Rare-Event Simulation for Markov-Modulated Heavy-Tailed Random Walks

Jose Blanchet and Henry Lam

In this paper, we develop efficient rare event simulation methodology for Markov modulated heavy-tailed random walks.

Model formulation and problem setup: Consider a random walk $S_n = \sum_{k=1}^{n} X_k, n = 1, 2, \ldots$ on $\mathbb{R}$ that is modulated by a Markov chain $\{Y_n : n = 1, 2, \ldots\}$ living on a complete separable metric space $\mathcal{Y}$. In particular, $X_n = f(Y_n, J_n)$ where $\{J_n, n = 1, 2, \ldots\}$ are i.i.d. r.v.'s living in some space $\mathcal{J}$ and $f : \mathcal{Y} \times \mathcal{J} \to \mathbb{R}$ is a deterministic function. We assume that X_n given Y_n satisfies, for fixed $Y_n = y$,

$$P(X_n > u|Y_n = y) \sim C(y)L(u)\,u^{-\alpha}$$

as $u \longrightarrow \infty$ where $L(ut)/L(u) \longrightarrow 1$ as $u \longrightarrow \infty$ for each $t > 0$ (i.e. $L(\cdot)$ is slowly varying at infinity). We assume that $C(\cdot)$ is uniformly bounded over $\mathcal{Y}$.

In order to facilitate the statement of our results we impose the following simplifying assumptions on $\{Y_n : n \geq 0\}$ (we will comment on the relaxation of these assumptions momentarily).

1) There exists a reference measure $\mu(\cdot)$ defined on the Borel sets of $\mathcal{Y}$, which are denoted by $\mathcal{B}(\mathcal{Y})$, such that for each $A \in \mathcal{B}(\mathcal{Y})$,

$$P(Y_{n+1} \in A|Y_n = y) = \int_A K(y, z)\,\mu(dz).$$

2) There exists $\lambda \in (0, 1)$ and a probability measure $v(\cdot)$ defined on $\mathcal{B}(\mathcal{Y})$ with density $v(\cdot)$ with respect to $\mu(\cdot)$ satisfying

$$\int_A K(y, z)\,\mu(dz) \geq \lambda v(A) = \lambda \int_A v(z)\,\mu(dz),$$

for each $A \in \mathcal{B}(\mathcal{Y})$

In applications typically $\mathcal{Y}$ is either finite, in which case $\mu(\cdot)$ is the counting measure or $\mathbb{R}^d$, in which case $\mu(\cdot)$ is often taken to be the Lebesgue measure. It follows from 1) and 2) that $\{Y_n : n \geq 0\}$ is a regenerative process. Indeed, it turns out, given a transition from $Y_0 = y_0$ to $Y_1 = y_1$, that a regeneration occurs at time 1 with probability $\lambda v(y_1)/K(y_0, y_1)$. Moreover, a generic time T between regenerations is geometrically distributed with parameter λ and therefore there exists $R > 1$ such that

$$\sup_{y \in \mathcal{Y}} E_y[R^T] < \infty. \tag{1}$$

It also follows that $\{Y_n : n \geq 0\}$ has a stationary distribution which we denote by $\pi(\cdot)$.

The Doeblin assumption encoded by means of 1) and 2) above is not crucial. Nevertheless, it allows to succinctly provide a model formulation leading to inequality (1), which is very important in the analysis behind our algorithm and results below. The results discussed below can be extended provided a suitable regenerative structure is present in $\{Y_n : n \geq 0\}$ so that inequality (1) holds.

We let the Markov modulated random walk $\{S_n : n \geq 0\}$ possess negative steady-state drift. That is, we assume $E_\pi X_n < 0$. Define $\tau_u = \inf\{k > 0 : S_k > u\}$ to be the first passage time above u. We are interested in the efficient estimation via simulation of the ruin probability

$$P(\tau_u < \infty \mid Y_0 = y) \tag{2}$$

when u is large i.e. the event is rare.

