AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 10(3): 5532–5551.
DOI: 10.3934/math.2025255
Received: 12 October 2024
Revised: 09 February 2025
Accepted: 24 February 2025
Published: 12 March 2025

Research article

Bounds of random star discrepancy for HSFC-based sampling

Xiaoda Xu*

School of Mathematics and Science, Suqian University, Suqian 223800, China

* **Correspondence:** Email: 23195@squ.edu.cn.

Abstract: This paper is dedicated to the estimation of the probabilistic upper bounds of star discrepancy for Hilbert's space filling curve (HSFC) sampling. The primary concept revolves around the stratified random sampling method, with the relaxation of the stringent requirement for a sampling number $N = m^d$ in jittered sampling. We leverage the benefits of this sampling method to achieve superior results compared to Monte Carlo (MC) sampling. We also provide applications of the main result, which pertain to weighted star discrepancy, L_2-discrepancy, integration approximation in certain function spaces and examples in finance.

Keywords: star discrepancy; stratified sampling; δ-covers; HSFC sampling; integration approximation
Mathematics Subject Classification: 11K38, 65C10, 65D30

1. Introduction

Among various techniques for solving multivariate integration problems, such as light transport evaluation in complex scenes, rendering applications, variance analysis [1–3] or risk management, derivatives pricing model, and portfolio optimization [4, 5], the sample mean method has proven to be one of the most efficient approaches. For a random variable X (which is randomly distributed in $[0,1]^d$), the canonical form of multivariate integration can be expressed as

$$I(f) = \mathbb{E}(f) = \int_{[0,1]^d} f(X)dX. \tag{1.1}$$

The method of sample mean can be utilized to provide an estimate for $I(f)$ through approximation.

$$\tilde{I}(f, \mathbf{P}) = \frac{1}{N}\sum_{i=1}^{N} f(x_i), \tag{1.2}$$

where $\mathbf{P} = \{x_1, \ldots, x_N\} \subset [0,1]^d$.

It is occasionally essential to estimate the approximation error, and the renowned Koksma-Hlawka inequality provides the precise upper bound for approximation.

$$\left| \int_{[0,1]^d} f(x)dx - \frac{1}{N} \sum_{t \in P_{N,d}} f(t) \right| \leq D_N^* (t_1, t_2, \ldots, t_N) V(f), \tag{1.3}$$

where $D_N^* (t_1, t_2, \ldots, t_N)$ is a star discrepancy of $P_{N,d}$ and $V(f)$ is the total variation of f in the sense of Hardy and Krause.

Improving the upper bounds of star discrepancy in a bounded total variation functional space leads to better approximation error, as indicated by (1.3).

The primary objective of star discrepancy research is to acquire a metric for the uniform distribution of the point set, as evident from its definition. The star discrepancy of a sampling set $P_{N,d} = \{t_1, t_2, \ldots, t_N\}$ is

$$D_N^* (t_1, t_2, \ldots, t_N) := \sup_{B \subset \mathscr{B}} \left| \frac{A(B; N; P_{N,d})}{N} - \lambda(B) \right|, \tag{1.4}$$

where $A(B; N; P_{N,d})$ denotes the number of points from $P_{N,d}$ that are located in rectangle B anchored at 0, and $\lambda(B)$ denotes the Lebesgue measure of B.

There are numerous constructions of so-called low discrepancy point sets, which exhibit a favorable convergence rate for the star discrepancy bounds. For a point set $\mathscr{P}$, the convergence order could reach $O((\ln N)^{\alpha_d}/N)$ for fixed dimension d as $N \to \infty$, where $\alpha_d \geq 0$ are constants depending on dimension d, see [6]. These point sets have numerous applications across various domains; see [4, 5, 7, 8].

In recent years, numerous researchers have been investigating pre-asymptotic star discrepancy bounds, which give helpful information for moderate values of sampling number N. In [9], Heinrich, Novak, et al. have obtained the result that the inverse of star discrepancy depends linearly on the dimension. This also implies a result of existence for the star discrepancy upper bound $c \cdot \frac{\sqrt{d}}{\sqrt{N}}$, where c is an unknown constant. In the work by Aistleitner [10], techniques involving δ-covers and dyadic chaining were employed to achieve a value of $c = 10$. The constant c has been improved to 2.53 in [11]. The results are obtained through a simple random sampling process (MC point set). From a practical applications perspective, the following probability bound can be derived:

$$D_N^*(X) \leq 5.7 \sqrt{4.9 - \frac{\ln(1-q)}{d}} \frac{\sqrt{d}}{\sqrt{N}}, \tag{1.5}$$

with probability at least q for MC point set X. This result has been improved by [11] through the utilization of refined δ-cover bounds.

Jittered sampling is an optimization technique for simple random sampling, utilizing $N = m^d$ sampling points. For the expected star discrepancy upper bound of jittered sampling, the former's result [12] is actually the following if we use the normalized star discrepancy:

$$\mathbb{E} D_N^*(X) \leq \frac{d}{m^{\frac{d}{2}+\frac{1}{2}}} \cdot \left(60.9984 \sqrt{\log(\frac{4em}{d})} + 180.5492 \right). \tag{1.6}$$

Jittered sampling is a special form of stratified sampling, and other forms of stratified sampling require more sophisticated design; see [13].

The main purpose of this paper is to provide a precise probabilistic upper bound for star discrepancy in a new stratified sampling approach (HSFC sampling). This result, on the one hand, improves the bound using an MC point set; on the other hand, it removes the condition $N = m^d$, which is the strict dependence of sampling number N on the dimension d.

The remainder of this manuscript is structured as follows. In Section 2, we provide an overview of some preliminary concepts. The main results of this study are outlined in Section 3. In Section 4, we present an application of the main result. Finally, in Section 5, we conclude the paper with a brief summary.

2. HSFC-based sampling and some estimations

2.1. Hilbert space filling curve-based sampling

We first introduce the Hilbert space filling curve sampling, which is abbreviated as HSFC sampling, and we mainly adopt the definition and notations in [14, 15]. HSFC sampling is actually a kind of stratified sampling formed by a special partition manner. The new application of HSFC sampling for quantile estimation is considered in [16].

Let a_i be the first $N = b^m$ points of the van der Corput sequence (van der Corput 1935) in base $b \geq 2, m = 0, 1, \ldots$ The integer $i - 1 \geq 0$ is written in base b as

$$i - 1 = \sum_{j=1}^{\infty} a_{ij} b^{j-1} \tag{2.1}$$

for $a_{ij} \in \{0, \ldots, b-1\}$. Then, a_i is defined by

$$a_i = \sum_{j=1}^{\infty} a_{ij} b^{-j}. \tag{2.2}$$

We will go over the van der Corput sequence in more detail. For example, we consider the base $b = 2$; we then describe the construction process.

Step 1. Choose a base: First, choose a base b. In this example, we use binary, i.e., base $b = 2$.

Step 2. List the sequence of natural numbers: Then, list the sequence of natural numbers $(1, 2, 3, ...)$ as the original sequence.

Step 3. Convert base representation: Convert each natural number to binary representation. Example: The binary representation of 1 is 1; the binary representation of 2 is 10; the binary representation of 3 is 11; the binary representation of 4 is 100; and so on ...

Step 4. Reverse numerical order: Reverses the numerical order in the b-ary representation of each number and gets a new sequence of numbers. In order to convert it to a decimal, we need to treat the reversed number sequence as a decimal where the decimal point is to the left of the leftmost digit. For instance, for a decimal number 13, whose binary is 1101, we first need to reverse the number sequence to get 1011, we see it as binary 0.1011, we need to follow the following steps to convert it to decimal. The conversion process is as follows: The first binary digit is 1 (the first digit after the decimal point), then it represents 2^{-1} or 0.5; the second binary bit is 0, which represents 2^{-2} or 0.25, but because it is 0, it is not added to the sum; the third binary bit is 1, which represents 2^{-3} or 0.125; the fourth binary bit is 1, which represents 2^{-4} or 0.0625. Add these values together to get the decimal number: $1 * 0.5 + 0 * 0.25 + 1 * 0.125 + 1 * 0.0625 = 0.6875$.

