EXPECTED $L_2$–DISCREPANCY BOUND FOR A CLASS OF NEW STRATIFIED SAMPLING MODELS

JUN XIAN, XIAODA XU

ABSTRACT. We introduce a class of convex equivolume partitions. Expected $L_2$–discrepancy are discussed under these partitions. There are two main results. First, under this kind of partitions, we generate random point sets with smaller expected $L_2$–discrepancy than classical jittered sampling for the same sampling number. Second, an explicit expected $L_2$–discrepancy upper bound under this kind of partitions is also given. Further, among these new partitions, there is optimal expected $L_2$–discrepancy upper bound.

1. INTRODUCTION

In real sampling processes, it is necessary to know how well-spread these sampling points are. One can select the sampling set randomly which has achieved successful applications in the field of Monte Carlo simulation, compressed sensing, image processing and learning theory [8, 10, 12, 25, 31, 34, 41]. The concept of discrepancy is a fundamental building block in the quantification of many point distributions problems. There is a list of interesting discrepancy measures, such as star discrepancy, extreme discrepancy, $G$–discrepancy, isotrope discrepancy, lattice discrepancy, and so on (see e.g., [21, 22]). Among them, $L_2$–discrepancy is the most widely studied.

$L_2$–discrepancy. $L_2$–discrepancy of a sampling set $P_{N,d} = \{t_1, t_2, \ldots, t_N\}$ is defined by

\begin{equation}
L_2(D_N, P_{N,d}) = \left( \int_{[0,1]^d} |\lambda([0,z]) - \frac{1}{N} \sum_{i=1}^{N} 1_{[0,z]}(t_i)|^2 dz \right)^{1/2},
\end{equation}

where $\lambda$ denotes the Lebesgue measure, $1_A$ denotes the characteristic function on set $A$. For the applications of $L_2$–discrepancy, see [15, 18].

In the definition of $L_2$–discrepancy, if we introduce the counting measure #, (1.1) can also be expressed as

\begin{equation}
L_2(D_N, P_{N,d}) = \left( \int_{[0,1]^d} |\lambda([0,z]) - \frac{1}{N} \sum_{i=1}^{N} 1_{[0,z]}(t_i)|^2 dz \right)^{1/2},
\end{equation}

\textit{Date:} April 20, 2022.

\textit{2010 Mathematics Subject Classification.} 65C10, 11K38, 65D30.

\textit{Key words and phrases.} Expected star discrepancy; Stratified sampling; Convex equivolume partitions.
\[
L_2(D_N, P_{N,d}) = \left( \int_{[0,1]^d} |\lambda([0,z]) - \frac{1}{N} \#(P_{N,d} \cap [0,z])|^2 dz \right)^{1/2},
\]

where \(\#(P_{N,d} \cap [0,z])\) denotes the number of points falling into the set \([0,z]\).

To simplify the expression of \(L_2\)-discrepancy, we employ the discrepancy function \(\Delta(P_{N,d},z)\) via:

\[
\Delta(P_{N,d}, z) = \lambda([0,z]) - \frac{1}{N} \#(P_{N,d} \cap [0,z]).
\]

Accordingly, the \(L_2\)-discrepancy can be extended to a fixed compact convex set \(K \subset \mathbb{R}^d\) with \(\lambda(K) > 0\), see [29]. Discrepancy function in (1.3) of a finite set of points \(P = \{x_1, x_2, \ldots, x_n\} \subset K\) is now given by

\[
\Delta(P, x) = \frac{\lambda([\infty, x] \cap K)}{\lambda(K)} - \frac{1}{N} \#(P \cap (\infty, x)).
\]

For fixed \(d\), the best known asymptotic upper bounds for discrepancy are of the form

\[
O\left(\frac{(\ln N)^{\alpha_d}}{N}\right),
\]

where \(\alpha_d \geq 0\) are constants depending on dimension \(d\). These involve special deterministic point set constructions, which are low discrepancy point sets. Examples of such point sets can be found in [14, 36]. For applications arising in computer graphics, quantitative finance and learning theory, see e.g., [1, 9, 32, 33].

