Quasi-Monte Carlo integration over non-cubical domains

Nilofar Mortazavi

A thesis presented for the Master degree of Mathematical Statistics

LUND UNIVERSITY

Center of Mathematical Sciences
Mathematical Statistics
Sweden

March 1, 2017

Quasi-Monte Carlo integration over non-cubical domains

Nilofar Mortazavi

Abstract

Monte Carlo (MQ) method is a powerful tool to approximate high dimensional integrals. The disadvantage of ordinary MQ is the slow convergence rate by cause of the essential randomness of this method. Computations of this convergence can lead to pure time consuming. Quasi-Monte Carlo (QMC) method yields considerably better results since this method is a deterministic alternative which uses Low-Discrepancy sequences instead of random samples . There are a multiplicity of big open problems in QMC-methods, problems partly arising from applications and partly arising from theory.
QMC are developed to integrate over unite cube, where it has much more accuracy than MC for integrands of bounded variation. Integration over more general spaces such as triangles, disks and Cartesian products of such spaces is more challenging for QMC. Nevertheless in real-world applications various problems are defined over such spaces.
The aim of this thesis is to provide a survey of a solution of such problems of numerical integration defined over non-cubical spaces. We present QMC and randomised QMC (RQMC) constructions in the triangle with a vanishing discrepancy based on the recent work of Basu 2016. The QMC construction is a version of the Van der Corput-Halton sequences specially made to the unit triangle. The attraction of scrambled net is replication based error estimation for QMC with slightly the same accuracy as QMC, and for smooth enough integrands.

Keywords: Quasi-Monte Carlo methods, Discrepancy, Low-discrepancy sequence, Scrambled net.

Acknowledgements

I would like to express my gratitude to my supervisor Professor Magnus Wiktorsson, for the support and the time he has dedicated to me.

Contents

List of Figures

Chapter 1

Introduction

The complication of numerical integration arises in applications from finance, physics, biology, computer graphics, and others, where one need to figure out some integral for instance an average value which may not be done analytically. Thus in such a case one need to resort to numerical methods. Monte Carlo (MC) sampling plays a key role in this case. In the following we consider the standardised problem of approaching an integral of the form

$$\int_{[0,1]^s} f(x)dx.$$

We approximate the integral by choosing quadrature points $x_0, ..., x_{N-1} \in [0,1)^s$, since the volume of the unit cube $[0,1]^s$ equals one, the value of this integral is just the expectation value of the function f, i.e.

$$\int_{[0,1]^s} f(x)dx \simeq \frac{1}{N}\sum_{i=0}^{N-1} f(x_i).$$

Now the question occurs which choice of the quadrature points $x_0, ..., x_{N-1}$ gives a significant result when we estimate the absolute value of the integration error, i.e. the expectation value of $\varepsilon_{N,f}$; [Dick et al. 2013]

$$\left|\int_{[0,1]^s} f(x)dx - \frac{1}{N}\sum_{i=0}^{N-1} f(x_i)\right| =: \varepsilon_{N,f}(x_0, ..., x_{N-1}).$$

Thus there is the following Theorem.

Theorem 1.0.1. *Let $f \in \mathcal{L}^2([0,1]^s)$, then $\forall N \in \mathbb{N}$ we have:*

$$\mathbb{E}(\varepsilon_{N,f}^2) = \frac{1}{N}\int_{[0,1]^s}\left(f(x) - \int_{[0,1]^s} f(y)dy\right)^2 dx = \frac{\sigma^2(f)}{N},$$

where

$$\sigma^2(f) := I_s(f^2) - (I_s(f))^2.$$

[Dick, Kuo, and Sloan 2013]

This Theorem says that the absolute value of the integration error is bounded by $\frac{\sigma(f)}{N}$, where $\sigma(f)$ is the standard deviation of f. Remark that the integration error does not depend on the dimension s (for some functions $\sigma(f)$ may be depended on s). Since $N^{-\frac{1}{2}} < N^{-\frac{1}{s}}$ for $s > 2$, accordingly for $s > 2$ it is better, on average, to use random points for approximating of the integration. This method of using random points $x_0, ..., x_{N-1}$ is called Monte Carlo(MC) sampling. [Boyle 1977, Lemieux 2009, Löbbe 2014]
MC-method can easily be implemented and with little additional work it can be used to estimate a broad set of different applications. However the MC-sampling has some disadvantages:

- The generation of random points is difficult.
- The method converges too slowly for some applications and some regularity of the integrand is not reflected by this method.

If we chose the quadrature points $\mathbf{x}_0, ..., \mathbf{x}_{N-1} \in [0,1]^s$ deterministically, then the algorithm $\frac{1}{N}\sum_{i=0}^{N-1} f(\mathbf{x}_i)$ is called a quasi-Monte Carlo (QMC) rule or a QMC algorithm. The main focus of this survey is on QMC-method that can be considered as a deterministic version of the ordinary MC-method. In the deterministic case the quadrature points are needed to be in some sense well-distributed in $[0,1)^s$. The idea is to use specially selected Low-Discrepancy Sequences (LDS) instead of the random numbers. QMC sampling is developed for problems of integration over the unit cube. Sampling over more complicated domains, such as the triangle or the discs is more difficult. Integration over such domains is relevant in graphical rendering, for some example of such a case we refer to Jensen et al. 2001.
The task of this thesis is to investigate such problems of numerical integration over non-cubical region $\mathcal{H}$. Namely the essential aim is to estimate:

$$\mu = \frac{1}{vol(\mathcal{H})} \int_{\mathcal{H}} f(\mathbf{x}) d\mathbf{x}$$

by equal weight

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{x}_i).$$

by suitably choice of the points $\mathbf{x}_i \in \mathcal{H}$ for $i = 1, ..., N$.
QMC consists of two center concepts, Discrepancy of point set and Variation of a function. In the following chapter we introduce the definition of discrepancy and a special Low-Discrepancy Sequence known as digital nets.
The thesis is structured in two particular parts. First we recall basic notions related to QMC-sampling and then the recent work of Basu 2016 is represented by means of numerical integration over non-cubical domains.

Chapter 2

Background on Quasi-Monte Carlo

In this chapter an overview of QMC over the unit cube is given. QMC is consisted of two center concepts, Discrepancy of a point set, and Variation of a function. The definition of discrepancy and focus on a special low-discrepancy sequence known as digital nets are described in sections 2.1 and 2.2. In section 2.3 we introduce the concept of variation of a function in the sense of Hardy and Krause. Then with these definitions we represent the Koksma-Hlawka inequality to get an error bound for QMC integration over $[0,1]^s$. To develop the ideas of randomization we describe a specific scrambling algorithm in the end of the chapter. We use all of the concepts introduced in this chapter to expand the ideas for the rest of the survey.

2.1 Definition of Discrepancy

QMC rule is improved upon MC by choosing $\mathbf{x}_i$ more uniformly distributed in $[0,1]^s$ than random samples usually are, where the uniformity measures via discrepancy. As an example we consider a one-dimensional function $f : [0,1] \to \mathbb{R}$ with continuous first derivative bounded on $[0,1]$. For subset B of $[0,1]$ we define $\chi_B(x)$ as the characteristic function of B, i.e.:

$$\chi_B(x) = \begin{cases} 1, & \text{if } \ x \in B \\ 0, & \text{if } \ x \notin B \end{cases}$$

For a point set $\mathcal{G}$ consisting of N points in the interval $[0,1]$, the function $\delta_{\mathcal{G}} : [0,1] \to \mathbb{R}$

$$\delta_{\mathcal{G}}(\alpha) := \frac{A([0,\alpha), N, \mathcal{G})}{N} - \alpha$$

is called the discrepancy function of the point set $\mathcal{G}$ at point α, where

$$A([0,\alpha), N, \mathcal{G}) := \sum_{i=0}^{N-1} \chi_{[0,\alpha)}(x_i)$$

is the number of points of the point set $\mathcal{G}$ which lie in the interval $[0,\alpha)$. [Dick and Pillichshammer 2010]
Lebesgue measure or the length of the interval $[0,\alpha)$ is α. Hence for a given $\alpha \in [0,1]$, the discrepancy function $\delta_{\mathcal{G}}(\alpha)$ measures the difference between the proportion of points of the set $\mathcal{G}$ in the interval $[0,\alpha)$ and the length of the interval $[0,\alpha)$. The

discrepancy function is small when the points $x_0, ..., x_{N-1}$ are smoothly spread over the interval $[0, 1]$.
The star discrepancy of the point set $\mathcal{H}$ consisting of N points in $[0, 1]^s$ at point $\gamma \in [0, 1]^s$,is defined as

$$D_N^*(\mathcal{H}) := \sup_{\gamma \in [0,1]^s} |\delta_{\mathcal{H}}(\gamma)|,$$

where the function $\delta_{\mathcal{H}}(\gamma) : [0, 1]^s \to \mathbb{R}$

$$\delta_{\mathcal{H}}(\gamma) := \frac{A([0, \gamma), N, \mathcal{H})}{N} - \prod_{i=1}^{s} x_i$$

denotes the s-dimensional discrepancy function of the point set $\mathcal{H}$ at point γ. Figure 2.1 illustrates the discrepancy of a set. The set of points is scattered uniformly if the absolute value of ration of the number of points lying in $[0, \gamma]$ and also the total number of points minus $vol([0, \gamma])$ are small. Because of this reason discrepancy is used to measure the uniformity of a set of points. [Fang and Wang 1994]

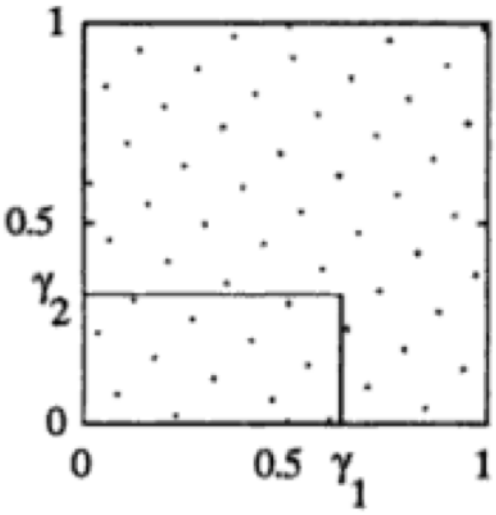


Figure 2.1: The illustration of Discrepancy. [Sorce: Fang and Wang 1994]

2.2 Digital nets

To describe the notion of digital nets we need a series of fundamental definition.

Definition 2.2.1. Suppose $b \geq 2$ be an integer. An s-dimensional, b-adic elementary box defines as an interval of the form

$$\prod_{i=1}^{s} \left[\frac{a_i}{b^{k_i}}, \frac{a_i + 1}{b^{k_i}} \right);$$

$\forall\ 1 \leq i < s$ with integers $1 \leq a_i < b^{k_i}$ and $k_i \geq 0$.

Definition 2.2.2. Suppose $b \geq 2$ be an integer. $\forall n \in \mathbb{N}_0$ with b-adic expansion $n = n_0 + n_1 b + n_2 b^2 + ...$ (this expansion is clearly finite) the (b-adic) radical inverse function $\phi_b : \mathbb{N}_0 \to [0, 1)$ is denoted as

$$\phi_b(n) = \frac{n_0}{b} + \frac{n_1}{b^2} + \frac{n_2}{b^3} + ...$$

Then the b-adic Van der Corput sequence can be defined as the one-dimensional sequence $S = (x_n)_{n\geq 0}$ with $x_n = \phi_b(n)$; $\forall n \in \mathbb{N}_0$. Halton's method is based on the b-adic representation of the natural number. [Dick and Pillichshammer 2010, Basu and Owen]

Example 2.2.1. *the first elements of the 2-adic Van der Corput sequence can be created as follows:*
First we take the natural numbers $0, 1, 2, ...$ *in base* $b = 2$*;*

$$0, 1_2, 10_2, 11_2, 100_2, 101_2, 110_2, ...$$

Then we apply the radical inverse function ϕ_2 *to each number, to get the sequence*

$$0, 0.1_2, 0.01_2, 0.11_2, 0.001_2, 0.101_2, 0.011_2, ...$$

which in fraction form is the sequence

$$0, \frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \frac{5}{8}, \frac{3}{8}, ... \blacksquare$$

Definition 2.2.3. For integers $b_1, ..., b_s \geq 2$, the Van der Corput-Halton sequence is the sequence $S = (x_n)_{n\geq 0}$ with $x_n = (\phi_{b_1}(n), ..., \phi_{b_s}(n))$, $\forall n \in \mathbb{N}_0$. Here ϕ_b denotes the b-adic radical inverse function as defined in Definition 2.2.2. The integers $b_1, ..., b_s$ are called the bases of the Van der Corput-Halton sequence. [Dick and Pillichshammer 2010]

Example 2.2.2. *The first points of the Van der Corput-Halton sequence for 2-dimensional in bases* $b_1 = 2$ *and* $b_2 = 3$*, are* $x_0 = (0, 0), x_1 = (1/2, 1/3), x_2 = (1/4, 2/3), x_3 = (3/4, 1/9), x_4 = (1/8, 4/9), ...$ *The figure 2.2 displays the first 1000 points of this sequence.* $\blacksquare$

Definition 2.2.4. For integers $0 \leq p \leq n$, the points $x_1, x_2, ..., x_{b^n} \in [0, 1]^s$ is called a (p, n, s)-net in base b if each s-dimensional b-adic box with volume b^{p-n} contains exactly b^p of the x_i.

The nets have low discrepancy because the unions of b-adic boxes can efficiently approximate boxes $[0, \alpha]$. Digital nets can obtain a discrepancy of $\mathcal{O}(\frac{(\log N)^{s-1}}{N})$. [Basu 2016]

Definition 2.2.5. For integer $p > 0$, the infinite sequence $x_1, x_2, ... \in [0, 1]^s$ is called a (p, s)-sequence in base b if the sub-sequence $x_{1+rb^n}, ..., x_{(r+1)b^n}$ is a (p, n, s)-net in base b, $\forall r \in \mathbb{Z} > 0$ and $n > p$, i.e. if the same value of p holds for all $n > 0$ then the digital net is a (p, s)-net. The Van der Corput sequence in base b is a $(0, 1)$-net in base b.

