

Monte Carlo tree search for generating vectors of lattice rules

Manoj Palani¹

(Received 15 December 2020; revised 10 January 2022)

Abstract

Lattice rules are widely studied in the context of quasi-Monte Carlo methods as a means to achieve a small integration error. The rules themselves are determined completely by so called generating vectors, so there is an interest in methods for constructing vectors that perform well. This article introduces a new component-wise construction of a generating vector using the principles of Monte Carlo tree search, with the goal of avoiding local optima. Error bounds are proven for the vectors obtained from this method, which are analogous to existing results for the popular component by component construction.

Contents

1 Introduction

C226

DOI:10.21914/anziamj.v62.16070, © Austral. Mathematical Soc. 2022. Published 2022-02-24, as part of the Proceedings of the 19th Biennial Computational Techniques and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to the DOI for this article.

1	<i>Introduction</i>	C226
2	Monte Carlo tree search	C228
2.1	Description	C228
2.2	Application to lattice rules	C228
3	Numerical Results	C237

1 Introduction

Lattice rules, belonging to the wider quasi-Monte Carlo family, are an approach to feasibly tackle numeric integration in higher dimensions. Our starting point is an integral over the s -dimensional unit cube with $s > 1$:

$$I_s(f) = \int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x}, \tag{1}$$

which rank-1 lattice rules approximate with an equal weight n -point cubature rule

$$Q_{n,s}(f) = \frac{1}{n} \sum_{k=0}^{n-1} f(\mathbf{t}_k), \tag{2}$$

where the points are deterministically given by

$$\mathbf{t}_k = \frac{k\mathbf{z}n}{n}, \tag{3}$$

for a *generating vector* $\mathbf{z} \in \mathbb{U}_n^s$, where \mathbb{U}_n is the multiplicative group of integers modulo n and xn denotes the remainder when x is divided by n , which here has the effect of keeping the points within the unit cube. In general, every lattice rule can be written as a multiple sum involving one or more generating vectors, with the minimal number of generating vectors required known as the rank of the rule. In this case, *rank-1* and the definition indicate that the rule is completely specified by a single generating vector, so finding a good lattice rule giving a low integration error is equivalent to finding a good generating vector.

Due to the size of the search space, an exhaustive search is not feasible as s and n grow, and the standard approach is to use the *component by component* (CBC) construction, popularised by Sloan and Restzov [7] and originally introduced by Korobov [4]. Favourable properties of this construction have since been proven [1, 5, among others] and a fast construction was proposed [6], all of which have cemented this method as a go-to for constructing rank one lattice rules.

For all its merits, an inherent drawback of the CBC construction is the *greedy* nature of the construction algorithm, which makes it susceptible to favouring local over global optima. For instance, suppose \mathbf{z}^* is a global optimiser of the full s dimensional error criterion and that the CBC method has constructed $(z_1^*, \dots, z_{d-1}^*)$ for some $d < s$. Then in the case that \mathbf{z}_d^* is not the optimiser of the d -dimensional error, the optimal solution \mathbf{z}_d^* will be passed over and \mathbf{z}^* cannot be obtained after this step in the construction.

Monte Carlo tree search (MCTS) appears a good candidate for addressing this concern as it takes ‘future’ outcomes into consideration as opposed to a strictly local optimality criterion. This article applies the principles of MCTS to the component wise construction of a generating vector to introduce a new method for constructing generating vectors of rank-1 lattice rules. We work in the ‘standard setting’ with the shift averaged worst case error criterion and weighted reproducing kernel Hilbert spaces, which were originally introduced by Sloan and Woźniakowski [8]. In particular, the shift averaged worst case error for a rank-1 lattice rule specified by generating vector \mathbf{z} in the weighted Sobolev space is

$$[e_{n,s}^{\text{sh}}(\mathbf{z})]^2 = \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}} \left(\frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in \mathbf{u}} \left[\omega \left(\frac{kz_j \bmod n}{n} \right) + \beta \right] - \beta^{|\mathbf{u}|} \right), \quad (4)$$

where $\gamma_{\mathbf{u}}$ are weights and $\omega(x) = x^2 - x + \frac{1}{6}$ (the second Bernoulli polynomial), with $\beta = c^2 - c + \frac{1}{3}$ for the anchored space with anchor \mathbf{c} , and $\beta = 0$ for the unanchored space.

2 Monte Carlo tree search

2.1 Description

Monte Carlo tree search is a heuristic search algorithm which is used in many applied computer science problems, most famously in computer programs for playing turn based games such as *Chess*, *Shogi* or *Go*. One of the most remarkable feats in this arena has been the victory of *Alpha Go*, Google's Go playing program, over the human world champion Lee Sedol in 2016. The sequential decision making in such games makes them readily visualisable as *trees*, with nodes representing a valid position or state in the game and decisions at each turn represented by branches extending out from parent to child nodes.