Although low discrepancy (deterministic) point sets are widely used in numerical integration, the simulation of many phenomena in the real world requires the introduction of random factors. Recently, a large amount of research investigating random sampling for different function spaces has emerged in [3, 4, 24], due to the simplicity, flexibility and effectiveness of the subject. Besides, in the field of discrepancy, probabilistic star discrepancy bounds for Monte Carlo point sets are considered in [2, 28], while centered discrepancy of random sampling and Latin hypercube random sampling are investigated in [23]. Motivated by these developments, we incorporate a random viewpoint into our study of star discrepancy to consider a special random sampling method, which is stratified sampling. Its special case is called jittered sampling that is formed by grid-based equivolume partition.

Some random sampling strategies, for example, simple random sampling, stratified sampling, Latin hypercube sampling, etc. are commonly used in the real sampling process, see [11, 35, 39]. Formers have made sufficient research on estimating the
expected discrepancy with random samples. For researches on expected star discrepancy of jittered sampling, we refer to [20, 38]. Both the upper and the lower bounds for the discrepancy of jittered sampling are given in [38], while the bounds in [20] improve them and remove the asymptotic requirement that $m$ is sufficiently large compared to dimensions $d$ (where $N = m^d$ means the number of subcubes of grid-based equivolume partition). Starting from the discrepancy itself, rather than estimating its bound. In [30], it is shown that jittered sampling construction gives rise to a set whose expected $L_p$-discrepancy is smaller than that of purely random points. Further, a theoretical conclusion that the jittered sampling does not have the minimal expected $L_2$-discrepancy among all stratified samples from convex equivolume partitions with the same number of points is presented in [29]. Our research will be carried out on the $d$-dimensional unit cube, which can be easily extended to a more general compact convex set. Studies on convex bodies are extensive, see [7, 26]. In the following, we shall construct a class of convex body partitions to analyze expected $L_2$-discrepancy, which turns out to provide better results than jittered sampling.

Throughout this paper, we adopt the idea of stratified random sampling to study $L_2$-discrepancy. First, we design an infinite family of partitions with partition parameter $0 \leq \theta \leq \frac{\pi}{2}$ that generates point sets with a smaller expected $L_2$-discrepancy than classical stratified sampling for sampling number $N = m^d$, which is,

$$\mathbb{E}(L_2^2(D_N, P_{\Omega_*})) \leq \mathbb{E}(L_2^2(D_N, P_{\Omega^*})), \quad$$

where $P_{\Omega_*}$ and $P_{\Omega^*}$ denote stratified samples generated by the new infinite family of partitions and grid-based equivolume partition respectively. The equal signs hold if and only if stratified sampling sets $P_{\Omega_*}$ are selected for jittered sampling set $P_{\Omega^*}$. Second, optimal expected $L_2$-discrepancy bound is also provided under this class of partitions. That is, they are better than the employment of jittered sampling. We obtain the following explicit estimation

$$\mathbb{E}(L_2^2(D_N, P_{\Omega_*})) \leq \frac{d}{N^{1+\frac{1}{3}}} + \frac{1}{N^3} \cdot \frac{1}{3^{d-2}} \cdot P(\theta),$$

where $P(\theta)$ is the function about partition $\theta$. Taking $\theta = arctan\frac{1}{2}$ and $\theta = 0$, we can obtain the upper bounds for optimal partition and grid-based equivolume partition respectively.

The rest of this paper is organized as follows. In Section 2 we present some preliminaries on stratified sampling and newly designed partition models. In Section 3 we provide comparisons of the expected $L_2$-discrepancy for stratified sampling under a kind of convex equivolume partitions. The explicit expected $L_2$-discrepancy upper bounds for these newly stratified models are also obtained. In Section 4 we
include the proofs of all theorems and lemmas. Finally, in Section 5 we conclude the paper.

