squared relative errors $\text{SRE}(\hat{\gamma}_1)$ and $\text{SRE}(\hat{\gamma}_2)$ do fulfill bound (5) in Table 2. Using (10) we find the bound $0.041 + 0.096 = 0.137$ for $\text{SRE}(\hat{\gamma}_1 + \hat{\gamma}_2)$. Assuming normality one can show that the implied conservative 95% confidence interval for γ overlaps with the narrow CMC confidence interval. However, one can show that the conservative confidence interval implied by the standard splitting technique is a significant underestimation.

Table 1: CMC estimates $\tilde{\gamma}_n$ for the subset probabilities, using $n = 10^7$ samples.

Probability	CMC estimate $\tilde{\gamma}_n$	$\text{SRE}(\tilde{\gamma}_n)$
$\gamma = \gamma_1 + \gamma_2$	3.50×10^{-3}	2.85×10^{-5}
γ_1	3.50×10^{-3}	2.85×10^{-5}
γ_2	2.50×10^{-6}	4.00×10^{-2}

We repeat the experiment for higher barriers by increasing μ, i.e. for smaller rare event probabilities. Table 4 shows that for barrier height $V(0) = 2.42$, $\text{SRE}(\hat{\gamma}_{12})$ is 3 times as high as bound (5) suggests. Table 5 and 6 show that $\text{SRE}(\hat{\gamma}_{12})$ exceeds the theoretical bound (5) by a factor up to 15 for even higher barriers. This illustrates that the accuracy of the splitting estimate may be much lower than expected when one does not assess separated rare event subsets in separated splitting simulations.

Table 2: Standard splitting estimator $\hat{\gamma}_{12}$ for γ and separated splitting estimator $\hat{\gamma}_1 + \hat{\gamma}_2$ for γ. The squared relative error $\text{SRE}(\hat{\gamma}_{12})$ exceeds bound (5), whereas $\text{SRE}(\hat{\gamma}_1)$ and $\text{SRE}(\hat{\gamma}_2)$ do not.

	Value	SRE($\cdot$)	Bound for SRE($\cdot$)
$\hat{\gamma}_{12}$	2.48×10^{-3}	0.069	0.031
$\hat{\gamma}_1$	3.58×10^{-3}	0.034	0.041
$\hat{\gamma}_2$	1.52×10^{-6}	0.066	0.096

Table 3: Efficiency loss using standard splitting: at the first level a significant number of successful sample paths approached the left interval 'in vain', since no sample path entered the left interval at the final level.

Level	Hits on left side	Hits on right side
1	67 (at $x = -3.92$)	33 (at $x = 0.88$)
2	8 (at $x = -4.78$)	92 (at $x = 2.17$)
3	0 (at $x = -5.63$)	100 (at $x = 3.44$)
4	0 (at $x = -6$)	100 (at $x = 4$)

Table 4: Experiment as in Table 2, but now $\mu = x_0 = 2.2$, thus raising the barrier height to $V(0) = 2.42$, and $b_1 = 4.4$, $b_2 = -6.6$. The squared relative error $\text{SRE}(\hat{\gamma}_{12})$ exceeds bound (5), whereas $\text{SRE}(\hat{\gamma}_1)$ and $\text{SRE}(\hat{\gamma}_2)$ do not.

	Value	SRE($\cdot$)	Bound for SRE($\cdot$)
$\hat{\gamma}_{12}$	5.28×10^{-4}	0.120	0.041
$\hat{\gamma}_1$	8.75×10^{-4}	0.043	0.052
$\hat{\gamma}_2$	6.18×10^{-8}	0.095	0.118

Table 5: Experiment as in Table 2, but now $\mu = x_0 = 2.5$, further raising the barrier height to $V(0) = 3.125$, and $b_1 = 5$, $b_2 = -6.5$. The squared relative error $\text{SRE}(\hat{\gamma}_{12})$ exceeds bound (5) by a factor 5.

	Value	SRE($\cdot$)	Bound for SRE($\cdot$)
$\hat{\gamma}_{12}$	2.56×10^{-5}	0.43	0.085
$\hat{\gamma}_1$	3.91×10^{-5}	0.076	0.074
$\hat{\gamma}_2$	1.79×10^{-6}	0.071	0.096

Table 6: Experiment as in Table 2, but now $\mu = x_0 = 3$, further raising the barrier height to $V(0) = 4.5$, and $b_1 = 6$, $b_2 = -7.2$. The squared relative error $\text{SRE}(\hat{\gamma}_{12})$ exceeds bound (5) by a factor 15.

	Value	SRE($\cdot$)	Bound for SRE($\cdot$)
$\hat{\gamma}_{12}$	3.89×10^{-7}	1.42	0.096
$\hat{\gamma}_1$	2.16×10^{-7}	0.121	0.107
$\hat{\gamma}_2$	3.01×10^{-7}	0.078	0.107

Finally, we should note that $\text{SRE}(\hat{\gamma}_1)$ is exceeding its bound as well in these last two experiments, although only to a moderate extent. To explain this, we point out that bound (5) is based on the product of the inequality

$$\text{SRE}(\hat{p}_k) \leq \frac{1}{r_k - 2} \tag{13}$$

over all levels k, where r_k is the number of hits at level k. The simulation results show that every transgression of the bound by $\text{SRE}(\hat{\gamma}_1)$ corresponds to violation of (13) for the level for which X_t has to traverse the interval $[0, \mu]$. For large μ, the drift $-\nabla V(X_t)$ of X_t rapidly changes at left side of this interval (see Figure 5). Hence even for a small time step the probability that X_t enters the next level depends sensitively on the entrance state of the previous level, whereby we cannot assume (4) anymore. Nevertheless, the overshoot of $\text{SRE}(\hat{\gamma}_1)$ is a factor 1.03 and 1.14, respectively, which is small compared to that of $\text{SRE}(\hat{\gamma}_{12})$ of a factor 5 and 15, respectively.

