Applications of Guaranteed Adaptive Quasi-Monte Carlo Algorithms

Lluís Antoni Jiménez Rugama
Joint work with Da Li (IIT), Fred J. Hickernell (IIT), Clémentine Prieur (Univ. Grenoble-Alpes), Elise Arnaud (Univ. Grenoble-Alpes), and Laurent Gilquin (Univ. Grenoble-Alpes)

Room 120, Bldg E1, Department of Applied Mathematics
Illinois Institute of Technology, Chicago, 60616 IL
Email: `ljimene1@hawk.iit.edu`

Monday 28$^{\text{th}}$ September, 2015

Outline

- Introduction
- Multivariate Normal Probability
- Sobol Indices
- Efficiency improvements
  - Importance Sampling
  - Control Variates

Outline

- Introduction—Overview of the GAIL quasi-Monte Carlo cubature algorithms
- Multivariate Normal Probability
- Sobol Indices
- Efficiency improvements
  - Importance Sampling
  - Control Variates

Numerical Integration Problem

Given ε_a and $\boldsymbol{x} \mapsto f(\boldsymbol{x})$, we want $\hat{I}$ such that

$$\left| \int_{[0,1)^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} - \hat{I}\left(\boldsymbol{x} \mapsto f(\boldsymbol{x}), \varepsilon_a\right) \right| \leqslant \varepsilon_a,$$

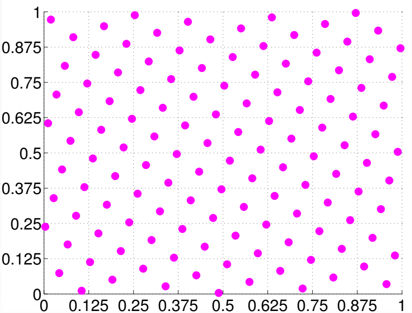
where

$$\hat{I}\left(\boldsymbol{x} \mapsto f(\boldsymbol{x}), \varepsilon_a\right) = \frac{1}{2^m} \sum_{i=0}^{2^m-1} f(\boldsymbol{z}_i \oplus \boldsymbol{\Delta}),$$

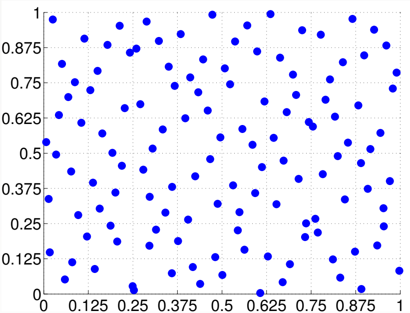
for some automatic and adaptive choice of m, $\{\boldsymbol{z}_i\}_{i=0}^{\infty} \in \left\{\begin{array}{c}\text{Lattice} \\ \text{Digital}\end{array}\right\}$ sequence, and

$$\text{cost}\left(\hat{I}\left(\boldsymbol{x} \mapsto f(\boldsymbol{x}), \varepsilon_a\right)\right) = \mathcal{O}\left((m + \$(f))2^m\right)$$

Examples of Sequences



Shifted rank-1 lattice sequence with generating vector $(1, 47)$.



Digitally shifted scrambled Sobol' sequence.

Adaptive Algorithm

The idea behind the results in Jiménez Rugama and Hickernell (2014) and Hickernell and Jiménez Rugama (2014) is that for all $f \in \mathcal{C}$,

$$\left| \int_{[0,1)^d} f(\boldsymbol{x})\, \mathrm{d}\boldsymbol{x} - \hat{I}\left(\boldsymbol{x} \mapsto f(\boldsymbol{x}), \varepsilon_a\right) \right| \leqslant a(r,m) \sum_{\kappa=\lfloor 2^{m-r-1} \rfloor}^{2^{m-r}-1} \left| \tilde{f}_{m,\kappa} \right|.$$

- $\tilde{f}_{m,\kappa} =$ discrete Fourier $\left\{ \begin{matrix} \text{Exponential} \\ \text{Walsh} \end{matrix} \right\}$ coefficients of f.
- $a(r,m) =$ inflation factor that depends on $\mathcal{C}$.