2. Preliminaries on stratified sampling and new partition models

Before introducing the main result, we list preliminaries used in this paper.

2.1. Stratified sampling. Stratified sampling is a special random sampling, that is different from simple random sampling, see Figure 1. The original sampling area is divided, and a uniformly distributed random sample point is selected in each subset of partitions. Jittered sampling is a special case of stratified sampling, involving grid-based equivolume partition. Explicitly, $[0, 1]^d$ is divided into $m^d$ axis parallel boxes $Q_i, 1 \leq i \leq N$, each with sides $\frac{1}{m}$, see Figure 2. Research on the jittered sampling are extensive, see [11, 20, 29, 30, 38].

![Simple random sampling](image1)

![Jittered sampling](image2)

**Figure 1.** Simple random sampling.

We now consider a rectangle $R = [0, x)$ (we shall call it the test set in the following) in $[0, 1]^d$ anchored at 0. For an isometric grid partition $\Omega = \{Q_1, Q_2, \ldots, Q_N\}$ of $[0, 1]^d$, we put

$$I_N := \{j : \partial R \cap Q_j \neq \emptyset\},$$

and

$$C_N := |I_N|,$$

which means the cardinality of the index set $I_N$. For $C_N$, it is easy to obtain

$$C_N \leq d \cdot N^{1-\frac{d}{2}}. \tag{2.1}$$
2.2. New partition models. In the end of this section, we design a class of partitions and construct it step by step. First, we consider the two-dimensional case.

Step one: a class of partitions design for two dimension.

Our designed equivolume partition is actually a special case of general equivolume partition (see Figure 3 for illustration in two dimensional case). For a grid-based equivolume partition in two dimension, we merge the two squares in the upper right corner to form a rectangle, then we use a series of straight line partitions to divide the rectangle into two equal-volume parts, which will be converted to a one-parameter model if we set the angle between the dividing line and horizontal line across the center $\theta$, where we suppose $0 \leq \theta \leq \frac{\pi}{2}$. From simple calculations, we can conclude the arbitrary straight line must pass through the center of the rectangle. For convenience of notation, we set this partition model $\Omega_{\sim} = (\Omega_1, \sim, \Omega_2, \sim, Q_3, \ldots, Q_N)$ in two dimensional case.
In the above one-parameter model, the case will be grid-based equivolume partition if we choose \( \theta = \frac{\pi}{2} \). The case \( \theta = \arctan \frac{1}{2} \) is introduced in \([29]\), see Figure 4 for two dimensional case. For notation convenience, we set this partition model \( \Omega = (\Omega_1, \Omega_2, Q_3, \ldots, Q_N) \) in two dimensional case.

![Figure 4. The partition for parameter \( \theta = \arctan \frac{1}{2} \) in two dimension](image)

The only difference between the new partition model and grid-based equivolume partition is to change two closed hypercubes into two special convex bodies, see illustration in Figure 5.

![Figure 5. Difference between two partition models.](image)

**Step two:** Suppose the original rectangle is \( I \), for the convenience of calculation, we set the lower left corner of the rectangle at the origin \((0,0)\) and the side length of the small square to 1. Now, consider \( I = [0, 2] \times [0, 1] \) and its two equivolume partitions \((\Omega_{1,\sim}, \Omega_{2,\sim})\) into two convex bodies with
\[ \Omega_{1,|} = [0, 1] \times [0, 1], \Omega_{1,\sim} = \text{conv}\{ (0, 0), (1 + \frac{\cot \theta}{2}, 0), (0, 1), (1 - \frac{\cot \theta}{2}, 1) \}, \]

where \( \text{conv} \) denotes the convex hull.

