Central Limit Theorems for Multilevel Monte Carlo Methods

Håkon Hoel and Sebastian Krumscheid

Keywords. Multilevel Monte Carlo, Central Limit Theorem

Abstract. In this work, we show that uniform integrability is not a necessary condition for central limit theorems (CLT) to hold for normalized multilevel Monte Carlo estimators, and we provide near optimal weaker conditions under which the CLT is achieved. In particular, if the variance decay rate dominates the computational cost rate (i.e., $\beta > \gamma$), we prove that the CLT always holds.

1. Introduction

The multilevel Monte Carlo (MLMC) method is a hierarchical sampling method which in many settings improves the computational efficiency of weak approximations by orders of magnitude. The method was independently introduced in the papers [1, 2] for the purpose of parametric integration and for approximations of observables of stochastic differential equations, respectively. MLMC methods have since been applied with considerable success in a vast range of stochastic problems, a collection of which can be found in the overview [3]. In this work we present near optimal conditions under which the normalized MLMC estimator converges in distribution to a standard normal distribution. Our result has applications in settings where the MLMC approximation error is measured in terms of probability of failure rather than the classical mean square error.

1.1. Main result. We consider the probability space $(\Omega, \mathcal{F}, P)$ and let $X \in L^2(\Omega)$ be a scalar random variable for which we seek the expectation $E[X]$, and let $(X_\ell)_{\ell=1}^{\infty} \subset L^2(\Omega)$ be a sequence of random variables satisfying the following:

Assumption 1.1. There exist positive rate constants $\alpha, \beta, \gamma$ with $\min(\beta, \gamma) \leq 2\alpha$ and a positive constant $c_\alpha > 0$ such that for all $\ell \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$

(i) $|E[X - X_\ell]| \leq c_\alpha e^{-\alpha \ell}$,
(ii) $V_\ell := \text{Var}(\Delta_\ell X) = O(\ell^{-\beta})$,
(iii) $C_\ell := \text{Cost}(\Delta_\ell X) = \Theta(\ell^{\gamma})$,

where $\Delta_\ell X := X_\ell - X_{\ell-1}$ with $X_{-1} := 0$. Here, $f(x_\ell) = O(y_\ell)$ means that there exists a constant $C > 0$ such that $|f(x_\ell)| < C |y_\ell|$ for all $\ell \in \mathbb{N}_0$, while $f(x_\ell) = \Theta(y_\ell)$ means that there exist constants $C > c > 0$, such that $c |y_\ell| < |f(x_\ell)| < C |y_\ell|$ for all $\ell \in \mathbb{N}_0$.

Definition 1.1 (MLMC estimator $A_{ML}$). The MLMC estimator $A_{ML}: (0, \infty) \rightarrow L^2(\mathcal{F}, \mathbb{P})$ applied to estimate the expectation of $X \in L^2(\Omega)$ based on the collection of random variables (r.v.) $(X_\ell) \subset L^2(\Omega)$ satisfying Assumption 1.1 is defined by

$$A_{ML}(\epsilon) = \sum_{\ell=0}^{L(\epsilon)} \sum_{i=1}^{M(\epsilon)} \Delta_\ell X^{i}_{\ell}$$
Here
\[ L^2(\Omega) \ni \Delta_{\ell}X^i = X^i_{\ell} - X^i_{\ell-1}, \quad \ell \in \mathbb{N}_0, \ i \in \mathbb{N} \]
denotes a sequence of independent r.v. and every subsequence \{\Delta_{\ell}X^i\}_i\) consist of independent and identically distributed (i.i.d.) r.v., the number of levels is
\[ L(\epsilon) := \max \left( \left\lceil \frac{\log(c_0\epsilon^{-1})}{\alpha} \right\rceil, 1 \right) \quad \epsilon > 0, \]
and the number of samples per level \(\ell = 0, 1, \ldots\) is
\[ M_{\ell}(\epsilon) := \max \left( \epsilon^2 \sqrt{\sum_{\ell=0}^{L(\epsilon)} C_{\ell}V_{\ell}}, 1 \right) \quad \epsilon > 0. \]

We will refer to
\[ \frac{A_{\text{ML}}(\epsilon) - \mathbb{E}[X_{L(\epsilon)}]}{\sqrt{\text{Var}(A_{\text{ML}}(\epsilon))}} \]
as the normalized MLMC estimator.

**Notation and conventions.** When confusion is not possible, we will use the following shorthands,
\[ A_{\text{ML}} := A_{\text{ML}}(\epsilon), \ M_{\ell} := M_{\ell}(\epsilon), \ L := L(\epsilon). \]

The following conventions will be employed throughout
\[ 0 \cdot (\pm \infty) = 0 \quad \text{and} \quad 0/0 = 0, \]
and we define the monotonically increasing sequence
\[ S_k := \sum_{\ell=0}^{k} \sqrt{V_{\ell}C_{\ell}}, \quad k \in \mathbb{N}_0. \]

Then the main result of this work can be stated as follows.