Step 5. Formation of the Van der Corput sequence: The above obtained decimal fractions in order, that is, the formation of the Van der Corput sequence. We obtain the Van der Corput sequence $0, 0.5, 0.75, 0.25, 0.625, 0.375, 0.875, \ldots$

The scrambled version of $a_1, a_2, \ldots, a_N$ is $x_1, x_2, \ldots, x_N$ written as

$$x_i = \sum_{j=1}^{\infty} x_{ij} b^{-j}, \tag{2.3}$$

where x_{ij} are defined through random permutations of the a_{ij}. These permutations depend on a_{ik}, for $k < j$. More precisely, $x_{i1} = \pi(a_{i1})$, $x_{i2} = \pi_{a_{i1}}(a_{i2})$ and generally for $j \geq 2$,

$$x_{ij} = \pi_{a_{i1} \ldots a_{ij-1}}(a_{ij}). \tag{2.4}$$

Each random permutation is uniformly distributed over the $b!$ permutations of $\{0, \ldots, b-1\}$ and is mutually independent of the others. Thanks to the nice property of nested uniform scrambling, the data values in the scrambled sequence can be reordered such that

$$x_i \sim U(I_i), \tag{2.5}$$

independently with

$$I_i = [\frac{i-1}{N}, \frac{i}{N}]$$

for $i = 1, 2, \ldots, N = b^m$.

Let

$$E_i = H(I_i) := \{H(x) | x \in I_i\},$$

where H is a mapping.

Then,

$$X_i = H(x_i) \sim U(E_i), i = 1, 2, 3, \ldots N = b^m$$

is the corresponding stratified samples. Set r_i be the diameter of E_i, according to the property of HSFC sampling, the following estimation holds:

$$r_i \leq 2\sqrt{d+3} \cdot N^{-\frac{1}{d}}. \tag{2.6}$$

The Van der Corput sequence plays a fundamental role in HSFC sampling, which is a one-dimensional set of low-discrepancy points. It is transformed into a uniformly distributed set of points in the interval [0, 1] by Owen scrambling and then mapped to the higher dimensions by Hilbert mapping, i.e., to obtain the HSFC sample.

2.2. Minkowski content

We use the definition of Minkowski content in [17], which provides convenience for analyzing the boundary characteristics of the test set B in (1.4), that is, for a set $\Omega \subset [0, 1]^d$,

$$\mathscr{M}(\partial\Omega) = \lim_{\epsilon \to 0} \frac{\lambda((\partial\Omega)_\epsilon)}{2\epsilon}, \tag{2.7}$$

where $(\partial\Omega)_\epsilon = \{x \in \mathbb{R}^d | dist(x, \partial\Omega) \leq \epsilon\}$. If $\mathscr{M}(\partial\Omega)$ exists and is finite, then $\partial\Omega$ is said to admit $(d-1)-$dimensional Minkowski content. If Ω is a convex set, then it is easy to see that $\partial\Omega$ admits $(d-1)-$dimensional Minkowski content; furthermore, $\mathscr{M}(\partial\Omega) \leq 2d$ as the outer surface area of a convex set in $[0,1]^d$ is bounded by the surface area of the unit cube $[0,1]^d$, which is $2d$.

2.3. δ-covers

To discretize the star discrepancy, we use the definition of δ-covers as in [18].

Definition 2.1. *For any $\delta \in (0,1]$, a finite set Γ of points in $[0,1)^d$ is called a δ-cover of $[0,1)^d$ if for every $y \in [0,1)^d$, there exist $x, z \in \Gamma \cup \{0\}$ such that $x \leq y \leq z$ and $\lambda([0,z]) - \lambda([0,x]) \leq \delta$. The number $\mathcal{N}(d,\delta)$ denotes the smallest cardinality of a δ-cover of $[0,1)^d$.*

From [19], combining with Stirling's formula, the following estimation for $\mathcal{N}(d,\delta)$ holds, that is, for any $d \geq 1$ and $\delta \in (0,1]$,

$$\mathcal{N}(d,\delta) \leq 2^d \cdot \frac{e^d}{\sqrt{2\pi d}} \cdot (\delta^{-1}+1)^d. \tag{2.8}$$

Furthermore, the following lemma provides convenience for estimating the star discrepancy with δ-covers.

Lemma 2.1. *[18] Let $P = \{p_1, p_2, \ldots, p_N\} \subset [0,1]^d$ and Γ be δ-covers, then,*

$$D_N^*(P) \leq D_\Gamma(P) + \delta, \tag{2.9}$$

where

$$D_\Gamma(P) := \max_{x\in\Gamma} |\lambda([0,x]) - \frac{\sum_{n=1}^N I_{[0,x]}(p_n)}{N}|. \tag{2.10}$$

2.4. Bernstein inequality

At the end of this section, we will restate the Bernstein inequality, which will be used in the estimation of star discrepancy bounds.

Lemma 2.2. *[20] Let $Z_1, \ldots, Z_N$ be independent random variables with expected values $\mathbb{E}(Z_j) = \mu_j$ and variances σ_j^2 for $j = 1, \ldots, N$. Assume $|Z_j - \mu_j| \leq C$ (C is a constant) for each j and set $\Sigma^2 := \sum_{j=1}^N \sigma_j^2$, then for any $\lambda \geq 0$,*

$$\mathbb{P}\left\{\Big|\sum_{j=1}^N [Z_j - \mu_j]\Big| \geq \lambda\right\} \leq 2\exp\left(-\frac{\lambda^2}{2\Sigma^2 + \frac{2}{3}C\lambda}\right).$$

3. Probabilistic star discrepancy bound for HSFC-based sampling

Theorem 3.1. *For integer number $b \geq 1$, $m \geq 1$ and $N = b^m$, then for the well-defined d-dimensional stratified samples $X_i \sim U(E_i)$, $i = 1, 2, \ldots, N = b^m$ in Section 2, we have*

$$D_N^*(X_1, X_2, \ldots, X_N) \leq \frac{6d^{\frac{3}{4}}}{N^{\frac{1}{2}+\frac{1}{2d}}} \cdot \sqrt{d\ln(N+1) + c(d,q)} + \frac{2c(d,q)}{3N} \tag{3.1}$$

with probability at least q, where

$$c(d, q) = \ln \frac{(2e)^d}{\sqrt{2\pi d} \cdot (1 - q)}.$$

Proof. Let A be a subset of E_i, then according to $X_i \sim U(E_i), 1 \le i \le N$, it follows that

$$\mathbb{P}(X_i \in A) = \frac{\lambda(A)}{\lambda(E_i)} = N\lambda(A). \tag{3.2}$$

Now, for an arbitrary d-dimensional rectangle $R = [0, x] \in [0, 1]^d$ with diameter κ, when x runs through the unit cube $[0, 1)^d$, we can assign two points y, z such that $y \le x \le z$ and $\lambda([0, z]) - \lambda([0, y]) \le \frac{1}{N}$. Let $R_0 = [0, y]$ and $R_1 = [0, z]$, then we have

$$R_0 \subseteq R \subset R_1,$$

and

$$\lambda(R_1) - \lambda(R_0) \le \frac{1}{N}. \tag{3.3}$$

We know that the diameter of R_0 is less than κ; we set it to κ_0. The diameter of R_1 is more than κ; we set it to κ_1. This forms a bracketing cover for the set R, and from (2.8) and (3.3), we can give the upper bound for the bracketing cover pair (R_0, R_1), which has a cardinality at most $2^{d-1} \frac{e^d}{\sqrt{2\pi d}}(N + 1)^d$. Besides, from Lemma 2.1, we obtain

$$D_N^*(Y_1, Y_2, \dots, Y_N; R) \le \max_{i=0,1} D_N^*(Y_1, Y_2, \dots, Y_N; R_i) + \frac{1}{N}. \tag{3.4}$$

For an anchored box R in $[0, 1]^d$, it is easy to check that R is representable as a disjoint union of $E_i's$ entirely contained in R and the union of l pieces, which are the intersections of some $E_j's$ and R, i.e.,

$$R = \bigcup_{i \in I} E_i \cup \bigcup_{j \in J} (E_j \cap \partial R), \tag{3.5}$$

where I and J denote the index-sets.

