Analysis of Asymptotic Relative Variance for Adaptive Multilevel Splitting Method

Qiming DU

1 Notations

In this section, we recall the notations used for the following calculations. We notice that the notations are mainly from [1] and [2]. In the framework of multilevel spitting method, we note p the probability $\mathbb{P}\big(S(X) > L_*\big)$. For adaptive multilevel splitting method, we decompose p such that, $p = \alpha^{n_0} r$, where $n_0 = \big\lfloor \frac{\log(p)}{\log(\alpha)} \big\rfloor$, $\alpha \in (0,1)$ and $r \in (\alpha, 1)$. Here, to simplify the application of Feynman-Kac model, we will let $n = n_0 + 1$. For all the $k = 0, 1, 2, \ldots, n_0 - 1$, we note $p_k = \alpha$ and $p_{n_0} = p_{n-1} = r$.
Therefore, the sequence of levels are noted as follow:

$$-\infty = L_{-1} < L_0 < \cdots < L_{n-1} < L_n = L_*$$

All the notations in the Feynman-Kac model are introduced from [3]. Particularly, we recall the definition of unnormed prediction (page 59 Definition 2.3.2) [3] which is useful for our further calculations.

$$\gamma_n(1) = \prod_{p=0}^{n-1} \eta_p(G_p) \qquad \text{with} \quad G_p = \mathbb{1}_{\mathcal{A}_{p+1}} \text{ where } \mathcal{A}_p := \big\{x \in \mathbb{R}^d : S(x) > L_p\big\}$$

By convention, we set $\gamma_0 = \eta_0$.

2 Proof for Lower Bound

Adopting the notations pf Proposition 9.4.1 on page 301 of [3], for the same purpose of A.2 of [2], we will have:

$$\mathrm{Var}\Big(\sqrt{N}\big(\hat{p} - p\big)\Big) \xrightarrow[N\to\infty]{\mathrm{d}} \mathbb{E}\big(W_n^\gamma\big) = \gamma_n^2(1) \sum_{k=0}^{n} \eta_k \Big(\big(\frac{Q_{k,n}(1)}{\eta_k Q_{k,n}(1)} - 1 \big)^2 \Big)$$

We remark that $Q_{n,n}$ in the last term of the sum is identity in the Feynman-Kac semi-group (page 88 Definition 2.7.2 [3]), i.e. $Q_{n,n}(x, \mathrm{d}y) = \delta_x(\mathrm{d}y)$.

$$\eta_n\Big(\big(\frac{Q_{n,n}(1)}{\eta_n Q_{n,n}(1)} - 1\big)^2\Big) = 0$$

In our context this can be written as follows:

$$\sigma^2 = \sum_{k=0}^{n-1} \eta_k\Big(\big(\frac{Q_{k,n}(1)}{\eta_k Q_{k,n}(1)} - 1\big)^2\Big) \tag{1}$$

We remark that when $k = n$, the term under the sum of (1) equals 0. By the simple calculation, we could also have an alternative version of σ^2:

$$\sigma^2 = \sum_{k=0}^{n-1} \left(\frac{\eta_k\big(Q_{k,n}(1)^2\big)}{\big(\eta_k Q_{k,n}(1)\big)^2} - 1 \right) \tag{2}$$

Now, for every $k = 0, 1, \ldots, n-1$ and every probability $\eta(\mathrm{d}x)$, we consider the following Boltzmann-Gibbs transformation:

$$\Psi_{G_k}(\eta(\mathrm{d}x))(\mathrm{d}x') = \frac{G_k(x)\eta(\mathrm{d}x)}{\eta(G_k)}$$

Since $G_k(x) \cdot Q_{k,n}(1)(x) = Q_{k,n}(1)(x)$ holds for every $k = 0, 1, \ldots, n-1$ and for every $x \in \mathbb{R}^{\mathrm{d}}$, one can readily check that

$$\eta_k Q_k, n(1) = \eta_k(G_k) \cdot \Psi_{G_k}(\eta_k)\big(Q_{k,n}(1)\big) = p_k \cdot \eta_{k+1} Q_{k,n}(1)$$

$$\eta_k\big(Q_k, n(1)^2\big) = \eta_k(G_k) \cdot \Psi_{G_k}(\eta_k)\big(Q_{k,n}(1)^2\big) = p_k \cdot \eta_{k+1}\big(Q_{k,n}(1)^2\big)$$

and

$$\eta_{k+1}\big(Q_{k,n}(1)\big) = \eta_{k+1}\big(Q_{k+1,n}(1)\big)$$

Therefore, we have

$$\frac{\eta_k\big(Q_{k,n}(1)^2\big)}{\big(\eta_k Q_{k,n}(1)\big)^2} = \frac{1}{p_k} \cdot \frac{\eta_{k+1}\big(Q_{k,n}(1)^2\big)}{\big(\eta_{k+1} Q_{k+1,n}(1)\big)^2}$$