**Theorem 1.1** (Main result). Let \(A_{\text{ML}}\) denote the MLMC estimator applied estimate the expectation of \(X \in L^2(\Omega)\) based on the collection of r.v. \(\{X_\ell\} \subset L^2(\Omega)\) satisfying Assumption 1.1. Suppose that \(V_0 > 0\) and further that

(i) if \(\beta = \gamma\), then \(\lim_{k \to \infty} S_k = \infty\) and

\[
\lim_{\ell \to \infty} \mathbf{1}(V_\ell > 0) \mathbb{E} \left[ \frac{\left| \Delta_{\ell}X - \mathbb{E}[\Delta_{\ell}X] \right|^2}{V_\ell} \mathbf{1}\left\{ \frac{|\Delta_{\ell}X - \mathbb{E}[\Delta_{\ell}X]|^2}{\nu S_{\ell}^2 \exp((2\alpha - \gamma)\ell)} > \nu S_{\ell}^2 \exp((2\alpha - \gamma)\ell) \right\} \right] = 0 \quad \forall \nu > 0,
\]

(ii) if \(\gamma > \beta\), then assume that \(\beta < 2\alpha\), equality (3) holds and that there exists an \(\nu \in [\beta, 2\alpha)\) such that \(\lim_{k \to \infty} S_k e^{(\nu - \gamma)k/2} > 1\).

Then the normalized MLMC estimator satisfies the central limit theorem (CLT), in the sense that
\[
\frac{A_{\text{ML}} - \mathbb{E}[X_L]}{\sqrt{\text{Var}(A_{\text{ML}})}} \xrightarrow{d} \mathcal{N}(0, 1) \quad \text{as} \quad \epsilon \downarrow 0.
\]

The main result follows from Theorems 2.5 and 2.6 below. We note that Theorem 1.1 implies that in settings with \(\beta > \gamma\) the (CLT) always holds for the normalized MLMC estimator.

**Remark 1.1.** For the setting \(\gamma > \beta\) and \(\beta = 2\alpha\), which we have not included in Theorem 1.1 one cannot impose reasonable assumptions to exclude \(M_\ell = \Theta(1)\) and \(V_\ell/\text{Var}(A_{\text{ML}}) = \Theta(1)\); cf. Example 2.1. This implies there are no reasonable ways to exclude cases for which a non-negligible contribution to the variance of the resulting MLMC estimator derives from a finite number of samples. Therefore, the central limit theorem is not relevant for this setting.
In literature, the CLT has been proved for the MLMC method through assuming (or verifying for the particular sequence of r.v. considered) either a Lyapunov condition \[4\], or uniform integrability \[5, 6, 7\], or a weaker higher moment decay rate \[8\] for the sequence \[\{1_{\{|\Delta_\ell X - E[\Delta_\ell X]|^2/V_\ell|^2 > x\}}\}_{\ell \in \mathbb{N}_0}\]. To show that this work extends the existing literature, we now provide an explicit example where Theorem 1.1 is valid although uniform integrability does not hold.

**Example 1.1.** Let \(\{X_\ell\}_{\ell=-1}^\infty\) denote the sequence of r.v. defined by

\[
X_\ell = \begin{cases} 
0, & \ell = -1, \\
\sum_{k=0}^\ell e^{k/4}1_{(\omega \in \Omega_k)}, & \ell \in \mathbb{N}_0,
\end{cases}
\]

where \(\Omega_k \in \mathcal{F}\) and \(\mathbb{P}(\Omega_k) = (1 - e^{-1})e^{-k}\) for \(k \in \mathbb{N}_0\) and \(\Omega_j \cap \Omega_k = \emptyset\) for all \(j \neq k\). Let further

\[
X = \sum_{k=0}^\infty e^{k/4}1_{(\omega \in \Omega_k)}.
\]

Then

\[
|E[X - X_\ell]| = \Theta_\ell(e^{-3\ell/4})
\]

and

\[
V_\ell = E \left[ \left( e^{\ell/4}1_{(\Omega_\ell)} - e^{-\ell/4}(1 - e^{-1}) \right)^2 \right] = \Theta_\ell(e^{-\ell/2}),
\]

yielding the respective decay rates \(\alpha = 3/4\) and \(\beta = 1/2\), cf. Assumption 1.1. Moreover, for any \(x > 1\)

\[
\lim_{\ell \to \infty} E \left[ \frac{|\Delta_\ell X - E[\Delta_\ell X]|^2}{V_\ell} \{ |\Delta_\ell X - E[\Delta_\ell X]|^2 > x \} \right] = 1,
\]

which implies that the sequence

\[
\frac{|\Delta_\ell X - E[\Delta_\ell X]|^2}{V_\ell}, \quad \ell = 0, 1, \ldots
\]

is not uniformly integrable. As \(V_0 > e^{-1}(1 - e^{-1}) > 0\), \(\beta < 2\alpha\), and

\[
S_k = \begin{cases} 
\Theta_k(1) & \text{if } \beta > \gamma, \\
\Theta_k(k) & \text{if } \beta = \gamma, \\
\Theta_k(e^{(\gamma-\beta)k/2}) & \text{if } \beta < \gamma,
\end{cases}
\]

the CLT \(4\) of Theorem 1.1 holds for all settings with \(\gamma \leq 1/2 = \beta\).