By the definition of Minkowski content, for any $\sigma > 2$, there exists ϵ_0 such that $\lambda((\partial R)_\epsilon) \le \sigma\epsilon\mathscr{M}(\partial R)$ whenever $\epsilon \le \epsilon_0$.

From (2.6), the diameter for each E_i is at most $2\sqrt{d + 3} \cdot N^{-\frac{1}{d}}$, we can assume $N > (\frac{2\sqrt{d+3}}{\epsilon_0})^d$, then, $2\sqrt{d + 3} \cdot N^{-\frac{1}{d}} := \epsilon < \epsilon_0$ and $\bigcup_{j \in J}(E_j \cap \partial R) \subseteq (\partial R)_\epsilon$, therefore,

$$|J| \le \frac{\lambda((\partial R)_\epsilon)}{\lambda(E_i)} \le \frac{\sigma\epsilon\mathscr{M}(\partial R)}{N^{-1}} = 2\sqrt{d + 3}\sigma\mathscr{M}(\partial R)N^{1-\frac{1}{d}}.$$

Without loss of generality, we can set $\sigma = 3$, combining with the fact $\mathscr{M}(\partial R) \le 2d$; it follows

$$|J| \le 12d\sqrt{d + 3} \cdot N^{1-\frac{1}{d}}. \tag{3.6}$$

The same argument (3.6) holds for test sets R_0 and R_1.

For R_0 or R_1, set

$$D_N^*(X_1, X_2, \ldots, X_N; R') = \max_{i=0,1} D_N^*(X_1, X_2, \ldots, X_N; R_i). \tag{3.7}$$

R' is also a test rectangle, which can be broken up into two parts:

$$R' = \bigcup_{k\in K} E_k \cup \bigcup_{l\in L} (E_l \cap R'), \tag{3.8}$$

and the cardinality of $R^{'} \subset [0,1)^d$ is at most $2^{d-1}\frac{e^d}{\sqrt{2\pi d}}(N+1)^d$ according to the δ-cover estimation.

Let

$$T = \bigcup_{l\in L} (E_l \cap R'), |L| = |\{1, 2, \ldots, l\}|. \tag{3.9}$$

If we define new random variables χ_j, $1 \le j \le l$ as follows:

$$\chi_l = \begin{cases} 1, & X_l \in E_l \cap R', \\ 0, & otherwise, \end{cases} \tag{3.10}$$

then, from the above discussions, we have

$$N \cdot D_N^* (X_1, X_2, \ldots, X_N; R') = N \cdot D_N^* (X_1, X_2, \ldots, X_N; T) = |\sum_{l=1}^{|L|} \chi_l - N(\sum_{l=1}^{|L|} \lambda(E_l \cap R'))|. \tag{3.11}$$

Since

$$\mathbb{P}(\chi_l = 1) = \frac{\lambda(E_l \cap R')}{\lambda(E_l)} = N \cdot \lambda(E_l \cap R'), \tag{3.12}$$

hence,

$$\mathbf{E}(\chi_l) = N \cdot \lambda(E_l \cap R'). \tag{3.13}$$

Thus, from (3.11) and (3.13), we obtain

$$N \cdot D_N^* (X_1, X_2, \ldots, X_N; R') = |\sum_{l=1}^{|L|} (\chi_l - \mathbb{E}(\chi_l))|. \tag{3.14}$$

Let $\sigma_l^2 = \mathbb{E}(\chi_l - \mathbb{E}(\chi_l))^2$ and $\Sigma^2 = (\sum_{l=1}^{|L|} \sigma_l^2)^{\frac{1}{2}}$, then we have

$$\Sigma^2 \le |L| \le 12d\sqrt{d+3} \cdot N^{1-\frac{1}{d}}. \tag{3.15}$$

Therefore, from Lemma 2.2, for every R', we have

$$\mathbb{P}\left(\Big| \sum_{l=1}^{|L|} (\chi_l - \mathbb{E}(\chi_l)) \Big| > \lambda \right) \le 2 \cdot \exp(-\frac{\lambda^2}{24d\sqrt{d+3} \cdot N^{1-\frac{1}{d}} + \frac{2\lambda}{3}}). \tag{3.16}$$

Let

$$\mathscr{B} = \bigcup_{R'} \left(\Big| \sum_{l=1}^{|L|} (\chi_l - \mathbb{E}(\chi_l))| > \lambda \right). \tag{3.17}$$

Then, using covering numbers, we have

$$\mathbb{P}(\mathscr{B}) \leq (2e)^d \cdot \frac{1}{\sqrt{2\pi d}} \cdot (N+1)^d \cdot \exp(-\frac{\lambda^2}{24d\sqrt{d+3} \cdot N^{1-\frac{1}{d}} + \frac{2\lambda}{3}}). \tag{3.18}$$

Let $A(d,q,N) = d\ln(2e) + d\ln(N+1) - \frac{\ln(2\pi d)}{2} - \ln(1-q)$, and we choose

$$\lambda = \sqrt{24d\sqrt{d+3} \cdot A(d,q,N) + \frac{A^2(d,q,N)}{9N^{1-\frac{1}{d}}}} \cdot N^{\frac{1}{2}-\frac{1}{2d}} + \frac{A(d,q,N)}{3}. \tag{3.19}$$

Put it into (3.18), we have

$$\mathbb{P}(\mathscr{B}) \leq 1-q. \tag{3.20}$$

Combining the above and (3.14), we obtain

$$D_N^*(X_1, X_2, \ldots, X_N; R') \leq \frac{\sqrt{24d\sqrt{d+3} \cdot A(d,q,N) + \frac{A^2(d,q,N)}{9N^{1-\frac{1}{d}}}}}{N^{\frac{1}{2}+\frac{1}{2d}}} + \frac{A(d,q,N)}{3N} \tag{3.21}$$

with probability at least q.

Thus, obviously, we have

$$\begin{aligned} \max_{i=0,1} D_N^*(X_1, X_2, \ldots, X_N; R_i) &\leq \frac{\sqrt{24d\sqrt{d+3} \cdot A(d,q,N) + \frac{A^2(d,q,N)}{9N^{1-\frac{1}{d}}}}}{N^{\frac{1}{2}+\frac{1}{2d}}} + \frac{A(d,q,N)}{3N} \\ &\leq \frac{\sqrt{24d\sqrt{d+3} \cdot A(d,q,N)}}{N^{\frac{1}{2}+\frac{1}{2d}}} + \frac{2A(d,q,N)}{3N} \end{aligned} \tag{3.22}$$

with probability at least q.

$$A(d,q,N) = \ln\frac{(2e)^d}{\sqrt{2\pi d} \cdot (1-q)} + d\ln(N+1) = c(d,q) + d\ln(N+1), \tag{3.23}$$

where

$$c(d,q) = \ln\frac{(2e)^d}{\sqrt{2\pi d} \cdot (1-q)}. \tag{3.24}$$

The proof is complete. □

Remark 3.1. *The convergence order of the probabilistic star discrepancy bound for HSFC stratified samples, as given by Theorem 3.1, is $O(N^{-\frac{1}{2}-\frac{1}{2d}} \cdot (\ln N)^{\frac{1}{2}})$. This can be compared with the asymptotic bounds using a crude Monte Carlo point set, which are $O(N^{-\frac{1}{2}})$. A comparison with the convergence order of the results in [11, 21] is also possible.*

4. Applications

4.1. Uniform integration approximation for functions in weighted function space

Many high-dimensional problems that arise in practical applications have low effective dimensions [22]; that is, they have different weights for different component function values. Therefore, the problem is abstracted as seeking uniform integral approximation errors in weighted Sobolev spaces.

