arXiv:2301.02882v2 [math.NA] 4 Sep 2023

面向不连续函数的 MLMC 技术

Michael B. Giles

摘要 多层蒙特卡洛（MLMC）方法通常适用于估计某个量的期望值，前提是该量是中间量的利普希茨函数；但若其为不连续函数，MLMC 校正项的方差可能会衰减得慢得多。本文回顾了在多种不同情境下可用于克服这一挑战的技术，并探讨了近期利用分支扩散或自适应抽样取得的进展。

1 引言

多层蒙特卡洛（MLMC）方法基于以下望远镜求和

$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$$

其中，$\widehat{P}_\ell$ 表示层级 l 上感兴趣输出量 P 的一个近似，其弱误差为 $\left|\mathbb{E}[\widehat{P}_\ell - P]\right|$，MLMC 方差为 $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$；随着层级 l 增加，两者均会减小，但相应的单样本计算成本会增加。

If $\widehat{Y}_\ell$ 的期望值为 $\mathbb{E}[Y_\ell]$，其方差为 V_ℓ，成本为 C_ℓ，那么，对于给定的目标均方根误差 ε，可以优化层级数 L 以及每个层级上的独立样本数 [13, 14]，从而得到总成本其近似等于 $\varepsilon^{-2}\left(\sum_{l=0}^{L} \sqrt{C_\ell V_\ell}\right)^2$。

Michael B. Giles 牛津大学数学研究所，Woodstock Rd, 牛津 OX2 6GG，英国 电子邮件：mike.giles@maths.ox.ac.uk

如果 $C_\ell V_\ell \to 0$ 如$l \to \infty$，那么成本主要由第0层的第一项决定，成本约为 $\varepsilon^{-2} C_0 V_0$，因此与 ε^{-2}成正比。

如果 $C_\ell V_\ell \to \text{const}$ 随$l \to \infty$变化，那么各层级的成本贡献几乎平均分布，成本约为 $\varepsilon^{-2} L^2 C_L V_L$，与 $\varepsilon^{-2} |\log \varepsilon|^2$ 成正比，前提是 $\mathbb{E}[\widehat{P}_\ell - P]$ 随l呈指数衰减。

更糟的是，如果 $C_\ell V_\ell \to \infty$ 随$l \to \infty$变化，那么成本主要由最细层级的贡献决定，因此约为 $\varepsilon^{-2} C_L V_L$；若 $\mathbb{E}[\widehat{P}_\ell - P] \sim 2^{-\alpha \ell}$、$V_\ell \sim 2^{-\beta \ell}$ 和 $C_\ell \sim 2^{\gamma \ell}$，则为 $O(\varepsilon^{-2-(\gamma-\beta)/\alpha})$。

在大多数多层蒙特卡洛应用中，P 是某些中间解量的光滑函数，或至少是利普希茨函数，例如随机微分方程的解、具有随机系数或随机初始/边界数据的偏微分方程的解，或内层条件期望的估计值。在这些情况下，我们通常有 $\beta \geq \gamma$，因此多层蒙特卡洛的复杂度为 $O(\varepsilon^{-2})$ 或 $O(\varepsilon^{-2} |\log \varepsilon|^2)$。

本文关注的是一小类但十分重要的应用，其中P 是不连续函数；因此，MLMC 方差 V_ℓ 的衰减速度可能会慢得多，导致复杂度落入第三类，即 $O(\varepsilon^{-2-(\gamma-\beta)/\alpha})$。

好消息是，多层蒙特卡罗社区内部已经开展了大量研究 [1] 来应对这一挑战。本文综述了已经开发出的各种方法，希望能够帮助研究人员在未来的应用中应对类似挑战。

为说明这些问题，我们首先详细介绍两个具体的应用挑战，正是它们推动了大量相关研究。随后，我们将讨论已经提出的众多方法，其中一些方法借鉴了计算灵敏度的相关文献中的思想（在数理金融文献中称为“希腊值”），其形式为 $\frac{\partial}{\partial \alpha} \mathbb{E}[f(\omega, \alpha)]$，所使用的是路径灵敏度方法 [25]（也称为无穷小扰动分析，简称 IPA [30]）或似然比方法[31], ; 另一些方法则借鉴了通过提高被积函数的光滑性来改善准蒙特卡洛积分收敛速度的方法 [1, 6]。

1.1 挑战 1：嵌套期望

假设 f 是一个标量函数，我们要估计嵌套期望 $\mathbb{E}[f(\mathbb{E}[Z|X])]$, , 其中外层期望是针对随机变量 X 求取的，并假设内层条件期望 $\mathbb{E}[Z|X]$ 在零附近具有有界密度。

一种非常简单的多层蒙特卡洛处理方法 [2] 在第l层使用 $M_\ell = 2^\ell M_0$ 个内层样本，因此第0层和更高层的估计量分别为

$$\widehat{Y}_0 = f(\overline{Z}^{(0,M_0)}), \quad \widehat{Y}_\ell = f(\overline{Z}^{(\ell,M_\ell)}) - f(\overline{Z}^{(\ell,M_{\ell-1})}),$$

[1] 参见 https://people.maths.ox.ac.uk/gilesm/mlmc community.html[2] 注意，如果 f 是光滑的，或至少是利普希茨连续的，那么最好使用“反对称”估计量 [8, 14, 15, 18], ; 但当 f 不连续时，这并不能提高收敛阶。

其中，$\overline{Z}_{(l,M_\ell)}$和 $\overline{Z}_{(l,M_{\ell-1})}$分别表示 M_ℓ和 $M_{\ell-1}$个相互独立的 Z样本的均值，且这些样本均以同一个 X [14, 15]值为条件。

如果 $\mathbb{V}[Z|X]$ 是有限的，且 f 是常数为 L_f 的利普希茨函数，那么

$$\mathbb{E}\left[\left(f(\overline{Z}^{(\ell,M_\ell)}) - f\left(\mathbb{E}[Z|X]\right)\right)^2 \mid X\right] \le L_f^2\,\mathbb{E}\left[\left(\overline{Z}^{(\ell,M_\ell)} - \mathbb{E}[Z|X]\right)^2 \mid X\right] = L_f^2 M_\ell^{-1}\mathbb{V}[Z|X],$$