**Step three:** We consider the translation and stretch of the rectangle \( I = [0, 2] \times [0, 1] \) into

\[ I' = [a_1, a_1 + 2b] \times [a_2, a_2 + b], \]

the above two dimensional case in Step one can then be extended to \( d \)-dimension as [29]. Consider \( d \)-dimensional cuboid

\[ (2.2) \quad I_d = I' \times \prod_{i=3}^{d} [a_i, a_i + b] \]

and its three equivolume partitions \( \Omega'_1 = (\Omega'_1, \Omega'_2) \) into two closed hypercubes, \( \Omega' = (\Omega'_1, \Omega'_2) \) into two closed, regular triangular hyperprisms and \( \Omega'_\sim = (\Omega'_{1,\sim}, \Omega'_{2,\sim}) \) into two closed, trapezoidal superconvex bodies with

\[ (2.3) \quad \Omega'_{1,|} = \prod_{i=1}^{d} [a_i, a_i + b], \]

\[ (2.4) \quad \Omega'_{1,\sim} = \text{conv}\{ (a_1, a_2), (a_1 + 2b, a_2), (a_1, a_2 + b) \} \times \prod_{i=3}^{d} [a_i, a_i + b], \]

and

\[ \Omega'_{1,\sim} = \text{conv}\{ (a_1, a_2), (a_1 + b + \frac{b \cdot \cot \theta}{2}, a_2), (a_1, a_2 + b), (a_1 + b - \frac{b \cdot \cot \theta}{2}, a_2 + b) \}
\]
\[ \times \prod_{i=3}^{d} [a_i, a_i + b], \]

where \( \text{conv} \) denotes the convex hull.

Just as grid-based partition \( N = m^d \), where \( m \) represents the number of partitions in each dimension and \( d \) denotes the dimensions. If we choose \( a_1 = \frac{m-2}{m}, a_2 = \frac{m-1}{m}, b = \frac{1}{m} \), then, through the construction method from step one to step three, we get a series of partitions (where we set \( 0 \leq \theta \leq \frac{\pi}{2} \)) that we call **local convex partition**, denoted by

\[ (2.5) \quad \Omega'_\sim = (\Omega'_{1,\sim}, \Omega'_{2,\sim}, Q_3 \ldots, Q_N). \]
Among the above local convex partition $\Omega^*_\sim$, if we choose the partition parameter $\theta = \frac{\pi}{2}$, isometric grid with partition number $m$ in each dimension is obtained, which we set

$$\Omega^*_\sim = (\Omega^*_1, \Omega^*_2, \ldots, \Omega^*_N).$$

Likewise, if we choose the partition parameter $\theta = \arctan \frac{1}{2}$, partition model in two dimensional case introduced above can then be extended to $d$ dimension, and we choose $a_1 = \frac{m-2}{m}, a_2 = \frac{m-1}{m}, b = \frac{1}{m}$ in (2.4), then this partition model is denoted by

$$\Omega^*_{\\setminus} = (\Omega^*_{1\setminus}, \Omega^*_{2\setminus}, \ldots, \Omega^*_{N}).$$
3. Expected $L_2$—discrepancy for stratified random sampling

In this section, comparisons of expected $L_2$—discrepancy under different partition models are obtained. Furthermore, we study expected $L_2$—discrepancy and several bounds are given under newly designed partition models.

3.1. Expected $L_2$—discrepancy under two partition models.

**Theorem 3.1.** Let $m,d \in \mathbb{N}$ with $m \geq d \geq 2, 0 \leq \theta \leq \frac{\pi}{2}$ and $N = m^d$. Stratified random $d$—dimension point sets $P_{\Omega^*_|}$ and $P_{\Omega^*_\setminus}$ are uniformly distributed in the grid-based stratified subsets of $\Omega^*_|$ and stratified subsets of $\Omega^*_\setminus$ respectively, then

\[
\mathbb{E}(L_2^2(D_N, P_{\Omega^*_\setminus})) \leq \mathbb{E}(L_2^2(D_N, P_{\Omega^*_|})).
\]

where $\Omega^*_|$, $\Omega^*_\setminus$ are defined in (2.5), (2.6) respectively and $\theta$ is the partition parameter related to $\Omega^*_\setminus$ as defined in Section 2.