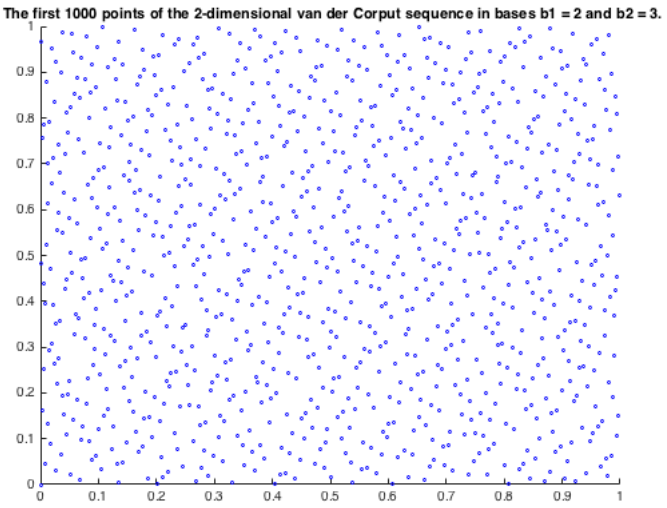


Figure 2.2: The first 1000 points of the 2-dimensional Van der Corput sequence in bases $b_1 = 2$ and $b_2 = 3$.

The (p, s)-sequences is called digital sequences which are extended versions of (p, n, s)-nets. They can obtain a discrepancy of $\mathcal{O}(\frac{(\log N)^s}{N})$, thus the Van der Corput-Halton is well uniformly scattered on $[0, 1]^s$. The discrepancy rate improves to $\mathcal{O}(\frac{(\log N)^{s-1}}{N})$ along the sub-sequence $N = \beta b^n$ for integers $n \geq 0$ and $1 \leq \beta < b$. [Dick and Pillichshammer 2010]
Using the framework of digital nets can allow us

- to provide the (p, n, s)-net or the (p, s)-sequence in a simple way (in the form of s matrices);
- to choose the quality parameter p in a rather fast way;
- to examine the point sets of high quality may be limited to the examination for the matrices mentioned above with certain properties.

2.3 Variation of function

Introductory text books on real analysis commonly cover the concept of total variation for functions of a single real variable. Hardly any of them declare much about multidimensional variation.

2.3.1 One-dimensional variation

Suppose $f(x)$ is a real valued function defined on the interval $[\alpha, \beta]$ where $-\infty < \alpha \leq \beta < \infty$. A ladder on the interval $[\alpha, \beta]$ is defined as a set $\mathcal{K}$ containing α and finitely many, possibly zero, values from the interval (α, β). The ladder $\mathcal{K}$ does not contain β except when $\alpha = \beta$. Obviously this case is decayed, but in some settings below it is harder to exclude it than to include it. Each element $k \in \mathcal{K}$ has

a successor element k_+, such that if $(k,\infty)\cap\mathcal{K}=\emptyset$ then $k_+=\beta$. Otherwise k_+ is the smallest element of $(k,\infty)\cap\mathcal{K}$. If the elements of $\mathcal{K}$ are formed into increasing order, i.e., $\alpha=k_0<k_1<...<k_n$, then the successor element of k_j is k_{j+1} for $j<n$ and it equals β for $j=n$. Note that the value k_+ depends on $\mathcal{K}$ but the notation will not make specific this dependence.
Let $\mathbb{K}$ denote the set of all ladders on the interval $[\alpha,\beta]$, hence the total variation of $f(x)$ on $[\alpha,\beta]$ is defined: [Owen 2005]

$$V(f;\alpha,\beta)=\sup_{\mathcal{K}\in\mathbb{K}}\sum_{k\in\mathcal{K}}|f(k_+)-f(k)|.$$

The total variation in the sense of Hardy and Krause is the most widely used definition for QMC.
To describe the total variation of a function in the sense of Hardy and Krause, we first introduce some notation.
Assume x^j is j-th component of $\mathbf{x}\in\mathbb{R}^s$ and $\mathbf{x}=(x^1,...,x^s)$,then for any $\mathbf{x},\mathbf{y}\in\mathbb{R}^s$, when $\mathbf{x}\le\mathbf{y}$ the hyperrectangle $[\mathbf{x},\mathbf{y}]$ is the set $\{\mathbf{z}\in\mathbb{R}^s|\mathbf{x}\le\mathbf{z}\le\mathbf{y}\}$. ($\forall\mathbf{x},\mathbf{y}\in\mathbb{R}^s$ it is written $\mathbf{x}<\mathbf{y}$ or $\mathbf{x}\le\mathbf{y}$ if the inequality holds for all s components.)
Also $[\mathbf{x},\mathbf{y})=\{\mathbf{z}\in\mathbb{R}^s|\mathbf{x}\le\mathbf{z}<\mathbf{y}\}$ and $(\mathbf{x},\mathbf{y}]$ and $(\mathbf{x},\mathbf{y})$ are defined in the same way. Then the s-dimensional volume, $\prod_{j=1}^s(y^j-x^j)$ of $[\mathbf{x},\mathbf{y}]$ is denoted $vol([\mathbf{x},\mathbf{y}])$.
For $u\subseteq\{1,...,s\}$, $|u|$ defines the cardinality of u and $-u$ the compliment of u with respect to the set $\{1,...,s\}$, so that $-u=\{1,...,s\}-u$. The expression $\mathbf{x}^u$ denotes a $|u|$-tuple of real values describing the components x^j for $j\in u$.
Assume that $u,v\subseteq\{1,...,s\}$ and $\mathbf{a},\mathbf{b}\in[\mathbf{x},\mathbf{y}]$ with $u\cap v=\emptyset$, The domain of $\mathbf{a}^u$ is the hyperrectangle $[\mathbf{x}^u,\mathbf{y}^u]$ and the domain of $\mathbf{b}^v$ is the hyperrectangle $[\mathbf{x}^v,\mathbf{y}^v]$. Then the notation $\mathbf{a}^u:\mathbf{b}^v$ introduces the point $\mathbf{c}\in[\mathbf{x}^{u\cup v},\mathbf{y}^{u\cup v}]$ with $c^j=a^j$ for $j\in u$ and $c^j=b^j$ for $j\in v$. The symbol $\mathbf{a}^u:\mathbf{b}^v$ is well defined for $\mathbf{a}^u\in[\mathbf{x}^u,\mathbf{y}^u]$ and $\mathbf{b}^v\in[\mathbf{x}^v,\mathbf{y}^v]$, when $u\cap v=\emptyset$.
Zero-dimensional domains and functions on them are of no interest in this survey. They however appear as special cases in some derivations. In the addition, $\mathbf{a}^\emptyset$ denotes the "zero-tuple". The Cartesian product of zero sets is the set consisting of the zero-tuple, and the volume of a zero-dimensional rectangle is $\prod_{i\in\emptyset}(b^i-a^i)=1$, since empty products are consistently taken to be one. A function f on $[\mathbf{x}^\emptyset,\mathbf{y}^\emptyset]$ is constant, with a value is denoted by $f()$.
Now we can represent the multi-dimensional variation using these notations.

2.3.2 Multi-dimensional variation

The s-fold alternating sum of function f defined over the interval $[\mathbf{x},\mathbf{y}]$ is: [Dick, Kuo, and Sloan 2013, Owen 2005]

$$\Delta(f;\mathbf{x},\mathbf{y})=\sum_{u\subseteq\{1,...,s\}}(-1)^{|u|}f(\mathbf{x}^u:\mathbf{y}^{-u}).$$

Note that the coefficient of $f(\mathbf{y})$ is one while it of $f(\mathbf{x})$ is $(-1)^s$. The alternating sums is well defined even when the inequality $\mathbf{x}\le\mathbf{y}$ does not hold.
Let $\mathcal{K}^j$, for any $j=1,...,s$ be a partition of $[x^j,y^j]$, then a multi-dimensional ladder on $[\mathbf{x},\mathbf{y}]$ has the form $\mathcal{K}=\prod_{j=1}^s\mathcal{K}^j$ and for $\mathbf{k}\in\mathcal{K}$, the successor point $\mathbf{k}_+$ can be

defined by taking ${k_+}^j$ to be the successor of k^j in $\mathcal{K}^j$. The variation of function f over $\mathcal{K}$ is:

$$V(f;\mathcal{K}) = V_{\mathcal{K}}(f) = \sum_{\mathbf{k}\in\mathcal{K}} |\Delta(f;\mathbf{k},\mathbf{k}_+)|.$$

A ladder is, with minor distinctness, that Clarkson and Adams 1933 call a "net". These nets include upper boundaries from b. Ladders are sets, which allows writing some manipulations economically.

Definition 2.3.1. The variation of function f on the hyperrectangle $[\mathbf{x},\mathbf{y}]$, in the sense of Hardy and Krause, is defined: [Owen 2005]

$$V_{HK}(f;\mathbf{x},\mathbf{y}) = V_{HK}(f) = \sum_{v\subset\{1,\ldots,s\}} V_{[\mathbf{x}^{-v},\mathbf{y}^{-v}]} f(\mathbf{z}^{-v};\mathbf{y}^v).$$

If $V_{HK}(f) < \infty$, then the function f has bounded variation in the sense of Hardy and Krause (BVHK). The definition of bounded variation in Hardy (1905)[Clarkson and Adams 1933] requires that $V_{[\mathbf{x}^{-v},\mathbf{y}^{-v}]} f(\mathbf{z}^{-v};\mathbf{t}^v) < \infty$ and $V_{[\mathbf{x},\mathbf{y}]} < \infty$ for all $0 < |v| < s$ and all $\mathbf{t}^v \in [\mathbf{x}^v,\mathbf{y}^v]$.

The most essential use of variation in QMC is in the Koksma-Hlawka inequality.

Theorem 2.3.1. *(Koksma-Hlawka inequality) Suppose $f : [0,1]^s \to \mathbb{R}$ is a function of bounded variation in the sense of Hardy and Krause. Then for any set of points $\mathbf{x}_1,\ldots,\mathbf{x}_N \in [0;1]^s$ with $N \geq 2$, the quadrature error has an upper bound equal to, see Niederreiter (1992),*

$$\left| \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_i) - \int_{[0,1]^s} f(\mathbf{x}) d(\mathbf{x}) \right| \leq D_N^* V_{HK}(f).$$

There are many constructions for which $D_N^* = \mathcal{O}(\frac{(logN)^{s-1}}{N})$ [Bratley, Fox, and Niederreiter 1994], hence when $V_{HK}(f) < \infty$, QMC is asymptotically much more accurate than MC.
The terms D_N^* and $V_{HK}(f)$ are difficult to compute, possibly much more than the integral of function f. Further in case that both the terms are known the Koksma-Hlawka bound is found to overestimate the true error of integration. Although the condition of $V_{HK}(f) < \infty$ is contrary because it is required that the function f be bounded, a condition that often disregarded in option pricing application. Thus the Koksma-Hlawka inequality as a practical error bound has a limited applicability. [Glasserman 2003]

2.4 Randomisation of digital net

In this section randomisation of digital net is considered. The purpose of this algorithm is to combine deterministic algorithms with random in such a way that one can reach the best features of the both methods. The improved rate of convergence is the advantage of a QMC-algorithm based on a digital net. On the other hand there are also some disadvantages compared to MC-algorithms, where the quadrature points are selected uniformly and i.i.d. in $[0,1)^s$, for instance:

- The first point of a digital net is always zero , which draws on problems in some applications. If the points in the interval $[0,1]^s$ need to be mapped to $\mathbb{R}^s$ such that the points are normally distributed, so the point zero gets mapped to $(-\infty, ..., -\infty)$.
- Estimation of integrals where the integrand has a irregularity, using deterministic samples in a QMC-algorithm is sometimes tricky. Uniform and i.i.d. random samples consistently avoid this matter (with high probability).
- Another involvement in applications is bias of the estimator $\frac{1}{N}\sum_{i=0}^{N-1} f(\mathbf{x}_i)$.
- A benefit of uniformly and i.i.d. chosen random samples over deterministically selected samples is that in the previous case a statistical estimate of the error by $\sqrt{\frac{1}{1-N}\sum_{i=0}^{N-1}(f(\mathbf{x}_i) - \frac{1}{N}\sum_{i=0}^{N-1} f(\mathbf{x}_i))^2}$ is obtainable. But for deterministic quadrature points such an estimate is not obtainable.

These problems can be avoided using randomised digital nets as quadrature points. Although uniform and i.i.d. random selection of the quadrature points give an unbiased estimator with standard deviation of order $\frac{1}{\sqrt{N}}$, thus they are inferior to QMC-sampling in related to the speed of convergence. [Dick and Pillichshammer 2010]
The aim of this section is to develop a combination of deterministic choices and random choices of the quadrature points to get the attractive features of both methods. To be more accurate, we want to obtain quadrature points $\mathbf{x}_0, ..., \mathbf{x}_{N-1}$ such that any $\mathbf{x}_i$ is uniform and i.i.d. and in the meantime the point set $\mathbf{x}_0, ..., \mathbf{x}_{N-1}$ has low discrepancy. This way we can obtain a statistical error estimate and avoid the problems introduced at the beginning. Numerous randomisation-methods are known. The simplest method of describing a randomisation in (p, n, s)-nets is by using a digital shift $\sigma \in [0,1)^s$ which distributes uniformly. We focus on the scrambling of digital nets as represented by Owen.