**2. Theory**

In this section we derive weak assumptions under which the normalized MLMC estimator \((A_{ML} - E[X_L])/\sqrt{\text{Var}(A_{ML})}\) converges in distribution to a standard normal as \(\epsilon \to 0\). The main tool used for verifying the CLT will be the Lindeberg condition, which in its classical formulation is an integrability condition for triangular arrays of independent random variables (r.v.) \(Y_{nm}\), with \(n \in \mathbb{N}\) and \(1 \leq m \leq n\); cf. \[9\]. However, in the multilevel setting it is more convenient to work with generalized triangular arrays of independent r.v. of the form \(Y_{nm}\), which for a fixed \(\epsilon > 0\) take possible non-zero elements within the set of indices \(1 \leq m \leq n(\epsilon)\), where \(n: (0, \infty) \to \mathbb{N}\) is a strictly decreasing function of \(\epsilon > 0\) with \(\lim_{\epsilon \to 0} n = \infty\).

The following theorem is a trivial extension of \[10\] from triangular arrays to generalized triangular arrays.
Theorem 2.1 (Lindeberg-Feller Theorem). For every $\epsilon > 0$, let $\{Y_{\epsilon m}\}, 1 \leq m \leq n(\epsilon)$ with $n: (0, \infty) \to \mathbb{N}$ and $\lim_{\epsilon \downarrow 0} n = \infty$ be a generalized triangular array of independent random variables that are centered and normalized, so that

\begin{equation}
E[Y_{\epsilon m}] = 0 \quad \text{and} \quad \sum_{m=1}^{n(\epsilon)} E[Y_{\epsilon m}^2] = 1 ,
\end{equation}

respectively. Then, the Lindeberg condition:

\begin{equation}
\lim_{\epsilon \downarrow 0} \sum_{m=1}^{n(\epsilon)} E\left[Y_{\epsilon m}^2 \mathbf{1}_{\{|Y_{\epsilon m}| > \nu\}}\right] = 0 \quad \forall \nu > 0 ,
\end{equation}

holds, if and only if

\begin{equation}
\sum_{m=1}^{n(\epsilon)} Y_{\epsilon m} \overset{d}{\to} \mathcal{N}(0, 1) \quad \text{as} \quad \epsilon \downarrow 0 \quad \text{and} \quad \lim_{\epsilon \downarrow 0} \max_{m \in \{1, 2, \ldots, n(\epsilon)\}} E[Y_{\epsilon m}^2] = 0 .
\end{equation}

We will refer to (7) as the extended CLT condition. By defining

\begin{equation}
n(\epsilon) := \sum_{\ell=0}^{L} M_{\ell},
\end{equation}

and

\begin{equation}
Y_{\epsilon m} := \begin{cases}
\frac{\Delta_{0} X^m - E[\Delta_{0} X]}{\sqrt{\text{Var}(A_{ML}^L) M_{0}}} & m \leq M_{0} \\
\frac{\Delta_{1} X^{m} - E[\Delta_{1} X]}{\sqrt{\text{Var}(A_{ML}^L) M_{1}}} & M_{0} < m \leq M_{0} + M_{1} \\
 & \vdots \\
\frac{\Delta_{L} X^{m} - E[\Delta_{L} X]}{\sqrt{\text{Var}(A_{ML}^L) M_{L}}} & n(\epsilon) - M_{L} < m \leq n(\epsilon),
\end{cases}
\end{equation}

the normalized MLMC estimator can be represented by generalized triangular-arrays as follows:

\begin{equation}
\frac{A_{ML} - E[X_L]}{\sqrt{\text{Var}(A_{ML})}} = \sum_{m=1}^{n(\epsilon)} Y_{\epsilon m} .
\end{equation}

Corollary 2.2. Let $A_{ML}$ denote the MLMC estimator applied to estimate the expectation of $X \in L^2(\Omega)$ based on the collection of r.v. $\{X_{\ell}\} \subset L^2(\Omega)$ satisfying Assumption 1.1. Suppose that $\text{Var}(A_{ML}) > 0$ for any $\epsilon > 0$. Then the normalized MLMC estimator (10) satisfies the extended CLT condition (7), if and only if for any $\nu > 0$,

\begin{equation}
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \frac{V_{\ell}}{\text{Var}(A_{ML}) M_{\ell}} \mathbb{E}\left[\frac{V_{\ell}}{V_{\ell}} \frac{[\Delta_{\ell} X - E[\Delta_{\ell} X]]^2}{\text{Var}(A_{ML}) M_{\ell}} \mathbf{1}_{\left\{|\Delta_{\ell} X - E[\Delta_{\ell} X]| > \frac{\text{Var}(A_{ML}) M_{\ell}}{V_{\ell}} \nu\right\}}\right] = 0 .
\end{equation}