**Remark 3.2.** In Theorem 3.1, as an infinite family of partitions is designed to generate point sets with a smaller expected $L_2$—discrepancy than classical stratified sampling (jittered sampling) for the same sampling number $N = m^d$. The equal signs on both sides of (3.2) hold if and only if when $\theta = 0$ or $\theta = \frac{\pi}{2}$.

**Corollary 3.3.** Let $N = m^d$ and $m,d \in \mathbb{N}$ with $m \geq d \geq 2$. Stratified random $d$—dimension point sets $P_{\Omega^*_|}$ and $P_{\Omega^*_\setminus}$ are uniformly distributed in $\Omega^*_|$ and $\Omega^*_\setminus$ respectively, then

\[
\mathbb{E}(L_2^2(D_N, P_{\Omega^*_\setminus})) < \mathbb{E}(L_2^2(D_N, P_{\Omega^*_|})).
\]

where $\Omega^*_|$ and $\Omega^*_\setminus$ defined in (2.6) and (2.7) respectively.
Remark 3.4. Actually, (3.2) holds if we choose parameter $\theta = \arctan \frac{1}{2}$ in Theorem 3.1. The Corollary 3.3 is main result in [29]. Obvious, the partition manner in [29] as Figure 4 is included in our new partition models as Figure 3.

3.2. Expected $L_2$—discrepancy upper bounds under the new partition models. In this subsection, expected $L_2$—discrepancy bounds under new partition models are given. Optimal result is also obtained under this class of partitions.

Theorem 3.5. Let $m, d \in \mathbb{N}$ with $m \geq d \geq 2, 0 \leq \theta \leq \frac{\pi}{2}$. Let $N = m^d$, the stratified random $d$—dimension point set $P_{\Omega^*}$ distributed in subsets of $\Omega^*$ defined in (2.5), then

$$(3.3) \quad \mathbb{E}(L_2(D_N, P_{\Omega^*})) \leq \frac{d}{N^{1+\frac{1}{d}}} + \frac{1}{N^3} \cdot \frac{1}{3^{d-2}} \cdot P(\theta),$$

where

$$P(\theta) = \begin{cases} \frac{2}{45} \tan^3 \theta + \frac{2}{15} \tan^2 \theta - \frac{\tan \theta}{6}, & 0 \leq \theta < \arctan \frac{1}{2}, \\ -\frac{2}{45}, & \theta = \arctan \frac{1}{2}, \\ -\frac{1}{24 \tan \theta} + \frac{1}{120 \tan^2 \theta} + \frac{1}{1440 \tan^3 \theta}, & \arctan \frac{1}{2} < \theta \leq \frac{\pi}{2}. \end{cases}$$

Remark 3.6. Noticing that in Theorem 3.5, $P(\theta)$ is a continuous function, decreases monotonically between 0 and $\arctan \frac{1}{2}$ and increases monotonically between $\arctan \frac{1}{2}$ and $\frac{\pi}{2}$, see Figure 9. Choose parameter $\theta = \frac{\pi}{2}$ in Theorem 3.5, then we are back to the case of classical jittered sampling. Furthermore, all of these local convex partitions with parameter $\theta \in (0, \frac{\pi}{2})$ obtain better upper bounds of expected $L_2$—discrepancy than the jittered sampling.
Corollary 3.7. Let $m, d \in \mathbb{N}$ with $m \geq d \geq 2$. Let $N = m^d$, the stratified random $d$-dimension point set $P_{\Omega^*}$ distributed in subsets of $\Omega^*$ defined in (2.7), then we obtain optimal expected $L_2$-discrepancy bound under new partition models

$$
E(L_2^2(D_N, P_{\Omega^*})) \leq \frac{d}{N^{1+\frac{1}{d}}} - \frac{2}{45} \cdot \frac{1}{N^3} \cdot \frac{1}{3d-2}.
$$