IID Monte Carlo Error Depends on All Fourier Coefficients

In the Monte Carlo case, if $I(f) = \int_{[0,1)^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$,

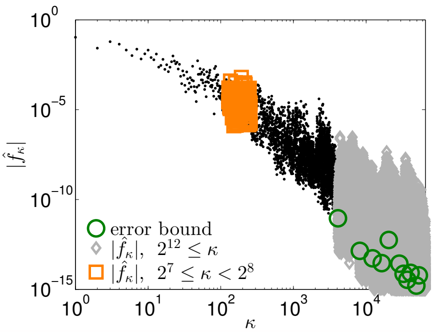
$$\mathbb{E}\left[\left(I(f) - \frac{1}{2^m}\sum_{i=0}^{2^m-1} f(\boldsymbol{z}_i)\right)^2\right] \leqslant \frac{1}{2^m}\overbrace{\sum_{\kappa\neq 0}\left|\hat{f}_\kappa\right|^2}^{\mathrm{Var}[f(\boldsymbol{Z})]}$$

- $\boldsymbol{z}_i$ are IID samples with distribution $U[0,1]^d$.
- $f(\boldsymbol{z}) = \sum_\kappa \hat{f}_\kappa \phi(\boldsymbol{z})$ is the Fourier decomposition of $f(\boldsymbol{z})$.

QMC Error Bounds Depend Only on Some Fourier Coefficients for $f \in \mathcal{C}$

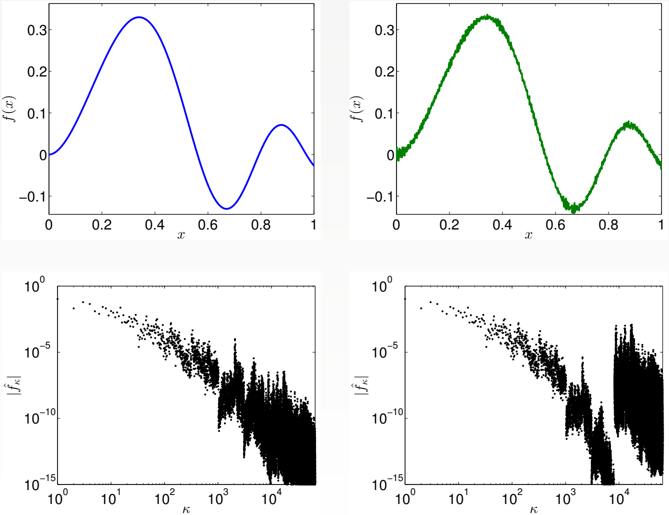
In the quasi-Monte Carlo case,

$$\left| I(f) - \frac{1}{2^m} \sum_{i=0}^{2^m-1} f(\boldsymbol{z}_i \oplus \boldsymbol{\Delta}) \right| \leqslant \sum \bigcirc \leqslant a(r,m) \sum_{\kappa=\lfloor 2^{m-r-1} \rfloor}^{2^{m-r}-1} \left| \tilde{f}_{m,\kappa} \right|$$



$$\mathcal{C} = \left\{ \begin{array}{l} \sum \square \text{ bounds } \sum \diamond \\ \sum \diamond \text{ bounds } \sum \bigcirc \end{array} \right\}$$

Inside and Outside $\mathcal{C}$



Outline

- Introduction
- Multivariate Normal Probability—Improvements by using Genz (1992) transformation
- Sobol Indices
- Efficiency improvements
  - Importance Sampling
  - Control Variates

Multivariate Normal Probability

Given $\Sigma \in \mathbb{R}^{d\times d}$ semi-positive definite, we are interested in approximating

$$\mathbb{P}\left(\boldsymbol{a} \leqslant \boldsymbol{X} \leqslant \boldsymbol{b}\right) = \int_{\boldsymbol{a}}^{\boldsymbol{b}} \frac{\mathrm{e}^{\boldsymbol{x}^t \Sigma^{-1} \boldsymbol{x}}}{(2\pi)^{d/2} \left|\Sigma\right|^{1/2}} \, \mathrm{d}\boldsymbol{x}.$$

This can be done by evaluating $\mathbb{P}\left(\boldsymbol{a} \leqslant \boldsymbol{X} \leqslant \boldsymbol{b}\right)$ as