Proof. For all $\epsilon > 0$, the triangular array representation (10) of the MLMC estimator obviously satisfies the centering and normalization conditions (5), and its elements are centered and mutually independent. By Theorem 2.1 the extended CLT condition thus holds, if and only if Lindeberg’s condition (6) holds. For any
ν > 0, here Lindeberg’s condition takes the form:

\[
\lim_{\epsilon \to 0} \sum_{m=1}^{n(\epsilon)} \mathbb{E} \left[ V_{\epsilon m}^2 \mathbf{1}_{\{Y_{\epsilon m} > \epsilon\}} \right] = \lim_{\epsilon \to 0} \sum_{\ell=0}^{L} \sum_{i=1}^{M_\ell} \text{Var}(A_{ML}) \mathbb{E} \left[ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{M_\ell^2 \text{Var}(A_{ML})} \mathbf{1} \left\{ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{\nu_\epsilon \text{Var}(A_{ML}) M_\ell^2} > \nu_\epsilon \right\} \right] \\
= \lim_{\epsilon \to 0} \sum_{\ell=0}^{L} \frac{V_\ell}{M_\ell \text{Var}(A_{ML})} \mathbb{E} \left[ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{V_\ell} \mathbf{1} \left\{ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{\nu_\epsilon \text{Var}(A_{ML}) M_\ell^2} > \nu_\epsilon \right\} \right].
\]

Assumption (11) does not provide any lower bound on the decay rate of the variance sequence \( \{V_\ell\} \), and, therefore, it alone is not sufficiently strong to ensure that Lindeberg’s condition (11) holds in general. The problem is that without any lower bound on \( V_\ell \), there are asymptotic settings where a non-negligible contribution to the variance of the MLMC estimator derives from a finite number of samples.

**Example 2.1.** Consider the setting where \( \beta < \min(\gamma, 2\alpha) \) and positive constants \( c_1, c_2 \) such that

\[ c_1 e^{-\alpha \ell} \leq V_\ell \leq c_2 e^{-\beta \ell} \]

Let there be an infinite subsequence \( \{k_i\} \subset \mathbb{N}_0 \) for which

\[ V_{k_i} = \Theta_{k_i}(e^{-2\alpha k_i}) \quad \text{and} \quad S_{k_i} = \Theta_{k_i}(e^{(\gamma - 2\alpha)k_i}). \]

Then equation (11) implies there exists \( c, C, \tilde{c}, \tilde{c} \in \mathbb{R}_+ \) such that for all \( y \in \{e > 0 \mid L(e) \in \{k_i\}\} \),

\[ M_{L(y)} < C, \]

and

\[ \tilde{c} \leq \max \left( \frac{V_{L(y)}}{M_{L(y)} \text{Var}(A_{ML}(y))}, \frac{M_{L(y)}^2 \text{Var}(A_{ML}(y))}{V_{L(y)}} \right) \leq \tilde{c}. \]

Hence, for any \( \nu < (2\tilde{c})^{-1} \),

\[
\lim_{\epsilon \to 0} \sum_{\ell=0}^{L} \frac{V_\ell}{M_\ell \text{Var}(A_{ML})} \mathbb{E} \left[ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{V_\ell} \mathbf{1} \left\{ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{\nu_\epsilon \text{Var}(A_{ML}) M_\ell^2} > \nu_\epsilon \right\} \right] \\
\geq \lim_{\epsilon \to 0} \sum_{\ell=0}^{L} \frac{V_\ell}{M_\ell \text{Var}(A_{ML})} \mathbb{E} \left[ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{V_\ell} \mathbf{1} \left\{ \frac{|\Delta_\ell X - \mathbb{E}[\Delta_\ell X]|^2}{\nu_\epsilon \text{Var}(A_{ML}) M_\ell^2} > \nu_\epsilon \right\} \right] \\
\geq \lim_{\ell \to \infty} \tilde{c} \mathbb{E} \left[ \frac{|\Delta_{k_\ell} X - \mathbb{E}[\Delta_{k_\ell} X]|^2}{V_{k_\ell}} \mathbf{1} \left\{ \frac{|\Delta_{k_\ell} X - \mathbb{E}[\Delta_{k_\ell} X]|^2}{\nu_\epsilon \text{Var}(A_{ML}) M_{k_\ell}^2} > \nu_\epsilon \right\} \right] > 0.
\]

Example 2.1 illustrates that Assumption (11) is not sufficiently strong to ensure condition (11). We therefore impose the following additional variance decay assumptions, which can be viewed as implicit weak lower bounds on the sequence \( \{V_\ell\} \).