Applications and related literature: Heavy-tailed random walks have received much attention because of its applications in areas such as insurance and queueing. For example, under the classical Cramer's model in risk theory the bankruptcy of insurance company can be formulated as the first passage of a random walk (see Asmussen (2000)), and statistical studies show that such random walk often exhibits heavy-tail features (see for example Embrechts, Klüppelberg and Mikosch (1997) and Adler, Feldman and Taqqu (1998)). In queueing theory, the maximum of heavy-tailed random walks arise in the calculation of steady-state waiting time distribution (see Asmussen (2003) and Whitt (2002)). In some insurance settings, Markov-modulated random walks arise in the case of changing economic environment (see for instance Delbaen and Haezendonck (1987), Paulsen (1993) and Paulsen and Gjessing (1997) for more on risk theory with random investment rates).

Despite its simple formulation, practically useful explicit expressions for the first passage probability (2) are analytically challenging. Large deviations theory is often employed to approximate the order of magnitude for these first passage probabilities (see Bucklew (1990) and Dembo and Zeitouni (1998) in light-tailed cases; Embrechts and Veraverbeke (1982), Mikosch and Samorodnitsky (2000) and most relevantly Foss and Zachary (2002) in heavy-tailed cases). As with any asymptotic approximation one can have a significant error in the prelimit (i.e. finite u). In addition, and very importantly, one might be interested in estimating conditional expectations given the event of ruin (such as the discounted deficit at the time of ruin). Because of this, simulation techniques, predominantly importance sampling, arise as a natural tool to sharpen the asymptotic approximations and to estimate associated conditional expectations given ruin.

In the light-tailed i.i.d. case (i.e. with no Markov modulation), Siegmund (1976) provided the first logarithmically efficient rare-event algorithm by applying exponential tilting to the walk according to the most likely path described by the associated large deviations theory. Subsequently, this idea of exponential tilting was studied in substantial generality (see Bucklew (1990), Heidelberger (1995), Asmussen et. al. (1994), Sadowsky (1996), Dieker and Mandjes (2005) and Dupuis and Wang (2004)). However, the classical rate function approach in large deviations fails for the case of heavy-tailed increments. Instead, rare events in this context are often triggered by an atypical big jump by one of the components while all the rest behave typically (see for example Borovkov and Borovkov (2001) and Rozovskii (1989, 1993)). This does not immediately suggest a rare-event algorithm, and this issue has been under much discussion in this decade. Asumssen et. al. (2000) and Bassamboo et. al. (2006) suggested some difficulties that arise in rare-event simulation of heavy-tailed systems. Asmussen and Binswanger (1997), Asmussen, Binswanger and Hojgaard (2000), Juneja and Shahabuddin (2002), Asmussen and Kroese (2006) proposed algorithms for the tail of the delay in an $M/G/1$ queue that involves regularly varying or Weibull service time, based on different importance sampling and conditional Monte Carlo ideas. Dupuis, Leder and Wang (2007) proposed the use of mixtures for the delay of an $M/G/1$ queue with regularly varying service times.

Connections to state-dependent importance sampling and Lyapunov techniques: Recently, Blanchet, Glynn and Liu (2007) further studied the mixture idea proposed by Dupuis, Leder and Wang (2007) for i.i.d. sum of regularly varying random walk. Their analysis builds on the work of Blanchet and Glynn (2008) who introduced the use of Lyapunov inequalities in order to construct a state-dependent algorithm that is provably strongly efficient for general light and heavy-tailed random walk. In this paper we extend the method of Blanchet, Glynn and Liu (2007) and Blanchet and Glynn (2008) for Markov-modulated walks. In particular, we introduce a state-dependent importance sampling mixture algorithm that exploits the regeneration property of the modulating process. At each time step, the activation of the change-of-measure is determined by the initial state of the cycle. If change-of-measure is activated, an occurrence of jump is determined by a biased coin toss whose bias depends on the current state. On the other hand, the jump sizes are adjusted according to the number of steps already taken in a regenerative cycle. This scheme is easy to implement and can be shown to be strongly efficient using Lyapunov methods (see Theorem 1 below).

One critical feature of our problem that is absent in Blanchet, Glynn and Liu (2007) and Blanchet and Glynn (2008) is the possibility of temporary (transient) positive drift of the random walk. Our proof techniques are also based on Lyapunov inequalities, which as mentioned before were introduced in Blanchet and Glynn (2008) and also applied in Blanchet, Glynn and Liu (2007). Nevertheless, because of the possibility of a transient positive drift in the random walk structure mentioned above (i.e. one could have $E_y X_i > 0$ for some y while $E_\pi X_i < 0$) the associated Lyapunov inequality cannot be tested in a single time step in our current situation. In the analysis of Blanchet and Glynn (2008) it is crucial to have a negative drift for the increment distributions throughout the course of the algorithm, so we have to overcome this situation in our current technical development.