The MCTS algorithm aims to find optimal decisions in search spaces by taking random samples in the decision space and progressing the search according to the results. Loosely speaking, at each stage it picks the choice which will give it the best chance of success in the future. This is achieved by 'playing out' each possible decision by randomly simulating a game to the end from that node, and updating beliefs about each node's chance of success based on those results.

2.2 Application to lattice rules

There are many similarities between the turn based game problem and the task of constructing a good generating vector for rank-1 lattice rules. Any component wise construction can be viewed as a tree search in which we want to find the optimal path through the tree; that is, the sequential choice of components which will give us a low worst case error at the end. A search of this tree in the vein of MCTS, where at each dimension we choose the component which gives us the best chance of this low final error, appears a sensible way of finding this path.

As mentioned previously, a reason to investigate a method which has some

consideration of the final outcome is the greedy nature of the CBC construction. In the game of chess, for instance, capturing an opponent's piece when it is available is not always the best choice, as it may result in a worse positioning of the player's pieces for upcoming turns.

Similarly, we can envisage situations where a component of a generating vector is chosen to satisfy the local optimality criterion that may prohibit otherwise good choices that would have been available for the full generating vector. It is with these insights then that we introduce the following construction.

1. Generate $\mathbf{t}^1, \dots, \mathbf{t}^R$ uniformly from \mathbb{U}_n^s . We refer to these as *tails* because their components are appended with vectors of smaller dimension to form a vector of dimension s . Let superscripts denote an index within this set and subscripts denote a component of a tail. That is, \mathbf{t}_d^r specifies the d th component of the r th tail \mathbf{t}^r .
2. For each $d = 1, \dots, s$ choose the component $z_d \in \mathbb{U}_n$ which has the lowest average squared worst case error over the tails. That is,

$$z_d^* = \operatorname{argmin}_{z_d \in \mathbb{U}_n} \frac{1}{R} \sum_{r=1}^R [e_{n,s}^{\text{sh}}(z_1^*, \dots, z_{d-1}^*, z_d, \mathbf{t}_{d+1}^r, \mathbf{t}_{d+2}^r, \dots, \mathbf{t}_s^r)]^2. \quad (5)$$

In analogy to MCTS, at each stage the randomly generated \mathbf{t}^r serve as random completions to the vector that has been constructed so far and we 'play out' each choice by looking at the full s -dimensional worst case error given a particular completion. Then we consider the component which gives the lowest average error across all the completions as being the one which gives us the best chance of success.

Naturally other choices for z_d^* are possible here. For example, at each step one could choose the component z_d^* which produces the minimum error across any of the tails. However, the choice of a simple average lends itself well to mathematical arguments in proofs. Note also that the use of R tails nominally increases the construction cost by a factor of R .

We now prove some error bounds of the described construction, which are probabilistic due to the inherent randomness. First we prove that in the simple case of a prime number of cubature points n and product weights, the construction has an expected squared worst case error smaller than the quasi-Monte Carlo (QMC) mean, which is analogous to Theorem 5.7 of Kuo and Sloan [2] .

Theorem 1. *Consider the weighted Sobolev space from (4) (either anchored or unanchored) with product weights $\gamma_1 \geq \gamma_2 \geq \dots > 0$ and prime*

$$n \geq \frac{\gamma_1}{6(1 + \gamma_1\beta)}. \tag{6}$$

If \mathbf{z}^* is constructed according to the above method, then

$$\mathbb{E} [e_{n,s}^{\text{sh}}(z_1^*, \dots, z_d^*, t_{d+1}, \dots, t_s)]^2 < \frac{1}{n} \left[\prod_{j=1}^s \left(\beta_j + \frac{\gamma_j}{6} \right) - \bar{\beta}_s \right], \tag{7}$$

for $d = 1, \dots, s$ where $\beta_j = 1 + \gamma_j\beta$, $\bar{\beta}_s = \prod_{j=1}^s \beta_j$ and the expectation is taken over $\mathbf{t} \in \{\mathbf{t}^1, \dots, \mathbf{t}^R\}$. The right hand side of (7) is known as the QMC mean.