Let $F_{d,1}$ be a Sobolev space; for functions $f \in F_{d,1}$, f is differentiable for each variable and has a finite L_1-module for its first order differential. For $d > 1$, the norm in $F_{d,1}$ is defined as

$$\|f\|_{F_{d,1}} = \|D^{\vec{1}} f\|_{L_1([0,1]^d)} = \int_{[0,1]^d} |D^{\vec{1}} f(x)| dx,$$

where $\vec{1} = [1, 1, \ldots, 1]$ and $D^{\vec{1}} = \partial^d / \partial x_1 \ldots \partial x_d$.

Then for weighted Sobolev space $F_{d,1,\gamma}$, its norm is

$$\|f\|_{F_{d,1,\gamma}} = \sum_{u \subseteq I_d} \gamma_{u,d} \int_{[0,1]^{|u|}} |\frac{\partial^{|u|}}{\partial x_u} f(x_u, 1)| dx_u. \tag{4.1}$$

Considering the problem of function approximation in $F_{d,1,\gamma}$ space, the sample mean method can still be used, that is,

$$I(f) = \int_{[0,1]^d} f(x) dx,$$

and sample mean function

$$\tilde{I}(f, \mathbf{P}) = \frac{1}{N} \sum_{i=1}^{N} f(x_i).$$

Consider the worst-case error

$$E_N(f) = \left| \int_{[0,1]^d} f(x) dx - \frac{1}{N} \sum_{t \in P_{N,d}} f(t) \right|,$$

then form Hlawka and Zaremba's identity [23], we have

$$e(E_N(f)) = \sup_{f \in F_{d,1,\gamma}, \|f\|_{F_{d,1,\gamma}} \leq 1} |I(f) - \tilde{I}(f, \mathbf{P})| = D^*_{N,\gamma}(t_1, t_2, \ldots, t_N).$$

For uniform integration approximation in weighted Sobolev spaces, we have the following theorem.

Theorem 4.1. *Let $f \in F_{d,1,\gamma}$ be functions in Sobolev space. Let integer $b \geq 1$, $m \geq 1$, and $N = b^m$. Let b, λ, λ_0, c be some integers, λ_0 be constants such that $b\lambda^2 \leq e^{2\lambda^2}$ holds for all $\lambda \geq \lambda_0$,$c = \max\{2, b, \lambda_0, \frac{1}{\log_2(2-\epsilon)}\}$, then for d-dimensional HSFC samples $X_i \sim U(E_i), i = 1, 2, \ldots, N = b^m$, we have*

$$\begin{aligned} &\sup_{f \in F_{d,1,\gamma}, \|f\|_{F_{d,1,\gamma}} \leq 1} |\int_{[0,1]^d} f(x) dx - \frac{1}{N} \sum_{j=1}^{N} f(X_j)| \\ &\leq \max_{\emptyset \neq u \subseteq I_d} \gamma_{u,d} [\frac{6|u|^{\frac{3}{4}}}{N^{\frac{1}{2}+\frac{1}{2|u|}}} \cdot \sqrt{|u| \ln(N+1) + c(|u|, \epsilon)} + \frac{2c(|u|, \epsilon)}{3N}] \end{aligned} \tag{4.2}$$

with probability at least ϵ, where

$$c(|u|,\epsilon)=\ln\frac{(2e)^{|u|}}{\sqrt{2\pi|u|}\cdot(1-\epsilon)}.$$

Proof. In Theorem 3.1, we choose probability $q=\epsilon=1-(b\lambda^2e^{-2\lambda^2})^d$, which holds for some positive constant b and for all $\lambda\geq\max\{1,b,\lambda_0\}$, where λ_0 is constant such that $b\lambda^2\leq e^{2\lambda^2}$ holds for all $\lambda\geq\lambda_0$, and we choose

$$\lambda=c\max\{1,\ \sqrt{(\ln d)/(\ln 2)}\},$$

and $c=\max\{2,b,\lambda_0\}$.

Let

$$c(|u|,\epsilon)=\ln\frac{(2e)^{|u|}}{\sqrt{2\pi|u|}\cdot(1-\epsilon)}.$$

For a given number of sampling points N and dimension d, we consider the following set:

$$A_d:=\{\mathbf{P}_{N,d}\subset[0,1]^d:D_N(\mathbf{P}_{N,d}(u))\leq[\frac{6|u|^{\frac{3}{4}}}{N^{\frac{1}{2}+\frac{1}{2|u|}}}\cdot\sqrt{|u|\ln(N+1)+c(|u|,\epsilon)}+\frac{2c(|u|,\epsilon)}{3N}],\quad\forall u\subseteq I_d,u\neq\emptyset\},$$

where $\mathbf{P}_{N,d}(u):=\{X_1(u),\dots,X_N(u)\}$. Besides, for $u\subseteq I_d,u\neq\emptyset$, we define

$$A_{u,d}:=\{\mathbf{P}_{N,d}\subset[0,1)^d:D_N(\mathbf{P}_{N,d}(u))\leq[\frac{6|u|^{\frac{3}{4}}}{N^{\frac{1}{2}+\frac{1}{2|u|}}}\cdot\sqrt{|u|\ln(N+1)+c(|u|,\epsilon)}+\frac{2c(|u|,\epsilon)}{3N}]\}.$$

Then we have

$$A_d=\bigcap_{\emptyset\neq u\subseteq I_d}A_{u,d}.$$

Hence,

$$\begin{aligned}\mathbb{P}(A_d)&=\mathbb{P}(\bigcap_{\emptyset\neq u\subseteq I_d}A_{u,d})=1-\mathbb{P}(\bigcup_{\emptyset\neq u\subseteq I_d}A_{u,d}^c)\\&\geq1-\sum_{\emptyset\neq u\subseteq I_d}\mathbb{P}(A_{u,d}^c)>1-\sum_{\emptyset\neq u\subseteq I_d}(b\lambda^2e^{-2\lambda^2})^{|u|}\\&=1-\sum_{u=1}^{d}\binom{d}{u}(b\lambda^2e^{-2\lambda^2})^u=2-(1+b\lambda^2e^{-2\lambda^2})^d.\end{aligned}$$

According to $\lambda=c\max\{1,\sqrt{\frac{\ln d}{\ln 2}}\}$ and $c=\max\{2,b,\lambda_0\}$, for all $d\geq2$ and $x=\frac{c^2}{\ln 2}>5$, we have $x^2\leq2^x\leq d^x$ and $\ln d\leq d^{x-1}$. Thus, if $x^2\ln d\leq d^{2x-1}$, then

$$\frac{c^3\ln d}{(\ln 2)d^{\frac{2c^2}{\ln 2}}}\leq\frac{\ln 2}{cd}.$$

Based on this inequality, we obtain a formula that holds for all $d \geq 2$,

$$\mathbb{P}(A_d) > 2 - (1 + b\lambda^2 e^{-2\lambda^2})^d \geq 2 - (1 + \frac{c^3 \ln d}{(\ln 2) d^{\frac{2c^2}{\ln 2}}})^d$$
$$\geq 2 - (1 + \frac{\ln 2}{cd})^d > 2 - e^{\frac{\ln 2}{c}} = 2 - 2^{1/c} \geq \epsilon.$$

Thus, for every $\emptyset \neq u \subseteq I_d$, we obtain

$$D_{N,\gamma}(t_1, t_2, \dots t_N) \leq \max_{\emptyset \neq u \subseteq I_d} \gamma_{u,d} [\frac{6|u|^{\frac{3}{4}}}{N^{\frac{1}{2} + \frac{1}{2|u|}}} \cdot \sqrt{|u| \ln(N+1) + c(|u|, \epsilon)} + \frac{2c(|u|, \epsilon)}{3N}]$$

with probability at least ϵ.