Statement of the algorithm and main result: We state our algorithm as follows. For simplicity in the notation let us assume that X_n given $Y_n = y$ has a continuous probability density function $f^{(y)}(x)$ with respect to the Lebesgue measure. Also denote $\bar{F}^{(y)}(x) = P(X_n > x|Y_n = y)$ and $F^{(y)}(x) = P(X_n \leq x|Y_n = y)$. Let us use $P((s_{n-1}, y_{n-1}), (s_n, y_n))$ to denote the original transition kernel, in other words,

$$P((s_{n-1}, y_{n-1}), (s_n, y_n)) = f^{(y_n)}(s_n - s_{n-1})\, K(y_{n-1}, y_n).$$

Similarly, we use $Q_k^s((s_{n-1}, y_{n-1}), (s_n, y_n))$ to denote the transition kernel under the underlying change-of-measure, to be described in a moment (s and k denote the dependency of Q on the initial Markov state of the current regenerative cycle and the step number within the current cycle; see (5) below).

We construct a so-called Lyapunov function taking the following form. The term Lyapunov function comes from the fact that, as it turns out, $h(\cdot)$ satisfies a Lyapunov-type inequality which we omit for brevity in our discussion here.

$$h(s_0, s_1) = \begin{cases} c(s_1 - s_0 + d)^{2(1-\alpha)} \wedge 1 & \text{for } s_0 < s_1 \\ 1 & \text{for } s_0 \geq s_1 \end{cases} \tag{3}$$

The function $h_u(\cdot)$ will guide the algorithm and also help to rigorously prove the efficiency of the procedure (see discussion below). We select $\kappa > 0$ satisfying $c(\kappa + d)^{2(1-\alpha)} = 1$ for some parameters d and c that are selected according to constraints indicated in Steps 1 to 4 below.

Also set a sequence of positive numbers $\{a_k\}$ such that

$$a := \sum_{k=1}^{\infty} a_k \prod_{j=1}^{k-1} (1 - a_j) < 1 \quad \text{and} \quad \inf_{k \geq 1} a_k > 0 \tag{4}$$

For example we can choose some $a_1 < 1$ and $a_k, k = 2, 3, \ldots$ all equal to a constant less than a_1.

We define cycles by the regenerative times taken by Y_k. Start at the initial state $(S_0, Y_0) = (w, z)$. Denote $s(l)$ and $y(l)$ as the values of S_n and Y_n at the beginning of cycle l ($s(0) = w, y(0) = z$). Also denote T_l as the l-th regenerative time for $l \geq 1$ ($T_0 = 0$). Suppose we are currently in cycle l. At step k within the current cycle, do the following:

Case 1: $u - s(l) > \frac{\kappa}{1-a}$

If $u - S_{T_{l-1}+k-1} > \kappa$, use the change of measure

$$\begin{aligned} & Q_k^{s(l)}((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k})) \\ = \; & p(s(l)) \cdot \frac{P((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))}{\bar{F}^{(Y_{T_{l-1}+k-1})}(a_k(u - S_{T_{l-1}+k-1}))} I(S_{T_{l-1}+k} > a_k(u - S_{T_{l-1}+k-1})) + \\ & (1 - p(s(l))) \cdot \frac{P((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))}{F^{(Y_{T_{l-1}+k-1})}(a_k(u - S_{T_{l-1}+k-1}))} I(S_{T_{l-1}+k} \leq a_k(u - S_{T_{l-1}+k-1})) \end{aligned} \quad (5)$$

Here a_k satisfy (4) and

$$p(w) = \frac{\theta}{u - w + r}$$

where θ and r are properly tuned parameters (see Steps 1 to 4 below).