Proof: We proceed by induction. First consider the base case $d = 1$. For any $z_1 \in \mathbb{U}_n$ and $\mathbf{t} \in \{\mathbf{t}^1, \dots, \mathbf{t}^R\}$, by the definition of $\mathbf{z}^* = (z_1^*, \dots, z_s^*)$ we have

$$\mathbb{E} [e_{n,s}^{\text{sh}}(z_1^*, t_2, \dots, t_s)]^2 \leq \mathbb{E}_{z_1} \mathbb{E} [e_{n,s}^{\text{sh}}(z_1, t_2, \dots, t_s)]^2 = \mathbb{E} [e_{n,s}^{\text{sh}}(t_1, \dots, t_s)]^2,$$

where the last step exploits the fact that each component of \mathbf{t} is identically distributed to z_1 as uniform on \mathbb{U}_n . Now

$$[e_{n,s}^{\text{sh}}(\mathbf{t})]^2 = -\bar{\beta}_s + \frac{1}{n} \prod_{j=1}^s \left(\beta_j + \frac{\gamma_j}{6} \right) + \frac{1}{n} \sum_{k=1}^{n-1} \prod_{j=1}^s [\beta_j + \gamma_j \omega_k(t_j)],$$

where

$$\omega_k(x) = \omega\left(\frac{kx \bmod n}{n}\right),$$

and ω is the second Bernoulli polynomial $x^2 - x + \frac{1}{6}$. So taking expectations,

$$\mathbb{E} [e_{n,s}^{\text{sh}}(\mathbf{t})]^2 = -\bar{\beta}_s + \frac{1}{n} \prod_{j=1}^s \left(\beta_j + \frac{\gamma_j}{6}\right) + \frac{1}{n} \sum_{k=1}^{n-1} \prod_{j=1}^s (\beta_j + \gamma_j \mathbb{E} [\omega_k(\mathbf{y})]),$$

where \mathbf{y} is generated uniformly from \mathbb{U}_n , because the components of \mathbf{t} are i.i.d uniform on \mathbb{U}_n . Then from equation (5.17) of Kuo and Sloan [2]

$$\mathbb{E} [\omega_k(\mathbf{y})] = \frac{1}{n-1} \sum_{y=1}^{n-1} \omega_k(y) = -\frac{1}{6n},$$

we get

$$\begin{aligned} \mathbb{E} [e_{n,s}^{\text{sh}}(\mathbf{t})]^2 &= -\bar{\beta}_s + \frac{1}{n} \prod_{j=1}^s \left(\beta_j + \frac{\gamma_j}{6}\right) + \frac{1}{n} \sum_{k=1}^{n-1} \prod_{j=1}^s \left(\beta_j - \frac{\gamma_j}{6n}\right) \\ &= -\bar{\beta}_s + \frac{1}{n} \prod_{j=1}^s \left(\beta_j + \frac{\gamma_j}{6}\right) + \frac{n-1}{n} \prod_{j=1}^s \left(\beta_j - \frac{\gamma_j}{6n}\right). \end{aligned}$$

Defining κ as the QMC mean,

$$\begin{aligned} \mathbb{E} [e_{n,s}^{\text{sh}}(z_1^*, t_2, \dots, t_s)]^2 - \kappa &\leq \left(-1 + \frac{1}{n}\right) \bar{\beta}_s + \frac{n-1}{n} \prod_{j=1}^s \left(\beta_j - \frac{\gamma_j}{6n}\right) \\ &= \frac{n-1}{n} \left[\prod_{j=1}^s \left(\beta_j - \frac{\gamma_j}{6n}\right) - \prod_{j=1}^s \beta_j \right]. \end{aligned}$$

As the function $x \mapsto x/[6(1+kx)]$ is increasing for all $k \in \mathbb{R}$, the assumption (6) on the weights ensures

$$n \geq \frac{\gamma_1}{6(1+\gamma_1\beta)} \geq \frac{\gamma_2}{6(1+\gamma_2\beta)} \geq \dots$$

so that $\beta_j = 1 + \gamma_j \beta \geq \frac{\gamma_j}{6n}$ for all j . As a result

$$\beta_j > \beta_j - \frac{\gamma_j}{6n} \geq 0,$$

for all j , which yields

$$\prod_{j=1}^s \left(\beta_j - \frac{\gamma_j}{6n} \right) < \prod_{j=1}^s \beta_j.$$

Hence

$$\mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, t_2, \dots, t_s) \right]^2 < \kappa,$$

which proves the $d = 1$ case.

Assume now the claim is true for dimension d . Then by the same argument employed earlier,

$$\begin{aligned} & \mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, \dots, z_d^*, z_{d+1}^*, t_{d+2}, \dots, t_s) \right]^2 \\ & \leq \mathbb{E}_{z_{d+1}} \mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, \dots, z_d^*, z_{d+1}, t_{d+2}, \dots, t_s) \right]^2 \\ & = \mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, \dots, z_d^*, t_{d+1}, t_{d+2}, \dots, t_s) \right]^2. \end{aligned}$$

But this last term is less than κ by the induction hypothesis. Hence the claim holds at dimension $d + 1$ and the result is proven by induction. ♠

We now look to prove an error bound for the unanchored Sobolev space, which applies for general weights and general n . We first prove some auxiliary lemmas that provide the estimates needed for this result.