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{a} \leqslant \boldsymbol{X} \leqslant \boldsymbol{b}\right) &= \int_{[0,1)^d} f_{\text{affine}}(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}, \quad x_j = a_j + (b_j - a_j)t_j \ , \\ &= \int_{[0,1)^{d-1}} f_{\text{Genz}}(\boldsymbol{t}) \, \mathrm{d}\boldsymbol{t}, \quad \text{using Genz (1992) transformation.}\end{aligned}$$

Multivariate Normal Probability in Dimension 8

For

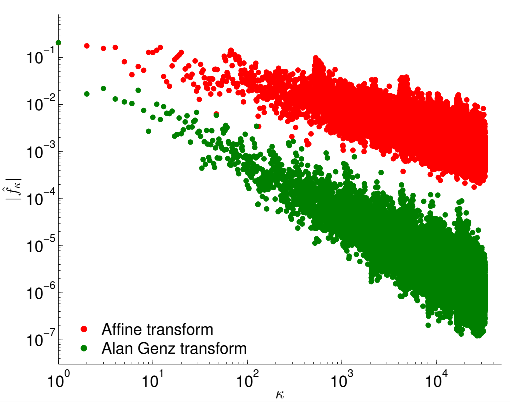
- $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^8$ chosen randomly,
- $\Sigma = 0.4\,\mathsf{I} + 0.6\,\mathbf{1}\mathbf{1}^T$,
- $\varepsilon_a = 10^{-4}$,

we found that $\mathbb{P}\left(\boldsymbol{a} \leqslant \boldsymbol{X} \leqslant \boldsymbol{b}\right)$ takes

- $8,388,608$ points and 13 seconds for f_{affine}.
- $8,192$ points and 0.05 seconds for f_{Genz}.

Multivariate Normal Probability in Dimension 8

$\left|\hat{f}_{\mathrm{affine},\kappa}\right|$ decreases much slower than $\left|\hat{f}_{\mathrm{Genz},\kappa}\right|$, leading to a smaller sample size needed.



Outline

- Introduction
- Multivariate Normal Probability
- Sobol Indices—Computing how important each dimension is
- Efficiency improvements
  - Importance Sampling
  - Control Variates

Computing Sobol Indices

For $\boldsymbol{X} \sim U[0,1]^d$, we are interested in estimating the *first-order* and *total-effect* indices,

$$S_j = \frac{\mathrm{Var}\left[\mathbb{E}\left(f(\boldsymbol{X})|\boldsymbol{X}_j\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)} = 1 - \frac{\mathbb{E}\left[\mathrm{Var}\left(f(\boldsymbol{X})|\boldsymbol{X}_j\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)},$$

$$S_j^{\text{tot}} = 1 - \frac{\mathrm{Var}\left[\mathbb{E}\left(f(\boldsymbol{X})|\boldsymbol{X}_{-j}\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)} = \frac{\mathbb{E}\left[\mathrm{Var}\left(f(\boldsymbol{X})|\boldsymbol{X}_{-j}\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)}.$$

satisfying $0 \leqslant S_j \leqslant S_j^{\text{tot}} \leqslant 1$.

Computing Sobol Indices

For $\boldsymbol{X} \sim U[0,1]^d$, we are interested in estimating the *first-order* and *total-effect* indices,

$$S_j = \frac{\mathrm{Var}\left[\mathbb{E}\left(f(\boldsymbol{X})|\boldsymbol{X}_j\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)} = 1 - \frac{\mathbb{E}\left[\mathrm{Var}\left(f(\boldsymbol{X})|\boldsymbol{X}_j\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)},$$

$$S_j^{\text{tot}} = 1 - \frac{\mathrm{Var}\left[\mathbb{E}\left(f(\boldsymbol{X})|\boldsymbol{X}_{-j}\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)} = \frac{\mathbb{E}\left[\mathrm{Var}\left(f(\boldsymbol{X})|\boldsymbol{X}_{-j}\right)\right]}{\mathrm{Var}\left(f(\boldsymbol{X})\right)}.$$

satisfying $0 \leqslant S_j \leqslant S_j^{\text{tot}} \leqslant 1$. The *first-order* index is composed by,

$$S_j = 1 - \frac{I^{(1)}}{I^{(2)} - I^{(3)}}, \quad \text{where} \begin{cases} I^{(1)} \text{ is a } 2d-1 \text{ dim. integral.} \\ I^{(2)} \text{ is a } d \text{ dim. integral.} \\ I^{(3)} \text{ is a } d \text{ dim. integral.} \end{cases}$$

Error bounds for S_j require more care than error bounds for $I^{(k)}$.