**Assumption 2.1.** For the rate triplet introduced in Assumption (11) assume that \( V_0 > 0 \) and

(i) if \( \beta = \gamma \), then \( \lim_{k \to \infty} S_k = \infty \),

(ii) if \( \gamma > \beta \), then we assume that \( \beta < 2\alpha \) and that there exists \( \nu \in [\beta, 2\alpha) \) such that

\[ \lim_{k \to \infty} S_k e^{(\nu - \gamma)k/2} > 1. \]
Lemma 2.3. Let $A_{ML}$ denote the MLMC estimator applied to estimate the expectation of $X \in L^2(\Omega)$ based on the collection of r.v. $\{X_\ell\} \subset L^2(\Omega)$ satisfying Assumption 1.1. If also Assumption 2.1 holds, then

$$\lim_{\epsilon \downarrow 0} \frac{\text{Var}(A_{ML})}{\epsilon^2} = 1.$$  

Proof. For any $\epsilon > 0$, it follows from equation (1) that

$$\frac{\text{Var}(A_{ML})}{\epsilon^2} = \sum_{\ell=0}^{L} \frac{V_\ell}{\epsilon^2 M_\ell} \leq \sum_{\ell=0}^{L} \frac{\sqrt{V_\ell C_\ell}}{S_L} = 1,$$

and by the mean value theorem there exists a constant $C > 0$ such that

$$\sum_{\ell=0}^{L} \frac{V_\ell}{\epsilon^2 M_\ell} \geq \sum_{\ell=0}^{L} \frac{1_{\{V_\ell > 0\}} V_\ell}{\epsilon^2 C_\ell S_L} \geq 1 - e^2 \sum_{\ell=0}^{L} \frac{C_\ell}{S_L^2} \geq 1 - C e^2 \frac{e^\gamma L}{S_L^2}.$$

To complete the proof, we thus have to verify that

$$\lim_{\epsilon \downarrow 0} \frac{e^2 e^\gamma L}{S_L^2} = 0.$$

If $\beta < \gamma$, then Assumption 2.1(ii) implies that

$$e^2 e^\gamma L \leq O(\epsilon^{2-2/\alpha}),$$

and since $\nu < 2\alpha$, the claim follows in this case. Similarly, if $\beta = \gamma$, then $e^2 e^\gamma L = O(1)$, and the claim follows from Assumption 1.1. Finally, if $\beta > \gamma$, then the assumption $\min(\beta, \gamma) \leq 2\alpha$ (cf. Assumption 1.1) implies $\gamma \leq 2\alpha$, and we have to consider two cases: (I) $\gamma < 2\alpha$ and (II) $\gamma = 2\alpha$. For case (I), equation (14) follows from $\lim_{\epsilon \downarrow 0} e^2 e^\gamma L = 0$ and $S_L \geq S_0 > 0$ for all $L \geq 0$. For case (II), we introduce

$$\hat{L} := \max \left( \left\lfloor \frac{4 \log(c_0 \epsilon^{-1})}{\gamma + \beta} \right\rfloor, 1 \right).$$

As $2\alpha = \gamma < \beta$, $\hat{L} \leq L$,

$$\frac{\text{Var}(A_{ML})}{\epsilon^2} \geq \sum_{\ell=0}^{\hat{L}} 1_{\{V_\ell > 0\}} \frac{V_\ell}{\sqrt{V_\ell S_L + \epsilon^2}} \geq \sum_{\ell=0}^{\hat{L}} 1_{\{V_\ell > 0\}} \frac{V_\ell}{\sqrt{V_\ell S_L + \epsilon^2}} \geq 1 - C e^2 e^\gamma \frac{e^\gamma L}{S_L^2},$$

and the result follows by $\lim_{\epsilon \downarrow 0} e^2 e^\gamma \hat{L} = 0$.  

Lemma 2.3 implies that we can reformulate Lindeberg’s condition for the MLMC estimator as follows:

Corollary 2.4. Let $A_{ML}$ denote the MLMC estimator applied estimate the expectation of $X \in L^2(\Omega)$ based on the collection of r.v. $\{X_\ell\} \subset L^2(\Omega)$ satisfying
Assumption 1.1. If Assumptions 1.1 and 2.1 hold, then the normalized MLMC estimator satisfies the extended CLT condition (11), if and only if for any $\nu > 0$,

$$
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \frac{\sqrt{C_{T}}}{S_{L}} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{V_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{V_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{V_{\ell}} \right\} = 0.
$$

Proof. From the proof of Lemma 2.3 it follows that there exists an $\epsilon > 0$ such that

$$
\frac{1}{2} \leq \frac{\text{Var}(A_{ML})}{\epsilon^{2}} \leq 1, \quad \forall \epsilon \in (0, \epsilon).
$$