Lemma 2. For $\mathbf{t} \in \mathbb{U}_n^{|\mathbf{u}|}$, $\mathbf{u} \subseteq \{1 : s\}$ and $\lambda \in (1/2, 1]$ we have

$$\left[\frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in \mathbf{u}} \omega_k(t_j) \right]^\lambda \leq \frac{1}{(2\pi^2)^{\lambda|\mathbf{u}|}} \sum_{\mathbf{h}_{\mathbf{u}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_{\mathbf{u}} \cdot \mathbf{t}_{\mathbf{u}} \equiv 0 \pmod n\}}. \quad (8)$$

Proof: For any $k = 0, 1, \dots, n - 1$ and $x \in [0, 1]$ the Fourier expansion of ω :

$$\omega_k(x) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z} \setminus \{0\}} \frac{e^{(2\pi i h k x)/n}}{h^2},$$

gives

$$\begin{aligned} \frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in \mathbf{u}} \omega_k(t_j) &= \frac{1}{(2\pi^2)^{|\mathbf{u}|}} \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\mathbf{h}_{\mathbf{u}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|}} \frac{e^{(2\pi k i \mathbf{h}_{\mathbf{u}} \cdot \mathbf{t}_{\mathbf{u}})/n}}{\prod_{j \in \mathbf{u}} h_j^2} \\ &= \frac{1}{(2\pi^2)^{|\mathbf{u}|}} \sum_{\mathbf{h}_{\mathbf{u}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} h_j^2} \mathbb{1}_{\{\mathbf{h}_{\mathbf{u}} \cdot \mathbf{t}_{\mathbf{u}} \equiv 0 \pmod{n}\}}, \end{aligned}$$

by the character property in equation (5.2) of Kuo and Sloan [2]. Thus the result follows upon exponentiating both sides and applying the inequality

$$\left(\sum_k a_k \right)^\lambda \leq \sum_k a_k^\lambda, \quad a_k \geq 0, \quad \lambda \in (0, 1]. \tag{9}$$



Lemma 3. Let $\mathbf{u} \subseteq \{1 : s\}$ and \mathbf{t} have a uniform distribution on \mathbb{U}_n^s . Then

$$\mathbb{E} \left[\sum_{\mathbf{h}_{\mathbf{u}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|}} \frac{1}{\prod_{j \in \mathbf{u}} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_{\mathbf{u}} \cdot \mathbf{t}_{\mathbf{u}} \equiv 0 \pmod{n}\}} \right] \leq \frac{2}{\varphi(n)} [2\zeta(2\lambda)]^{|\mathbf{u}|}, \tag{10}$$

where φ is Euler's totient function

$$\varphi(n) := |\mathbb{U}_n|, \tag{11}$$

and ζ is the Riemann zeta function

$$\zeta(x) := \sum_{h=1}^{\infty} \frac{1}{h^x}, \quad x > 1. \tag{12}$$

Proof: Let $\mathbf{n} \mid \mathbf{h}_u$ denote $\mathbf{n} \mid \mathbf{h}_j$ for all $j \in u$ and write $\mathbf{n} \nmid \mathbf{h}_u$ to mean there exists $\ell \in u$ such that $\mathbf{n} \nmid \mathbf{h}_\ell$. Expanding

$$\begin{aligned} & \mathbb{E} \left[\sum_{\mathbf{h}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{\mathbf{n}}\}} \right] \\ &= \frac{1}{\varphi(\mathbf{n})^{|u|}} \sum_{\mathbf{t}_u \in \mathbb{U}_n^{|u|}} \sum_{\mathbf{h}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{\mathbf{n}}\}} \\ &= \frac{1}{\varphi(\mathbf{n})^{|u|}} \left(\sum_{\mathbf{t}_u \in \mathbb{U}_n^{|u|}} \sum_{\mathbf{h}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{\mathbf{n}}, \text{ and } \mathbf{n} \mid \mathbf{h}_u\}} \right) \\ & \quad + \frac{1}{\varphi(\mathbf{n})^{|u|}} \left(\sum_{\mathbf{t}_u \in \mathbb{U}_n^{|u|}} \sum_{\mathbf{h}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{\mathbf{n}}, \text{ and } \mathbf{n} \nmid \mathbf{h}_u\}} \right). \end{aligned}$$

We now consider these two terms (call them S_1 and S_2) separately.