If $u - S_{k-1} \leq \kappa$, set

$$Q_k^{s(l)}((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k})) = P((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))$$

Case 2: $\kappa < u - s(l) \leq \frac{\kappa}{1-a}$

Proceed the same as Case 1 but now $\{a_k\}$ are replaced by $\{a_k(s(l))\}_{k \geq 1}$ chosen so that (4) is still satisfied and

$$(1 - a(s(l)))(u - s(l)) > \kappa$$

We use the same θ and r.

Case 3: $u - s(l) \leq \kappa$

No importance sampling is applied for the whole cycle i.e.

$$Q_k^{s(l)}((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k})) = P((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))$$

Moreover, in all these cases, the following stopping rule applies:

Case 1: If the next regenerative time T_{l+1} is hit, we enter a new cycle.

Case 2: If first passage occurs, stop the algorithm.

Case 3: If a jump occurs i.e. the time

$$\tilde{\tau}_l := \inf\left\{k \geq 0 : u - S_{T_{l-1}+k-1} > \kappa,\ S_{T_{l-1}+k} > a_k(u - S_{T_{l-1}+k-1})\right\}$$

is hit, deactivate the importance sampler and continue with the original kernel.

If first passage occurs, output the likelihood ratio

$$L = \prod_{l=1}^{N} \prod_{k=1}^{T_l \wedge \tau_u} \frac{P((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))}{Q_k^{s(l)}((S_{T_{l-1}+k-1}, Y_{T_{l-1}+k-1}), (S_{T_{l-1}+k}, Y_{T_{l-1}+k}))} \tag{6}$$

where N is the number of cycles enacted.

Parameter Selection:

1. Pick any $d > 0$. Say $d = 1$.
2. Set $r = (1-a)d$
3. Define

$$A(d,a) = 2(\alpha - 1) \inf_{x \geq 0} \frac{((1-a)x + d)^{1-2\alpha}}{(x+d)^{1-2\alpha}}$$
$$B(d,a) = (1-a) \sup_{v \geq 0} \frac{\sup_{x>v, z \in \mathcal{Y}} \bar{F}^{(z)}((1-a)x)(x+d) \ \ \sup_{z \in \mathcal{Y}} \bar{F}^{(z)}((1-a)v)}{(v+d)^{1-2\alpha}}$$

and

$$\tilde{A}(d) = \inf_{a(w): u-w > \kappa} A(d, a(w))$$
$$\tilde{B}(d) = \sup_{a(w): u-w > \kappa} B(d, a(w))$$

Choose θ small enough and c large enough so that

$$\left(1 - \frac{\theta}{r}\right)^{-1} \leq R$$

and

$$\frac{\tilde{B}(d)}{\theta c} + \frac{\theta R \sup_{y \in \mathcal{Y}} E_y[T(y) R^{T(y)}]}{1-a} + \tilde{A}(d) \sup_{y \in \mathcal{Y}} E_y[S_{T(y)}] \leq 0$$

are satisfied.

4. Pick κ satisfying

$$c(\kappa + d)^{2(1-\alpha)} = 1$$

We demonstrate through examples in the paper how to set these parameters. The algorithm as well as the guidance in choosing the parameters is set in such a way that strong efficiency is maintained i.e. the coefficient of variation (ratio of standard deviation to mean) of the estimator remains bounded as $u \nearrow \infty$. Our main theoretical result is summarized in the next theorem.

Theorem 1. *The function $h : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ defined in (3) satisfies that*

$$\textit{2nd Moment of the Estimator } L \textit{ defined in (6)} \leq h(0,u)$$

Moreover, we have

$$\limsup_{u \to \infty} \frac{h(0,u)}{P(\tau_u < \infty | Y_0 = y)^2} < \infty.$$

for every $y \in \mathcal{Y}$. Therefore, the estimator given by the algorithm above is strongly efficient.

References

1. Adler, J., Feldman, R., and Taqqu, M., (1998) *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Birkhauser, Boston.

2. Asmussen, S. (2000), *Ruin Probabilities*, World Scientific.

3. Asmussen, S., Binswanger, K. (1997), Simulation of ruin probabilities for subexponential claims, *Ast. Bulletin*, **27**(2), 297–318.

4. Asmussen, S., Binswanger, K., and Hojgaard, B. (2000), Rare event simulation for heavy-tailed distributions. *Bernoulli* **6**(2), 303-322.