5 CONCLUSION

Estimating the probability of a rare event set consisting of multiple separated subsets by use of a splitting technique is challenging since the most likely path to the rare event set may be very different from the most likely path to an intermediate level. We proposed a separated splitting technique to mitigate the difficulties posed by this path deviation by estimating the probability of hitting each subset individually. For each subset, a modified splitting simulation replicates sample paths that approach this subset, while killing those that enter a subset that is already addressed in a previous simulation.

We compared the separated splitting technique with a standard splitting technique by estimating the probability of entering either of two separated intervals on the real line. Choosing a stochastic process with a potential barrier close to one of the intervals, we show that the standard splitting technique replicates sample paths that approach the subset that is not reached in the end. This efficiency loss — quantified as the squared relative error of the estimator — is observed to be significantly higher when using standard splitting than when using separated splitting. This difference is shown to become more severe for smaller rare event probabilities, illustrating the advantage of the separated splitting technique.

REFERENCES

Amrein, M., and H. R. Künsch. 2011. "A Variant of Importance Splitting for Rare Event Estimation: Fixed Number of Successes". *ACM Transactions on Modeling and Computer Simulation* 21 (2): Article 13.

Billinton, R., and W. Li. 1994. *Reliability assessment of electrical power systems using Monte Carlo methods*. Springer.

de Boer, P.-T. 2000. *Analysis and efficient simulation of queueing models of telecommunication systems*. Ph. D. thesis.

Dean, T., and P. Dupuis. 2009. "Splitting for rare event simulation: A large deviation approach to design and analysis". *Stochastic Processes and Their Applications* 119 (2): 562–587.

Del Moral, P., and J. Garnier. 2005. "Genealogical particle analysis of rare events". *The Annals of Applied Probability* 15 (4): 2496–2534.

Garvels, M. 2000. *The Splitting Method in Rare Event Simulation*. Ph. D. thesis, Universiteit Twente.

Garvels, M. J., J.-K. C. Van Ommeren, and D. P. Kroese. 2002. "On the importance function in splitting simulation". *European Transactions on Telecommunications* 13 (4): 363–371.

Gillespie, D. T. 1996. "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral". *Physical Review E* 54 (2): 2084–2091.

Glasserman, P., P. Heidelberger, P. Shahabuddin, and T. Zajic. 1998. "A large deviations perspective on the efficiency of multilevel splitting". *IEEE Transactions on Automatic Control* 43 (12): 1666–1679.

Glasserman, P., P. Heidelberger, P. Shahabuddin, and T. Zajic. 1999. "Multilevel Splitting for Estimating Rare Event Probabilities". *Operations Research* 47 (4): 585–600.

Grainger, J. J., and W. D. Stevenson. 1994. *Power system analysis*. McGraw-Hill.

L'Ecuyer, P., V. Demers, and B. Tuffin. 2006. "Splitting for rare-event simulation". In *Proceedings of the 2006 Winter Simulation Conference*, edited by L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, 137–148. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Rubino, G., and B. Tuffin. 2009. *Rare event simulation using Monte Carlo methods*. John Wiley & Sons.

Villén-Altamirano, M., and J. Villén-Altamirano. 2006. "On the efficiency of RESTART for multidimensional state systems". *ACM Transactions on Modeling and Computer Simulation* 16 (3): 251–279.

Wadman, W., D. Crommelin, and J. Frank. 2013. "Applying a splitting technique to estimate electrical grid reliability". In *Proceedings of the 2013 Winter Simulation Conference*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 577–588. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wang, S.-P., A. Chen, C.-W. Liu, C.-H. Chen, and J. Shortle. 2011. "Rare-event splitting simulation for analysis of power system blackouts". In *Proceedings of the 2011 IEEE Power and Energy Society General Meeting*, Paper 678. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

WANDER WADMAN is a PhD student in the Scientific Computing group at CWI Amsterdam. He received a master's degree in Applied Mathematics at Delft University of Technology. His PhD topic is the stochastic modeling of electric power systems, and his research interests include stochastic modeling of dynamical systems, power flow problems and Monte Carlo methods with a focus on rare event simulation. E-mail: w.wadman@cwi.nl, homepage: www.cwi.nl/~wadman/.

DAAN CROMMELIN is leader of the Scientific Computing research group at CWI Amsterdam and a professor at the Korteweg-de Vries Institute for Mathematics of the University of Amsterdam. He received his PhD from Utrecht University in 2003 and worked as a postdoc at New York University's Courant Institute, before joining CWI in 2006. His research interests include stochastic and computational methods for multiscale dynamical systems, rare event simulation, statistical inference, applications in atmosphereocean-climate science and energy networks. E-mail: daan.crommelin@cwi.nl, homepage: www.cwi.nl/~dtc.

JASON FRANK is professor of Numerical Analysis at Utrecht University. He received his PhD in numerical analysis from Delft University of Technology. From 2000 to 2013 we was employed by CWI Amsterdam as researcher and (from 2007) leader of the research group Computational Dynamics. From 2010 to 2013 he held the chair of Numerical Analysis and Dynamical Systems at the University of Amsterdam. His research addresses geometric numerical methods for partial differential equations, with applications in geophysical fluid dynamics and nonlinear waves. E-mail: j.e.frank@uu.nl, homepage: www.staff.science.uu.nl/~frank011/.