Consequently, for any $\epsilon \in (0, \epsilon)$ and any $\nu > 0$ we have that

$$
\sum_{\ell=0}^{L} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{\epsilon^{2} M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{\epsilon^{2} M_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{\epsilon^{2} M_{\ell}} \right\} \geq 0.
$$

as well as

$$
\sum_{\ell=0}^{L} \frac{1}{\text{Var}(A_{ML})} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{M_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{M_{\ell}} \right\} \leq 2 \sum_{\ell=0}^{L} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{\epsilon^{2} M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{\epsilon^{2} M_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{\epsilon^{2} M_{\ell}} \right\}.
$$

These upper and lower bounds imply that Lindeberg’s condition (11) is equivalent to the following condition: for any $\nu > 0$ it holds that

$$
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{\epsilon^{2} M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{\epsilon^{2} M_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{\epsilon^{2} M_{\ell}} \right\} = 0.
$$

Following similar steps as those leading to inequality (13), we further note that for sufficiently small $\epsilon > 0$,

$$
\sum_{\ell=0}^{L} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{\epsilon^{2} M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{\epsilon^{2} M_{\ell}} > \nu \right\} \leq 2 \sum_{\ell=0}^{L} \mathbb{E} \left[ \frac{\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]}{\epsilon^{2} M_{\ell}} \right] \mathbb{I} \left\{ \frac{|\Delta_{\ell} X - [\Delta_{\ell} X]|^{2}}{\epsilon^{2} M_{\ell}} > \frac{\nu \cdot \mathbb{E}[(\Delta_{\ell} X)^{2}]}{\epsilon^{2} M_{\ell}} \right\} + \rho(\epsilon),
$$

where the mapping $\rho: \mathbb{R} \to \mathbb{R}$, satisfying $\lim_{\epsilon \downarrow 0} \rho(\epsilon) = 0$, can be derived as in the proof of Lemma 2.3.

In settings with $\beta > \gamma$, the geometric decay of the sequence $\{\sqrt{C_{T}}/V_{\ell}\}_{\ell \geq 0}$ turns out to be sufficient to prove that the extended CLT condition holds.

**Theorem 2.5.** Let $A_{ML}$ denote the MLMC estimator applied to estimate the expectation of $X \in L^{2}(\Omega)$ based on the collection of r.v. $\{X_{\ell}\} \subset L^{2}(\Omega)$ satisfying Assumption 1.1. If $\beta > \gamma$ and if Assumption 1.1 holds (the only relevant condition being $V_{0} > 0$), then the extended CLT condition (11) is satisfied for the normalized MLMC estimator.

**Proof.** We prove this result by verifying that condition (15) holds. It follows from Assumption 2.1 that

$$
S := \lim_{k \to \infty} S_{k} = \lim_{k \to \infty} \sum_{\ell=0}^{k} \sqrt{V_{\ell} C_{T}} \leq c \lim_{k \to \infty} \sum_{\ell=0}^{k} e^{(\gamma-\beta)\ell/2} < \infty.
$$
Furthermore, as the sequence \( \{S_L\}_{L \geq 0} \) is monotonically increasing, it is contained in the bounded set \([S_0, S]\) with \(S_0 > 0\). Consequently, Lindeberg’s condition \((17)\) is equivalent to:

\[
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \sqrt{V_{\ell}C_{\ell}} \mathbf{1}(V_{\ell} > 0) \mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] = 0, \quad \forall \nu > 0.
\]

For a fixed \(\nu > 0\), introduce the sequence of functions \(\{f_\epsilon\}_{\epsilon > 0}\), where \(f_\epsilon : [0, S] \to [0, 1]\) is defined by

\[
f_\epsilon(x) = \begin{cases} 
\mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] & \text{if } 0 \leq x < S_0 , \\
\mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] & \text{if } S_0 \leq x < S_1 , \\
\vdots & \\
\mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{L}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{L}} > \epsilon^2 \frac{M_2^2}{V_{L}} \right\} \right] & \text{if } S_{L-1} \leq x < S_L , \\
0 & \text{if } S_L \leq x \leq S .
\end{cases}
\]

For any \(\epsilon > 0\), one thus has that

\[
\sum_{\ell=0}^{L} \sqrt{V_{\ell}C_{\ell}} \mathbf{1}(V_{\ell} > 0) \mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] = \int_0^S f_\epsilon(x) dx .
\]

Since \(|f_\epsilon| \leq 1_{\{0, S\}}\) for all \(\epsilon > 0\), the dominated convergence theorem implies that

\[
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \sqrt{V_{\ell}C_{\ell}} \mathbf{1}(V_{\ell} > 0) \mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] = \int_0^S \lim_{\epsilon \downarrow 0} f_\epsilon(x) dx .
\]

It remains to verify that

\[(17)\]

\[
\int_0^S \lim_{\epsilon \downarrow 0} f_\epsilon(x) dx = 0.
\]

