

# Multilevel Monte Carlo Methoden

Hauptseminar Wissenschaftliches Rechnen

Hendrik Kleikamp

5. Juli 2017

---

## 1 Monte Carlo und Multilevel Monte Carlo

In vielen Anwendungsbereichen, beispielsweise der Finanzmathematik, ist es notwendig, den Erwartungswert einer (reellen) Zufallsvariablen  $P$  zu berechnen. Eine Standardmethode hierfür ist die Monte Carlo Simulation, welche auf dem Schwachen Gesetz der großen Zahlen beruht. Hierbei werden  $N$  Simulationen  $S^{(1)}, \dots, S^{(N)}$ , sogenannte Samples, aus dem zugrundeliegenden Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, \mathbb{P})$  berechnet und jeweils mittels  $P$  ausgewertet. Für  $\mathbb{E}[P]$  ergibt sich mit  $P^{(i)} := P(S^{(i)})$  der Monte Carlo Schätzer

$$X := \frac{1}{N} \sum_{i=1}^N P^{(i)}. \quad (1)$$

Der Schätzer  $X$  ist zwar erwartungstreu, das heißt es ist  $\mathbb{E}[X] = \mathbb{E}[P]$ , jedoch sind sehr viele Samples notwendig, um eine ausreichende Genauigkeit zu erhalten.

Die Multilevel Monte Carlo Methode versucht nun, den Rechenaufwand bei gleichbleibender Genauigkeit zu verringern. Hierfür wird die Zufallsvariable  $P$  durch eine Folge  $\{P_l\}_{l \in \mathbb{N}}$  ersetzt, die  $P$  mit wachsendem Level  $l$  immer besser approximiert. Zudem wird davon ausgegangen, dass die Kosten zur Berechnung eines Samples mit wachsendem Level ebenfalls steigen. Beim Multilevel Monte Carlo Ansatz werden die Vorteile von kleinen und großen Leveln kombiniert. Samples lassen sich in kleinen Leveln mit relativ geringem Rechenaufwand berechnen, während  $P_l$  für großes  $l$  eine bessere Approximation an  $P$  darstellt als für kleines  $l$ . Anstatt  $\mathbb{E}[P]$  zu berechnen, soll nur  $\mathbb{E}[P_L]$  für  $L$  hinreichend groß als Approximation an  $\mathbb{E}[P]$  berechnet werden. Unter Verwendung der Linearität des Erwartungswertes ergibt sich

$$\mathbb{E}[P_L] = \mathbb{E}[P_0] + \sum_{l=1}^L \mathbb{E}[P_l - P_{l-1}]. \quad (2)$$

Um  $\mathbb{E}[P_0]$  und  $\mathbb{E}[P_l - P_{l-1}]$  für  $l = 1, \dots, L$  zu nähern, werden meist Monte Carlo Schätzer wie in Formel (1) genutzt. Dies führt für  $\mathbb{E}[P_L] \approx \mathbb{E}[P]$  auf den Schätzer

$$Y := \frac{1}{N_0} \sum_{i=1}^{N_0} P_0^{(0,i)} + \sum_{l=1}^L \left( \frac{1}{N_l} \sum_{i=1}^{N_l} P_l^{(l,i)} - P_{l-1}^{(l,i)} \right). \quad (3)$$

In jedem Level werden hierbei stochastisch unabhängige Samples verwendet, wobei für die Korrekturdifferenzen  $P_l^{(l,i)} - P_{l-1}^{(l,i)}$  jeweils dieselben  $N_l$  Samples bezüglich  $P_l$  und  $P_{l-1}$  ausgewertet werden.

Für den Schätzer  $Y$  gilt nun  $\mathbb{E}[Y] = \mathbb{E}[P_L]$  und  $\mathbb{V}[Y] = \sum_{l=0}^L \frac{1}{N_l} \mathbb{V}[P_l - P_{l-1}]$ , wenn  $P_{-1} \equiv 0$  gesetzt wird. Damit ergibt sich der mean squared error ( $MSE$ ) des Multilevel Monte Carlo Ansatzes als

$$MSE = \mathbb{E}[(Y - \mathbb{E}[P])^2] = \mathbb{V}[Y] + (\mathbb{E}[Y] - \mathbb{E}[P])^2 = \underbrace{\sum_{l=0}^L \frac{1}{N_l} \mathbb{V}[P_l - P_{l-1}]}_{\text{Stichproben-Fehler}} + \underbrace{(\mathbb{E}[P_L - P])^2}_{\text{Approximationsfehler}}. \quad (4)$$

Das Ziel der theoretischen Untersuchungen ist es, einen  $MSE$  von  $\epsilon^2$  für eine vorgegebene Genauigkeit  $\epsilon > 0$  zu erreichen. Hierzu wird versucht, sowohl den Stichproben-Fehler als auch den Approximationsfehler kleiner als  $\epsilon^2/2$  zu halten. Unter gewissen Voraussetzungen an Varianz und Rechenkosten in den einzelnen Leveln kann der gewünschte  $MSE$  garantiert und eine Abschätzung für die Gesamtkosten des Schätzers  $Y$  angegeben werden. Dies zeigt das Theorem im folgenden Abschnitt.