因此 $\mathbb{E}[\widehat{Y}^2_l \,|X] \le 2\,L_f^2(M_\ell^{-1} + M_{\ell-1}^{-1})\mathbb{V}[Z|X]$ 对于 l>0 成立。如果 $\mathbb{V}[Z|X]$ 一致有界，则可得 $V_\ell = O(M^{-1}l)$。如果每个 Z 条件样本的成本为 $O(1)$，那么 $C_\ell = O(M_\ell)$，因此复杂度为 $O(\varepsilon^{-2}|\log\ \varepsilon|^2)$。

遗憾的是，当 f 是由 $H(x)$=0（若 x<0）和 $H(x)$=1（若 x≥0）定义的 Heaviside阶跃函数 H 时，情况就差得多了。这种情况在许多应用中都会出现，因为 $\mathbb{P}\left[\mathbb{E}[Z|X] > K\right] = \mathbb{E}\left[H(\mathbb{E}[Z|X] - K)\right]$,，因此它对应于条件期望超过某个阈值 K 的概率，而这在风险计算中是一个非常重要的量。

若 K=0和 $E[Z|X]$在零附近具有有界密度，则 $|E[Z|X]| = O(M^{-1/2}\,l)$ 的概率为 $O(M^{-1/2}\,l)$，在这种情况下，由于 $Z_{(l,M_\ell)}$为正且 $Z_{(l,M_{\ell-1})}$为负，或反之，发生这种情况的概率为 $O(1)$。因此，$V_\ell \approx O(M^{-1/2}\,l)$，复杂度约为$O(\varepsilon^{-5/2})$ [19]。

这一挑战是第7节的主要动机，也出现在第3节的背景中。

1.2 挑战 2：不连续支付函数

对于标量随机微分方程

$$\mathrm{d}S_t = a(S_t)\,\mathrm{d}t + b(S_t)\,\mathrm{d}W_t, \tag{1}$$

其输出关注量为 $P \equiv f(S_T)$，标准估计量为

$$\widehat{Y}_\ell = \widehat{P}_\ell - \widehat{P}_{\ell-1}$$

$\widehat{\ }$ $\widehat{\ }$其中使用相同的布朗运动 W_t 计算 P_ℓ 和 $P_{\ell-1}$，但采用不同的均匀时间步长 h_ℓ 和 $h_{\ell-1}$。

如果 f是常数为 L_f 的利普希茨函数，那么

$$V_\ell \;\le\; \mathbb{E}\left[(\widehat{P}_\ell - \widehat{P}_{\ell-1})^2\right] \;\le\; L_f^2\,\mathbb{E}\left[(\widehat{S}_\ell - \widehat{S}_{\ell-1})^2\right]$$

$\widehat{\ }$其中 S_ℓ 是对 S_T的第 l 层数值近似。因此，根据标准的强收敛结果 [28]，我们有 $V_\ell = O(h_\ell)$，这适用于欧拉-丸山离

散化的随机微分方程；对于一阶 Milstein 离散化，则有 $V_\ell = O(h_\ell^2)$。两种情况下，成本 C_ℓ 均为 $O(h_\ell^{-1})$，因此多层蒙特卡洛的复杂度分别为 $O(\varepsilon^{-2}|\log\ \varepsilon|^2)$ 和 $O(\varepsilon^{-2})$，

在数理金融中，数字看涨期权的收益为 0 或 1，具体取决于 S_T 是否低于或高于执行价 K，因此收益函数可写为 $f(S_T) = H(S_T - K)$。多层蒙特卡洛的问题在于，如果粗路径与细路径跨越执行价，也就是说它们位于执行价的不同一侧，那么两条路径之间即使只有很小的差异，也可能导致收益相差 ± 1。

使用随机微分方程的粗略近似（详细分析见 [4, 22]），$\widehat{S}_\ell$ 细路径/粗路径配对中有 $O(h^{1/2}\ \iota)$ 的比例会跨越执行价，因此 $V_\ell = O(h^{1/2}\ \iota)$，复杂度由此为 $O(\varepsilon^{-5/2})$。

类似地，使用 Milstein 近似可得 $\widehat{S}_\ell - \widehat{S}_{\ell-1} = O(h_\ell)$，因此 $V_\ell = O(h_\ell)$。这显然有所改善，复杂度为 $O(\varepsilon^{-2}|\log\ \varepsilon|^2)$，但仍存在一个问题：大多数多层蒙特卡洛样本 Y_ℓ 在较细层级上为零，因此峰度为 $O(h_\ell^{-1})$，这会导致实际估算 V_ℓ 时出现问题，进而难以确定层级 l 上应使用的样本数 N_ℓ。此外，多维随机微分方程的 Milstein 离散化通常需要模拟 L´évy 面积，这也带来困难，不过可以通过使用反向估计量 [24] 来解决。

这一挑战是第 2、4、5 和 6 节的主要动机，在第 3 和第 7 节中也会出现。

2 显式平滑

在数理金融中，通过路径敏感度分析（或路径导数法，IPA）计算称为希腊值的参数敏感度时，要求收益函数 f 连续且分段光滑。数字期权显然存在这一问题，一种标准做法是将 Heaviside 阶跃函数 H 替换为平滑近似 $H_\delta(x) \equiv g(x/\delta)$，其中 $g(x) \to 0$ 为 $x \to -\infty$，而 $g(x) \to 1$ 为 $x \to +\infty$，从而将不连续性平滑到宽度为 $O(\delta)$ 的区间内。

出于金融方面的考虑，人们通常倾向于采用单侧平滑，例如图 1 中黄色所示的分段线性近似。这种单侧近似会引入弱误差，或称偏差，其为 $O(\delta)$。若将其用于多层蒙特卡洛，则对于最终落入斜坡区域的 $O(\delta)$ 比例的路径，$H'_\delta(S_T) = \delta^{-1}$，因此 $V_\ell = O(\delta \times (\delta^{-1})^2) = O(\delta^{-1})$。所以，$\delta$ 的最优取值需要在偏差与方差之间权衡。