The proof is completed. □

4.2. Integral approximation on Borel convex subsets

The classical Koksma-Hlawka inequality does not apply to functions with simple discontinuities, and thus a generalized Koksma-Hlawka-type inequality is proposed, which applies to a piecewise smooth function $f \cdot \mathbf{1}_\Omega$, where f is smooth and Ω is a Borel convex subset of $[0,1]^d$, see [24]. An approximation error in the space of piecewise smooth functions will be given below using a star-discrepancy bound for stratified sampling. First, there is the following lemma:

Lemma 4.1. *[24] Let f be a piecewise smooth function defined on $[0,1]^d$ and Ω be a Borel convex subset of $[0,1]^d$. Then for the set of samples $x_1, x_2, \dots, x_N$ in $[0,1]^d$, there are*

$$|\frac{\sum_{n=1}^{N} (f \cdot \mathbf{1}_\Omega)(x_n)}{N} - \int_\Omega f(x) dx| \leq D_N^\Omega(x_1, x_2, \dots, x_N) \cdot V(f), \tag{4.3}$$

where

$$D_N^\Omega(x_1, x_2, \dots, x_N) = 2^d \sup_{A \subseteq [0,1]^d} |\frac{\sum_{n=1}^{N} \mathbf{1}_{\Omega \cap A}(x_n)}{N} - \lambda_d(\Omega \cap A)|, \tag{4.4}$$

and

$$V(f) = \sum_{u \subset \{1,2,\dots,d\}} 2^{d-|u|} \int_{[0,1]^d} |\frac{\partial^{|u|}}{\partial x_u} f(x)| dx. \tag{4.5}$$

The symbol $\frac{\partial^{|u|}}{\partial x_n} f(x)$ is the partial derivative of f with respect to the component x indexed at u, and the upper bound is taken over all axis-parallel rectangles A.

Theorem 4.2. *For integers $b \geq 1$, $m \geq 1$, and $N = b^m$, let f be a piecewise smooth function defined on $[0,1]^d$ and Ω be a Borel convex subset of $[0,1]^d$. Then for a sample of Hilbert space-filling curves $X_i \sim U(E_i)$, $i = 1, 2, \dots, N = b^m$ in $[0,1]^d$, there are*

$$\mathbb{P}\Big(|\frac{\sum_{n=1}^{N} (f \cdot \mathbf{1}_\Omega)(X_n)}{N} - \int_\Omega f(x) dx| \leq 2^d D_N^*(X, q) \cdot V(f)\Big) > q, \tag{4.6}$$

where $D_N^*(X, q)$ denotes the upper bound of $D_N^*(X)$ with probability at least q.

Example 1. Comparison of integral approximation errors for HSFC sampling on simplex

For $\epsilon > 0$, let Σ be a simplex, namely,

$$\Sigma = \{(x_1, x_2, \ldots, x_d) \in [0,1]^d : x_1 \geq \ldots \geq x_d \geq \epsilon, 1 - x_1 - \ldots - x_d \geq \epsilon\}.$$

Define

$$f(x_1, x_2, \ldots, x_d) = \frac{1}{x_1 x_2 \ldots x_d}(1 - x_1 - x_2 - \ldots - x_d).$$

Easy to show that

$$\sum_{|\alpha| \leq d} \int_{\Sigma} |(\frac{\partial}{\partial x})^{\alpha} f(x)| dx \leq \epsilon^{-d}.$$

According to Theorem 4.2, for HSFC samples X_i, $i = 1, 2, \ldots, N$, and all the convex subsets Ω_0 contained in Σ, we have

$$\mathbb{P}\Big(|\frac{\sum_{n=1}^{N}(f \cdot \mathbf{1}_{\Omega_0})(X_n)}{N} - \int_{\Omega_0} f(x)dx| \leq \epsilon^{-d} \cdot 2^d \cdot D_N^*(X, q)\Big) > q, \tag{4.7}$$

where $D_N^*(X, q)$ the upper bound of $D_N^*(X)$ holds with probability at least q.

From (1.5), for MC point set X, we have

$$\mathbb{E}D_N^*(X) \leq 12.62 \cdot \frac{\sqrt{d}}{\sqrt{N}}.$$

For jittered sampling set Y with condition $N = m^d$, we have, which is proved in [12],

$$\mathbb{E}D_N^*(Y) \leq \frac{d}{m^{\frac{d}{2}+\frac{1}{2}}} \cdot \left(60.9984\sqrt{\log(\frac{4em}{d})} + 180.5492\right). \tag{4.8}$$

For HSFC sampling set Z, we have

$$\mathbb{E}D_N^*(Z) \leq \frac{6d^{\frac{3}{4}}}{N^{\frac{1}{2}+\frac{1}{2d}}} \cdot \sqrt{d\ln(N+1) + c(d)} + \frac{2c(d)}{3N}, \tag{4.9}$$

where

$$c(d) = \ln\frac{(2e)^d}{\sqrt{2\pi d}}.$$

Set

$$E_N(f, X) = \mathbb{E}(|\frac{\sum_{n=1}^{N}(f \cdot \mathbf{1}_{\Omega_0})(X_n)}{N} - \int_{\Omega_0} f(x)dx|),$$

we compare this quantity for different sampling sets (HSFC sampling point set, jittered sampling set, MC point set), choose parameter $\epsilon = \frac{1}{2}$ in the simplex, and then we have

$$E_N(f, X) \leq 12.62 \cdot \frac{\sqrt{d}}{\sqrt{m^d}},$$

$$E_N(f, Y) \leq \frac{d}{m^{\frac{d}{2}+\frac{1}{2}}} \cdot \left(60.9984\sqrt{\log(\frac{4em}{d})} + 180.5492\right), \tag{4.10}$$

$$E_N(f, Z) \leq \frac{6d^{\frac{3}{4}}}{m^{\frac{d}{2}+\frac{1}{2}}} \cdot \sqrt{d\ln(m^d+1) + \ln\frac{(2e)^d}{\sqrt{2\pi d}}} + \frac{2\ln\frac{(2e)^d}{\sqrt{2\pi d}}}{3m^d}. \tag{4.11}$$

Remark 4.1. *From the numerical examples above, which are Tables 1–3, and Figures 1–3, it can be seen that HSFC sampling as well as jittered sampling exhibit better integral approximation than MC sampling when the sample size is sufficiently large, and HSFC sampling does not need to have large samples to obtain results superior to jittered sampling and MC sampling; the advantages of using HSFC sampling are demonstrated here.*

Table 1. Approximation errors in $d = 3$.

m	$E_N(f, X)$	$E_N(f, Y)$	$E_N(f, Z)$
10	0.69123	8.8839	0.67714
50	0.061825	0.38357	0.034105
100	0.021858	0.098587	0.0091831
500	0.0019551	0.0041715	0.00042217
1000	0.00069123	0.0010655	0.00011093