For the first term, as $\mathbf{n} \mid \mathbf{h}_u$ guarantees $\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{\mathbf{n}}$ for all \mathbf{t} , the average over \mathbf{t}_u vanishes to leave

$$S_1 = \sum_{\mathbf{h}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{n} \mid \mathbf{h}_u\}} = \sum_{\mathbf{m}_u \in (\mathbb{Z} \setminus \{0\})^{|u|}} \frac{1}{\mathbf{n}^{2\lambda} \prod_{j \in u} m_j^{2\lambda}}.$$

Writing this as a multiple sum and distributing the terms of the product accordingly,

$$S_1 = \frac{1}{\mathbf{n}^{2\lambda}} \sum_{m_{u_1} \in \mathbb{Z}_0} \frac{1}{m_{u_1}^{2\lambda}} \cdots \sum_{m_{u_{|u|}} \in \mathbb{Z}_0} \frac{1}{m_{u_{|u|}}^{2\lambda}} = \frac{1}{\mathbf{n}^{2\lambda}} \left(\sum_{m \in \mathbb{Z}_0} \frac{1}{m^{2\lambda}} \right)^{|u|} = \frac{[2\zeta(2\lambda)]^{|u|}}{\mathbf{n}^{2\lambda}}.$$

As $\varphi(\mathbf{n}) \leq \mathbf{n} - 1 < \mathbf{n}$, we have $\frac{1}{\varphi(\mathbf{n})} > \frac{1}{\mathbf{n}} > \frac{1}{\mathbf{n}^\lambda}$. Hence

$$S_1 \leq \frac{(2\zeta(2\lambda))^{|u|}}{\varphi(\mathbf{n})}.$$

Next, in S_2 , $\mathbf{n} \nmid \mathbf{h}_u$ says that there exists an $\ell \in u$ such that $\mathbf{n} \nmid h_\ell$. Thus

$$S_2 = \frac{1}{\varphi(\mathbf{n})^{|\mathbf{u}|}} \sum_{c=1}^{n-1} \left(\sum_{t_\ell \in \mathbb{U}_n} \sum_{\substack{h_\ell \in \mathbb{Z} \setminus \{0\} \\ h_\ell \equiv -ct_\ell^{-1} \pmod{n}}} \frac{1}{h_\ell^{2\lambda}} \right) \\ \times \left(\sum_{\mathbf{t}_{u \setminus \{\ell\}} \in \mathbb{U}_n^{|\mathbf{u}|-1}} \sum_{\substack{\mathbf{h}_{u \setminus \{\ell\}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|-1} \\ \mathbf{h}_{u \setminus \{\ell\}} \cdot \mathbf{t}_{u \setminus \{\ell\}} \equiv c \pmod{n}}} \frac{1}{\prod_{j \in u \setminus \{\ell\}} h_j^{2\lambda}} \right).$$

Now an identical argument to that in Theorem 5.8 of Kuo and Sloan [2] shows that for fixed $1 \leq c \leq n-1$ we have

$$\sum_{t_\ell \in \mathbb{U}_n} \sum_{\substack{h_\ell \in \mathbb{Z} \setminus \{0\} \\ h_\ell \equiv -ct_\ell^{-1} \pmod{n}}} \frac{1}{h_\ell^{2\lambda}} \leq 2\zeta(2\lambda).$$

Thus

$$S_2 \leq \frac{1}{\varphi(\mathbf{n})^{|\mathbf{u}|}} 2\zeta(2\lambda) \sum_{\mathbf{t}_{u \setminus \{\ell\}} \in \mathbb{U}_n^{|\mathbf{u}|-1}} \sum_{\substack{\mathbf{h}_{u \setminus \{\ell\}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|-1} \\ \mathbf{h}_{u \setminus \{\ell\}} \cdot \mathbf{t}_{u \setminus \{\ell\}} \not\equiv 0 \pmod{n}}} \frac{1}{\prod_{j \in u \setminus \{\ell\}} h_j^{2\lambda}} \\ \leq \frac{1}{\varphi(\mathbf{n})^{|\mathbf{u}|}} 2\zeta(2\lambda) \sum_{\mathbf{t}_{u \setminus \{\ell\}} \in \mathbb{U}_n^{|\mathbf{u}|-1}} \sum_{\mathbf{h}_{u \setminus \{\ell\}} \in (\mathbb{Z} \setminus \{0\})^{|\mathbf{u}|-1}} \frac{1}{\prod_{j \in u \setminus \{\ell\}} h_j^{2\lambda}} \\ = \frac{[2\zeta(2\lambda)]^{|\mathbf{u}|}}{\varphi(\mathbf{n})},$$

where the crude bounding of the restricted sum with the unrestricted sum applies because each term in the sum is non negative. Consequently

$$S_1 + S_2 \leq \frac{2}{\varphi(\mathbf{n})} [2\zeta(2\lambda)]^{|\mathbf{u}|}.$$



We are now ready to prove a second error bound, which shows that the algorithm (5) is expected to yield a convergence rate arbitrarily close to $\mathcal{O}(n^{-1})$ (this is analogous to Theorem 5.8 of Kuo and Sloan [2] for the standard component-by-component construction).