Bratley et al. (1992)

$$f(\boldsymbol{X}) = \sum_{i=1}^{6} (-1)^i \prod_{j=1}^{i} \boldsymbol{X}_j = -\boldsymbol{X}_1 + \boldsymbol{X}_1\boldsymbol{X}_2 - \boldsymbol{X}_1\boldsymbol{X}_2\boldsymbol{X}_3 + \dots$$

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
S_j	0.653	0.179	0.037	0.013	0.001	0.001
S_j^{tot}	0.740	0.266	0.077	0.034	0.007	0.007

Computation time was less than 2 seconds for $\varepsilon_a = 10^{-3}$.

Multivariate Normal Probability

$$\mathbb{P}\left(\boldsymbol{a} \leqslant \boldsymbol{X} \leqslant \boldsymbol{b}\right) = \int_{\boldsymbol{a}}^{\boldsymbol{b}} \frac{\mathrm{e}^{\boldsymbol{x}^t \Sigma^{-1} \boldsymbol{x}}}{(2\pi)^{d/2} \left|\Sigma\right|^{1/2}} \, \mathrm{d}\boldsymbol{x}.$$

f_{affine}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$
S_j	0.003	0.020	0.008	0.000	0.025	0.021	0.014	0.028
S_j^{tot}	0.174	0.467	0.265	0.048	0.535	0.497	0.407	0.570

f_{Genz}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$
S_j	0.379	0.229	0.036	0.000	0.000	0.000	0.000
S_j^{tot}	0.713	0.550	0.097	0.000	0.001	0.000	0.000

Arithmetic Mean Asian Call Option

For the payoff of an arithmetic mean Asian option with $S_0 = 100$, $K = 200$, $r = 2\%$, $\sigma = 50\%$, $T = 1$, generating the Brownian motion using PCA with $d = 6$ time steps, and $\varepsilon_a = 10^{-3}$,

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
S_j	0.992	0.000	0.000	0.000	0.000	0.000
S_j^{tot}	1.000	0.007	0.002	0.001	0.001	0.000

Outline

- Introduction
- Multivariate Normal Probability
- Sobol Indices
- Efficiency improvements
  - Importance Sampling—How is importance sampling affecting the adaptive algorithm
  - Control Variates

Arithmetic Mean Asian Call Option

Suppose that the price of an option is $\mathbb{E}\left[f_{\text{payoff}}(\boldsymbol{Z})\right]$ where $\boldsymbol{Z} \sim N(\boldsymbol{0}, \Sigma)$. Choose $\boldsymbol{h}$ such that $\boldsymbol{Z} = \boldsymbol{h}(\boldsymbol{X})$ with $\boldsymbol{X} \sim U[0,1]^d$. Then,

$$\text{Price} = \int_{[0,1)^d} \underbrace{f_{\text{payoff}}\left(\boldsymbol{h}(\boldsymbol{x})\right)}_{f(\boldsymbol{x})} \mathrm{d}\boldsymbol{x} =$$

$$\int_{[0,1)^d} \underbrace{f_{\text{payoff}}\left(\boldsymbol{\mu} + \boldsymbol{h}(\boldsymbol{x})\right) \mathrm{e}^{-\boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}/2 + \boldsymbol{h}(\boldsymbol{x}))}}_{f(\boldsymbol{x},\boldsymbol{\mu})} \mathrm{d}\boldsymbol{x}.$$

The price of the previous arithmetic mean Asian option, now considering $d = 52$ time steps, takes

- $131,072$ points and 1.6 seconds for $\boldsymbol{\mu} = \boldsymbol{0}$.
- $16,384$ points and 0.2 seconds for $\boldsymbol{\mu} = 1.5(T/d, 2T/d, \dots, T)$.