2.4.1 Owen's scrambling algorithm

Owen's scrambling algorithm is easiest defined for some generic point $\mathbf{x} = (x_1, ..., x_s) \in [0,1)^s$ where $x_i = x_{i,1}b^{-1} + x_{i,2}b^{-2} + ...$ is the base b expansion of x_i. The scrambled point can be denoted by $\mathbf{y} \in [0,1)^s$, with $\mathbf{y} = (y_1, ..., y_s)$ and $y_i = y_{i,1}b^{-1} + y_{i,2}b^{-2} + ...$. The point $\mathbf{y}$ can be obtained by applying permutations to any digit of any coordinate of $\mathbf{x}$. The permutation applied to digit $x_{i,j}$ depends on the previous digits $x_{i,k}$ for $1 \leq k < j$. Particularly, $y_{i,1} = \pi_{i(x_{i,1})},\ y_{i,2} = \pi_{i,x_{i,1}}(x_{i,2}),\ y_{i,3} = \pi_{i,x_{i,1},x_{i,2}}(x_{i,3}), ...$, and generally

$$y_{i,j} = \pi_{i,x_{i,1},...,x_{i,j-1}}(x_{i,j})$$

where $\pi_{i,x_{i,1},...,x_{i,j-1}}$ is a random permutation of $\{0, ..., b-1\}$ and each random permutation is uniformly distributed on the set of $b!$ permutations of $\{0, ..., b-1\}$. Owen uses the term *nested uniform scrambling* for this procedure, where *nested* defines the dependence of the permutation for the digit $i > 1$ on the values of digits $i > j$, and *uniform* defines the use of all $b!$ possible permutations.[Niederreiter and Shiue 2012]
A nested uniform scramble of $\mathbf{x}$ applies independent nested uniform scrambles to each s components of $\mathbf{x}$. We consider that permutations with different indices are

independent from each other and that any permutation has the same probability. This way the scrambled point $\mathbf{y}$ is uniformly distributed in $[0,1)^s$.
To explain Owen's scrambling, let for $1 \leq i \leq s$,

$$\mathbf{\Pi}_i = \{\pi_{i,x_{i,1},...,x_{i,j-1}}; j \in \mathbb{N},\ x_{i,1},...,x_{i,j-1} \in \{0,...,b-1\}\},$$

where for $j = 1$ we assume that $\pi_{i,x_{i,1},...,x_{i,j-1}} = \pi_i$, be a given set of permutations and $\mathbf{P}_i = (\mathbf{\Pi}_1,...,\mathbf{\Pi}_s)$. Then, using these permutations to digital point $\mathbf{x} \in [0,1)^s$ when applying Owen's scrambling, we define $\mathbf{y} = \mathbf{x}_{\mathbf{P}_i}$. [Dick and Pillichshammer 2010]

Let $\mathbf{x} \in [0,1)^s$ and let $\mathbf{y}$ be the result of a Owen's scrambling of $\mathbf{x}$. Then $\mathbf{y} \sim \mathbb{U}[0,1)^s$. [Basu and Owen]

If the sequence $\mathbf{x}_1,...,\mathbf{x}_m$ forms a (p,n,s)-net in base b, then the scrambled points $\mathbf{y}_i$ are a (p,n,s)-net in base b with probability one. In addition if $\mathbf{x}_i$ is a (p,s)-sequence in base b, then the scrambled points $\mathbf{y}_i$ form also a (p,s)-sequence in base b with probability one. [Basu 2016]

In scrambled net quadrature we estimate $\mu = \int_{[0,1)^s} f(\mathbf{y})d\mathbf{y}$ by

$$\hat{\mu} = \frac{1}{N}\sum_{i=1}^{N} f(\mathbf{y}_i)$$

where $\mathbf{y}_i$ are the scrambled points of $\mathbf{x}_i$. For $f \in \mathcal{L}^1[0,1)^s$ we have $\mathbb{E}(\hat{\mu}) = \mu$, which follows from proposition 1. When $f \in \mathcal{L}^2[0,1)^s$ the variance of $\hat{\mu}$ can be estimated using independent random replications of the scrambled nets. If $V_{HK}(f) < \infty$ then from the Koksma-Hlawka inequality we can obtain

$$Var(\hat{\mu}) = \mathcal{O}(\frac{\log(N)^{2(s-1)}}{N^2}) = \mathcal{O}(N^{-2+\varepsilon}).$$

Following Theorem mentions that the scrambled net has the potential to improve accuracy.

Theorem 2.4.1. *For $f : [0,1]^s \to \mathbb{R}$ with continuous $\dfrac{\partial^s}{\partial_{y_1}...\partial_{y_s}} f$ suppose $\mathbf{y}_i$ are a nested uniform scramble of the first $N = \lambda b^n$ points of a (p,s)-sequence in base b, for $\lambda \in \{1,2,...,b-1\}$, then:*

$$Var(\hat{\mu}) = \mathcal{O}(\frac{\log(N)^{(s-1)}}{N^3}) = \mathcal{O}(N^{-3+\varepsilon}).$$

[Owen 1997b, Owen 2008]

The Theorem below establishes that smoothness and bounded variation are not necessary for the scrambled nets to attain a better rate than MC-method.

Theorem 2.4.2. *Let* $\mathbf{y}_1, ..., \mathbf{y}_N$ *be a nested uniform scramble of a* (p, n, s)*-net in base* b *and let* $f \in \mathcal{L}^2[0,1)^s$*. Then:*

$$Var(\hat{\mu}) = o(\frac{1}{N})$$

as $N \to \infty$*.*

Note that the term $\log(N)^{(s-1)}$ is not necessarily small compared to N^3 for proper sizes of N and large s. For proof see Owen 1998.

Theorem 2.4.3. *Assume* $\mathbf{y}_1, ..., \mathbf{y}_N$ *be a nested uniform scramble of a* (p, n, s)*-net in base* b *and let* $f \in \mathcal{L}^2[0,1)^s$ *with variance* σ^2 *where* $\mathbf{y} \sim \mathbb{U}[0,1]^s$*. Then:*

$$Var(\hat{\mu}) \leq b^p \left(\frac{b+1}{b-1}\right)^{s-1} \frac{\sigma^2}{N}.$$

If $p = 0$*, then* $Var(\hat{\mu}) \leq e\frac{\sigma^2}{N} = 2.718\frac{\sigma^2}{N}$*.*

Proof. The first result can be found in [Owen 1998] and the second is in [Owen 1997a]. □

Chapter 3

Low-Discrepancy Constructions in Simplex

As we mentioned before most research in the field of QMC-sampling focuses on sampling from the unit cube. Namely many problems in computer graphics, are described via quadrature over the unit triangle. QMC integration over the simplex have been developed by Pillards and Cools 2005 and Brandolini et al. 2013a.
In this chapter we consider numerical integration over a triangular region, using QMC-sampling. Commonly integrals of this form arise in graphical rendering. Usually sampling form these domains can be approached by applying a mapping $\phi : [0,1]^s \to D$ where D is the domain of interest. The mapping defines such that if $\mathbf{x} \sim \mathbb{U}([0,1]^s)$ then $\phi(\mathbf{x}) \sim \mathbb{U}(D)$. Generally there are various choices for such mappings, and the dimension of D does not necessarily equal to the dimension s of the cube. Using such a mapping we are able to generate QMC points $\mathbf{x}_i \in [0,1]^s$ where $\phi(\mathbf{x}_i)$ can be used as sample points in D.

This approach gives us an estimation $\mu = \int_D g(\mathbf{x})d\mathbf{x}$ by $\frac{1}{N}\sum_{i=1}^{N} f(\mathbf{x}_i)$ where $f(\mathbf{x}) = g(\phi(\mathbf{x}))$. But the problem is that the composite function $f = g \circ \phi$ may not be well adapted to QMC, it can have singularities, cusps, or discontinuities. These factors may decline the performance of QMC or at least they make it more challenging to analyze QMC's performance.
Pillards and Cools 2005 investigated QMC integration over the simplex. They noticed that the Koksma–Hlawka bound might be applied using the variation of the composite function $g \circ \phi$ and the discrepancy of the original points $\mathbf{x}_i$. They also constructed a measure of variation for functions on the simplex, a corresponding discrepancy measure for points inside the simplex, and a Koksma–Hlawka bound using these two features. But they did not give conditions for that variation to be finite.
Brandolini et al. 2013b also introduce a version of the Koksma–Hlawka inequality for the simplex. They construct a measure of variation for the simplex and a discrepancy measure for points in the simplex.
Neither Pillards and Cools 2005 nor Brandolini et al. 2013a define a sequence of points with vanishing discrepancy.
In this chapter we describe the recent results obtained by Basu and Owen to construct the points in the triangle. It is an extensible digital construction of the Van der Corput sequence and uses a recursive partitioning of the triangle. They have

combined the theorems of Chen and Travaglini 2007 and Brandolini et al. 2013a to show that their points have vanishing discrepancy.
This chapter is organized as following. In Section 3.1 we introduce results from the literature which are needed along with notation to define those results. Section 3.2 presents the Van der Corput sequence from the unit interval to an arbitrary triangle. In Section 3.3 we show that triangular Van der Corput points give integral estimates with vanishing error when the integrand is slightly Riemann integrable over the triangle. Section 3.4 contains final discussion.

3.1 Background

Here we use the same notation as we presented in Chapter 2, in Section 2.3 and define some of previous results.
Suppose A, B, and C are three non-collinear points in $\mathbb{R}^s$. Using these points we determine the non-degenerate triangle

$$\Delta(A, B, C) = \{w_1A + w_2B + w_3C | \min(w_1, w_2, w_3) \geq 0, w_1 + w_2 + w_3 = 1\}.$$

Using the following triangle leads to less challenging expressions and calculations. For the most part the simplex is defined via the corners $(0, 0, 1)^T$, $(0, 1, 0)^T$ and $(1, 0, 0)^T$. Although for some determination the points may be scaled so that the triangle has unit area. Sometimes the triangle can be scaled to have area equal to the number of points in a quadrature rule. Pillards and Cools 2005 worked with the right-angle triangle defined by $A = (0, 0)^T$, $B = (0, 1)^T$ and $C = (1, 1)^T$.

3.1.1 Koksma-Hlawka inequality

In the triangle discrepancy measures through equidistribution over trapezoidal subsets. Brandolini et al. 2013a show that triangular Van der Corput sequence has trapezoidal discrepancy of $\mathcal{O}(N^{-\frac{1}{2}})$. Following the idea of Section 2.1 we define the notions of discrepancy for quadrature problems over a set $\mathcal{H}$. We take $\mathcal{H}$ to be a bounded Borel subset of $\mathbb{R}^s$, following Brandolini et al. 2013a, and $vol(.)$ denotes s-dimensional Lebesgue measure. If a linear flat subset[1] $\mathcal{F}$ of $\mathbb{R}^s$ contains $\mathcal{H}$ then we define volumes as Lebesgue measure with respect to the lowest-dimensional such $\mathcal{F}$. We suppose that $vol(\mathcal{H}) > 0$ to avoid uninteresting cases. Let $\Psi = (\mathbf{x}_0, ..., \mathbf{x}_{N-1})$; $\forall N > 1$ be a list of (not necessarily distinct) points in $\mathbb{R}^s$. For a subset $\mathcal{M}$ of $\mathbb{R}^s$ we define the counting function

$$A(\mathcal{M}, N, \Psi) := \sum_{i=0}^{N-1} \chi_{\mathcal{M}}(\mathbf{x}_i).$$

Then the signed discrepancy of Ψ at the measurable set $\mathcal{M} \subset \mathbb{R}^s$ is [Matousek 1999 and Dick, Kuo, and Sloan 2013]

$$\delta_N(\mathcal{M}; \Psi, \mathcal{H}) := \frac{vol(\mathcal{M} \cap \mathcal{H})}{vol(\mathcal{M})} - \frac{A(\mathcal{M}, N, \Psi)}{N}.$$

[1] A subset $\mathcal{D}$ of $\mathbb{R}^s$ is called flat if for any X and Y in $\mathcal{D}$, then $\mathcal{D}$ contains the line through X and Y. [Hoffman 2013]

The additive property of the signed discrepancy states that if $\mathcal{M}_1 \cap \mathcal{M}_2 = \emptyset$ then

$$\delta_N(\mathcal{M}_1 \cup \mathcal{M}_2; \Psi, \mathcal{H}) = \delta_N(\mathcal{M}_1; \Psi, \mathcal{H}) + \delta_N(\mathcal{M}_2; \Psi, \mathcal{H}).$$

Note that $\delta_N(\emptyset; \Psi, \mathcal{H}) = 0$.

The absolute discrepancy of points Ψ for a class $\mathcal{K}$ of measurable subsets of $\mathcal{H}$ is defined

$$D_N(\mathcal{K}; \Psi, \mathcal{H}) := \sup_{\mathcal{M} \subset \mathcal{K}} D_N(\mathcal{M}; \Psi, \mathcal{H}).$$

where $D_N(\mathcal{M}; \Psi, \mathcal{H}) := |\delta_N(\mathcal{M}; \Psi, \mathcal{H})|$.
For general $\mathcal{H}$, Ψ can be extended by all integer shifts, which is through assuming all $\mathbf{x}_i + m \in \mathcal{H}$ for $i = 1, ..., N$ and $m \in \mathbb{Z}^s$. Since $\mathcal{H}$ is bounded, this extension still contains finitely many points. The extended counting function is defined

$$\bar{A}(\mathcal{M}, N, \Psi) := \sum_{m \in \mathbb{Z}^s} \sum_{i=0}^{N-1} \chi_{\mathcal{M}}(\mathbf{x}_i + m).$$

hence we take

$$\bar{D}_N(\mathcal{K}; \Psi, \mathcal{H}) := \sup_{\mathcal{M} \subset \mathcal{K}} \bar{D}_N(\mathcal{M}; \Psi, \mathcal{H}).$$

where $\bar{D}_N(\mathcal{M}; \Psi, \mathcal{H}) := |\bar{\delta}_N(\mathcal{M}; \Psi, \mathcal{H})|$, and

$$\bar{\delta}_N(\mathcal{M}; \Psi, \mathcal{H}) := \frac{vol(\mathcal{M} \cap \mathcal{H})}{vol(\mathcal{M})} - \frac{\bar{A}(\mathcal{M}, N, \Psi)}{N}.$$

We mention that $\bar{A}(\mathcal{M}, N, \Psi)$ is divided by N, and not the number of extended points lying in $\mathcal{H}$.
A Koksma-Hlawka inequality for compact simplex is provided by Brandolini et al. 2013a. Their variation measure sums integrals over all faces of all dimensions of the compact manifolds which is not similar to the ordinary Koksma-Hlawka inequality. Suppose $\mathcal{M}$ is a closed simplex in $\mathbb{R}^s$, and $V_0, ..., V_s$ are its vertices. Let $w_1^k, ..., w_s^k$, $\forall k = 0, ..., s$ be the vectors joining the vertex V_k with the other vertices, in any order. Consider W_k as the matrix with columns $w_1^k, ..., w_s^k$. Call $\mathcal{H}_k$ the parallelepiped determined by the vectors $w_1^k, ..., w_s^k$ and the vertex V_k. Certainly, let $\mathcal{M}_a^k$ be the a-dimensional face of $\mathcal{M}$ parallel to the directions $a_1 w_1^k, ..., a_s w_s^k$, for every multi-index $a \in \{0, 1\}^s$. In order to figure out a Koksma-Hlawka inequality for the triangle from the Koksma-Hlawka inequality for parallelepipeds, it satisfies decomposing the characteristic function of the simplex $\mathcal{M}$ into a weighted sum of characteristic functions of the parallelepipeds $\mathcal{H}_k$.