## 2 Hauptresultat

**Theorem** (MLMC [Giles, 2015]). Sei  $P$  eine Zufallsvariable und  $P_l$  eine numerische Approximation an  $P$  im Level  $l$ . Angenommen, es existieren unabhängige Schätzer  $Y_l$  basierend auf  $N_l$  Monte Carlo Samples, jeweils mit erwarteten Kosten  $C_l$  und einer Varianz  $V_l$ , sowie positive Konstanten  $\alpha, \beta, \gamma, c_1, c_2, c_3$  mit  $\alpha \geq \frac{1}{2} \min(\beta, \gamma)$  und

1.  $|\mathbb{E}[P_l - P]| \leq c_1 2^{-\alpha l}$
2.  $\mathbb{E}[Y_l] = \begin{cases} \mathbb{E}[P_0], & l = 0 \\ \mathbb{E}[P_l - P_{l-1}], & l > 0 \end{cases}$
3.  $V_l \leq c_2 2^{-\beta l}$
4.  $C_l \leq c_3 2^{\gamma l}$ .

Dann gibt es eine Konstante  $c_4 > 0$ , sodass für alle  $\epsilon < e^{-1}$  Zahlen  $L$  und  $N_l$  existieren, für die der Multilevel-Schätzer

$$Y := \sum_{l=0}^L Y_l$$

einen Fehler von

$$MSE = \mathbb{E}[(Y - \mathbb{E}[P])^2] < \epsilon^2$$

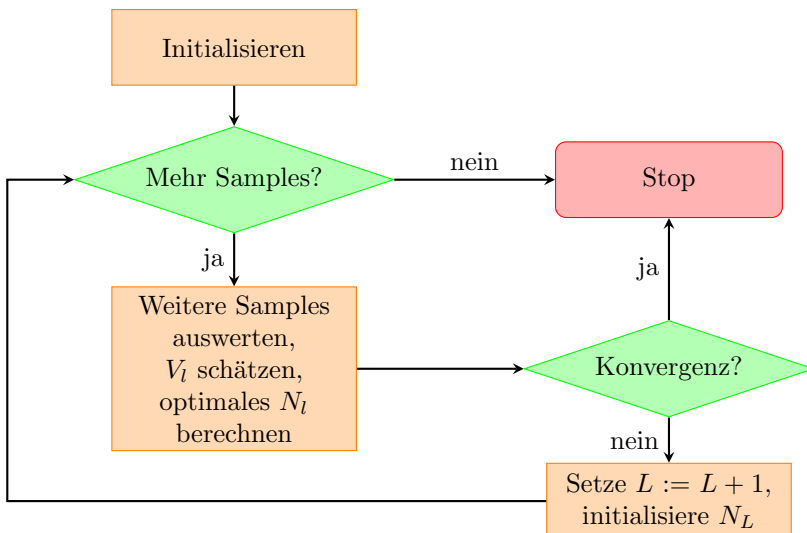
besitzt und der Rechenaufwand  $C$  insgesamt beschränkt ist durch

$$\mathbb{E}[C] \leq \begin{cases} c_4 \epsilon^{-2}, & \beta > \gamma \\ c_4 \epsilon^{-2} (\log \epsilon)^2, & \beta = \gamma \\ c_4 \epsilon^{-2 - (\gamma - \beta)/\alpha}, & \beta < \gamma. \end{cases}$$

Ein Beweis des Theorems findet sich in dem Paper „Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients“ von Cliffe et al. (2011).

*Bemerkung.* Das Theorem setzt eine geometrische Level-Sequenz voraus, das heißt die Kosten wachsen exponentiell mit dem Level, während der Fehler  $|\mathbb{E}[P_l - P]|$  und die Varianz  $\mathbb{V}[P_l - P_{l-1}]$  exponentiell mit wachsendem Level sinken. Nur durch diese Annahmen kann eine Abschätzung für den Rechenaufwand angegeben werden. Es ist wichtig zu beachten, dass das Theorem keine genaueren Angaben zur Konstruktion der Schätzer  $Y_l$  macht, diese müssen einzig und allein den korrekten Erwartungswert besitzen (vgl. Annahme 2 des MLMC-Theorems). Eine andere Wahl von  $Y_l$  als der Naheliegenden aus Formel (1) ist also durchaus denkbar, solange diese die Voraussetzungen des Theorems erfüllt. Die Idee alternativer Versionen der  $Y_l$  ist es, die Varianz  $V_l$  weiter zu reduzieren und somit eine bessere Konvergenzgeschwindigkeit zu erhalten.  $\square$

## 3 Implementierung



### Hinweise zur Implementierung:

- Die Konstanten  $c_1$  und  $c_2$  aus dem MLMC-Theorem sind meist nicht bekannt.
- Sind  $\alpha$  und  $\beta$  aus dem MLMC-Theorem nicht bekannt, so müssen diese im Algorithmus geschätzt werden.
- Die Genauigkeit der Schätzer für  $V_l$  ( $l = 0, \dots, L$ ) hängt von der Anzahl der verwendeten Samples im jeweiligen Level ab.
- Der Algorithmus ist heuristisch, ein  $MSE$  von  $\epsilon^2$  ist nicht garantiert.