5. Asmussen, S. and Kroese, D. P. (2006), Improved algorithms for rare event simulation with heavy tails, *Adv. Appl. Prob.* **38**(2).

6. Asmussen, S., Rubinstein, R. Y., and Wang, C. L. (1994), Regenerative rare event simulation via likelihood ratios, *J. Appl. Prob.* **31**(3), 797-815.

7. Blanchet, J. and Glynn, P. (2008), Efficient rare-event simulation for the maximum of a random walk with heavy-tailed increments, *Ann. Appl. Prob.*, **18**4, 1351-1378.

8. Blanchet, J., Glynn, P. and Liu, J. C. (2007), Fluid heuristics, Lyapunov bounds and efficient importance sampling for a heavy-tailed G/G/1 queue, *QUESTA*, **57**(2-3), 99–113.

9. Bassamboo, A., Juneja, S. and Zeevi, A. (2006), On the efficiency loss of state-independent importance sampling in the presence of heavy-tails, *Oper. Res. Lett.*, **35**(2), 521–531.

10. Borovkov, A. A. and Borovkov, K. A. (2001), On probabilities of large deviations for random walks. I. Regularly varying distribution tails. *Theory Prob. Appl.* **49**(2), 189-205.

11. Bucklew, J. (1990), *Large Deviation Techniques in Decision, Simulation, and Estimation*, Wiley, New York.

12. Delbaen, F. and Haezendonck, J. (1987), Classical risk theory in an economic environment, *Insurance Math. Econ.*, **6**, 85-116.

13. Dembo, A. and Zeitouni O. (1998) *Large Deviation Techniques and Applications*, 2nd ed., Springer, New York.

14. Dieker, A. B. and Mandjes, M. (2005), On asymptotically efficient simulation of large deviation probabilities, *Adv. in Appl. Probab.* **37**(2), 539–552.

15. Dupuis, P. and Wang, H. (2004), Importance sampling, large deviations, and differential games, *Stoc. Stoc. Reports*, **76**(6), 481-508.

16. Dupuis, P., Leder, K., and Wang, H. (2007), Importance sampling for sums of random variables with regularly varying tails, *ACM TOMACS*, **17**(3), Article 14.

17. Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, New York.

18. Embrechts, P. and Veraverbeke, N. (1982), Estimates for the probability of ruin with special emphasis on the possibility of large claims, *Insurance Math. Econ.*, **1**, 55-72.

19. Foss, S., Zachary, S. (2002), Asymptotics for the maximum of a modulated random walk with heavy-tailed increments, *Analytic Methods in Applied Probability (in memory of Fridrih Karpelevich)*, AMS Translations, Series 2, 207, 37-52.

20. Heidelberger, P. (1995), Fast simulation of rare events in queueing and reliability models, *ACM TOMACS*, **5**(1), 43-85.

21. Juneja, S. and Shahabuddin, P. (2002), Simulating heavy-tailed processes using delayed hazard rate twisting, *ACM TOMACS* **12**(2), 94-118.

22. Mikosch, T. and Samorodnitsky, G. (2000), The supremum of a negative drift random walk with dependent heavy-tailed steps, *Ann. Appl. Prob.*, **10**(3), 1025-1064.

23. Paulsen, J. (1993), Risk theory in a stochastic economic environment, *Stoc. Proc. Appl.*, **46**(2), 327-361.

24. Paulsen, J. and Gjessing, H. K. (1997), Ruin Theory with Stochastic Return on Investments, *Adv. Appl. Prob.*, **29**(4), 965-985.

25. Rozovskii, L. V. (1989), Probabilities of large deviations of sums of independent random variables with common distribution function in the domain of attraction of the normal law, *Theory Prob. Appl.* **34**(4), 625-644.

26. Rozovskii, L. V. (1993), Probabilities of large deviations on the whole axis, *Theory Prob. Appl.*, **38**(1), 53-79.

27. Sadowsky, J. S. (1996), On Monte Carlo estimation of large deviations probabilities, *Ann. Appl. Prob.* **6**(2), 399-422.

28. Whitt, W. (2002), *Stochastic-Process Limits*, Springer.

29. Siegmund, D. (1976), Importance sampling in the Monte Carlo study of sequential tests, *Ann. Statist.*, **4**, 673-684.