Lemma 3.1.1. *There is a constant C_s that only depends on the dimension s. Then for every simplex $\mathcal{M}$ there are smooth functions $\phi_0, ..., \phi_s$ which satisfy the following conditions:*

i) $\phi_k(V_k) = 1$, *for every* $k = 0, ..., s$, *and the open half space determined by the faces of* $\mathcal{M}$ *opposite to* V_k *contains* $\sup(\phi_k)$.

ii) *For every* $x \in \mathcal{M}$, *we have* $\sum_{k=0}^{s} \phi_k(x) = 1$.

iii) *For any multi-index* $a \in \{0,1\}^s$ *and* $k = 0, ..., s$,

$$\sup_{x \in \mathcal{M}} \left| \left(\frac{\partial}{\partial w^k} \right)^a \phi_k(x) \right| \leq C_s.$$

[Brandolini et al. 2013a]

Theorem 3.1.2. *Assume a smooth compact s-dimensional simplex* $\mathcal{M}$, *let* f *be a smooth function on* $\mathbb{R}^s$, *and* $\Psi = \{x_i + m;\ i = 1, ..., N, m \in \mathbb{Z}^s\}$ *a periodic distribution of points. Then the total variation of the function* f *in the simplex* $\mathcal{M}$ *is defined*

$$\mathcal{V}_{\mathcal{M}}(f) = C_s \sum_{k=0}^{s} \sum_{a \in \{0,1\}^s} \sum_{b \leq a} \frac{1}{|\mathcal{M}_a^k|} \int_{\mathcal{M}_a^k} \left| \left(\frac{\partial}{\partial w^k} \right)^b f(x) \right| dx.$$

A multi-index b *is less than or equal to another multi-index* a *if* $b_i \leq a_i,\ \forall i = 1, ..., s$.

The discrepancy of Ψ with respect to the $s+1$ parallelepipeds associated with the simplex $\mathcal{M}$ is

$$D(\mathcal{M}, \Psi) := \max_{k=0,...,s} D(\mathcal{H}_k, \Psi).$$

Hence

$$\left| \int_{\mathcal{M}} f(x)dx - \frac{1}{N} \sum_{y \in \mathcal{M} \cap \Psi}^{\bigstar} f(y) \right| \leq D(\mathcal{M}, \Psi)\mathcal{V}_{\mathcal{M}}(f).$$

Where the notation $\sum_{y \in \mathcal{M} \cap \Psi}^{\bigstar} f(y)$ means that if y belongs to a i-dimensional face of the simplex $\mathcal{M}$, then $f(y)$ should be replaced by $2^{i-s} f(y)$. Note that the integration over $\mathcal{M}_a^k$ is determined with respect to the a-dimensional Lebesgue surface measure.

Basu and Owen present Brandolini et al. discrepancy measure for the case of a triangle with corners A, B and C. Let $\mathcal{T}_{a,b,C}$ be the parallelogram for real values a and b, defined by the point C with vectors $b(B - C)$ and $a(A - C)$. Figure 3.1 displays such parallelogram with vertices C, D, F and E.

Assume

$$\mathcal{M}_C = \{\mathcal{T}_{a,b,C};\ 0 < b < \|B - C\|,\ 0 < a < \|A - C\|\}.$$

$\mathcal{M}_A$ and $\mathcal{M}_B$ can correspondingly be defined. Thus the parallelogram discrepancy of points Ψ in $\mathcal{H} = \Delta(A, B, C)$ is

$$D_N^P(\Psi, \mathcal{H}) := D_N(\mathcal{M}_P; \Psi, \mathcal{H})$$

where $\mathcal{M}_P = \mathcal{M}_A \cup \mathcal{M}_B \cup \mathcal{M}_C$.

A discrepancy for simplex is also defined by Pillards and Cools 2005. Their $\mathcal{H}$ for simplex with three vertices, is the triangle $T_{PC} = \Delta((0,0)^T, (0,1)^T, (1,1)^T)$. Then the discrepancy is measured

$$D_N^{PC}(\Psi, T_{PC}) := D_N(\mathcal{M}_\tau; \Psi, T_{PC})$$

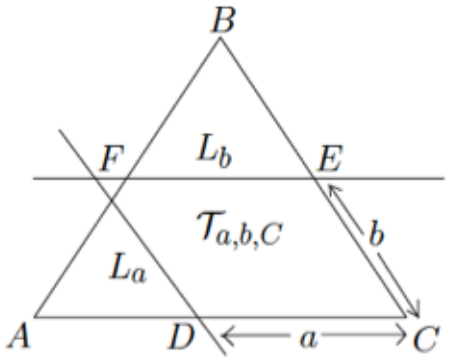


Figure 3.1: The construction of the parallelogram [Source: Basu 2016].

where $\mathcal{M}_\tau = \{[0, \alpha);\ \alpha \in [0,1)^2\}$.

Following Lemma mentions that a sequence with vanishing parallel discrepancy has also vanishing discrepancy in the sense of Pillands and Cools.

Lemma 3.1.3. *For $N > 1$, let Ψ be the list of points $\mathbf{x}_1, ..., \mathbf{x}_N \in T_{PC}$. Then*

$$D_N^{PC}(\Psi, T_{PC}) \leq 2D_N^P(\Psi, T_{PC});$$

and

$$\bar{D}_N^{PC}(\Psi, T_{PC}) \leq 2\bar{D}_N^P(\Psi, T_{PC}).$$

[Basu 2016]

When sample points $\mathbf{x}_i$ are on the boundary $\mathcal{H}$ then the numerical treatment of the points is different. Suppose $\mathcal{H}$ is a closed polytope in $\mathbb{R}^s$ not placing in a flat of dimension $s-1$ or less. Then the weight function is

$$w_{\mathcal{H}}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \notin \mathcal{H} \\ 1, & \mathbf{x} \in \text{ the interior of } \mathcal{H} \\ 2^{i-s}, & \mathbf{x} \in \text{a } i\text{-dimensional face of } \mathcal{H} \end{cases}.$$

The integer i is the smallest dimension of any face of $\mathcal{H}$ that $\mathbf{x}$ is contained in. Basu and Owen work instead with the relative interior of $\mathcal{H}$ where $\mathcal{H}$ lies in a lower dimensional flat and replace s by the smallest containing dimension. GivenΨ with the list of points $\mathbf{x}_1, ..., \mathbf{x}_N$ and the function f on $\mathcal{H}$ they define

$$\sum_{\mathcal{H},\Psi}^{\star} f := \sum_{j=1}^{N} \sum_{m \in \mathbb{Z}^s} f(\mathbf{x}_j + m) w_{\mathcal{H}}(\mathbf{x}_j + m).$$

Now the Theorem of Brandolini et al. 2013a can be specialized to the triangle.

Theorem 3.1.4. *Let $\mathcal{H} = \Delta(\mathbf{A}, \mathbf{B}, \mathbf{C})$ be a non- degenerate triangle in $\mathbb{R}^s$. Given Ψ with the list of points $\mathbf{x}_1, ..., \mathbf{x}_N$ in $\mathbb{R}^s$, then*

$$\left| \int_{\mathcal{H}} f(\mathbf{x}) dx - \frac{1}{N} \sum_{\mathcal{H},\Psi}^{\star} f \right| \leq \bar{D}_N^P(\mathcal{H}, \Psi) \mathcal{V}_{\mathcal{H}}(f).$$

[Basu 2016]

3.1.2 Transformation

Assuming two non-degenerate triangles $\Delta(\mathbf{A}, \mathbf{B}, \mathbf{C}) \subset \mathbb{R}^s$ and $\Delta(\acute{\mathbf{A}}, \acute{\mathbf{B}}, \acute{\mathbf{C}}) \subset \mathbb{R}^d$, there exists a linear mapping $M : \mathbb{R}^d \to \mathbb{R}^s$ such that $\mathbf{A} = M\acute{\mathbf{A}}, \mathbf{B} = M\acute{\mathbf{B}}$ and $\mathbf{C} = M\acute{\mathbf{C}}$. A transformation of points $\mathbf{x}_i$ to $M\mathbf{x}_i$ is called $M\Psi$, then $D_N^P(\Psi, \Delta(\acute{\mathbf{A}}, \acute{\mathbf{B}}, \acute{\mathbf{C}})) = D_N^P(M\Psi, \Delta(\mathbf{A}, \mathbf{B}, \mathbf{C}))$, but this equality does not hold for D_N^{PC} since a linear transformation may map anchored boxes onto parallelepipeds.

3.2 Geometric Van der Corput sequences

To generate a triangular Van der Corput sequence Basu and Owen worked with a base 4 recursive partitioning of the triangle. They first partition the triangle into 4 congruent subsets. As it is displayed in the leftmost panel in Figure 3.2 they appoint base 4 digits 0 through 3 to the sub-triangles with 0 in the center and the others subject to an arbitrary selection. In a similar way every such triangle can be divided again, as presented in the second panel.

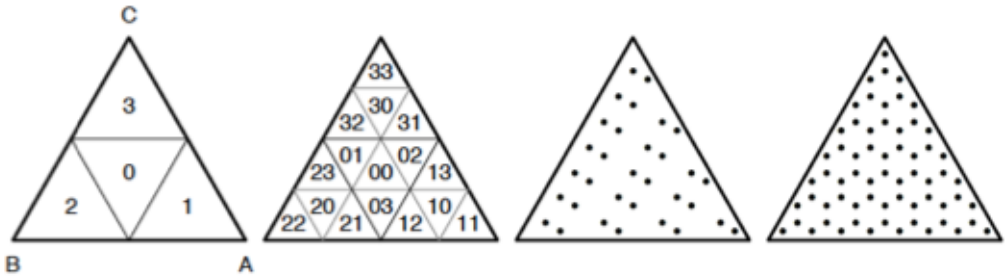


Figure 3.2: A partitioning of $\Delta(A, B, C)$ into 4 and then 16 congruent sub-triangles, and the first 32 triangular Van der Corput points followed by the first 64. [Source: Basu 2016].

The digital construction that they use works by improving the construction of Van der Corput sampling (1935) from the interval $[0, 1)$ to the triangle. In Van der Corput sampling of the unit interval the integer $i \geq 0$ can be written in the base $b > 2$ as $\sum_{k\geq 1} d_k b^{k-1}$, when $d_k = d_k(i) \in \{0, 1, ..., b-1\}$. Thus i maps to $x_i = \sum_{k\geq 1} d_k b^{-k}$ and the discrepancy of the points $x_i \in [0, 1), \forall\ i = 0, ..., N-1$ are $\mathcal{O}(\frac{\log(N)}{N})$. So the integer $i \geq 0$ in base 4 can be presented as $\sum_{k\geq 1}^{K_i} d_k 4^{k-1}$, when $d_k = d_k(i) \in \{0, 1, 2, 3\}$ and $K_i = (\log_4(i) + 1)$.
Given triangle T the integer i is mapped to the point $f_T(i) \in T$ by first identifying a sub-triangle of T corresponding to d_1, say $T(d_1)$. Then corresponding to d_2 within $T(d_1)$ getting the sub-triangle $T(d_1, d_2) = (T(d_1))(d_2)$, and so on. Through this process we map the integer i to the triangle $T(d_1, d_2, ..., d_{K_i})$. The center point of the triangle $T(d_1, d_2, ..., d_{K_i})$ is the point $f_T(i)$, which is the arithmetic average of the triangle's vertices. Notice that the triangle $T(d_1, d_2, ..., d_{K_i}, 0, 0, ..., 0)$ also has center $f_T(i)$, and as the number of zeros beyond d_{K_i} is increased , all the three

corners of the resulting triangle converge to the point $f_T(i)$. When $i = 0$, then the point $f_T(0)$ is the center of the original triangle T.
Let we formally specify which sub-triangle of $T(d_1)$ we mean by $T(d_1, d_2)$ where $d_2 \neq 0$. For an arbitrary triangle $T = \Delta(\mathbf{A}, \mathbf{B}, \mathbf{C})$, and for $d_k = d_k(i) \in \{0, 1, 2, 3\}$ then the sub-triangle of T is defined as follows

$$T(d_k) = \begin{cases} \Delta\left(\frac{\mathbf{B}+\mathbf{C}}{2}, \frac{\mathbf{A}+\mathbf{C}}{2}, \frac{\mathbf{A}+\mathbf{B}}{2}\right), & d_k = 0 \\ \Delta\left(\mathbf{A}, \frac{\mathbf{A}+\mathbf{B}}{2}, \frac{\mathbf{A}+\mathbf{C}}{2}\right), & d_k = 1 \\ \Delta\left(\frac{\mathbf{A}+\mathbf{B}}{2}, \mathbf{B}, \frac{\mathbf{B}+\mathbf{C}}{2}\right), & d_k = 2 \\ \Delta\left(\frac{\mathbf{A}+\mathbf{C}}{2}, \frac{\mathbf{B}+\mathbf{C}}{2}, \mathbf{C}\right), & d_k = 3 \end{cases}.$$

In figure 3.2 this pattern is shown. If the triangle T is introduced by a vector of the three corner points A, B, and C then each component can be represented as follows

$$T(0) = \frac{\mathbf{A}+\mathbf{B}+\mathbf{C}}{2} - \frac{T}{2}, T(1) = \frac{\mathbf{A}+T}{2}, T(2) = \frac{\mathbf{B}+T}{2}, T(3) = \frac{\mathbf{C}+T}{2}.$$

Using this construction Owen and Basu define an infinite sequence of $f_T(i) \in T$ for integers $i \geq 0$. Although in case of an N point rule they take $\mathbf{x}_i = f_T(i-1)$ for $i = 1, ..., N$.
Some desirable features of this triangular Van der Corput sequence are described. It can be extended, for example if N points are sampled and we find that we need M points more simply the next M points can be taken in the sequence. It is balanced, which means if $N = 4^k$ so we get the centers of a symmetric triangulation as displayed by the last panel in figure 3.2. In case that the sample is not a multiple of 4^k, there is still reasonable balance, as shown by the third panel in figure 3.2. There exist 32 points of which the second 16 points get in the gaps left by the first 16 points. [Basu 2016]

3.2.1 Discrepancy of Triangular Van der Corput sequence

Here we state some developments on the parallel discrepancy of the triangular Van der Corput sequence. Since this discrepancy is similar for any triangle, we work with an equilateral triangle Δ_e of unit area to reduce computing areas and counting points of discrepancy calculations.

Theorem 3.2.1. *For an integer $k \geq 0$ and non-degenerate triangle $\mathcal{H} = \Delta(A, B, C)$, suppose Ψ contains $\mathbf{x}_i = f_{\mathcal{H}}(i-1), \forall i = 1, ..., N = 4^k$. Then*

$$D_N^P(\Psi, \mathcal{H}) = \begin{cases} \frac{7}{9} & if\ N = 1 \\ \frac{2}{3\sqrt{N}} - \frac{1}{9N} & otherwise \end{cases}.$$

[Basu 2016]

This theorem can be proved by consideration of several sub-cases. For a proof of the theorem we refer to Basu 2016. If the nested uniform digit scrambling of Owen 1995 is applyed to the base 4 digits of $i-1$, then $\mathbf{x}_i = f_{\mathcal{H}}(i-1), \forall i = 1, ..., N = 4^k$ are uniformly and independent distributed within their sub-triangles. In this case if the first derivative of f is bounded on Δ_e then the triangular Van der Corput sequence has trapezoidal discrepancy of $\mathcal{O}(N^{-\frac{1}{2}})$.