Outline

- Introduction
- Multivariate Normal Probability
- Sobol Indices
- Efficiency improvements
  - Importance Sampling
  - Control Variates—Alternative control variates based on the adaptive algorithm

Control Variates Introduction

If $I(g) = \int_{[0,1)^d} g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ is known, then

$$\int_{[0,1)^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \int_{[0,1)^d} f(\boldsymbol{x}) + \beta(I(g) - g(\boldsymbol{x}))\,\mathrm{d}\boldsymbol{x},$$

and, based on our cubatures, we choose β such that

$$\sum_{\kappa=\lfloor 2^{m-1}\rfloor}^{2^m-1} \left|\hat{f}_\kappa - \beta\hat{g}_\kappa\right| \overset{m\to\infty}{\Longrightarrow} 0 \quad \text{faster than} \quad \sum_{\kappa=\lfloor 2^{m-1}\rfloor}^{2^m-1} \left|\hat{f}_\kappa\right| \overset{m\to\infty}{\Longrightarrow} 0.$$

This choice is not necessarily

$$\beta = \frac{\mathrm{cov}(f(\boldsymbol{X}), g(\boldsymbol{X}))}{\mathrm{Var}(g(\boldsymbol{X}))} = \operatorname*{argmin}_b \sum_{\kappa=0}^{\infty} \left|\hat{f}_\kappa - b\hat{g}_\kappa\right|^2,$$

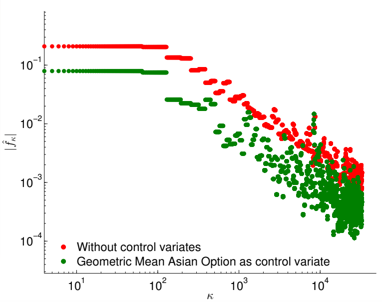
as noted by Hickernell et al. (2005).

Arithmetic Mean Asian Call Option

For the same configuration as in the importance sampling example and using the geometric mean Asian call option payoff as control variate, computing the price takes

- $131,072$ points and 2.2 seconds without control variates.
- $32,768$ points and 0.9 seconds with control variates.

$\left|\hat{f}_\kappa\right|$ are smaller than $\left|\hat{f}_\kappa - \beta\hat{g}_\kappa\right|$, leading to a smaller sample size needed.



Discussion

- These examples are a work in progress.
- We can handle error tolerances of the form $\max(\varepsilon_a, \varepsilon_r\,|I|)$.
- Examples can inform the default parameters that define $\mathcal{C}$.

Future work includes

- Extending the Sobol Indices algorithm to any set of dimensions.
- Designing an adaptive and automatic multi-level quasi-Monte Carlo algorithm.
- Adapting the conditions that define the cone $\mathcal{C}$ when the data-based necessary conditions for $f \in \mathcal{C}$ are violated.

References I

Bratley, P., B. L. Fox, and H. Niederreiter. 1992. *Implementation and tests of low-discrepancy sequences*, ACM Trans. Model. Comput. Simul. **2**, 195–213.

Genz, A. 1992. *Numerical computation of multivariate normal probabilities*, J. Comput. Graph. Statist. **1**, 141–150.

Hickernell, F. J. and Ll. A. Jiménez Rugama. 2014. *Reliable adaptive cubature using digital sequences*. submitted for publication, arXiv:1410.8615 [math.NA].

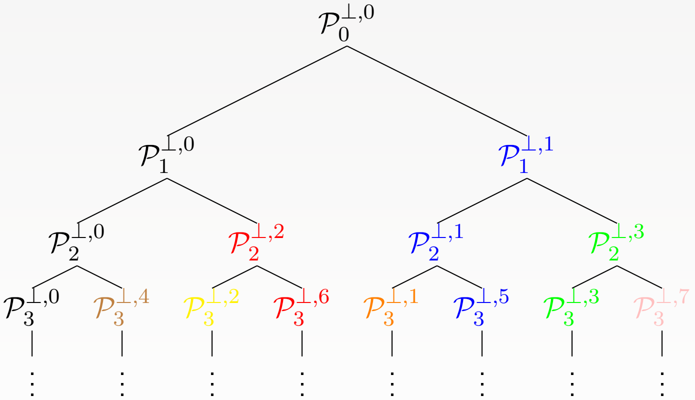
Hickernell, F. J., C. Lemieux, and A. B. Owen. 2005. *Control variates for quasi-Monte Carlo*, Statist. Sci. **20**, 1–31.

I.M., Sobol. 2001. *Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates*, Mathematics and Computers in Simulation (MATCOM) **55**, no. 1, 271–280.