通过使平滑关于 $x = 0$ 呈反对称，可以减小偏差，使得 $H_\delta(x) - H(x) = -(H_\delta(-x) - H(-x))$，例如选择 $g(x) \equiv \Phi(x)$，如图 1 中橙色所示。若 S_T 具有光滑概率密度 $\rho(S)$，则弱误差为

$$\int_{-\infty}^{\infty} \left(H_\delta(S-K) - H(S-K)\right) \rho(S)\, \mathrm{d}S = \delta \int_{-\infty}^{\infty} \left(g(x) - H(x)\right)\, \rho(K+x\delta)\, \mathrm{d}x$$

对 $\rho(K+x\delta)$ 进行泰勒级数展开，可得渐近误差展开

$$a_1 \rho(K)\, \delta + a_2 \rho'(K)\, \delta^2 + a_3 \rho''(K)\, \delta^3 + a_4 \rho'''(K)\, \delta^4 + O(\delta^5)$$

其中

$$a_k = \int_{-\infty}^{\infty} x^{k-1} \left(g(x) - H(x)\right)\, \mathrm{d}x.$$

如果 $g(x) - H(x) = -\left(g(-x) - H(-x)\right)$，那么 $a_1 = a_3 = 0$，并且

$$a_2 = 2 \int_0^{\infty} x(g(x) - 1)\, \mathrm{d}x, \quad a_4 = 2 \int_0^{\infty} x^3 (g(x) - 1)\, \mathrm{d}x.$$

如果 $g(x)$ 是单调的，那么 $a_2 \neq 0$；但通过考虑诸如 $g(x) = (4/3)\,\Phi(x) - (1/3)\,\Phi(2x)$ 这样的非单调函数，可以设定 $a_2 = 0$，使弱误差为 $O(\delta^4)$。因此，为了使整体精度达到 $O(\varepsilon)$，我们需要 $\delta = O(\varepsilon^{1/4})$；而在最粗层级上，$V_\ell = O(\delta^{-1}) = O(\varepsilon^{-1/4})$，所以在整体成本由最粗层级的成本主导的最佳情况下，总体复杂度约为 $O(\varepsilon^{-2-1/4})$。

Giles、Nagapetyan 和 Ritter [23] 采用显式平滑来估计累积分布函数（CDF）。对于标量随机变量 X，为了估计 $C(x) = \mathbb{P}(X < x) = \mathbb{E}[H(x-X)]$，他们采用的方法是使用多层蒙特卡洛估计一组样条点 x_j 处的 $C(x_j)$，然后用三次样条对这些值进行插值。总体而言，除了多层蒙特卡洛采样误差，他们的方法还平衡了三种弱误差：最细层级上的随机微分方程离散化误差、由 H_δ 导致的平滑误差，以及三次样条插值误差。

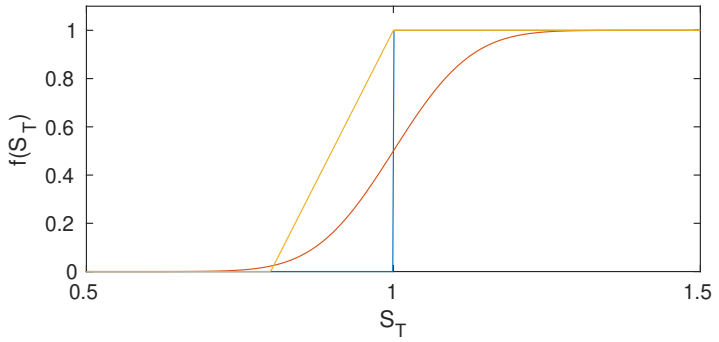


图 1 数字看涨期权的 Heaviside 阶跃函数的两个显式平滑版本

3 积分/微分与 Malliavin 演算

Krumscheid & Nobile [29]采用了略有不同的方法来估计累积分布函数，尤其是在风险估计方面。以恒等式

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathbb{E}\left[\,\max(0,x-S_T)\,\right] = \mathbb{E}[\,H(x-S_T)\,]$$

他们使用多层蒙特卡洛估计 $\mathbb{E}[\max(0,\ x_j-S_T)\ ]$对应一组样条点 x_j 的值，再用三次样条对这些值进行插值，然后对样条求导，从而近似得到累积分布函数 $C(x)$。这样可以避免因平滑 Heaviside函数而引入额外的弱误差，但对三次样条求导会放大样条数据中的噪声。

类似地，Altmayer 和 Neuenkirch [2]使用 Malliavin 演算中的分部积分，处理基于 Heston随机波动率随机微分方程解的不连续收益。他们观察到，从渐近意义上看，这种方法降低了较细层级上的多层蒙特卡洛方差，但会增加粗层级上的方差。为此，他们将收益拆分为两部分：一部分是平滑部分，采用标准多层蒙特卡洛方法处理；另一部分是紧支撑不连续部分，则采用 Malliavin 多层蒙特卡洛方法处理。

Malliavin演算最初是为计算灵敏度而发展起来的，因此这是利用灵敏度计算相关文献来开发改进的多层蒙特卡罗算法的另一个例子。

4 条件期望

使用随机微分方程的一阶 Milstein离散化时，改进数字期权的 MLMC 方差的一种方法是将最后一个时间步改为欧拉-丸山近似，然后对最后一条细路径的布朗增量 ΔW [12, 17] 求条件期望。

对于标量随机微分方程 (1) 的细路径近似，其时间步长为h_ℓ，最终时刻 T 的路径值 S_T 为

$$\widehat{S}_T^f = \widehat{S}_{T-h_\ell}^f + a(\widehat{S}_{T-h_\ell}^f)\,h_\ell + b(\widehat{S}_{T-h_\ell}^f)\,\Delta W_N,$$

因此，数字看涨期权的条件期望值为

$$\widehat{P}_\ell^f \;=\; \mathbb{E}\left[H(\widehat{S}_T^f-K)\mid \widehat{S}_{T-h_\ell}^f\right] \;=\; \Phi\left(\frac{\widehat{S}_{T-h_\ell}^f + a(\widehat{S}_{T-h_\ell}^f)\,h_\ell - K}{b(\widehat{S}_{T-h_\ell}^f)\,\sqrt{h_\ell}}\right).$$