Remark 3.8. The optimal expected $L_2$-discrepancy bound under this class of partitions is obtained at $\theta = \arctan \frac{1}{2}$ in Theorem 3.5. An upper bound on the expected $L_p$-discrepancy is derived by acceptance-rejection sampler using stratified inputs under the implicit constants in [42]. Our results give explicit expected $L_2$-discrepancy bounds under a class of new partitions, which our order is the same with [42].

3.3. Some Examples. This subsection presents some examples of expected $L_2$-discrepancy bounds under different sampling models for $N = m^d$. The cases of $\theta = \arctan \frac{1}{2}$ and $\theta = \frac{\pi}{4}$ acquire better result than that of jittered sampling.

Example 1. Expected bound of stratified sampling set for $\theta = 0$

$$
E(L_2^2(D_N)) \leq \frac{d}{N^{1+\frac{1}{d}}}.
$$
**Example 2.** Expected bound of stratified sampling set for $\theta = \frac{\pi}{4}$

$$
\mathbb{E}(L_2^2(D_N)) \leq \frac{d}{N^{1+\frac{1}{d}}} - \frac{47}{1440 \cdot 3^{d-2} \cdot N^3}.
$$

**Example 3.** Expected bound of stratified sampling set for $\theta = \frac{\pi}{2}$

$$
\mathbb{E}(L_2^2(D_N)) \leq \frac{d}{N^{1+\frac{1}{d}}}. 
$$
Figure 12. Stratified sampling for $\theta = \frac{\pi}{2}$

**Example 4.** Expected bound of stratified sampling set for $\theta = \arctan \frac{1}{2}$

\[
\mathbb{E}(L_2^2(D_N)) \leq \frac{d}{N^{1+\frac{1}{d}}} - \frac{2}{45 \cdot 3^{d-2} \cdot N^3}.
\]

Figure 13. Stratified sampling for $\theta = \arctan \frac{1}{2}$

4. **Proofs**

In this section, we present the proofs of Theorem 3.1 and 3.5. The following lemma reveals the expected $L_2$-discrepancy quantitative relationship between the two partition models $\Omega\sim$ and $\Omega\mid$.

**Lemma 4.1.** For two equivolume partitions $\Omega\sim = (\Omega_1\sim, \Omega_2\sim, Q_3, \ldots, Q_N)$ and $\Omega\mid = \{Q_1, Q_2, Q_3, \ldots, Q_N\}$ as defined in (2.5) and (2.6) respectively, we have
\[ \mathbb{E}L_2^2(D_N, P_{\Omega}) - \mathbb{E}L_2^2(D_N, P_{\Omega'}) = \begin{cases} \frac{1}{N^3} \cdot \frac{1}{3^{d-2}} \cdot P_1(\theta), & 0 \leq \theta < \arctan \frac{1}{2}, \\ -2 \cdot \frac{1}{45} \cdot \frac{1}{N^3} \cdot \frac{1}{3^{d-2}}, & \theta = \arctan \frac{1}{2}, \\ \frac{1}{N^3} \cdot \frac{1}{3^{d-2}} \cdot P_2(\theta), & \arctan \frac{1}{2} < \theta \leq \frac{\pi}{2}. \end{cases} \]

where
\[ P_1(\theta) = \frac{2}{45} \tan^3 \theta + \frac{2}{15} \tan^2 \theta - \frac{\tan \theta}{6} , \]

and
\[ P_2(\theta) = -\frac{1}{24 \tan \theta} + \frac{1}{120 \tan^2 \theta} + \frac{1}{1440 \tan^3 \theta}. \]