Theorem 4. *For $d = 1, \dots, s$ if z_1^*, \dots, z_{d-1}^* have been chosen according to the construction (5), then*

$$\mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, \dots, z_d^*, t_{d+1}, \dots, t_s) \right]^2 \leq \left[\frac{2}{\varphi(n)} \sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u^\lambda \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^\lambda} \right)^{|u|} \right]^{1/\lambda} \quad (13)$$

for all $\lambda \in (\frac{1}{2}, 1]$.

Proof: The proof is by induction. For the base case $d = 1$, we see that

$$\mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1^*, t_2, \dots, t_s) \right]^2 \leq \mathbb{E}_{z_1} \mathbb{E} \left[e_{n,s}^{\text{sh}}(z_1, t_2, \dots, t_s) \right]^2 = \mathbb{E} \left[e_{n,s}^{\text{sh}}(t_1, \dots, t_s) \right]^2.$$

Hence, since the function $x \mapsto x^\lambda$ is non decreasing and applying the inequality (9),

$$\mathbb{E}^\lambda \left[e_{n,s}^{\text{sh}}(z_1^*, t_2, \dots, t_s) \right]^2 \leq \mathbb{E}^\lambda \left[e_{n,s}^{\text{sh}}(t) \right]^2 \leq \mathbb{E} \left(\left[e_{n,s}^{\text{sh}}(t) \right]^2 \right)^\lambda.$$

Applying (9) again and moving the expectation inside,

$$\begin{aligned} \mathbb{E} \left(\left[e_{n,s}^{\text{sh}}(t) \right]^2 \right)^\lambda &= \mathbb{E} \left(\sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u^\lambda \frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in u} \omega_k(t_j) \right)^\lambda \\ &\leq \sum_{\emptyset \neq u \subseteq \{1:s\}} \gamma_u^\lambda \mathbb{E} \left(\frac{1}{n} \sum_{k=0}^{n-1} \prod_{j \in u} \omega_k(t_j) \right)^\lambda. \end{aligned}$$

We now apply Lemmas (2) and (3) in turn to obtain

$$\mathbb{E} \left(\left[e_{n,s}^{\text{sh}}(t) \right]^2 \right)^\lambda \leq \sum_{\emptyset \neq u \subseteq \{1:s\}} \frac{\gamma_u^\lambda}{(2\pi^2)^{|u|\lambda}} \mathbb{E} \left[\sum_{\mathbf{h}_u \in Z_0^{|u|}} \frac{1}{\prod_{j \in u} h_j^{2\lambda}} \mathbb{1}_{\{\mathbf{h}_u \cdot \mathbf{t}_u \equiv 0 \pmod{n}\}} \right]$$

$$\leq \frac{2}{\varphi(\mathbf{n})} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \right)^{|\mathbf{u}|},$$

so the claim indeed holds for the $\mathbf{d} = 1$ case.

Now we assume the statement (13) holds at dimension \mathbf{d} and consider

$$\mathbb{E} \left[\mathbf{e}_{\mathbf{n},s}^{\text{sh}}(z_1^*, \dots, z_{\mathbf{d}}^*, z_{\mathbf{d}+1}^*, \mathbf{t}_{\mathbf{d}+2}, \dots, \mathbf{t}_s) \right]^2.$$

Using the previous argument,

$$\begin{aligned} & \mathbb{E} \left[\mathbf{e}_{\mathbf{n},s}^{\text{sh}}(z_1^*, \dots, z_{\mathbf{d}}^*, z_{\mathbf{d}+1}^*, \mathbf{t}_{\mathbf{d}+2}, \dots, \mathbf{t}_s) \right]^2 \\ & \leq \mathbb{E}_{z_{\mathbf{d}+1}} \mathbb{E} \left[\mathbf{e}_{\mathbf{n},s}^{\text{sh}}(z_1^*, \dots, z_{\mathbf{d}}^*, z_{\mathbf{d}+1}, \mathbf{t}_{\mathbf{d}+2}, \dots, \mathbf{t}_s) \right]^2 \\ & \leq \mathbb{E} \left[\mathbf{e}_{\mathbf{n},s}^{\text{sh}}(z_1^*, \dots, z_{\mathbf{d}}^*, \mathbf{t}_{\mathbf{d}+1}, \mathbf{t}_{\mathbf{d}+2}, \dots, \mathbf{t}_s) \right]^2. \end{aligned}$$

But this is nothing more than the left hand side of the statement in the previous dimension $\mathbf{d} - 1$, which satisfies the induction hypothesis. Hence

$$\mathbb{E} \left[\mathbf{e}_{\mathbf{n},s}^{\text{sh}}(z_1^*, \dots, z_{\mathbf{d}}^*, z_{\mathbf{d}+1}^*, \mathbf{t}_{\mathbf{d}+2}, \dots, \mathbf{t}_s) \right]^2 \leq \left[\frac{2}{\varphi(\mathbf{n})} \sum_{\emptyset \neq \mathbf{u} \subseteq \{1:s\}} \gamma_{\mathbf{u}}^{\lambda} \left(\frac{2\zeta(2\lambda)}{(2\pi^2)^{\lambda}} \right)^{|\mathbf{u}|} \right]^{1/\lambda},$$

and our result is true by mathematical induction. ♠

The error bound (13) holds for all $\lambda \in (1/2, 1]$, so we obtain the optimal convergence close to $\mathcal{O}(\mathbf{n}^{-1})$ as $\lambda \rightarrow 1/2$. The case $\lambda = 1/2$ is not possible as the \mathbf{p} -series defined by the Riemann zeta function only converges for $\lambda > 1/2$.