## 4 Multi-Index Monte Carlo

Die im Abschnitt 1 vorgestellte Multilevel Monte Carlo Methode beschränkt sich auf eindimensionale Levelindizes  $l$ . In einigen Anwendungen, die beispielsweise die Lösung von Differentialgleichungen voraussetzen, können mehrdimensionale Level von Vorteil sein. Hierdurch werden unterschiedliche Diskretisierungsweiten in unterschiedliche Koordinatenrichtungen ermöglicht.

Um die Multilevel Teleskopsumme aus Formel (2) auf diesen Fall zu erweitern, wird zunächst für  $l = (l_i)_{i=1}^d \in \mathbb{N}^d$  die Differenz

$$\Delta_i P_l = \begin{cases} P_l - P_{l-e_i}, & l_i > 0 \\ P_l, & l_i = 0 \end{cases}$$

definiert, wobei  $e_i$  den Einheitsvektor in Koordinatenrichtung  $i$  bezeichnet.

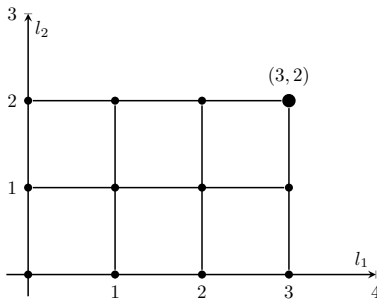
Sei zudem

$$\Delta P_l = \left( \prod_{i=1}^d \Delta_i \right) P_l.$$

Mit  $I = \{l \in \mathbb{N}^d : l \geq 0\}$  ergibt sich für  $\mathbb{E}[P]$  der Schätzer

$$Y := \sum_{l \in I} \mathbb{E}[\Delta P_l].$$

Als Beispiel betrachten wir eine andere Wahl von  $I$ . Es sei  $I = \{(l_1, l_2) \in \mathbb{N}^2 : 0 \leq l_1 \leq 3 \text{ und } 0 \leq l_2 \leq 2\}$ , dies entspricht dem folgenden Rechteck:



In diesem Fall ist  $\Delta P_{(3,2)} = \Delta_1 \Delta_2 P_{(3,2)} = P_{(3,2)} - P_{(2,2)} - P_{(3,1)} + P_{(2,1)}$  und  $Y = \mathbb{E}[P_{(3,2)}]$ .

Die hier gezeigte Wahl von  $I$  ist im Allgemeinen nicht optimal. Haji-Ali et al. konnten zeigen, dass die optimale Wahl der verwendeten Levelindizes ähnlich zur Konstruktion dünner Gitter ist (siehe [4]).

Es ergibt sich

$$I = \{l \in \mathbb{N}^d : l \geq 0, \langle l, n \rangle \leq L\}$$

für eine Richtung  $n \in \mathbb{N}^d, n > 0$ , und ein  $L \in \mathbb{N}$ .

Im Zweidimensionalen entspricht dies einem Dreieck, wobei zwei Seiten parallel zu den Koordinatenachsen durch den Ursprung verlaufen und die dritte Seite orthogonal zu  $n$  steht.

Das MLMC-Theorem lässt sich in leicht abgewandelter Form auch auf die Multi-Index Monte Carlo Methode übertragen (siehe [1] und [4]).

## 5 Literaturhinweise

### Literatur zur Multilevel Monte Carlo Methode und ihren Anwendungen

- [1] Michael B. Giles. „Multilevel Monte Carlo methods“. In: *Acta Numerica* 24 (2015), 259–328.
- [2] Sascha Petzold. „Multi-Level Monte-Carlo-Methoden und ihre Anwendung in der Finanzmathematik“. Masterarbeit. Mathematisches Institut Universität Bayreuth, 2016.

### Literatur zu weiterführenden Forschungsthemen

- [3] Michael B. Giles und Ben J. Waterhouse. „Multilevel quasi-Monte Carlo path simulation“. In: *Advanced Financial Modelling, Radon Series on Computational and Applied Mathematics* (2009), S. 165–181.
- [4] Abdul-Lateef Haji-Ali, Fabio Nobile und Raúl Tempone. „Multi-index Monte Carlo: when sparsity meets sampling“. In: *Numerische Mathematik* 132.4 (2016), S. 767–806.
- [5] Vincent Lemaire, Gilles Pagès u. a. „Multilevel Richardson–Romberg extrapolation“. In: *Bernoulli* 23.4A (2017), S. 2643–2692.
- [6] Chang-han Rhee und Peter W. Glynn. „A new approach to unbiased estimation for SDE’s“. In: *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference. 2012, S. 17.
- [7] Ferran Vidal-Codina u. a. „A model and variance reduction method for computing statistical outputs of stochastic elliptic partial differential equations“. In: *Journal of Computational Physics* 297 (2015), S. 700–720.