Consider two cases: (I) \(\gamma < 2\alpha\) and (II) \(\gamma = 2\alpha\). For case (I), note that

\[
\lim_{\epsilon \downarrow 0} \min_{\ell \in \{0, 1, \ldots, L\}} \frac{2\epsilon^2 M_2^2}{V_{\ell}} \geq \lim_{\epsilon \downarrow 0} \min_{\ell \in \{0, 1, \ldots, L\}} \epsilon^2 \frac{S_2^2}{C_{\ell}} \geq \lim_{\epsilon \downarrow 0} \epsilon^2 \frac{S_2^2}{C_L} = \infty ,
\]

since \(C_L = \Theta_{\ell}(e^{-\gamma/\alpha})\). For any \(x \in [0, S]\), say \(x \in [S_{\ell-1}, S_\ell]\) for some \(0 \leq \ell \leq L\) for which \(V_{\ell} > 0\), the dominated convergence theorem then implies that

\[(18)\]

\[
\lim_{\epsilon \downarrow 0} f_\epsilon(x) = \mathbb{E} \left[ \lim_{\epsilon \downarrow 0} \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] = 0.
\]

Consequently,

\[
\lim_{\epsilon \downarrow 0} \sum_{\ell=0}^{L} \sqrt{V_{\ell}C_{\ell}} \mathbf{1}(V_{\ell} > 0) \mathbb{E} \left[ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} \mathbf{1}\left\{ \frac{[\Delta_{\ell} X - \mathbb{E}[\Delta_{\ell} X]]^2}{V_{\ell}} > \epsilon^2 \frac{M_2^2}{V_{\ell}} \right\} \right] = 0 ,
\]

and \((17)\) follows. For case (II), we introduce

\[
\tilde{L} = \max \left( \frac{4 \log(\epsilon_0 \epsilon^{-1})}{\gamma + \beta} , 1 \right).
\]

Since \(\tilde{L} \leq L\) and

\[
\int_{S_L}^S \lim_{\epsilon \downarrow 0} f_\epsilon dx \leq \lim_{\epsilon \downarrow 0} \int_{S_L}^S dx = \lim_{\epsilon \downarrow 0} (S_L - S_L) = 0 ,
\]

\[
\int_{S_L}^S \lim_{\epsilon \downarrow 0} f_\epsilon dx = \lim_{\epsilon \downarrow 0} \int_{S_L}^S dx = \lim_{\epsilon \downarrow 0} (S_L - S_L) = 0 ,
\]

\[
\int_{S_L}^S \lim_{\epsilon \downarrow 0} f_\epsilon dx = \lim_{\epsilon \downarrow 0} \int_{S_L}^S dx = \lim_{\epsilon \downarrow 0} (S_L - S_L) = 0 ,
\]

\[
\int_{S_L}^S \lim_{\epsilon \downarrow 0} f_\epsilon dx = \lim_{\epsilon \downarrow 0} \int_{S_L}^S dx = \lim_{\epsilon \downarrow 0} (S_L - S_L) = 0 .
\]
we have that
\[ \int_0^S \lim_{\epsilon \downarrow 0} f_\epsilon(x) \, dx \leq \int_0^{S_L} \lim_{\epsilon \downarrow 0} f_\epsilon(x) \, dx. \]
Moreover,
\[ \lim_{\epsilon \downarrow 0} \min_{\ell \in \{0, \ldots, L\}} \frac{\epsilon^2 M^2_\ell}{V_\ell} \geq \lim_{\epsilon \downarrow 0} \epsilon^{-2} \frac{S^2_\ell}{C_\ell} = \infty, \]
since \( C_\ell = \Theta_\ell (\epsilon^{-\gamma/(\gamma + \beta)}) \) and \( 2\gamma/(\gamma + \beta) < 1 \). From (18) it then follows that \( \lim_{\epsilon \downarrow 0} f_\epsilon(x) = 0 \) for all \( x \in [0, S_L] \), so that (7) holds.

As the above argument is valid for any fixed \( \nu > 0 \), we have proved that Lindeberg’s condition holds. □

We conclude the paper by treating the case \( \gamma \geq \beta \).

**Theorem 2.6.** Let \( A_{MLC} \) denote the MLMC estimator applied to estimate the expectation of \( X \in L^2(\Omega) \) based on the collection of r.v. \( \{X_\ell\} \subset L^2(\Omega) \) satisfying Assumption [11, 12]. Suppose that \( \beta \geq \gamma \), Assumption (7) holds, and
\[ \lim_{\ell \to \infty} 1_{\{V_\ell > 0\}} E \left[ \frac{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid^2}{V_\ell} 1\{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid > \nu \gamma \alpha \beta \} 1_{\{$S_0, S_1, \ldots, S_{L-1}\}$} \right] = 0 \]
holds for any \( \nu > 0 \). Then the extended CLT condition (7) is satisfied for the normalized MLMC estimator.