3.2.2 Riemann Integral over the triangle

Many books such as Marsden and Hoffman 1993 or Ash and Doleans-Dade 2000 contain the definition of Riemann Integral of a bounded function on a set in $\mathbb{R}^2$. Assume T is a non-degenerate closed triangle in $\mathbb{R}^2$. For $k \geq 0$ and $N = 4^k$, take $T_{k,1}, ..., T_{k,N}$ as a partition of T into N congruent triangles the same as T. If f is a bounded function defined on T. Then f is called Riemann integrable on T if the following limit exists for each $\mathbf{x}_{k,i} \in T_{k,i}$,

$$\lim_{k\to\infty} \frac{1}{4^k} \sum_{i=1}^{4^k} f(\mathbf{x}_{k,i}) = \mu$$

where $\mu \in \mathbb{R}$. Thus $\int_T f(\mathbf{x})d\mathbf{x}$ is equal to $\mu \times vol(T)$.
For any $k \geq 1$ and $i = 1, ..., 4^k$ we define $m_{k,i} = \inf\{f(\mathbf{x})|\ \mathbf{x} \in T_{k,i}\}$ and $M_{k,i} = \sup\{f(\mathbf{x})|\ \mathbf{x} \in T_{k,i}\}$, then we say f is Riemann integrable if and only if

$$\lim_{k\to\infty} \sum_{i=1}^{4^k} \frac{M_{k,i} - m_{k,i}}{4^k} = 0.$$

Theorem 1.6.6 of Ash and Doleans-Dade 2000 modifies an argument that the function f is Riemann integrable if it is continuous and bounded nearly everywhere on T. If the Riemann integral is obtained then it matches the Lebesgue integral.

Lemma 3.2.2. *For $N \geq 1$, let Ψ be the list of points $x_1, ..., x_N \in T$, where T is a triangle. Suppose Ψ and T have parallel discrepancy $D_N^P(\Psi, T)$. Given S as a sub-triangle of T with sides parallel to those of T. Then*

$$D_N(S, \Psi) \leq 6D_N^P(\Psi, T).$$

Proof can be found in Basu 2016.

Theorem 3.2.3. *Assume f is a Riemann integrable function on a non-degenerate triangle $\mathcal{H}$. Given $\Psi_N = (\mathbf{x}_{N,1}, ..., \mathbf{x}_{N,N})$ for $\mathbf{x}_{N,i} \in \mathcal{H}, i = 0, ..., N$ if $\lim_{N\to\infty} D_N^P(\Psi_N, \mathcal{H}) = 0$ then*

$$\lim_{N\to\infty} \frac{vol(\mathcal{H})}{N} \sum_{i=1}^{N} f(\mathbf{x}_{N,i}) = \int_{\mathcal{H}} f(\mathbf{x})d\mathbf{x}.$$

For proof we refer to Basu 2016.

3.2.3 Discussion

The Van der Corput construction is extendable and we can randomize the digits in it. If f is continuously differentiable integrand, then in Owen 1995 for $N = 4^k$, the randomization will attain integral estimates with a root mean squared error $\mathcal{O}(N^{-1})$. Another advantage of randomized Van der Corput points is that they are not strictly periodic. It is regularly advantageous in computer graphics, where unwanted visual artifacts may be produced by periodic points.

Chapter 4

Scrambled Digital Geometric Nets

The essential motivation in this Chapter is the application of equidistributed points for numerical integration with QMC samplings. We define the QMC algorithms over product spaces of the form $\mathcal{X}^s$ when $\mathcal{X}$ is a bounded set of dimension d. The cases with $d = 2$ such as triangles is especially interested. The Van der Corput construction as presented in previous Chapter is generalized from the unit triangle to some other sets and it is also replaced by digital nets in dimension s, to obtain QMC-sample in $\mathcal{X}^s$. As we mentioned before the attraction of digital nets is that we can randomize them to the extend of estimating the quadrature error via independent replications of the estimate. Compared to unrandomized QMC, these randomizations reduced the error by about $\mathcal{O}(N^{-\frac{1}{2}})$, where the integrand is smooth enough. For a survey of randomized QMC (RQMC) to improve the efficiency of simulations in finance see Lemieux 2004.

Here the aim is to study QMC and RQMC estimates, which builds on the recent work by Basu and Owen 2015a. Namely we want to estimate

$$\mu = \frac{1}{vol(\mathcal{X})^s} \int_{\mathcal{X}^s} f(\mathbf{x}) d\mathbf{x}$$

by equal weight

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} f(\mathbf{x}_i), \text{ where } \mathbf{x}_i = \phi(\mathbf{y}_i)$$

$\mathbf{y}_i$ are random points in $[0,1]^s$, and $[0,1)$ is mapped component-wisely into $\mathcal{X}$ by the transformation ϕ. Whenever we integrate over $\mathcal{X}$ to simplify several expressions we consider $vol(\mathcal{X}) = 1$.

The outline of this Chapter is as follows. Section 4.1, describes recursive geometric splits of a domain $\mathcal{X} \subset \mathbb{R}^s$ and geometric Van der Corput points based on them. In Section 4.2 we generalize those constructions to Cartesian products domains for $s \geq 1$. Section 4.3 presents the analysis of variance of the Cartesian products of such sets. In Section 4.4 we study the smoothness conditions. Section 4.5 investigates the effect of transformations τ on variation. We describe mappings from the unite cube to some specific simplices in Section 4.6. Section 4.7 of this Chapter defines non-uniform transformations including importance sampling. Conclusions are drawn in the last Section.

4.1 Splits and geometric Van der Corput points

4.1.1 Recursive geometric splits

First we need to define splitting sets. In fact splits are similar to partitions, except that the empty intersections among their parts are not required.

Definition 4.1.1. Suppose $\mathcal{X} \subset \mathbb{R}^s$ has finite and positive volume. A b-fold split of $\mathcal{X}$ is a collection of Borel sets $\mathcal{X}_\alpha$ for with $\mathcal{X} = \bigcup_{\alpha=0}^{b-1} \mathcal{X}_\alpha$, $vol(\mathcal{X}_\alpha) = \dfrac{vol(\mathcal{X})}{b}$ for $\alpha \in \mathbb{Z}_b$, and $vol(\mathcal{X}_\alpha \cap \mathcal{X}_{\acute{\alpha}}) = 0$ for $0 \le \alpha < \acute{\alpha} < b$. [Basu and Owen]

In this survey in the cases of interest, for $\alpha \neq \acute{\alpha}$, each overlap between $\mathcal{X}_\alpha$ and $\mathcal{X}_{\acute{\alpha}}$ lands on the boundaries of these sets. In QMC the unit interval $[0, 1)$ is partitioned into sub-intervals $\left[\dfrac{\alpha}{b}, \dfrac{\alpha+1}{b}\right)$. Approaching $\mathcal{X} = [0, 1]$ requires exceptions that the rightmost interval is closed and all others are half-open. In general for closed sets $\mathcal{X}$ it can be difficult to keep path of which subsets have which parts of their boundaries. Using splits we are allowed for instance to divide a closed triangle into three congruent closed triangles as displays in Figure 4.1. Basu and Owen prefer to use a randomization under which the probability of any sample point appearing on a split boundary, is zero.

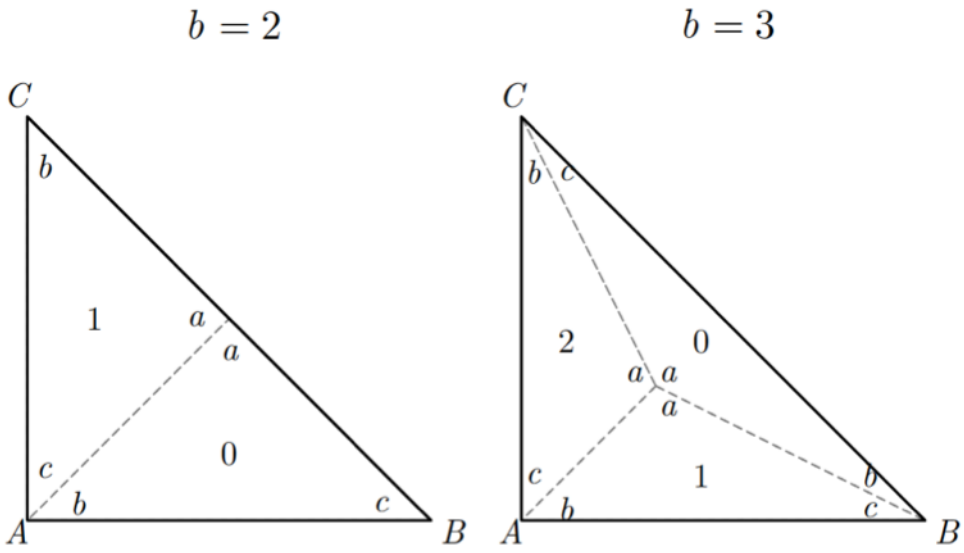


Figure 4.1: Splits of a triangle $\mathcal{X}$ for bases $b = 2, 3$. The sub-triangles $\mathcal{X}_\alpha$ are labeled by the digit $\alpha \in \mathbb{Z}_b$. [Source: Basu 2016].

Figure 4.1 presents a triangle $\mathcal{X} = ABC$ split into sub-triangles. The left panel has $b = 2$ sub-triangles and the right panel has $b = 3$. In case of $b = 2$, the vertex A is connected to the midpoint of the opposite edge. The subset ABC_0 is the one containing vertex B. In the both cases, the new A is the mean of the old B and C. Through an algebraic description we define the new ABC using lower case abc, (base 2 and digit 0);

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix}.$$

Definition 4.1.2. Suppose $\mathcal{X} \subset \mathbb{R}^s$ has finite and positive volume. A recursive b-fold split of $\mathcal{X}$ is a collection of $\mathbf{Y}$ of sets consisting of $\mathcal{X}$ and strictly one b-fold split of each set in the collection. The member of $\mathbf{Y}$ is called cell. [Basu and Owen]

A member of a recursive split of $\mathcal{X}$ lies at level $r \geq 1$ of the recursive split, if it occurs after r splits of $\mathcal{X}$. The subset of $\mathcal{X}_{\alpha_1}$ corresponding to $\alpha = \alpha_2$ is the cell $\mathcal{X}_{\alpha_1,\alpha_2}$ and similarly $\mathcal{X}_{\alpha_1,\alpha_2,\ldots,\alpha_r}$ is an arbitrary cell at level $r \geq 2$ for $\alpha_i \in \mathbb{Z}_b$. For example the original set $\mathcal{X}$ is at level 0, and the cells $\mathcal{X}_0, \ldots, \mathcal{X}_{b-1}$ are at level 1. Since we need to specify all of the cells in a split $\mathbf{Y}$, we take $t = \sum_{i=1}^{r} \alpha_i b^{i-1} \in \mathbb{Z}_{b^r}$ where $\mathcal{X}_{(r,t)} = \mathcal{X}_{\alpha_1,\ldots,\alpha_r}$. Thus the cells in the split $\mathbf{Y}$ are $\mathcal{X}_{(r,t)}$ when $r \in \mathbb{N}$, $t \in \mathbb{Z}_{b^r}$ and $\mathcal{X}_{(0,0)} = \mathcal{X}$.

The first few levels of recursive splits for each of the splits from Figure 4.1 is shown in Figure 4.2.

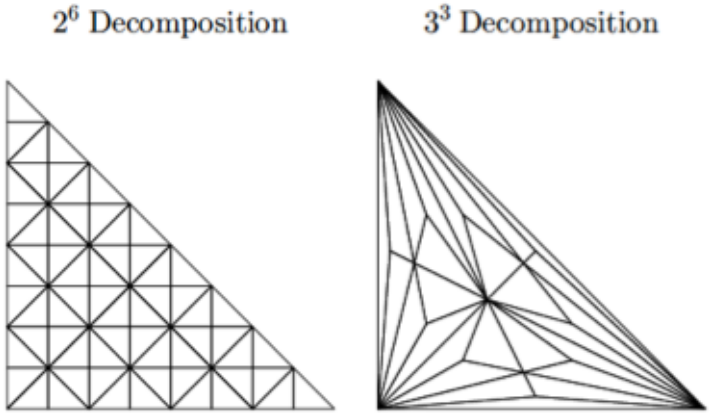


Figure 4.2: The base b splits from Figure 4.1 transmitted to $r = 6$ or 3 levels. [Source: Basu 2016].

4.1.2 Geometric Van der Corput sequence

Given a recursive splitting of a set $\mathcal{X}$ in base b, a geometric Van der Corput sequence for the set $\mathcal{X}$ can be constructed. We write the integer i in base b as $i = \sum_{r=1}^{\infty} \alpha_r(i) b^{r-1}$, and then to this integer i a sequence of sets is defined as follows

$$\mathcal{X}_{i:R} = \mathcal{X}_{\alpha_1(i),\ldots,\alpha_R(i)}.$$

Hence $\mathbf{x}_i$ is any point in $\bigcap_{R=1}^{\infty} \mathcal{X}_{i:R}$ and the volume of $\mathcal{X}_{i:R}$ is b^{-R} that converges to zero as $R \to \infty$. For the constructions that we are interested in, any $\mathbf{x}_i$ is a uniquely determined point. Like the base 3 decomposition in Figure 4.2, for decompositions without nice aspect ratios, some of the sequences converge to a line segment. For example given $i \in \{0, 1, 2\}$, then $\alpha_r(i) = 0$ for $r \geq 1$ and the infinite tail of zeros tends to a point $\mathbf{x}_i$ on one of the sides of the triangle. [Basu and Owen] We use the notion of a sequence of sets converging nicely to a point, to get a unique limit $\mathbf{x}_i$, which is a version from Stromberg 1994.

Definition 4.1.3. The sequence $\mathcal{S}_r \in \mathbb{R}^s$ of Borel sets for $r \in \mathbb{N}$ converges nicely to $\mathbf{x} \in \mathbb{R}^s$ as $r \to \infty$ if there is $\gamma < \infty$ and s-dimensional cubes $\mathcal{C}_r$ such that $\mathbf{x} \in \mathcal{C}_r$, $\mathcal{S}_r \subseteq \mathcal{C}_r, 0 < vol(\mathcal{C}_r) \leq \gamma vol(\mathcal{S}_r)$, and $\lim_{r\to\infty} diam(\mathcal{S}_r) = 0$.