类似地，对于时间步长为 $h_{\ell-1} = 2h_\ell$ 的粗路径，最后一个粗时间步的布朗增量是细路径最后两个布朗增量之和，ΔW_{N-1}+ΔW_N，因此

$$\widehat{S}_T^c = \widehat{S}_{T-h_{\ell-1}}^c + a(\widehat{S}_{T-h_{\ell-1}}^c)\, h_{\ell-1} + b(\widehat{S}_{T-h_{\ell-1}}^c)\ (\Delta W_{N-1}+\Delta W_N)\,,$$

由此可得

$$\begin{aligned}\widehat{P}_{\ell-1}^c &= \mathbb{E}\left[H(\widehat{S}_T^c-K) \mid \widehat{S}_{T-h_{\ell-1}}^c, \Delta W_{N-1}\right] \\ &= \Phi\left(\frac{\widehat{S}_{T-h_{\ell-1}}^c + a(\widehat{S}_{T-h_{\ell-1}}^c)\, h_{\ell-1} + b(\widehat{S}_{T-h_{\ell-1}}^c)\, \Delta W_{N-1} - K}{b(\widehat{S}_{T-h_{\ell-1}}^c)\, \sqrt{h_\ell}}\right).\end{aligned}$$

对于 $\widehat{Y}_\ell \equiv \widehat{P}_\ell - \widehat{P}_{\ell-1}$，数值分析 [17] 证明了 $V_\ell \approx O(h^{3/2} l)$，因此多层蒙特卡洛的复杂度为 $O(\varepsilon^{-2})$。从直观上看，这是因为路径落在执行价 K 的 $O(h^{1/2} l)$ 范围内的概率为 $O(h^{1/2}\ l)$，而对于这些

$$\widehat{S}_{T-h_{\ell-1}}^f - \widehat{S}_{T-h_{\ell-1}}^c = O(h_\ell), \quad \frac{\partial \widehat{P}}{\partial \widehat{S}} = O(h_\ell^{-1/2}) \quad \Longrightarrow \quad \widehat{P}_\ell - \widehat{P}_{\ell-1} = O(h_\ell^{1/2}),$$

因此 $V_\ell \approx O(h^{1/2}\ l \times (h^{1/2}\ l)^2) = O(h^{3/2}\ l)$。此外，峰度改进为$O(h^{-1/2}\ l)$。遗憾的是，当对整个路径使用 Euler-Maruyama离散化时，这种方法并无帮助，因为 $\widehat{S}_{T_h_{\ell-1}}^f - \widehat{S}_{T-h_{\ell-1}}^c = O(h^{1/2} l)$，因此$\frown \quad \frown P_\ell - P_{\ell-1} = O(1)$

这种条件期望的用法是一种标准技术，可用于平滑收益，从而进行 IPA/路径灵敏度计算 [25]。另一个例子是敲出障碍期权：如果路径跌破某个特定值，期权就会被敲出。在这种情况下，可以根据离散时间步上计算得到的路径近似，计算发生这种情况的概率，从而平滑收益 [25]。同样，在使用一阶 Milstein离散化时，这种方法对多层蒙特卡洛效果很好 [12, 17],，但对欧拉-丸山离散化没有帮助。

Achtsis、Cools 和 Nuyens [1]，以及 Bayer、Siebenmorgen 和 Tempone [6] 引入了另一种条件期望平滑方法，以提高准蒙特卡洛计算的收敛性；随后，Bayer、Ben Hammouda 和 Tempone [5] 使用该方法降低数字期权的多层蒙特卡洛方差。

在最简单的形式中，他们将数值模拟的随机输入拆分为一个标量 Z 和其余部分 Z_r，并将所需的多层蒙特卡洛层级 l 期望表示为

$$\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = \mathbb{E}\left[\ \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \mid Z_r]\ \right]$$

并且请注意，在许多金融应用中，可以采用一种方式进行这种拆分 $\widehat{\quad}$ $\widehat{\quad}$，使条件期望 $\mathbb{E}[P_\ell \mid Z_r]$，$\mathbb{E}[P_{\ell-1} \mid Z_r]$成为关于 Z_r 的光滑函数；当 Z 仅存在一个不连续点时，还可以通过解析方法或一维数值求积非常精确地计算这些函数。

对于标量随机微分方程，Z 可以是驱动布朗运动的终值，此时 Z_r 表示构造布朗桥所需的其他正态随机变量，用于生成布朗增量。

5 测度变换

另一种使用 Milstein 离散化处理数字期权的方法是采用测度变换 [9, 14], ，这种方法与用于敏感性分析 [31]的似然比方法（LRM）有关。

对于细路径和粗路径，我们都有关于 $\sim S_T$ 的条件高斯分布，其均值和方差略有不同，如图 2 所示。因此，我们可以进行测度变换，将它们变换为均值为 μ、方差为 σ^2 的同一个高斯分布，如图 2 所示；随后，从这个共同的高斯分布中为两条路径选取相同的样本 $\widehat{S}_T$。

由此得到的多层蒙特卡洛估计量为

$$\widehat{Y}_\ell = f(\widehat{S}_T)\,(R_\ell - R_{\ell-1})$$

其中 R_ℓ, $R_{\ell-1}$ 分别是细路径和粗路径对应的拉东-尼科迪姆导数。对于标量随机微分方程 (1)， R_ℓ 为

$$R_\ell = \frac{\sigma}{b(\widehat{S}^f_{T-h_\ell})\sqrt{h_\ell}} \; \frac{\exp\left(-(\widehat{S}_T - \widehat{S}^f_{T-h_\ell} - a(\widehat{S}^f_{T-h_\ell})\,h_\ell)^2 \,/\, (2\,b^2(\widehat{S}^f_{T-h_\ell})\,h_\ell)\right)}{\exp\left(-(\widehat{S}_T-\mu)^2 \,/\, (2\sigma^2)\right)}$$