**Proof.** From (11) and \( C_\ell = \Theta_\ell (\epsilon^{-\gamma}) \) it follows that there exists a \( c > 0 \) such that
\[ 1_{\{V_\ell > 0\}} \frac{\epsilon^2 M^2_\ell}{V_\ell} \geq 1_{\{V_\ell > 0\}} \frac{\epsilon^{-2} S^2_\ell}{C_\ell} \geq 1_{\{V_\ell \geq 0\}} c \epsilon(2\alpha - \gamma) \ell S^2_\ell. \]

Consequently,
\[ \sum_{\ell = 0}^L \sqrt{V_\ell C_\ell} \epsilon \left[ \frac{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid^2}{V_\ell} 1\{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid \leq c \epsilon S^2_\ell \} \right] \]
\[ \leq \sum_{\ell = 0}^L \sqrt{V_\ell C_\ell} \epsilon \left[ \frac{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid^2}{V_\ell} 1\{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid \leq \nu \epsilon c (2\alpha - \gamma) \ell S^2_\ell \} \right]. \]

Introduce the infinite matrix \( A = (a_{k\ell}) \) where
\[ a_{k\ell} := 1_{\{\ell \leq k\}} \sqrt{V_\ell C_\ell} \frac{1}{S_k}, \quad k, \ell \in \mathbb{N}_0 \]
and the sequence
\[ x_\ell := 1_{\{V_\ell > 0\}} E \left[ \frac{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid^2}{V_\ell} 1\{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid \leq \nu \epsilon c (2\alpha - \gamma) \ell S^2_\ell \} \right], \quad \ell \in \mathbb{N}_0. \]

Since
\[ \lim_{k \to \infty} a_{k\ell} = 0 \quad \forall \ell \in \mathbb{N}_0, \quad \lim_{k \to \infty} \sum_{\ell = 0}^\infty a_{k\ell} = 1, \quad \text{and} \quad \sum_{\ell = 0}^\infty |a_{k\ell}| \leq 1 \quad \forall k \in \mathbb{N}_0, \]
the matrix \( A \) defines a regular summability method and the Silverman–Toeplitz theorem [11, 12] yields that
\[ \lim_{\epsilon \downarrow 0} \sum_{\ell = 0}^L \sqrt{V_\ell C_\ell} \epsilon \left[ \frac{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid^2}{V_\ell} 1\{\mid \Delta_\ell X - \mathbb{E}[\Delta_\ell X] \mid \leq \nu \epsilon c (2\alpha - \gamma) \ell S^2_\ell \} \right] \]
\[ = \lim_{k \to \infty} \sum_{\ell = 0}^\infty a_{k\ell} x_\ell = \lim_{\ell \to \infty} x_\ell = 0, \]
which completes the proof. □
References

[1] S. Heinrich, Monte Carlo complexity of global solution of integral equations, J. Complexity 14 (2) (1998) 151–175.
[2] M. B. Giles, Multilevel Monte Carlo path simulation, Oper. Res. 56 (3) (2008) 607–617.
[3] M. B. Giles, Multilevel Monte Carlo methods, Acta Numer. 24 (2015) 259–328.
[4] H. Hoel, E. Von Schwerin, A. Szepessy, R. Tempone, Implementation and analysis of an adaptive multilevel monte carlo algorithm, Monte Carlo Methods and Applications 20 (1) (2014) 1–41.
[5] M. Ben Alaya, A. Kebaier, Central limit theorem for the multilevel Monte Carlo Euler method, Ann. Appl. Probab. 25 (1) (2015) 211–234.
[6] S. Dereich, S. Li, Multilevel Monte Carlo for Lévy-driven SDEs: central limit theorems for adaptive Euler schemes, Ann. Appl. Probab. 26 (1) (2016) 136–185.
[7] D. Giorgi, V. Lemaire, G. Pagès, Limit theorems for weighted and regular multilevel estimators, Monte Carlo Methods Appl. 23 (1) (2017) 43–70.
[8] N. Collier, A.-L. Haji-Ali, F. Nobile, E. von Schwerin, R. Tempone, A continuation multilevel Monte Carlo algorithm, BIT 55 (2) (2015) 399–432.
[9] R. Durrett, Probability: theory and examples, 2nd Edition, Duxbury Press, Belmont, CA, 1996.
[10] A. Klenke, Probability theory, 2nd Edition, Universitext, Springer, London, 2014.
[11] O. Toeplitz, Über allgemeine lineare Mittelbildungen, Prace matematyczno-fizyczne 22 (1) (1911) 113–119.
[12] E. Kreyszig, Introductory functional analysis with applications, Vol. 1, Wiley New York, 1978.

(H. Hoel) Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden
E-mail address, Corresponding author: hhakon@chalmers.se, haakonahl@gmail.com

(S. Krumscheid) Calcul Scientifique et Quantification de l’Incertitude (CSQI), Institute of Mathematics, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland
E-mail address: sebastian.krumscheid@epfl.ch