We combine these to the equations (1) and (2) to have the form we desired.

$$\begin{aligned}
\sigma^2 &= \sum_{k=0}^{n-1} \frac{1-p_k}{p_k} + \sum_{k=0}^{n-1} \frac{1}{p_k} \left(\frac{\eta_{k+1}\big(Q_{k,n}(1)^2\big)}{\big(\eta_{k+1} Q_{k+1,n}(1)\big)^2} - 1 \right) \\
&= \sum_{k=0}^{n-1} \frac{1-p_k}{p_k} + \sum_{k=0}^{n-1} \frac{1}{p_k} \eta_{k+1}\Big(\big(\frac{Q_{k,n}(1)}{\eta_{k+1} Q_{k+1,n}(1)} - 1\big)^2\Big) \\
&\geq \sum_{k=0}^{n-1} \frac{1-p_k}{p_k}
\end{aligned}$$

3 An Upper Bound

As is shown on above, we failed to decompose the σ^2 recursively. Therefore, to obtain the inequality in the form:

$$\eta_k\big(Q_{k-1,n}(1)^2\big) \leq \mathcal{C} \cdot \eta_k\big(Q_{k,n}(1)^2\big)$$

where $\mathcal{C}$ is some positive constant. We will make some assumptions over the mixing properties of transition kernels M_k. One possible assumption is that, for any $k \in \{0, 1, \cdots, n-1\}$ and any test function f, we have:

$$\forall x \in \mathbb{R}^{\mathrm{d}} \qquad M_k(f)(x) \leq C_k \cdot f(x)$$

In our context, f is a density function if $M_k(x, \mathrm{d}y)$ is absolutely continuous. To have more intuition of the choice of C_k, let us check an easy example. We consider a discrete markov transition where M_k can be represented by a matrix and function f is a vector. We supose that all the eigenvalues of M_k are real. In this case, we could choose C_k as the largest eigenvalue of matrix M_k.
In order to the calculations, we will introduce some new notations:

$$E_k := \frac{\eta_k\big(Q_{k,n}(1)^2\big)}{\big(\eta_k Q_{k,n}(1)\big)^2} \qquad \tilde{E}_k := \frac{\eta_k\big(Q_{k-1,n}(1)^2\big)}{\big(\eta_k Q_{k,n}(1)\big)^2}$$

We have seen in the last section that:

$$E_k = \frac{1}{p_k}\tilde{E}_{k+1}$$

By the assumption above, we have also:

$$\tilde{E}_{k+1} \leq C_k^2 \cdot E_{k+1}$$

Then, by combining the two equations above, we finally obtain an inequality:

$$E_k \leq \beta_k \cdot E_{k+1} \qquad \text{where } \beta_k = \frac{C_k^2}{p_k}$$

It's important to remark that we set all the $E_k = 1$ and $\beta_k = 1$ for $k \geq n$.

$$\begin{aligned}
\sigma^2 &= \sum_{k=0}^{n-1}\big(E_k - 1\big) \\
&\leq -n + \sum_{k=0}^{n-1} \beta_k \cdot E_{k+1} \\
&\vdots \\
&\leq -n + \sum_{k=0}^{n-1} \beta_k \cdot \beta_{k+1} \cdots \beta_{k+(n-1)} \cdot 1
\end{aligned}$$

If we could choose a constant C, such that $C_k = C$ for every $k = 0, 1, \ldots, n-1$, we will obtain the upper bound below:

$$\sigma^2 \leq \frac{C^2}{r} \cdot \left(\frac{1 - \big(\frac{C^2}{\alpha}\big)^n}{1 - \big(\frac{C^2}{\alpha}\big)}\right) - n$$

We could see that the quality of the upper bound are guaranteed by the choice of constant C_k. And it is clear that there are some bound for C, because we have seen that there exists a lower bound of σ^2.

References

[1] Frederic Cerou and Arnaud Guyader. Fluctuation analysis of adaptive multilevel splitting. *arXiv preprint arXiv:1408.6366*, 2014.

[2] F. Cérou, P. Del Moral, T. Furon, and A. Guyader. Sequential monte carlo for rare event estimation. *Statistics and Computing*, 22(3):795–808, 2012.

[3] Pierre Del Moral. *Feynman-Kac formulae: genealogical and interacting particle systems with applications.* Series: Probability & Applications Springer Verlag, 2004.