并且 $R_{\ell-1}$ 的定义类似。可以证明，差值 $R_\ell - R_{\ell-1}$ 约为 $O(h^{1/2}l)$，这意味着 $V_\ell \approx O(h_\ell)$。为降低方差，我们注意到拉东-尼科迪姆导数的条件期望始终为 1，即 $\mathbb{E}[R_\ell \mid \widehat{S}^f_T{}_{h_\ell}] = \mathbb{E}[R_{\ell-1} \mid \widehat{S}^c_{T-h_{\ell-1}}, \Delta W_{N-1}] = 1$，因此我们可以更改 $\widehat{Y}_\ell$ 的定义，使其为

$$\widehat{Y}_\ell \;=\; \Big(f(\widehat{S}_T) - f(\mu)\Big)\,(R_\ell - R_{\ell-1})$$

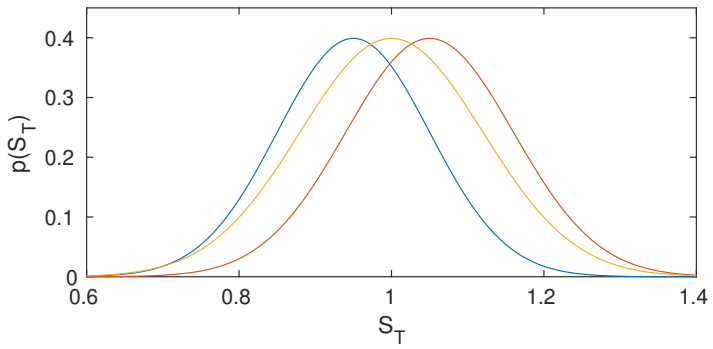


图 2 粗路径和细路径的条件高斯分布，以及第三个共同分布

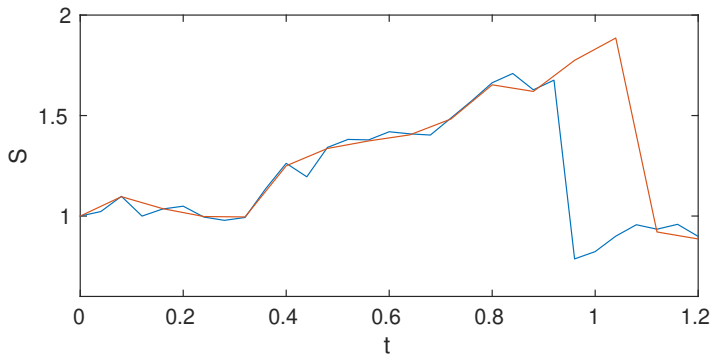


图 3 粗、细跳扩散路径在 $T = 1$之前和之后发生跳跃的示意图。

$\widehat{}$，而不改变其期望值。现在，只有当 S_T 和 μ 位于执行价 K 的两侧时，该估计量才非零；这种情况在粗路径/细路径中所占的比例为 $O(h^{1/2}l)$。因此，新的 MLMC 方差 V_ℓ 约为 $O(h^{3/2}l)$，与使用解析条件期望时一样。

这种方法的优点是，在许多情况下无法计算解析条件期望 [9, 14]时，它仍能在多维情形下发挥良好效果。不过，同样地，对于完整路径的欧拉-丸山离散化，它并无助益，因为这会得到 $R_\ell - R_{\ell-1} = O(1)$。

Xia[33, 34] 较早地在一次 MLMC 计算中，对具有路径依赖跳跃率λ的 Merton 型跳扩散随机微分方程$S,\ t$使用了测度变换。正如图 3 所示，这一应用的难点在于，粗路径和细路径会在不同时间发生跳跃；其中一条路径可能恰好在最终时刻 T之前跳跃，而另一条则可能在之后跳跃，从而导致计算值 $f(S_T)$之间出现较大差异。路径依赖跳跃率通过 Glasserman 和 Merener [26],的稀疏化技术处理：以均匀速率 $\lambda_{sup} > \lambda(S,\ t)$对可能的跳跃时刻进行过采样，然后使用接受／拒绝步骤筛选出实际跳跃时刻。Xia 通过测度变换对此进行了改进，确保细路径和粗路径作出相同的接受／拒绝判断，从而在同一时刻发生跳跃。这样便得到如下形式的估计量

$$\widehat{Y}_\ell = \widehat{P}_\ell R_\ell - \widehat{P}_{\ell-1} R_{\ell-1}.$$

与跳跃时刻之间对随机微分方程采用一阶 Milstein离散化相结合后，对于标准看跌或看涨期权等 Lipschitz 支付函数，可得到 $V_\ell = O(h_\ell^2)$ [33, 34]。

6 分裂

回到由随机微分方程解所产生的数字期权这一难题，第三种方法是使用路径分裂，对第 4 [14] 节中引入的条件期望进行无偏估计。

这是一般分裂技术 [3] 的一种变体。如图4所示，该方法先进行标准的细路径模拟，直到最终时刻 T之前一个时间步，然后对最后一个时间步进行多次独立模拟，并对每次模拟的收益取平均，以近似条件期望。粗路径也采用相同的做法，只不过每次分裂都使用与细路径倒数第二个时间步相同的 ΔW_{N-1}。

由于路径在分裂时刻之前的计算成本为$O\left(h_\ell^{-1}\right)$，因此可以使用多达 $O\left(h_\ell^{-1}\right)$ 次分裂，而不会显著增加路径成本。如果使用M_ℓ 次分裂，那么标准的分裂方差分析得出

$$\mathbb{V}[\widehat{Y}_\ell] = \mathbb{V}\left[\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \mid \{\Delta W_n\}_{n<N}]\right] + M_\ell^{-1}\mathbb{E}\left[\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \mid \{\Delta W_n\}_{n<N}]\right].$$

如前所述，$\mathbb{V}\left[\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \mid \{\Delta W_n\}_{n<N}]\right] = O\left(h^{3/2} l\right)$，同理可以证明 $\mathbb{E}\left[\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1} \mid \{\Delta W_n\}_{n<N}]\right] = O\left(h_\ell\right)$。因此，选择使 M_ℓ介于 $O\left(h_\ell^{-1}\right)$ 和 $O\left(h^{-1/2} l\right)$ 之间，即可在不显著增加每个样本计算成本的情况下，获得分裂带来的益处。