Notice that a sequence of sets which converges nicely to $\mathbf{x}$ cannot also converge nicely to any $\mathbf{y} \neq \mathbf{x}$. Generally the following condition is assumed.

Definition 4.1.4. A recursive split $\mathbf{Y}$ in base b is convergent if for each infinite sequence $\alpha_1, \alpha_2, \ldots \in \mathbb{Z}_b$, the cells $\mathcal{X}_{\alpha_1,\ldots,\alpha_R}$ as $R \to \infty$ converges nicely to a point which is denoted by $\lim_{R\to\infty} \mathcal{X}_{\alpha_1,\ldots,\alpha_R}$. [Basu and Owen]

In a geometric Van der Corput sequence, a convergent recursive split is taken and $\mathbf{x}_i$ is defined as follows

$$\mathbf{x}_i = \lim_{M\to\infty} \mathcal{X}_{\alpha_{1(i-1)},\ldots,\alpha_{R(i-1)}\overbrace{0,0,\ldots,0}^{M}}.$$

when $R \neq 0$ is the last digit in the expansion of $i-1$ and there exist $M \geq 1$ zeros above. For triangular splits of base $b = 2$, $\mathbf{x}_i$ is an interior but not a central point and for base 3, the recursive split is not convergent.

Definition 4.1.5. Suppose $\mathbf{Y}$ is a recursive split of $\mathcal{X} \subset \mathbb{R}^s$ in base b. Then $\mathbf{Y}$ satisfies the sphericity condition (a variation of sphericity is called 'circularity')) if there is a constant $C < \infty$ that for all cells $\mathcal{X}_{\alpha_1,\ldots,\alpha_r}$ in $\mathbf{Y}$ we have $diam(\mathcal{X}_{\alpha_1,\ldots,\alpha_r}) \leq Cb^{-\frac{r}{s}}$. [Basu 2016]

Here we mention that under sphericity assumption a recursive split is necessarily convergent. The lowest value for the constant C is 1 for $s = 1$ so the cells are intervals but usually C is greater than 1. Hence without loss of generality we postulate that $1 \leq C < \infty$.

Definition 4.1.6. Let $\mathbf{Y}$ be a a convergent recursive split of $\mathcal{X} \subset \mathbb{R}^s$ in base b. Then the $\mathbf{Y}$-transformation of $[0,1)$ is a function $\phi : [0,1) \to \mathcal{X}$ such that $\phi(x) = \lim_{R\to\infty} \mathcal{X}_{x_1,\ldots,x_R}$, where x has the base b description $0.x_1x_2\ldots$ and if x has two descriptions, so we use the one with trailing zeros. [Basu and Owen]

4.2 Digital geometric nets and Scrambled geometric nets

Given a bounded set $\mathcal{X} \subset \mathbb{R}^s$ with finite nonzero volume. Digital geometric nets in $\mathcal{X}^d$ is defined via splittings. The set $\{1,\ldots,d\}$ for $d \in \mathbb{N}$ is denoted by $1:d$ and for $i \in 1:d$ the bounded sets $\mathcal{X}^{(i)} \subset \mathbb{R}^{s_i}$ with volume 1 are given. We represent the complement $1:d-u$ for the sets of indices $u \subseteq 1:s$ by $-u$ and the cardinality of u by $|u|$. $\mathcal{X}^u$ denotes the Cartesian product of $\mathcal{X}^{(i)}$ for $i \in u$. Any vector $\mathbf{x} \in \mathcal{X}^{1:d}$ contains components of the form $\mathbf{x}_i \in \mathcal{X}^i$. $\mathbf{x}^u$ defines the vector in $\mathcal{X}^u$ with components $\mathbf{x}_i$ for $i \in u$. Any point in $\mathcal{X}^{1:d}$ has components of the form $\sum_{i=1}^d s_i$, which is written as $(\mathbf{x}_1,\ldots,\mathbf{x}_s) = \mathbf{x}$.

Definition 4.2.1. Let $\mathbf{Y}_i$ be a recursive split of $\mathcal{X}^{(i)}$ in base b for $i = 1, \ldots, d$. The cells of $\mathbf{Y}_i$ is denoted by $\mathcal{X}_{i,(k,t)}$ for $k \in \mathbb{N}$ and $t \in \mathbb{Z}_{b^k}$. Thus a b-adic cell for those splits is a Cartesian product

$$\prod_{i=1}^{d} \mathcal{X}_{i,(k_i,t_i)}$$

where $k_i \geq 0$ and $t_i \in \mathbb{Z}_{b^{k_i}}$. [Basu 2016]

Definition 4.2.2. For $i \in 1 : d$, let $\mathcal{X}^{(i)} \subset \mathbb{R}^{s_i}$ have volume one. Suppose $\mathbf{Y}_i$ is a recursive split of $\mathcal{X}^{(i)}$ in base b. Then for integers $0 \leq p \leq n$ the points $\mathbf{x}_1, \ldots, \mathbf{x}_{b^n} \in \mathcal{X}^{1:s}$ are a geometric (p, n, s)-net in base b if each b-adic cell of volume b^{p-n} contains exactly b^p of the $\mathbf{x}_j$. (p, n, s)-net in base b is called a weak geometric net if each b-adic cell of volume b^{p-n} contains at least b^p of the $\mathbf{x}_j$. [Basu and Owen]

Suppose the sequence $\mathbf{x}_1, \ldots, \mathbf{x}_m$ form a (p, n, s)-net in base b and $\mathbf{y}_1, ..., \mathbf{y}_m$ form a nested uniform scramble of them. Let $\mathbf{Y}_i$ be a recursive split of the unit volume set $\mathcal{X}^{(i)} \subset \mathbb{R}^{s_i}$ in base b with transformation ϕ. Then $\mathbf{x}_j = \phi(\mathbf{y}_j)$(componentwise) is a geometric (p, n, s)-net in base b and $\mathbf{z}_j = \phi(\mathbf{x}_j)$(componentwise) is a weak geometric (p, n, s)-net in base b with probability one. For proof see Basu 2016.

4.2.1 Preservation of Measure

Given a function $\phi : [0, 1) \to \mathcal{X}$, where $\mathcal{X} \subset \mathbb{R}^s$, which maps points in unit interval to $\mathcal{X}$ according to the convergent recursive split $\mathbf{Y}$, we show that this function preserves the uniform distribution. Here we need to differentiate Lebesgue measures of different dimensions, so to that end we take λ for Lebesgue measure in $\mathbb{R}$ and λ_s for Lebesgue measure in $\mathbb{R}^s$.

Let $\mathbf{Y}$ be a recursive split of $\mathcal{X} \subset \mathbb{R}^s$, with $vol(\mathcal{X}) = 1$ in base $b \geq 1$. Assume ϕ is a $\mathbf{Y}$-transformation of the interval $[0, 1)$ and $B \subset \mathcal{X}$ is a Borel set. Hence

$$\lambda(\phi^{-1}(B)) = \lambda_s(B),$$

where $\phi^{-1}(B) = \{x \in B;\ \phi(x) \in B\}$.
Proof of this proposition can be found in Basu 2016.

Following proposition represents this uniformity under scrambling.

For $i \in 1 : d$, let $\mathcal{X}^{(i)} \subset \mathbb{R}^{s_i}$ have volume one. Suppose $\mathbf{Y}_i$ is a recursive split of $\mathcal{X}^{(i)}$ in bases $b_i \geq 2$ with corresponding transformations ϕ_i. If y_i is a base b_i nested uniform scramble of x_i for $\mathbf{x} \in [0, 1)^d$, then $\phi(\mathbf{x}) = (\phi^{-1}(y_1), \ldots, \phi^{-1}(y_d)) \sim \mathbb{U}(\mathcal{X}^{1:d})$.

Geometric scrambled nets do not require that we use the same base to define both the transformations and the digital net. Even part of the basic properties of scrambled nets can be applied for geometric scrambled nets without requiring smoothness of the integrand.

Theorem 4.2.1. *For $i \in 1 : d$, let $\mathcal{X}^{(i)} \subset \mathbb{R}^{s_i}$ have volume one. Suppose $\mathbf{Y}_i$ is a convergent recursive split of $\mathcal{X}^{(i)}$ in bases $b_i \geq 2$ with corresponding transformations*

ϕ_i. Let $\mathbf{y}_1, \ldots, \mathbf{y}_N$ be a nested uniform scramble of a (p, n, s)-net in base $b \geq 2$ and let $f \in \mathcal{L}^2(\mathcal{X}^{1:d})$. If $\mathbf{x}_i = \phi(\mathbf{y}_i)$, then:

$$Var(\hat{\mu}) = o(\frac{1}{N})$$

as $N \to \infty$. [Basu and Owen]

Proof. Since $f \in \mathcal{L}^2(\mathcal{X}^{1:d})$ so $f \circ \phi \in \mathcal{L}^2[0,1]^d$. Therefore Theorem 2.4.2 can be applied. □

Theorem 4.2.2. *Assume $\mathbf{y}_1, \ldots, \mathbf{y}_N$ be a nested uniform scramble of a (p, n, s)-net in base $b \geq 2$ and let $f \in \mathcal{L}^2(\mathcal{X}^{1:d})$ and $var(f(\mathbf{x})) = \sigma^2$ where $\mathbf{x} \sim \mathbb{U}(\mathcal{X}^{1:d})$. Then:*

$$Var(\hat{\mu}) \leq b^p \left(\frac{b+1}{b-1}\right)^{d-1} \frac{\sigma^2}{N}.$$

If $p = 0$, then $Var(\hat{\mu}) \leq e\frac{\sigma^2}{N} = 2.718\frac{\sigma^2}{N}$.

Proof. Once again $f \in \mathcal{L}^2(\mathcal{X}^{1:d})$ so $f \circ \phi \in \mathcal{L}^2[0,1]^d$. Thus Theorem 2.4.3 can be applied. □

4.3 Analysis of Variance (ANOVA) of $\mathcal{X}^{1:d}$

It is well known that QMC-sampling looses the better accuracy in high dimension. Hence it is fundamental to release the most important (in statistical sense) components or to diminish the nominal dimension of the problem by means of ANOVA considerations. Where $f \in \mathcal{L}^2(\mathcal{X}^{1:d})$ then f can be written into the sum of orthogonal functions each of them defined in a different subset $u \subseteq 1:d$, which only depends on the variables in any of these subsets

$$f(\mathbf{x}) = \sum_{u \subseteq 1:d} f_u(\mathbf{x}).$$

Now if $|u|$ denote the cardinality of u and $\sigma^2 = \int_{\mathcal{X}^{1:d}} (f(\mathbf{x}) - \mu)^2 d(\mathbf{x})$, $\sigma_u^2 = \int_{\mathcal{X}^{1:d}} f_u(\mathbf{x})^2 d\mathbf{x}$, $\sigma_\emptyset^2 = 0$, then assuming $\sigma < \infty$ and $|u| > 0$ it holds

$$\sigma^2 = \sum_{u \subseteq 1:d} \sigma_u^2.$$

This equation partitions the total variance into parts corresponding to any subset $u \subseteq 1:d$. [Rutherford 2012]

4.4 Smoothness

Some smoothness of the integrand require to reach the main results for scrambling geometric nets. Assume f is a real-valued function on $\mathcal{X} \subset \mathbb{R}^m$, where the dimension $m = d \times s$, for an d-fold tensor product of a s-dimensional domain. For $v \subseteq 1:m$, the mixed partial derivative of f is taken once with respect to x_i for any $i \in v$ denotes by $\partial^v f$. Since differentiating a function zero times leaves it unchanged we have $\partial^\emptyset f = f$.

Definition 4.4.1. Suppose $\mathcal{X} \subset \mathbb{R}^m$ for $m \in \mathbb{N}$. The function $f : \mathcal{X} \to \mathbb{R}$ is smooth if $\partial^{1:m} f$ is continuous on $\mathcal{X}$. [Basu and Owen]

Basu and Owen prove that under smoothness and sphericity assumptions, the variance of averages over scrambled geometric nets is $\mathcal{O}\left(\frac{\log(N)^{s-1}}{N^{1+\frac{2}{s}}}\right)$.

4.5 Impact of Transformations on Variation

Monte Carlo sampling over a simplex domain $\mathcal{X} \subset \mathbb{R}^s$ is usually done by obtaining a uniformity preserving transformation $\tau : [0,1]^m \to \mathcal{X}$, where it yields $\mathbf{x} = \tau(\mathbf{y}) \sim \mathbb{U}(\mathcal{X})$. When $\mathbf{y} \sim \mathbb{U}([0,1]^m)$, then

$$\frac{1}{vol(\mathcal{X})} \int_{\mathcal{X}} f(\mathbf{x}) d\mathbf{x} = \int_{[0,1]^m} f(\tau(\mathbf{y})) d\mathbf{y}. \ (1)$$

So we estimate $\mu = \int_{\mathcal{X}} f(\mathbf{x}) d\mathbf{x}$ by

$$\frac{vol(\mathcal{X})}{N} \sum_{i=1}^{N} f(\tau(\mathbf{y}_i))$$

for $\mathbf{y}_i \overset{\text{iid}}{\sim} \mathbb{U}([0,1]^s)$, and often for simplicity we take $vol(\mathcal{X}) = 1$.

For QMC sampling of such regions a very standard approach is to employ the same transformation as in MC, and replace independent random variables $\mathbf{y}_i$ by QMC or randomized QMC (RQMC) points.

4.5.1 Necessary and Sufficient Conditions

QMC sampling gives an error rate $\mathcal{O}\left(\frac{\log(N)^{s-1}}{N}\right)$ if the function f is of bounded variation in the sense of Hardy and Krause (BVHK). For scrambled nets, a sort of RQMC, to reach a root mean squared error of order $\mathcal{O}\left(\frac{\log(N)^{\frac{s-1}{2}}}{N^{\frac{3}{2}}}\right)$ the function f has to be smooth in the sense as follows

$$\|\partial^v f\|_2^2 := \int (\partial^v f(\mathbf{x}))^2 d\mathbf{x} < \infty, \ \forall \ v \subseteq 1 : s. \ (\star)$$

[Dick and Pillichshammer 2010]

Following the notations from Section 2.3.1 we describe the sufficient conditions under which $f \circ \tau$ is smooth for all $f \in C^m(\mathcal{X})$ in the sense of equation $(\star)$. Similarly we define conditions on τ such that $f \circ \tau \in$ BVHK for all $f \in C^m(\mathcal{X})$.