4.1. **Proof of Lemma 4.1**. For equivolume partition \( \Omega_{0,\sim} = (\Omega_{1,\sim}, \Omega_{2,\sim}) \) of \( I \) (the same argument if we replace \( \Omega_{0,\sim} \) with \( \Omega_{0,|} \)), from [Proposition 2] in [29], which is, for an equivolume partition \( \Omega = \{\Omega_1, \Omega_2, \ldots, \Omega_N\} \) of a compact convex set \( K \subset \mathbb{R}^d \) with \( \lambda(K) > 0 \), \( P_\Omega \) is the corresponding stratified sampling set, then

\[ \mathbb{E}L_2^2(D_N, P_{\Omega'}) = \frac{1}{N^2 \lambda(K)} \sum_{i=1}^{N} \int_K q_i(x)(1 - q_i(x))dx, \]

where
\[ q_i(x) = \frac{\lambda(\Omega_i \cap [0, x])}{\lambda(\Omega_i)}. \]

Through simple derivation, it follows that

\[ \mathbb{E}L_2^2(D_N, P_{\Omega_{0,\sim}}) = \frac{1}{8} \sum_{i=1}^{2} \int_I q_i(x)(1 - q_i(x))dx, \]

and

\[ q_i(x) = \frac{\lambda(\Omega_i, \sim \cap [0, x])}{\lambda(\Omega_i, \sim)} = \lambda(\Omega_i, \sim \cap [0, x]). \]

Conclusion (4.4) is equivalent to the following
We first consider parameter $\arctan \frac{1}{2} \leq \theta \leq \frac{\pi}{2}$, then we define the following two functions for simplicity of the expression.

\[
F(x) = \frac{1}{2} \cdot [(x_1 - 1) \tan \theta + x_2 - \frac{1}{2}] \cdot [(x_1 - 1) + (x_2 - \frac{1}{2}) \cdot \cot \theta],
\]

and

\[
G(x) = x_1x_2 - x_2 - \frac{\cot \theta}{2}x_2 + \frac{1}{2}x_2^2 \cdot \cot \theta,
\]

where $x = (x_1, x_2)$.

Furthermore, for $\Omega_{0,i} = (\Omega_{1,i}, \Omega_{2,i})$ defined in Step two of Section 2.2, (4.5) implies

\[
q_{1,i}(x) = \begin{cases} 
  x_1x_2, & x \in \Omega_{1,i} \\
  x_2, & x \in \Omega_{2,i},
\end{cases}
\]

and

\[
q_{2,i}(x) = \begin{cases} 
  0, & x \in \Omega_{1,i} \\
  (x_1 - 1)x_2, & x \in \Omega_{2,i},
\end{cases}
\]

Besides,

\[
q_{1,\sim}(x) = \begin{cases} 
  x_1x_2, & x \in \Omega_{1,\sim} \\
  x_1x_2 - F(x), & x \in \Omega_{2,\sim,1} \\
  x_1x_2 - G(x), & x \in \Omega_{2,\sim,2},
\end{cases}
\]

and

\[
q_{2,\sim}(x) = \begin{cases} 
  0, & x \in \Omega_{1,\sim} \\
  F(x), & x \in \Omega_{2,\sim,1} \\
  G(x), & x \in \Omega_{2,\sim,2},
\end{cases}
\]

where $\Omega_{1,\sim}, \Omega_{2,\sim}$ denote subsets of partition $\Omega_{0,\sim}$. In the following, we shall continue to divide subsets $\Omega_{1,\sim} = \{\Omega_{1,\sim,1, \Omega_{1,\sim,2}}\}$ and $\Omega_{2,\sim} = \{\Omega_{2,\sim,1, \Omega_{2,\sim,2}}\}$ to facilitate calculation. See Figures 14 to 15.