Table 2. Approximation errors in $d = 4$.

m	$E_N(f, X)$	$E_N(f, Y)$	$E_N(f, Z)$
50	0.010096	0.071444	0.0079025
100	0.002524	0.012999	0.0015069
500	0.000101	0.000246	3.1047e-05
8000	3.9437e-07	2.6086e-07	3.6183e-08
10000	2.524e-07	1.5018e-07	2.0959e-08

Table 3. Approximation errors in $d = 5$.

m	$E_N(f, X)$	$E_N(f, Y)$	$E_N(f, Z)$
100	0.00028219	0.0016103	0.00022148
300	1.8103e-05	6.2191e-05	9.0803e-06
500	5.048e-06	1.3673e-05	2.0435e-06
8000	4.9297e-09	3.6245e-09	5.9609e-10
10000	2.8219e-09	1.8666e-09	3.0885e-10

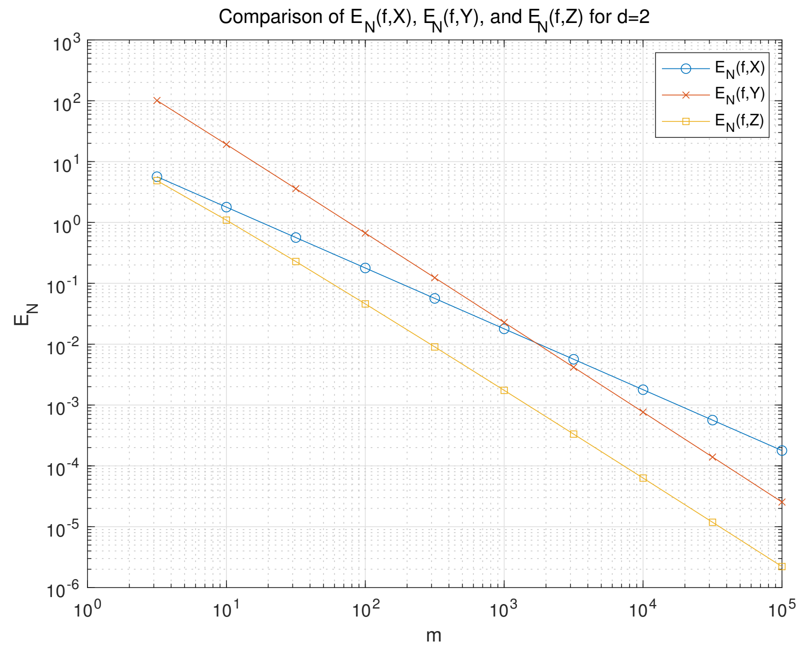

Figure 1. Comparison of different sampling set in $d = 2$.

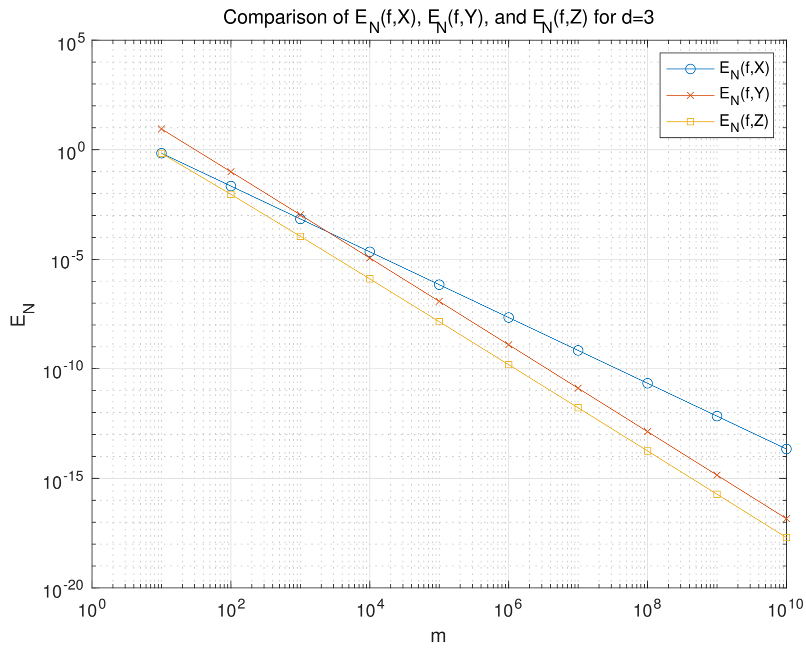

Figure 2. Comparison of different sampling set in $d = 3$.

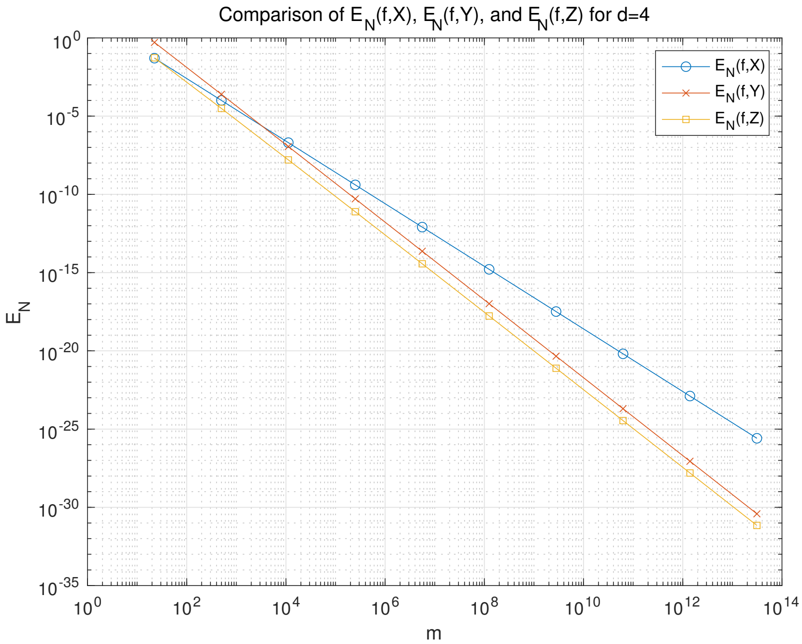

Figure 3. Comparison of different sampling set in $d = 4$.

Example 2. Comparison in the field of finance

For the option pricing problem, the price of an option can be expressed as the following expectation (Black-Scholes model):

$$C = e^{-rT} \cdot \mathbb{E}[\max(S_T - K, 0)],$$

where S_T denotes the geometric Brownian motion model, which is used to model the stochastic evolution of the underlying asset price. Its mathematical expression is

$$S_T = S_0 \cdot \exp\left((r - 0.5\sigma^2)T + \sigma\sqrt{T} \cdot Z\right),$$

where S_0 is the initial price of the underlying asset, r is the risk-free interest rate, σ is the volatility of the underlying asset, T is the time to maturity of the option, Z is a standard normal random variable, and K is the strike price of the option.

Indeed, if C denotes a multiple integral, then the sample mean method can be employed to give the approximation,

$$C_{appro} = e^{-rT} \cdot \frac{1}{N} \sum_{i=1}^{N} \max(S_T^{(i)} - K, 0),$$

where for samples $P = \{P_1, P_2, \ldots, P_N\}$ in $[0, 1]^d$,

$$S_T^{(i)} = S_0 \cdot \exp\left((r - 0.5\sigma^2)T + \sigma\sqrt{T} \cdot \Phi^{-1}(P_i)\right),$$

$\Phi^{-1}(x)$ denotes the inverse cumulative distribution function (ICDF).

Then for different samples X, Y, Z, we are able to give the approximation error. Choose $S_0 = 100, K = 100, r = 0.05, \sigma = 0.2, T = 1, d = 5$. We give a comparison of the approximation errors under the three sampling sets; see Figure 4.