Jiménez Rugama, Ll. A. and F. J. Hickernell. 2014. *Adaptive multidimensional integration based on rank-1 lattices*. submitted for publication, arXiv:1411.1966.

Owen, Art B. 2013. *Variance components and generalized Sobol' indices*, SIAM/ASA Journal on Uncertainty Quantification **1**, no. 1, 19–41.

Mapping Properties for Embedded Sequences



$\tilde{\boldsymbol{k}}(0)$	$\tilde{\boldsymbol{k}}(1)$	$\tilde{\boldsymbol{k}}(2)$	$\tilde{\boldsymbol{k}}(3)$	$\tilde{\boldsymbol{k}}(4)$	$\tilde{\boldsymbol{k}}(5)$	$\tilde{\boldsymbol{k}}(6)$	$\tilde{\boldsymbol{k}}(7)$

Alan Genz's Transform

Be $\Sigma = CC'$ the Cholesky decomposition. Furthermore,

$$a_1' = a_1/c_{11}, \qquad b_1' = b_1/c_{11}, \qquad q_1 = \phi(a_1'), \qquad e_1 = \phi(b_1')$$

where ϕ is the univariate standard normal probability distribution function. If for $i = 2, \dots, d$ we recursively define,

$$y_{i-1} = \phi^{-1}(q_{i-1} + w_{i-1}(e_{i-1} - q_{i-1})), \tag{1}$$

$$a_i' = \frac{a_i - \sum_{j=1}^{i-1} c_{ij} y_j}{c_{ii}}, \qquad b_i' = \frac{b_i - \sum_{j=1}^{i-1} c_{ij} y_j}{c_{ii}}, \tag{2}$$

$$q_i = \phi(a_i'), \qquad e_i = \phi(b_i'). \tag{3}$$

Sobol Indices Error Bound

We identified $\widehat{S}$ such that

$$\widehat{S}_j(b_1,\ldots,b_4,c_1,\ldots,c_4) \leqslant S_j(a_1,\ldots,a_4) \leqslant \widehat{S}_j(c_1,\ldots,c_4,b_1,\ldots,b_4)$$

for all $b_\ell \leqslant a_\ell \leqslant c_\ell$. Then, when one has

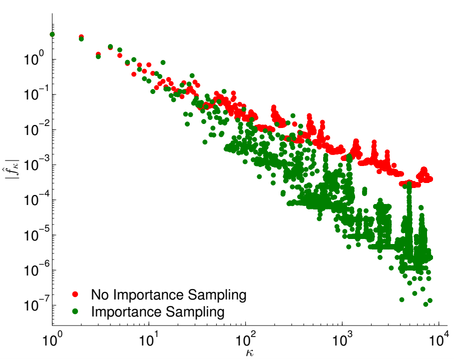
$$\widehat{S}_j(c_1,\ldots,c_4,b_1,\ldots,b_4) - \widehat{S}_j(b_1,\ldots,b_4,c_1,\ldots,c_4) \leqslant 2\varepsilon,$$

one may stop and choose

$$S_j(a_1,\ldots,a_4) \approx \frac{1}{2}\left(\widehat{S}_j(b_1,\ldots,b_4,c_1,\ldots,c_4) + \widehat{S}_j(c_1,\ldots,c_4,b_1,\ldots,b_4)\right)$$

European Call Option

Using as importance measure $N(0,4)$, the European Call with $S_0 = 100$, $K = 100$, $r = 2\%$, $\sigma = 5\%$, and $T = 2$ has discrete coefficients:



Alternative Optimal β

Based on our algorithm, we can find β minimizing,

$$\left| \int_{[0,1)^d} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \frac{1}{2^m} \sum_{i=0}^{2^m-1} f(\boldsymbol{z}_i \oplus \boldsymbol{\Delta}) + \beta(I(g) - g(\boldsymbol{z}_i \oplus \boldsymbol{\Delta})) \right|$$

$$\leqslant a(r,m) \sum_{\kappa=\lfloor 2^{m-r-1} \rfloor}^{2^{m-r}-1} \left| \tilde{f}_{m,\kappa} - \beta \tilde{g}_\kappa \right|.$$

Note that $\beta = 0$ corresponds to the initial problem and that this problem has an optimal solution.

Arithmetic Mean Asian Call Option