此外，还可以使用精度更高的 Milstein离散化来处理最后一个时间步，而不必切换到欧拉-丸山离散化。Burgos [9, 10] 详细介绍了相关分析，并且对各种金融期权进行了路径敏感度分析，也采用了相同的方法。

Giles和Bernal [16] 也将分裂用于停时扩散过程所产生的Feynman-Kac泛函，即当解路径离开指定区域时终止的随机微分方程模拟。这里的问题是，当细路径离开时，存在一个 $O(h^{1/2} l)$

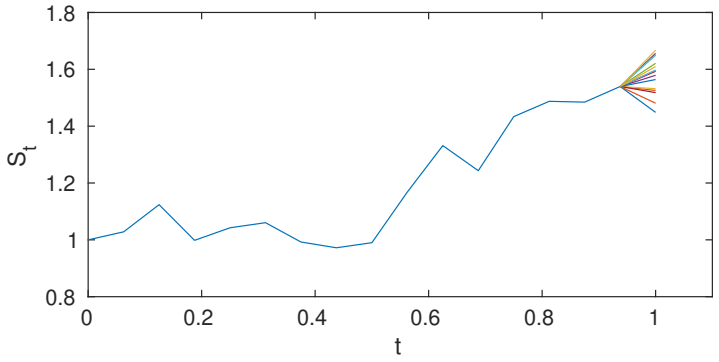


图 4 在最后一个时间步进行路径分裂以估计条件期望

的概率，使得相应的粗路径直到晚得多才离开。为解决这一问题，通过将粗路径分裂为$O(h^{-1/2}\,l)$个相互独立的子模拟来估计条件期望；每个子模拟都会持续进行，直到离开定义域。V_ℓ从 $O(h^{1/2}l)$提升至约 $O(h_\ell)$，而每个样本的成本并未显著增加，最终实现的多层蒙特卡洛复杂度为$O(\varepsilon^{-2}|\log\ \varepsilon|^3)$。

迄今介绍的三种方法（条件期望、测度变换、分裂）在使用 Euler-Maruyama 离散化时都无济于事。为此，Giles 和 Haji-Ali 最近开发了一种新方法 [20]。

该方法同样采用分裂，但受到分支扩散模拟的启发，会在多个确定性时刻进行分裂，如图 5 所示。图中展示了一组分裂路径的逻辑结构。这里考虑的是在单位时间区间上的模拟。一对细路径/粗路径被计算至时间$t\ =\ 1/2$，此时细时间步的数量为 $1\ 2h_\ell^{-1}$。随后，该模拟被分裂为两个相互独立的模拟，分别运行至时间 $t\ =\ 3/4$；这两个模拟合计增加了 $1\ 2h_\ell^{-1}$个细时间步。之后还会在 $t = 3/4$、 $t = 7/8$ 等时刻进一步分裂，直至只剩一个粗时间步时进行最后一次分裂。

模拟的细时间步总数为 $O(h_\ell^{-1}|\log h_\ell|)$，因此与原方法仅使用一对细路径/粗路径相比，计算成本只略有增加。Y_ℓ 定义为所有最终路径对应数值 $P_\ell - P_{\ell-1}$ 的平均值，并且可以证明，其方差为 $O(h_\ell)$，与利普希茨支付函数 [20]具有相同的渐近收敛阶。峰度也有所改善，因此，这项技术充分解决了使用欧拉-丸山离散化的多层蒙特卡洛来估计数字期权价值这一难题。

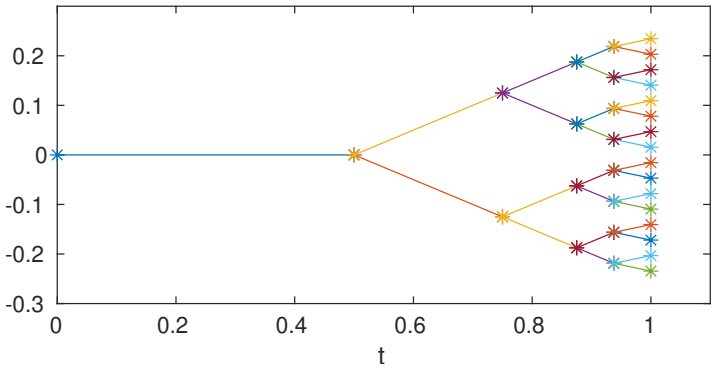


图 5 利用重复路径分裂估计条件期望

7 自适应抽样

现在我们回到估计嵌套期望 $\mathbb{E}\left[\, H(\mathbb{E}[Z|X]) \,\right]$这一难题，并注意到，只有当内层条件期望 $\mathbb{E}[Z|X]$ 接近零时，我们才需要对其进行准确估计。这一观察是 Broadie、Du 和 Moallemi [7] 在标准蒙特卡洛过程中开发自适应抽样方法的基础。随后，Giles 和 Haji-Ali [19] 将其扩展为结合多层蒙特卡洛的自适应抽样方法，定义层级l 上的内层样本数 M_ℓ 为

- $M_\ell = 2^\ell M_0$ inner samples when $|\mathbb{E}[Z|X]| \gg \sqrt{\mathbb{V}[Z|X]/(2^\ell M_0)}$
  这是层级l上使用的最小样本数。 $\sqrt{\mathbb{V}[Z|X]/(2^\ell M_0)}$ 是 $\mathbb{E}[Z|X]$ 的蒙特卡罗估计的标准差，因此这个不等式意味着，该样本数足以使我们非常确信估计的符号正确。
- $M_l = 4_l M_0$内层样本，当 $|\mathbb{E}[Z|X]| = O(\sqrt{\mathbb{V}[Z|X]/(4_l M_0)})$
  这是层级l上使用的最大样本数。在这种情况下， $\mathbb{E}[Z|X]$ 的估计可能会出现符号错误，但这种情况只会在 $|\mathbb{E}[Z|X]| = O(2^{-\ell})$ 时发生，其概率为 $O(2^{-\ell})$。同样，在该区域使用更多样本所增加的总成本为 $O(2^{-\ell} \times 4^\ell) = O(2^\ell)$，因此不会显著增加总体平均成本。
- $2_l M_0 < M_l < 4_l M_0$ 中间值
  在此区域中，选择的样本数量足以确保对 $\mathbb{E}[Z|X]$ 的估计符号正确，同时总成本为 $O(2^\ell)$。