3 Numerical Results

We consider the application of this construction to the problem of pricing an Asian option, which is well investigated within the QMC literature [e.g., 3] and consider how our construction performs in comparison to the vectors

generated by the usual CBC method, and across the Brownian Bridge and Principal Components covariance matrix factorisations which are frequently used for QMC.

In producing the generating vectors, parameters for the problem and the underlying function spaces were set. Firstly, $R = 50$ was used in the new construction. Next, it was suggested by Sloan and Wang [9] that optimal weights for lattice rules applied to pricing a path dependent option could be of the form $\gamma_j = \alpha^j$, so we consider the product weights $\gamma_j = 0.5^j$ (and in the unanchored space for simplicity).

We use the single asset path dependent formulation set out by Kuo et al. [3], which assumes that the price of the stock underlying the option follows a geometric Brownian motion with $d = 256$ equally spaced time steps in the time interval $[0, T]$ with $T = 1$, so that the stock price at time $t_j = jT/d$ is

$$S_j = S_0 \exp \left[\left(r - \frac{1}{2} \sigma^2 \right) t_j + \sigma \mathbf{w}_j \right], \tag{14}$$

where $S_0 = 100$ is the stock price at time 0, $r = 0.1$ is the risk free interest rate, $\sigma = 0.2$ is the volatility of the stock and \mathbf{w} corresponds to a Brownian path which is normally distributed with mean zero and covariance matrix $\Sigma = [\min(t_i, t_j)]_{i,j=1}^d$. The ‘fair price’ or premium of the option is approximated by

$$\frac{1}{n} \sum_{i=0}^{n-1} g(\mathbf{A} \Phi^{-1}(\mathbf{x}^{(i)})) \tag{15}$$

where

$$g(\mathbf{w}) = \max \left(\frac{1}{d} \sum_{j=1}^d S_{t_j} - 100, 0 \right), \tag{16}$$

$\Sigma = \mathbf{A} \mathbf{A}^T$ and Φ^{-1} denotes the inverse normal cumulative distribution function applied element-wise to a vector or matrix, with the $\mathbf{x}^{(i)} \in [0, 1]^s$ being the sample points used.

We apply random shifting to the QMC lattice point sets with 35 random shifts. Figures (1) and (2) compare the error estimates for different numbers of points.

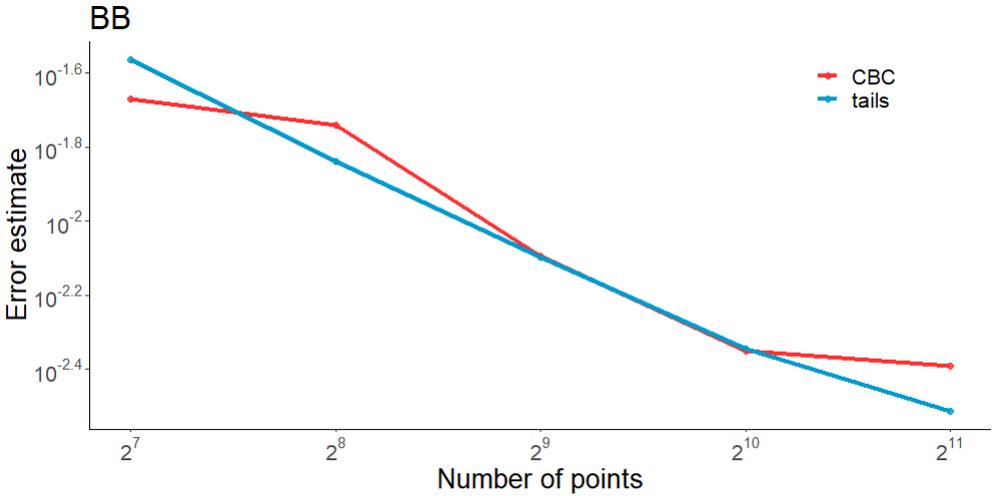


Figure 1: Comparison of error estimates with the Brownian Bridge factorisation of the covariance matrix.

These results show that the lattice point set using the generating vector constructed with the new method performs similarly to CBC for this problem. In such cases, the additional computational cost of the new construction compared to CBC makes the traditional construction a better choice. The performance could potentially improve with a larger number of tails, as 50 seems small relative to the size of \mathbb{U}_n^{256} for the choices of n here, but this would require greater computational resources.