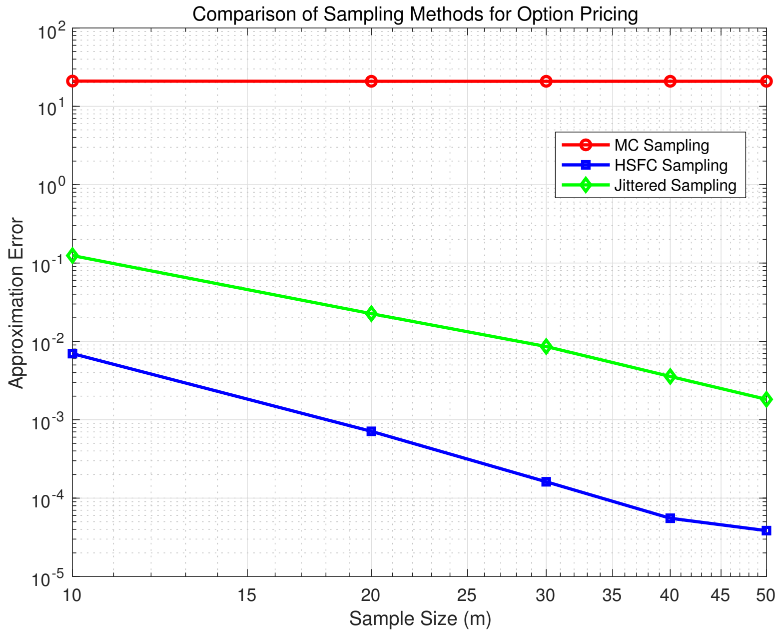

Figure 4. Comparison of approximation for different sampling sets.

4.3. Mean square error estimation for integral approximation in special function spaces

We mainly consider the function space $\mathcal{H}^{\mathbf{1}}(K)$ defined in [25].

Let

$$\mathcal{H}^{\mathbf{1}} := \mathcal{H}^{(1,1,\ldots,1)}([0,1]^d)$$

be the Sobolev spaces on $[0,1]^d$, $\forall f \in \mathcal{H}^{\mathbf{1}}$, we have

$$\frac{\partial^d}{\partial x} f(x) \in \mathcal{H}([0,1]^d),$$

where $\partial x = \partial x_1 \partial x_2 \ldots, \partial x_d$, $\mathcal{H}([0,1]^d)$ denotes the Hilbert space. Then for $f, g \in \mathcal{H}^{\mathbf{1}}$, we define the following inner product for the Hilbert space $\mathcal{H}([0,1]^d)$:

$$\langle f, g \rangle_{\mathcal{H}^1} = \int_{[0,1]^d} \frac{\partial^d f}{\partial x}(t) \frac{\partial^d g}{\partial x}(t) dt. \tag{4.12}$$

Further, we set $\|f\|_{\mathcal{H}^1} = \langle f, f \rangle_{\mathcal{H}^1}^{1/2}$ be the norm induced by the inner product defined in (4.12). Next, we define a reproducing kernel in $\mathcal{H}^{\mathbf{1}}$, given by

$$K(x,y) = \int_{[0,1]^d} \mathbf{1}_{(x,1]}(t) \mathbf{1}_{(y,1]}(t) dt, \tag{4.13}$$

where $x = (x_1, x_2, \ldots, x_d), y = (y_1, y_2, \ldots, y_d)$, $(x, 1] = \prod_{i=1}^d (x_i, 1], (y, 1] = \prod_{i=1}^d (y_i, 1]$, and $\mathbf{1}_A$ denotes the characteristic function on set A. $\mathcal{H}^{\mathbf{1}}(K)$ denotes the Sobolev space $\mathcal{H}^{\mathbf{1}}$ equipped with a reproducing kernel function $K(x,y)$ defined in (4.13). Correspondingly, in (4.12), we define $\langle f, g \rangle_{\mathcal{H}^1} = \langle f, g \rangle_{\mathcal{H}^1(K)}$.

It is easy to check that for the kernel function defined in (4.13), the reproducing property is satisfied, that is,

$$\langle f, K(\cdot, y)\rangle_{\mathcal{H}^1(K)} = \int_{[0,1]^d} \frac{\partial^d f}{\partial x}(t)\frac{\partial^d K(x,y)}{\partial x}(t)dt = f(y).$$

Then, the main tools are the L_2-discrepancy, we first give the definition of L_p-discrepancy, where we can choose $p = 2$.

L_p**-discrepancy:** For a sampling set $P_{N,d} = \{t_1, t_2, \ldots, t_N\}$, L_p-discrepancy is defined by

$$L_P(D_N, P_{N,d}) = \Big(\int_{[0,1]^d} |z_1 z_2 \ldots z_d - \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{[0,z)}(t_i)|^p dz\Big)^{1/p},$$

where $1 \leq p < \infty$, $\mathbf{1}_A$ denotes the characteristic function on set A. In [26], authors gave the following approximation for different sampling sets.

Theorem 4.3. *[26] For d-dimension samples $X_1, X_2, X_3, \ldots, X_N$ are uniformly distributed in an isometric grid partition $\{Q_1, Q_2, \ldots, Q_N\}$ of the unit cube $[0, 1]^d$, integers $d, m \geq 2$ and $N \in \mathbb{N}$ such that $N = m^d$, then we have*

$$\mathbb{E}\Big[\Big|\frac{1}{N}\sum_{n=1}^{N} f(X_n) - \int_{[0,1]^d} f(x)dx\Big|^2\Big] \leq \frac{d}{N^{1+\frac{1}{d}}} \cdot \|f\|_{\mathcal{H}^1(K)}, \tag{4.14}$$

where f is a function in Sobolev space $\mathcal{H}^1(K)$.

Theorem 4.4. *[26] For the HSFC-based sampling, d-dimensional samples $X'_1, X'_2, X'_3, \ldots, X'_N$ uniformly distributed in $E_1, E_2, E_3, \ldots, E_N$, we have*

$$\mathbb{E}\Big[\Big|\frac{1}{N}\sum_{n=1}^{N} f(X'_n) - \int_{[0,1]^d} f(x)dx\Big|^2\Big] \leq \frac{2d\sqrt{d+3}}{N^{1+\frac{1}{d}}} \cdot \|f\|_{\mathcal{H}^1(K)}, \tag{4.15}$$

where f is a function in Sobolev space $\mathcal{H}^1(K)$.

Theorem 4.5. *[26] Let $x_1, x_2, \ldots, x_N$ be points from MC samples in $[0, 1]^d$, then we have*

$$\mathbb{E}\Big[\Big|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_{[0,1]^d} f(z)dz\Big|^2\Big] \leq \frac{d^{\frac{3}{2}}}{N} \cdot \|f\|_{\mathcal{H}^1(K)}, \tag{4.16}$$

where f is a function in $\mathcal{H}^1(K)$.

By comparing the mean square error of the three different sampling methods, we find that HSFC sampling and jittered sampling are superior to Monte Carlo sampling, and if the same sampling points are selected, the upper bound of jittered sampling is slightly better than HSFC sampling, but still the same order of convergence, whereas HSFC sampling has the advantage that it does not require jittered sampling of $N = m^d$ number of samples.

Example 3. Choose

$$f(x) = \frac{1}{\pi^d}\prod_{i=1}^{d} \sin(\pi x_i), \tag{4.17}$$

according to the definition, we have

$$\|f\|_{\mathcal{H}^1(K)} = \frac{1}{2^{\frac{d}{2}}}.$$

We then could achieve the following comparison; see Figure 5 for $d = 3$.