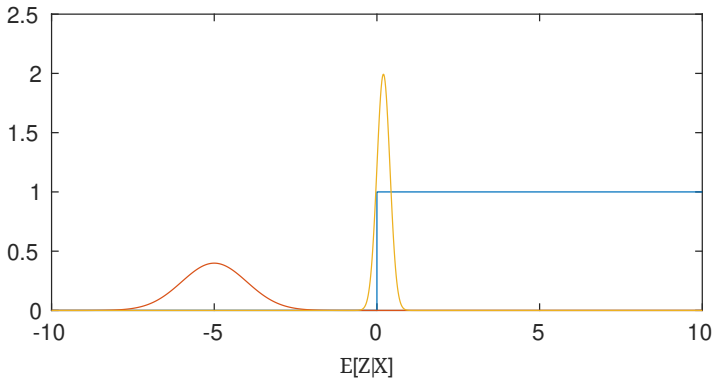


图 6 两个条件期望的误差分布：i）只需少量样本即可确保符号正确（左），以及 ii）即使使用大量样本也无法确保符号正确（中）。蓝线表示海维赛德阶跃函数。

总体而言，这种自适应抽样方法带来了 $C_\ell \sim 2^\ell$、 $V_\ell \sim 2^{-\ell}$，因此复杂度约为 $O(\varepsilon^{-2})$ [19]。不过，峰度为 $O(2^\ell)$，因为只有 $O(2^{-\ell})$ 比例的外层样本对 $\widehat{Y}_\ell$ 产生非零值。

Haji-Ali、Spence 和 Teckentrup [27] 进一步将其推广，用于估计以下形式的量：

$$\mathbb{P}[G \in \Omega] \equiv \mathbb{E}[\mathbf{1}_{G\in\Omega}]$$

其中 G 是一个无法直接抽样的 d 维随机变量。在他们的论文中，尤其考虑了本文中的两个挑战。对于单位时间区间上采用 Euler-Maruyama 离散化的数字期权，自适应抽样会调整层级l 所用的时间步，使得

- $h_l = 2^{-l}$ 当 $|\widehat{S}_l - K|$ 相较于路径近似的强误差较大时
- $h_l = 4^{-l}$ 当 $|\widehat{S}_l - K|$ 与强误差处于同一量级时
- $2^{-l} < h_l < 4^{-l}$ 对于中间取值

当需要将时间步从初始值$h_\ell = 2^{-\ell}$细化以进行自适应过程时，会使用布朗桥构造。自适应再次得出 $C_\ell \sim 2^\ell$、 $V_\ell \sim 2^{-\ell}$，因此复杂度约为 $O(\varepsilon^{-2})$，但峰度仍然很高 [27]。

在早期研究中，Elfverson、Hellman 和 Malqvist [11] 考虑了对$\mathbb{E}[H(X)]$ 的估计，其中 X 无法精确采样，但存在一系列精度和成本逐渐增加的近似 X'_0， X'_1， X'_2， ..... X。受偏微分方程应用的启发，其截断误差表现良好，因此$|X'_j - X|$具有一致的几何界；他们的方法在层级l 上使用

$$\widehat{X}_\ell = X'_j, \quad j = \min\{\ell, \min j : |\widehat{X}'_j - X| < |X|\}$$

并取得了同样良好的多层蒙特卡洛效益。这一思路与 Haji-Ali等人的研究基本相同，但要求 $|X'_j - X|$满足一致界，比 Haji-Ali等人针对某些 $q>2$所要求的 $\mathbb{E}[\,|X'_j - X|^q]$界限严格得多。

最后补充一点，Haji-Ali、Spence 和 Teckentrup 的分析可以推广到示性函数与利普希茨函数的乘积， $\mathbb{E}[\mathbf{1}_{G\in\Omega} f(S)]$,因此也能处理障碍期权。此外，Haji-Ali 和 Spence 将自适应抽样方法拓展到了数理金融中出现的极具挑战性的三重嵌套期望 [21]。他们采用 Rhee 和 Glynn 的随机化多层蒙特卡洛处理方法 [32] ，既处理底层随机微分方程的时间离散化，也处理内层条件期望的抽样，从而实现了约为 $O(\varepsilon^{-2})$ 的总体复杂度。对于如此困难的应用而言，这一结果非常出色。

8 结论

值得再次强调的是，在大多数多层蒙特卡洛应用中，所关注的输出量是中间模拟量的利普希茨函数，因此，中间量良好的强收敛会自动带来 MLMC 方差的良好收敛速率 V_ℓ。

对于函数不连续的应用，本文指出，已有大量文献提出了各种不同方法，以降低 MLMC 方差，并尝试恢复最优的 $O(\varepsilon^{-2})$ 复杂度。值得注意的是，其中许多方法借鉴了蒙特卡洛敏感性分析中的思路，而不连续泛函同样会给这类分析带来问题。希望这篇综述能为今后在其他新兴应用领域面临类似挑战的研究人员提供帮助。

致谢 本文基于我与众多学生、博士后及其他合作者开展的研究，在此向他们表示感谢。研究经费由英国工程与自然科学研究委员会通过 EP/E031455/1、EP/H05183X/1 和 EP/P020720/2 项目提供，此外还获得了香港创新科技署（InnoHK 项目 CIMDA）的资助。本文是在访问德克萨斯大学奥斯汀分校奥登计算工程与科学研究所期间撰写的，感谢接待我的各位给予热情款待。