Acknowledgements

The content of this article was part of a thesis written under the supervision of Josef Dick. Many thanks also to Josef for helpful suggestions which improved the presentation of this material.

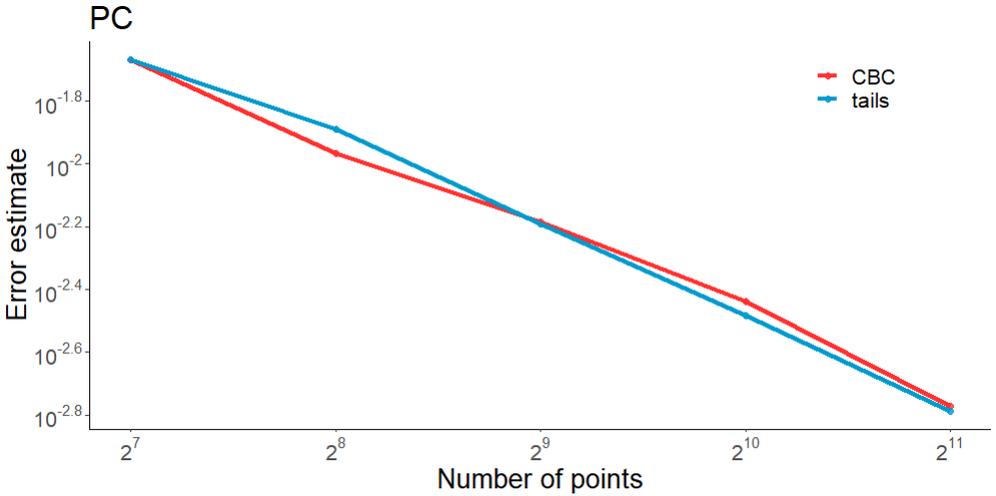


Figure 2: Comparison of error estimates with the Principal Components factorisation of the covariance matrix.

References

- [1] J. Dick. “On the convergence rate of the component-by-component construction of good lattice rules”. In: *J. Complex.* 20 (2004), pp. 493–522. DOI: [10.1016/j.jco.2003.11.008](https://doi.org/10.1016/j.jco.2003.11.008) (cit. on p. [C227](#)).
- [2] J. Dick, F. Y. Kuo, and I. H. Sloan. “High-dimensional integration: The quasi-Monte Carlo way”. In: *Acta Numer.* 22 (2013), pp. 133–288. DOI: [10.1017/S0962492913000044](https://doi.org/10.1017/S0962492913000044) (cit. on pp. [C230](#), [C231](#), [C233](#), [C235](#), [C236](#)).
- [3] M. Giles, F. Y. Kuo, I. H. Sloan, and B. J. Waterhouse. “Quasi-Monte Carlo for finance applications”. In: *ANZIAM J.* 50 (2008), pp. C308–C323. DOI: [10.21914/anziamj.v50i0.1440](https://doi.org/10.21914/anziamj.v50i0.1440) (cit. on pp. [C237](#), [C238](#)).

- [4] N. M. Korobov. “Approximate evaluation of repeated integrals”. In: *Doklady Akademii Nauk SSSR* 124 (1959), pp. 1207–1210 (cit. on p. [C227](#)).
- [5] F. Y. Kuo. “Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces”. In: *J. Complex.* 19 (2003), pp. 301–320. DOI: [10.1016/S0885-064X\(03\)00006-2](#) (cit. on p. [C227](#)).
- [6] D. Nuyens and R. Cools. “Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces”. In: *Math. Comput.* 75 (2006), pp. 903–920. DOI: [10.1090/S0025-5718-06-01785-6](#) (cit. on p. [C227](#)).
- [7] I. H. Sloan and A. V. Restzov. “Component-by-component construction of good lattice rules”. In: *Math. Comput.* 71 (2002), pp. 263–273. DOI: [10.1090/S0025-5718-01-01342-4](#) (cit. on p. [C227](#)).
- [8] I. H. Sloan and H. Woźniakowski. “When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals?” In: *J. Complex.* 14 (1998), pp. 1–33. DOI: [10.1006/jcom.1997.0463](#) (cit. on p. [C227](#)).
- [9] X. Wang and I. H. Sloan. “Efficient weighted lattice rules with applications to finance”. In: *SIAM J. Sci. Comput.* 28 (2006), pp. 728–750. DOI: [10.1137/S1064827502418197](#) (cit. on p. [C238](#)).

Author address

1. **Manoj Palani**, School of Mathematics and Statistics, University of New South Wales, NSW 2000, AUSTRALIA.
<mailto:m.palani@unswalumni.com>
orcid:[0000-0001-9217-5171](#)