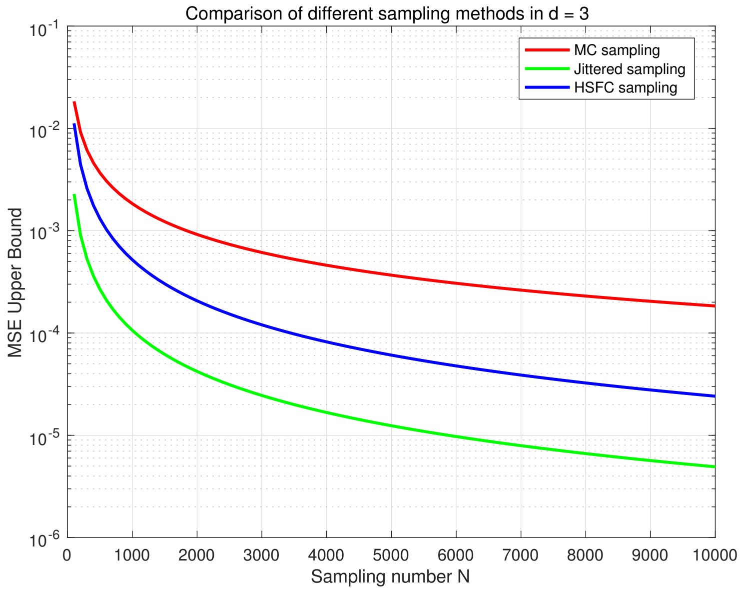

Figure 5. Comparison of Mean square errors for different sampling set.

5. Conclusions

The above analysis reveals that the convergence of the star discrepancy bound for HSFC-based sampling is $O(N^{-\frac{1}{2}-\frac{1}{2d}} \cdot (\ln N)^{\frac{1}{2}})$, which aligns with the rate achieved using jittered sampling sets and surpasses the rate obtained using the classical MC method. The stringent requirement for the sampling number $N = m^d$ in the jittered case is eliminated, thereby enhancing the applicability of this new stratified sampling method in higher dimensions. While our current findings are conservative, a more favorable convergence rate of the upper bound leads to improved integration approximation. However, a direct comparison of random star discrepancy sizes under different stratified sampling models remains unresolved.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The author was partially supported by the Natural Science Foundation of Suqian University (2024XRC017). This work was partly done at Suqian University, and the author thanks Suqian University for providing research support.

Conflict of interest

The author declares no conflict of interest.

References

1. M. Pharr, W. Jakob, G. Humphreys, *Physically based rendering: from theory to implementation*, Morgan Kaufmann Publishers Inc., 2016.
2. R. Ramamoorthi, J. Anderson, M. Meyer, D. Nowrouzezahrai, A theory of Monte Carlo visibility sampling, *ACM Trans. Graph.*, **31** (2012), 1–16. https://doi.org/10.1145/2231816.223181
3. A. Pilleboue, G. Singh, D. Coeurjolly, M. Kazhdan, V. Ostromoukhov, Variance analysis for Monte Carlo integration, *ACM Trans. Graph.*, **34** (2015), 1–14. https://doi.org/10.1145/2766930
4. Y. Lai, *Monte Carlo and quasi-Monte Carlo methods and their applications*, Ph.D. Dissertation, Claremont Graduate University, 1999.
5. Y. Z. Lai, Intermediate rank lattice rules and applications to finance, *Appl. Numer. Math.*, **59** (2009), 1–20. https://doi.org/10.1016/j.apnum.2007.11.024
6. J. Dick, F. Pillichshammer, *Digital nets and sequences*, Cambridge University Press, 2010.
7. C. Cervellera, M. Muselli, Deterministic design for neural network learning: an approach based on discrepancy, *IEEE Trans. Neural Netw.*, **15** (2004), 533–544. https://doi.org/10.1109/TNN.2004.824413
8. A. G. M. Ahmed, H. Perrier, D. Coeurjolly, V. Ostromoukhov, J. W. Guo, D. M. Yan, et al., Low-discrepancy blue noise sampling, *ACM Trans. Graph.*, **35** (2016), 1–13. https://doi.org/10.1145/2980179.2980218
9. S. Heinrich, E. Novak, G. W. Wasilkowski, H. Woźniakowski, The inverse of the star-discrepancy depends linearly on the dimension, *Acta Arith.*, **96** (2001), 279–302.
10. C. Aistleitner, Covering numbers, dyadic chaining and discrepancy, *J. Complexity*, **27** (2011), 531–540. https://doi.org/10.1016/j.jco.2011.03.001
11. M. Gnewuch, N. Hebbinghaus, Discrepancy bounds for a class of negatively dependent random points including Latin hypercube samples, *Ann. Appl. Probab.*, **31** (2021), 1944–1965.
12. B. Doerr, A sharp discrepancy bound for jittered sampling, *Math. Comp.*, **91** (2022), 1871–1892. https://doi.org/10.1090/mcom/3727
13. N. Chopin, H. J. Wang, M. Gerber, Adaptive stratified Monte Carlo using decision trees, 2025, arXiv: 2501.04842.
14. Z. J. He, A. B. Owen, Extensible grids: uniform sampling on a space filling curve, *J. R. Stat. Soc. Ser. B*, **78** (2016), 917–931. https://doi.org/10.1111/rssb.12132
15. Z. J. He, L. J. Zhu, Asymptotic normality of extensible grid sampling, *Stat. Comput.*, **29** (2019), 53–65. https://doi.org/10.1007/s11222-017-9794-y
16. J. Y. Tan, Z. J. He, X. Q. Wang, Extensible grid sampling for quantile estimation, *Math. Comp.*, **94** (2025), 763–800.

17. L. Ambrosio, A. Colesanti, E. Villa, Outer Minkowski content for some classes of closed sets, *Math. Ann.*, **342** (2008), 727–748. https://doi.org/10.1007/s00208-008-0254-z

18. B. Doerr, M. Gnewuch, A. Srivastav, Bounds and constructions for the star-discrepancy via δ-covers, *J. Complexity*, **21** (2005), 691–709. https://doi.org/10.1016/j.jco.2005.05.002

19. M. Gnewuch, Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy, *J. Complexity*, **24** (2008), 154–172. https://doi.org/10.1016/j.jco.2007.08.003

20. F. Cucker, D. X. Zhou, *Learning theory: an approximation theory viewpoint*, Cambridge University Press, 2007.

21. C. Aistleitner, M. Hofer, Probabilistic discrepancy bound for Monte Carlo point sets, *Math. Comp.*, **83** (2014), 1373–1381. https://doi.org/10.1090/S0025-5718-2013-02773-1

22. F. Y. Kuo, I. H. Sloan, Lifting the curse of dimensionality, *Not. AMS*, **52** (2005), 1320–1328.

23. E. Novak, H. Woźniakowski, *Tractability of multivariate problems, Volume II: Standard information for functionals*, European Mathematical Society, 2010. https://doi.org/10.4171/084

24. L. Brandolini, L. Colzani, G. Gigante, G. Travaglini, On the Koksma-Hlawka inequality, *J. Complexity*, **29** (2013), 158–172. https://doi.org/10.1016/j.jco.2012.10.003

25. J. Dick, F. Pillichshammer, Discrepancy theory and quasi-Monte Carlo integration, In: *A panorama of discrepancy theory*, Cham: Springer, 2014, 539–619. https://doi.org/10.1007/978-3-319-04696-9_9

26. X. D. Xu, D. Q. Han, Z. Y. Li, X. Q. Lin, Z. D. Qi, L. Zhang, Expected integration approximation under general equal measure partition, *Results Appl. Math.*, **21** (2024), 100419. https://doi.org/10.1016/j.rinam.2023.100419

AIMS Press