参考文献

1. N. Achtsis、R. Cools 和 D. Nuyens。LT 方法下障碍期权定价的条件抽样。*SIAM* 金融数学杂志，4:327–352（2013年） 2. M. Altmayer 和 A. Neuenkirch。利用 Malliavin 分部积分，对广义 Heston 模型中的不连续收益进行多层蒙特卡洛求积。*SIAM* 金融数学杂志，6(1):22–52（2015年） 3. S. Asmussen 和 P.W. Glynn。随机模拟。Springer，纽约（2007年） 4. R. Avikainen。随机微分方程的不规则泛函与欧拉格式。《金融与随机》，13(3):381–401（2009年） 5. C. Bayer、C. Ben Hammouda 和 R. Tempone。多层蒙特卡洛结合数值平滑，实现稳健、高效的期权定价与密度估计。*arXiv* 预印本：*2003.05708*（2020年） 6. C. Bayer、M. Siebenmorgen 和 R. Tempone。对收益进行平滑以高效计算篮子期权价格。《定量金融》，18(3):491–505（2018年） 7. M. Broadie、Y. Du 和 C.C. Moallemi。通过嵌套序贯模拟高效估计风险。《管理科学》，57(6):1172–1194（2011年） 8. K. Bujok、B. Hambly 和 C. Reisinger。伯努利随机变量泛函的多层模拟及其在篮子信用衍生品中的应用。《应用概率的方法与计算》，17(3):579–604（2015年） 9. S. Burgos。利用多层蒙特卡洛计算希腊值。哲学博士论文，牛津大学（2014年） https://ora.ox.ac.uk/objects/uuid:6453a93b-9daf-4bfe-8c77-9cd6802f77dd 10. S. Burgos 和 M.B. Giles。利用多层路径模拟计算希腊值。载于 L. Plaskota 和 H. Wo´zniakowski 编，《2010年蒙特卡洛与准蒙特卡洛方法》，第 281–296 页。Springer（2012年） 11. D. Elfverson、F. Hellman 和 A. M˚alqvist。一种计算失效概率的多层蒙特卡洛方法。*SIAM/ASA* 不确定性量化杂志，4(1):312–330（2016年）

12. M.B. Giles。利用 Milstein 格式改进多层蒙特卡洛收敛性。载于 A. Keller、S. Heinrich 和 H. Niederreiter 编，《2006年蒙特卡罗与准蒙特卡罗方法》，第 343–358 页。Springer（2008年） 13. M.B. Giles。多层蒙特卡洛路径模拟。《运筹学》，56(3):607–617（2008年） 14. M.B. Giles。多层蒙特卡洛方法。《数值学报》，24:259–328（2015年） 15. M.B. Giles。嵌套期望的多层蒙特卡洛。载于《当代计算数学——庆祝 Ian Sloan 80 岁生日》。Springer（2018年） 16. M.B. Giles 和 F. Bernal。停时扩散过程的期望退出时间及其他泛函的多层估计。*SIAM/ASA* 不确定性量化杂志，6(4):1454–1474（2018年） 17. M.B. Giles、K. Debrabant 和 A. R¨oßler。利用 Milstein 离散化分析多层蒙特卡洛路径模拟。《离散与连续动力系统：*B*辑》，24(8):3881–3903（2019年） 18. M.B. Giles 和 T. Goda。不确定性下的决策：使用多层蒙特卡洛高效估计期望完美信息部分价值。《统计与计算》，29(4):739–751（2019年） 19. M.B. Giles 和 A.-L. Haji-Ali。用于高效风险估计的多层嵌套模拟。*SIAM/ASA* 不确定性量化杂志，7(2):497–525（2019年） 20. M.B. Giles 和 A.-L. Haji-Ali。数字期权的多层路径分支。*arXiv* 预印本：*2209.03017*（2022年） 21. M.B. Giles、A.-L. Haji-Ali 和 J. Spence。信用估值调整的高效风险估计。*arXiv* 预印本：*2301.05886*（2023年） 22. M.B. Giles、D.J. Higham 和 X. Mao。分析非全局 Lipschitz 收益期权的多层蒙特卡洛。《金融与随机》，13(3):403–413（2009年） 23. M.B. Giles、T. Nagapetyan 和 K. Ritter。分布函数与密度函数的多层蒙特卡洛逼近。*SIAM/ASA* 不确定性量化杂志，3(1):267–295（2015年） 24. M.B. Giles 和 L. Szpruch。无需模拟 L´evy 面积的多维随机微分方程反变量多层蒙特卡洛估计。《应用概率年刊》，24(4):1585–1620（2014年） 25. P. Glasserman。《金融工程中的蒙特卡罗方法》。Springer，纽约（2004年） 26. P. Glasserman 和 N. Merener。状态依赖强度跳跃扩散过程离散化格式的收敛性。《伦敦皇家学会会刊 *A* 辑》，460:111–127（2004年） 27. A.-L. Haji-Ali、J. Spence 和 A. Teckentrup。概率的自适应多层蒙特卡洛。《*SIAM* 数值分析杂志》，60(4):2125–2149（2022年） 28. P.E. Kloeden 和 E. Platen。《随机微分方程的数值解》。Springer，柏林（1992年） 29. S. Krumscheid 和 F. Nobile。函数的多层蒙特卡洛逼近。*SIAM/ASA* 不确定性量化杂志，6(3):1256–1293（2018年） 30. P. L’Ecuyer。路径导数法、分数法和似然比梯度估计技术的统一观点。《管理科学》，36(11):1364–1383（1990年） 31. P. L’Ecuyer。似然比导数估计量中导数与期望的交换问题。《管理科学》，41(4):738–748（1995年） 32. C.-H. Rhee 和 P.W. Glynn。随机微分方程模型的无偏估计与平方根收敛。《运筹学》，63(5):1026–1043（2015年） 33. Y. Xia。跳跃过程的多层蒙特卡洛。哲学博士论文，牛津大学（2014年） https://ora.ox.ac.uk/objects/uuid:7bc8e98a-0216-4551-a1f3-1b318e514ee8/ 34. Y. Xia 和 M.B. Giles。跳跃扩散随机微分方程的多层路径模拟。载于 L. Plaskota 和 H. Wo´zniakowski 编，《2010年蒙特卡洛与准蒙特卡洛方法》，第 695–708 页。Springer（2012年）