Theorem 4.5.1. *For $\tau : [0,1]^m \to \mathcal{X}$, where $\tau(\mathbf{y}) = (\tau_1(\mathbf{y}), \ldots, \tau_s(\mathbf{y}))$ and $\mathcal{X}$ is a bounded and close subset of $\mathbb{R}^d$, assume that $\partial^{1:m}\tau_i$ exists for all $i = 1, \ldots, s$. If*

$$\int_{[0,1]^s} \prod_{j=1}^{d} |\partial^{l_j} \tau_{k_j}(\mathbf{y})|^2 d\mathbf{y} < \infty$$

holds for all nonempty $u \subseteq 1:m$, $d \subseteq 1:|u|$, for all disjoint l_j with union u and for all distinct $k_j \in 1:s$. Hence for all $f \in C^m(\mathcal{X})$, $f \circ \tau$ is smooth in the sense of equation $(\star)$. [Basu 2016]

This equation gives us sufficient conditions under which $f \circ \tau$ is smooth for all $f \in C^m(\mathcal{X})$.
Following Theorem illustrates conditions on f and τ so that $f \circ \tau$ belongs to BVHK for all $f \in C^m(\mathcal{X})$.

Theorem 4.5.2. *For $\tau : [0,1]^m \to \mathcal{X}$, where $\tau(\mathbf{y}) = (\tau_1(\mathbf{y}), \ldots, \tau_s(\mathbf{y}))$ and $\mathcal{X}$ is a bounded and close subset of $\mathbb{R}^s$, assume that $\partial^{1:m}\tau_i$ exists for all $i = 1, \ldots, s$. If*

$$\int_{[0,1]^{|u|}} \prod_{j=1}^{d} |\partial^{l_j} \tau_{k_j}(\mathbf{y}^u : 1^{-u})| d\mathbf{y}^u < \infty$$

holds for all nonempty $u \subseteq 1:m$, $d \subseteq 1:|u|$, for all disjoint l_j with union u and for all distinct $k_j \in 1:s$. Hence for all $f \in C^m(\mathcal{X})$, $f \circ \tau \in BVHK$. [Basu 2016]

For QMC analysis on spheres, Brauchart et al. 2014 describe similarly that higher dimensions requires greater smoothness to control a worst case quadrature error.

4.6 Uniformity preserving mappings

Uniformity transformations satisfy equation (1). In this Section we represent the uniformity preserving mappings from Fang and Wang 1994. They render transformations from $[0,1]^s$ to some particular domains for quadrature problems. Here we consider the following domains

$$A_s = \{(x_1, \ldots, x_s); 0 \le x_1 \le \ldots \le x_s \le 1\}$$
$$V_s = \{(x_1, \ldots, x_s) \in \mathbb{R}_s^+; x_1 + \ldots + x_s \le 1\}$$
$$T_s = \{(x_1, \ldots, x_s) \in \mathbb{R}_s^+; x_1 + \ldots + x_s = 1\}$$

where $\mathbb{R}_s^+$ is the non-negative part of $\mathbb{R}_s$. A_s, V_s and T_{s+1} are all s-dimensional simplices. Then we show that all those mappings have components τ in BVHK and none of them have all mixed partial derivatives in $\mathcal{L}^2$. Thus they are appropriated to QMC, but they are not smooth enough to take advantage of the improved rate for RQMC versus QMC. These transformations allow us to study them directly because they have a separable character.

4.6.1 Mapping from the unit cube to A_s

The map $\tau = (\tau_1, \ldots, \tau_s)$ is defined by $\tau_i(\mathbf{y}) = \prod_{k=i}^{s} y_k^{\frac{1}{k}}$ for $i = 1, \ldots, s$. The partial mixed derivative of τ_1 is

$$(\partial^{1:s}\tau_1)^2 = \prod_{k=1}^{s} \frac{1}{k^2} y_k^{\frac{2}{k-2}}$$

which is diverged on integrating with respect to y_2, so $\partial^{1:s}\tau_1 \notin \mathcal{L}^2$. It can be shown that τ satisfies the BVHK conditions of Theorem 4.5.2. Given any nonempty $l \subseteq 1:s$ where for $k \in l$ if $k < i$ then $\partial^l \tau_i = 0$. Thus we suppose that $l \subseteq i:s$

so we have

$$\partial^l \tau_i = \prod_{k\in l} \frac{1}{k} y_k^{\frac{1}{k-1}} \prod_{k\in i:s-l} y_k^{\frac{1}{k}}.$$

Then this τ satisfies the conditions of Theorem 4.5.2 since all the powers of each y_k are above -1.

4.6.2 Mapping from the unit cube to V_s

Suppose $s \geq 2$ to get a dimension of at least 2, then the map τ is defined as follows

$$\tau_i = y_1^{s^{-1}} \prod_{k=2}^{i} y_k^{(s-k+1)^{-1}} (1 - y_{i+1}^{(s-i)^{-1}}) \quad \text{for} \quad i = 1, \ldots, s-1$$

$$\tau_s = y_1^{s^{-1}} \prod_{k=2}^{s} y_k^{(s-k+1)^{-1}}$$

The mixed partial derivative of τ_s is

$$\partial^{1:s}\tau_s = s^{-1} y_1^{s^{-1}-1} \prod_{k=2}^{s} (s-k+1)^{-1} y_k^{(s-k+1)^{-1}-1} = (s!)^{-1} y_1^{(\frac{s-1}{s})^{-1}} y_2^{(\frac{s-2}{s-1})^{-1}} \ldots y_{s-1}^{(\frac{1}{2})^{-1}}.$$

Noticing that the integral with respect to y_{s-1} obviously $\partial^{1:s}\tau_s$ is not in $\mathcal{L}^2$. This transformation satisfies Theorem 4.5.2 following the same argument as previous Section.

4.6.3 Mapping from $[0,1]^{s-1}$ to T_s

In this case we assume $s \geq 3$ and define the map τ as follows

$$\tau_i = \prod_{k=1}^{i-1} y_k^{(s-k)^{-1}} (1 - y_i^{(s-i)^{-1}}) \quad \text{for} \quad i = 1, \ldots, s-1$$

$$\tau_s = \prod_{k=1}^{s-1} y_k^{(s-k)^{-1}}$$

The mixed partial derivative of τ_s is

$$\partial^{1:(s-1)}\tau_s = ((s-1)!)^{-1} y_1^{(\frac{s-2}{s-1})^{-1}} y_2^{(\frac{s-3}{s-2})^{-1}} \ldots y_{s-2}^{(\frac{1}{2})^{-1}} y_{s-1}.$$

$\partial^{1:(s-1)}\tau_s$ is not in $\mathcal{L}^2$ and following the same argument as before this transformation also satisfies Theorem 4.5.2.

4.7 Non-uniformity preserving Transformations

In this Section we consider importance sampling methods to integrate with respect to a non-uniform measure on $\mathcal{X}$. Importance sampling QMC for the simplex is introduced and it is shown that some importance sampling methods give the $\mathcal{O}(N^{-\frac{3}{2}+\epsilon})$ rate for RMSE on the simplex.

4.7.1 Importance Sampling QMC for Simplex

Let $\tau : [0,1]^s \to \mathcal{X}$, where $\tau(\mathbf{y}) = \mathbf{x},\ \mathbf{y} = (y_1, \ldots, y_s)$ be a mapping in the simplex

$$A_s = \{(x_1, \ldots, x_s); 0 \le x_1 \le ... \le x_s \le 1\}.$$

The mapping is defined by

$$x_i = \tau_i(\mathbf{y}) = \prod_{k \ge i} y_k^{c_k}$$

for constant $c_k > 0$.

In the case of the uniformity preserving mapping from Fang and Wang 1994, c_k is $\frac{1}{k}$. Since the Jacobian matrix for this transformation is upper triangular so the the Jacobian determinant is

$$J(\mathbf{y}) = \prod_{i=1}^{s} \frac{\partial x_i}{\partial y_i} = \prod_{i=1}^{s} c_i y_i^{c_i - 1} \prod_{k>i} y_k^{c_k} = C \prod_{i=1}^{s} y_i^{ic_i - 1}$$

where $C = \prod_i c_i$.

Then $0 \le J(\mathbf{y}) \le C$ and the average of $J(\mathbf{y})$ equals $\frac{1}{vol(A_s)} = \frac{1}{s!}$. This choice of Fang and Wang 1994 determines $J(\mathbf{y}) = \frac{1}{s!}$ for all $\mathbf{y}$. Then the RQMC estimate of

$$\mu = s! \int_{A_s} f(\mathbf{x}) d\mathbf{x} = s! \int_{[0,1]^s} f(\tau(\mathbf{y})) J(\mathbf{y}) d\mathbf{y}$$

is $\hat{\mu} = \frac{s!}{N} \sum_{i=1}^{N} f(\tau(\mathbf{y}_i)) J(\mathbf{y}_i)$.

If we ignore the $s!$ factor, the integrand on $[0,1]^s$ is now $\tilde{f}(\mathbf{y})) = f(\tau(\mathbf{y}))J(\mathbf{y})$, and $\partial^v \tilde{f} = \sum_{w \subseteq v} \partial^w (f \circ \tau) \times \partial^{v-w} J$. In this case the definition of τ_i makes it suitable to work with a simple function class consisting of integrands of the form $\prod_{i=1}^{s} x_i^{a_i}$ for real values $a_i \ge 0$.

Theorem 4.7.1. *Suppose $f(\mathbf{x}) = \prod_{i=1}^{s} x_i^{a_i},\ i = 1, \ldots, s$ for $\mathbf{x} \in A_s$ and $a_i \ge 0$. For $\mathbf{y} \in [0,1]^s$, let $x_i = \tau_i(\mathbf{y}) = \prod_{k \ge i} y_k^{c_k}$ and the Jacobian be $J(\mathbf{y}) = \prod_{i=1}^{s} y_i^{ic_i - 1}$ for positive values of c_i. Then $\partial^u f(\mathbf{x}(\mathbf{y}))J(\mathbf{y}) \in \mathcal{L}^2([0,1]^s)$ for all $u \subseteq 1:s$ and all a_i if and only if $c_i > \frac{3}{2i}$ holds for $i = 1, \ldots, s$.*

Proof. Suppose $A_k = \sum_{i \le k} a_k$ and $C = \prod_i c_i$, then

$$\tilde{f}(\mathbf{y})) = f(\tau(\mathbf{y}))J(\mathbf{y}) = C \prod_{k=1}^{s} \mathbf{y}^{kc_k - 1 + c_k A_k}.$$

Let $(\partial^v \tilde{f}(\mathbf{y})))^2$ for all $v \subseteq 1:s$ be defined as follows

$$C^2 \prod_{k \in v} (kc_k - 1 + c_k A_k)^2 \mathbf{y}^{2(kc_k - 2 + c_k A_k)} \prod_{k \in -v} \mathbf{y}^{2(kc_k - 1 + c_k A_k)}. \quad (2)$$

For all a_i the coefficient $(kc_k - 1 + c_k A_k)$ cannot vanish, hence (1) has a finite integral for all a_i if and only if $\forall\, i = 1, \ldots, s$ and all a_i, $2(ic_i - 2 + c_i A_i) > -1$, which easily holds if all $c_i > \dfrac{3}{2i}$. Conversely, let $c_i \leq \dfrac{3}{2i}$ for some i, then we may choose $A_i = 0$ and $v = \{i\}$ and see that (2) is not integrable. [Basu and Owen 2015b] □

Theorem 4.7.1 shows that RQMC can obtain the $\mathcal{O}(N^{-\frac{3}{2}+\epsilon})$ rate for functions of the form $\prod_{i=1}^{s} x_i^{a_i}$ on the simplex A_s. That rate can be extended to linear combinations of finitely many such functions, including polynomials. If we choose $c_i = \dfrac{3}{2i} + \kappa$ for some small $\kappa > 0$ then for fixed s we have $s!J(\mathbf{y}) = (3/2)s + O(\kappa)$. Thus there exists a dimension effect. The integrand becomes more spiky as s increases. It can be expected that the lead constant in the error bound grows exponentially with s. For $s = 1$, Theorem 4.7.1 requires $c_1 > 3/2$ when in fact RQMC attains the $\mathcal{O}(N^{-\frac{3}{2}+\epsilon})$ RMSE with $c_1 = 1$ in that case. The reason for this difference is that the theorem covers more complicated integrands like $x_1^{1/2}$ whose derivative does not belong to $\mathcal{L}^2$. If we work with polynomials taking only $a_i \in \mathbb{N}$, then the choice $c_k = 1/k$ zeros out (2) when $A_k = 0$. So the smallest nonzero A_k is 1 and we need to impose the coefficient $2(ic_i - 2 + c_i A_i) > -1$. This can be simplified to $A_k > k/2$ that can only be ensured for $k = 1$ and thus the Fang and Wang choice $c_k = 1/k$ will not obtain the RQMC rate for polynomials when $s > 2$.

In MC-sampling, the impact of nonuniform importance sampling sometimes measures via an effective sample size, see Kong, Liu, and Wong 1994. For $J(\mathbf{y})$ the effective sample size can be denoted as the nominal one multiplied by

$$\frac{(\int J(\mathbf{y})d\mathbf{y})^2}{\int J(\mathbf{y})^2 d\mathbf{y}}.$$

Given $c_i = \dfrac{3}{2i}$ this factor is $(\dfrac{8}{9})^s$, that is corresponding to a mild exponential decreasing in effectiveness for MC-sampling. It seems to be as yet no suitable measure of effective sample size for RQMC.

4.8 Conclusion

Quasi-Monte Carlo methods aim to achieve the convergence rate by using low discrepancy sequences. These sequences render better stratification and has theoretical error bounds of size $\mathcal{O}(N^{-1}(\log N)^s)$. [Moskowitz and Caflisch 1996] Essentially RQMC was proposed to give a more reasonable estimation of the integration error, because the Koksma-Hlawka bound for QMC determines unconfined error estimates for many practical applications. Following the approach in Basu 2016, it has been shown that a sequence of Quasi-Monte Carlo methods exists for integration over d-fold products of simplices which has a Monte Carlo rate of convergence. Basu and Owen 2015a give constructions for particular case of a single triangle with a higher rate of convergence than Monte Carlo. To obtain a variance rate of $\mathcal{O}\left(\dfrac{\log(N)^{s-1}}{N^{1+\frac{2}{s}}}\right)$,

Basu and Owen generalized a scrambled net method to construct points on the product of d, s-dimensional spaces. Removing the dependency on d needs much more work.

Chapter 5

Implementation

In the numerical examples, two simple 2-dimensional integrals with bounded and unbounded integrands are considered and the aim is to compare the integration errors using MC-, QMC- and RQMC-methods. Implementation of the RQMC-algorithm according to Basu 2016 is quiet complicated which is a drawback from a practical point of view. Thus in the following Section a Matlab implementation of the RQMC-algorithm on unite cube is presented and then we use the RQMC-algorithm to estimate the following integrals. Later a uniformity transformation is used to map from the unit cube to the right-angle triangle defined by $A = (0,0)^T, B = (0,1)^T$ and $C = (1,1)^T$.

5.1 Determination of Scrambling-algorithm

Due to create a nested uniform scrambling which was discussed in Section 2.4.1 we use *stream* as the random generator, and then we implement a point set randomization that performs Owen's nested uniform scrambling. To randomize QMC, one can take a QMC-sequence say x_i and transforms it into random points y_i such that y_i retain a QMC property. The easiest way to achieve it, is to have each $y_i \sim \mathbb{U}[0,1)^s$,

$$y_i = x_i + u \ (mod\, 1)$$

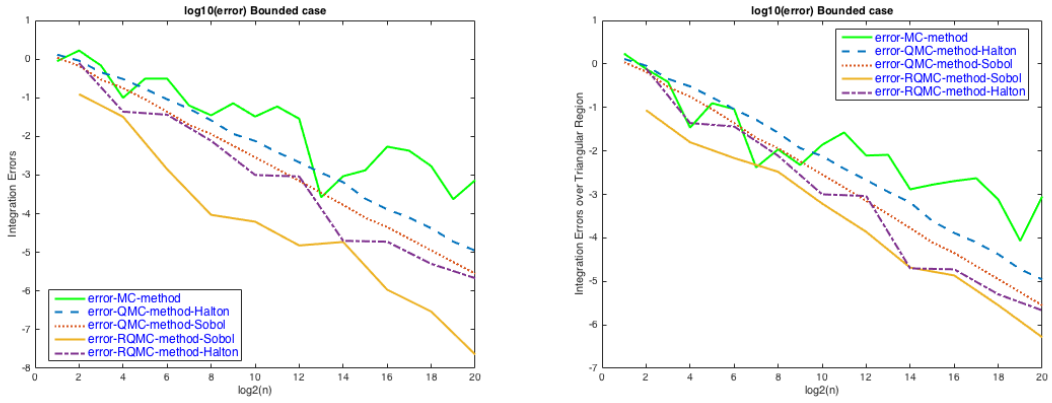
where $u \sim \mathbb{U}[0,1)^s$. The idea of this algorithm is to scramble $n_1(k)$ bits of a low-discrepancy sequence with $2^k = n_1(k) \times n_2(k)$ points when the number of scrambling is $n_2(k)$. The value of $n_2(k)$ can be any positive integer as long as we are able to find a proper Linear Congruential Generators (LCG) for it. Since the common pseudo-random number generators are LCGs with both power-of-two and prime moduli so one possible choice can be $n_1(k) = n_1(k) = 2^j$. In this approach using LCGs as scramblers is the main step.

5.2 Integration of a bounded integrand

In this Section we will compare the integration errors of a bounded integrand using MC-, QMC- and RQMC-methods. Suppose we want to estimate the following integration

$$\mathfrak{J} = \int_0^1 \int_0^1 \mathrm{e}^{(x+y)}\, dx\, dy$$

which analytically can be computed and the true value equals $(\mathrm{e}-1)^2$, so we are able to determine the integration errors. Note that the true integration result over the triangle ABC is the same as the true integration result over unit cube.



(a) (b)

Figure 5.1: Logarithmic Plots of Relative Integration Errors of a bounded integrand over Unit Cube and Triangle ABC.

Figures 5.1a and 5.1b show the integration errors of MC, QMC and RQMC over unit cube and the triangle ABC, where n is the number of points used for each estimators. In the both cases the performance of estimation using QMC vs MC can be pointed out when n increases. The Koksma–Hlawka inequality is proved by Hlawka for Riemann integrable[1] function f and $x_1, \ldots, x_N \in [0,1]^s$

$$|\hat{\mathfrak{J}} - \mathfrak{J}| \leq V_{HK}(f) D^*(x_1, \ldots, x_N)$$

and many constructions assure that $D^*(x_1, \ldots, x_N) = \mathcal{O}(N^{-1+\epsilon})$ for all $\epsilon > 0$. Hence when $V_{HK}(f) < \infty$ Hlawka's Theorem satisfies the asymptotic superior of QMC vs MC for the both domains. [Niederreiter and Talay 2006]

RQMC-method is used to provide a more sensible estimation of the integration error, since the Koksma-Hlawka bound for QMC-method for many practical applications offers unconfined error estimations. As the Figures display the integration errors of QMC- and RQMC- methods are quite different. Using RQMC-methods we are allowed to reduce variance of the integration error but the consistency between QMC errors and RQMC errors is not always promised, which is shown for such simple integrals over unit cube and triangle ABC .

[1]

Theorem 5.2.1. *A bounded function on a compact interval is Riemann integrable if and only if Lebesgue measure of the set of points at which it is discontinuous is zero.[Wheeden 2015]*

Since Lebesgue measure of such a set is zero, so this Theorem denotes that each bounded function with a countable set of discontinuities is Riemann integrable.

5.3 Integration of an unbounded integrand

In this Section we deal with the integrations of an unbounded integrand over unit cube and triangle ABC using MC-, QMC- and RQMC-methods. Suppose the following integral

$$\mathfrak{I} = \int_0^1 \int_0^1 \mathrm{e}^{(\mathrm{norminv}(x)+\mathrm{norminv}(y))}\, dx\, dy.$$

where norminv computes confidence interval for P using a normal approximation to the distribution of the estimate $\hat{\mu} + q\hat{\sigma}$, when q is the P-th quantile from a normal distribution $\mathcal{N}(0,1)$. The definition of the normal inverse function in terms of the normal cdf is as follows

$$x = F^{-1}(p|\mu,\sigma) = \{x : F(x|\mu,\sigma) = p\}$$

where

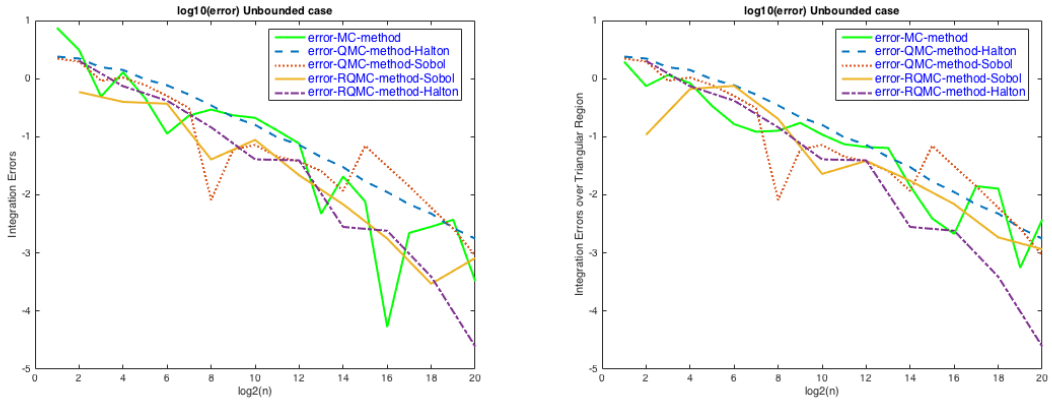
$$p = F(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathrm{e}^{\frac{-(t-\mu)^2}{2\sigma^2}}\, dt$$

This integral can moreover be calculated analytically

$$\mathfrak{I} = (\int_0^1 \mathrm{e}^{\mathrm{norminv}(x)}\, dx)^2 \simeq (\mathrm{e}^{1/2})^2 = \mathrm{e}.$$

This way we are allowed to determine the integration errors over the both domains. Note that the true integration result over the right-angle triangle ABC defined by $A = (0,0)^T, B = (0,1)^T$ and $C = (1,1)^T$ is the same as the true integration result over unit cube.

Figures 5.2a and 5.2b display the integration errors over $[0,1]^2$ and triangle ABC using MC-, QMC- and RQMC-methods, where n is the number of points used for each estimators.



(a) (b)

Figure 5.2: Logarithmic Plots of Relative Integration Errors of an unbounded integrand over Unit Cube and Triangle ABC.

From the Figures it can be observed that the integration errors of MC, QMC and RQMC are totally different for the both regions. An unbounded function certainly has infinite variation in the sense of Hardy and Krause but for some unbounded integrands QMC obtain a $\mathcal{O}(N^{-1+\epsilon})$ of convergence. For others QMC errors diverges to infinity as N goes to infinity even these integrands assure $D^*(x_1, \ldots, x_N) = \mathcal{O}(N^{-1+\epsilon})$ for all $\epsilon > 0$. [Niederreiter and Talay 2006] Hence here the asymptotic inferior of QMC vs MC can be seen for the both integrals. The magnitude of the improvement using RQMC-sampling varies for the both cases depending on the number $N = 2^n$ of points used and the type of low-discrepancy sequence used.

Whole bibliography

Ash, Robert B, and Catherine Doleans-Dade. 2000. *Probability and measure theory*. Academic Press.

Basu, Kinjal. 2016. "Quasi-Monte Carlo Methods in non-Cubical Spaces". PhD thesis, Stanford University.

Basu, Kinjal, and Art B Owen. "Scrambled geometric net integration over general product spaces". *Foundations of Computational Mathematics*: 1–30.

— . 2015a. "Low discrepancy constructions in the triangle". *SIAM Journal on Numerical Analysis* 53 (2): 743–761.

— . 2015b. "Transformations and Hardy-Krause variation". *arXiv preprint arXiv:1512.02713*.

Boyle, Phelim P. 1977. "Options: A monte carlo approach". *Journal of financial economics* 4 (3): 323–338.

Brandolini, Luca, et al. 2013a. "A Koksma–Hlawka inequality for simplices". In *Trends in Harmonic Analysis*, 33–46. Springer.

— . 2013b. "On the Koksma–Hlawka inequality". *Journal of Complexity* 29 (2): 158–172.

Bratley, Paul, Bennett L Fox, and Harald Niederreiter. 1994. "Programs to generate Niederreiter's low-discrepancy sequences". *ACM Transactions on Mathematical Software (TOMS)* 20 (4): 494–495.

Brauchart, J, et al. 2014. "QMC designs: optimal order Quasi Monte Carlo integration schemes on the sphere". *Mathematics of computation* 83 (290): 2821–2851.

Chen, WWL, and Giancarlo Travaglini. 2007. "Discrepancy with respect to convex polygons". *Journal of complexity* 23 (4): 662–672.

Clarkson, James A, and C Raymond Adams. 1933. "On definitions of bounded variation for functions of two variables". *Transactions of the American Mathematical Society* 35 (4): 824–854.

Dick, Josef, Frances Y Kuo, and Ian H Sloan. 2013. "High-dimensional integration: the quasi-Monte Carlo way". *Acta Numerica* 22:133–288.

Dick, Josef, and Friedrich Pillichshammer. 2010. *Digital nets and sequences: Discrepancy Theory and Quasi–Monte Carlo Integration*. Cambridge University Press.

Dick, Josef, et al. 2013. *Monte Carlo and Quasi-Monte Carlo Methods 2012*. Springer.

Fang and Wang. 1994. *Number-Theoretic Methods in Statistics*. CRC Press.

Glasserman, Paul. 2003. *Monte Carlo methods in financial engineering*. Springer Science & Business Media.

Hoffman, Kenneth. 2013. *Analysis in Euclidean space*. Courier Corporation.

Jensen, Hendrik Wann, et al. 2001. "State of the art in Monte Carlo ray tracing for realistic image synthesis". *SIGGRAPH 2001 Course Notes* 2.

Kong, Augustine, Jun S Liu, and Wing Hung Wong. 1994. "Sequential imputations and Bayesian missing data problems". *Journal of the American statistical association* 89 (425): 278–288.

Lemieux, Christiane. 2004. "Randomized quasi-Monte Carlo: A tool for improving the efficiency of simulations in finance". In *Simulation Conference, 2004. Proceedings of the 2004 Winter*, 2:1565–1573. IEEE.

— . 2009. *Monte Carlo and quasi-Monte Carlo sampling*. Springer Science & Business Media.

Löbbe, Thomas. 2014. "Probabilistic Star Discrepancy Bounds for Lacunary Point Sets". *arXiv preprint arXiv:1408.2220*.

Marsden, Jerrold E, and Michael J Hoffman. 1993. *Elementary classical analysis*. Macmillan.

Matousek, Jiri. 1999. *Geometric Discrepancy: An Illustrated Guide. Vol. 18 of Algorithms and Combinatorics*.

Moskowitz, Bradley, and Russel E Caflisch. 1996. "Smoothness and dimension reduction in quasi-Monte Carlo methods". *Mathematical and Computer Modelling* 23 (8): 37–54.

Niederreiter, Harald, and Peter J Shiue. 2012. *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing: Proceedings of a conference at the University of Nevada, Las Vegas, Nevada, USA, June 23–25, 1994*. Vol. 106. Springer Science & Business Media.

Niederreiter, Harald, and Denis Talay. 2006. *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Springer.

Owen, Art B. 1995. "Randomly permuted (t, m, s)-nets and (t, s)-sequences". In *Monte Carlo and quasi-Monte Carlo methods in scientific computing*, 299–317. Springer.

— . 1997a. "Monte Carlo variance of scrambled net quadrature". *SIAM Journal on Numerical Analysis* 34 (5): 1884–1910.

— . 1997b. "Scrambled net variance for integrals of smooth functions". *The Annals of Statistics*: 1541–1562.

— . 1998. "Scrambling Sobol'and Niederreiter–Xing points". *Journal of Complexity* 14 (4): 466–489.

— . 2005. "Multidimensional variation for quasi-Monte Carlo". In *International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday*, 49–74.

Owen, Art B, et al. 2008. "Local antithetic sampling with scrambled nets". *The Annals of Statistics* 36 (5): 2319–2343.

Pillards, Tim, and Ronald Cools. 2005. "Transforming low-discrepancy sequences from a cube to a simplex". *Journal of computational and applied mathematics* 174 (1): 29–42.

Rutherford, Andrew. 2012. *ANOVA and ANCOVA: A GLM Approachs*. John Wiley Sons.

Stromberg, Karl. 1994. *Probability for analysts*. CRC Press.

Wheeden, Richard L. 2015. *Measure and integral: an introduction to real analysis*. Vol. 308. CRC Press.