Therefore, for $\theta = \frac{\pi}{2}$, we introduce two symbols $B_{1,i}, B_{2,i}$ and have

\[
(4.6) \quad B_{1,i} = \int_I q_{1,i}^2(x)dx = \int_{\Omega_{1,i}} x_1^2x_2^2dx + \int_{\Omega_{2,i}} x_2^2dx = \frac{1}{9} + \frac{1}{3} = \frac{4}{9},
\]

and
\[ B_{2,\|} = \int_I q_{2,\|}^2(x) \, dx = \int_{\Omega_{2,\|}} (x_1 - 1)^2 x_2^2 \, dx = \frac{1}{9}. \]

Thus,

\[ 8 \mathbb{E}(L^2(P_{\Omega_{0,\|}})) = 1 - (B_{1,\|} + B_{2,\|}) = \frac{4}{9}. \]

Furthermore, we introduce \( B_{1,\sim} \) and \( B_{2,\sim} \), then

\[ B_{1,\sim} = \int_I q_{1,\sim}^2(x) \, dx = \int_{\Omega_{1,\sim}} x_1^2 x_2^2 \, dx + \int_{\Omega_{2,\sim,1}} (x_1 x_2 - F(x))^2 \, dx \]

\[ + \int_{\Omega_{2,\sim,2}} (x_1 x_2 - G(x))^2 \, dx, \]

and

\[ B_{2,\sim} = \int_I q_{2,\sim}^2(x) \, dx = \int_{\Omega_{2,\sim,1}} F^2(x) \, dx + \int_{\Omega_{2,\sim,2}} G^2(x) \, dx. \]

We divide our calculation in three steps. First, we compute \( \int_{\Omega_{1,\sim}} x_1^2 x_2^2 \, dx \), see Figure 14 for illustration.

\[ \int_{\Omega_{1,\sim,1}} x_1^2 x_2^2 \, dx = \int_0^{1 - \cot \theta} x_1^2 dx_1 \cdot \int_0^1 x_2^2 dx_2 = \frac{(2 - \cot \theta)^2}{72}. \]
\[ \int_{\Omega_{1,\sim,2}} x_1^2 x_2^2 d\mathbf{x} = \int_{1 - \frac{\cot \theta}{2}}^{1 + \frac{\cot \theta}{2}} x_1^2 dx_1 \cdot \int_{0}^{(1 - x_1) \cdot \tan \theta + \frac{1}{2}} x_2^2 dx_2 = \frac{60 \tan^2 \theta - 36 \tan \theta + 7}{720 \tan^3 \theta}. \]

Therefore, (4.10) and (4.11) imply

\[ \int_{\Omega_{1,\sim}} x_1^2 x_2^2 d\mathbf{x} = \int_{\Omega_{1,\sim,1}} x_1^2 x_2^2 d\mathbf{x} + \int_{\Omega_{1,\sim,2}} x_1^2 x_2^2 d\mathbf{x} = -\frac{1}{12 \tan \theta} + \frac{1}{30 \tan^2 \theta} - \frac{1}{240 \tan^3 \theta} + \frac{1}{9}. \]

Second, we compute \( \int_{\Omega_{2,\sim,1}} (x_1 x_2 - F(x))^2 d\mathbf{x} \) and \( \int_{\Omega_{2,\sim,2}} (x_1 x_2 - G(x))^2 d\mathbf{x} \).

\[ \int_{\Omega_{2,\sim,1}} (x_1 x_2 - F(x))^2 d\mathbf{x} = \int_{1 - \frac{\cot \theta}{2}}^{1} \int_{0}^{1} (x_1 x_2 - F(x))^2 x_2 dx_2 dx_1 = \frac{180 \tan^2 \theta - 12 \tan \theta + 5}{720 \tan^3 \theta}, \]

\[ \int_{\Omega_{2,\sim,2}} (x_1 x_2 - G(x))^2 d\mathbf{x} = \int_{1 + \frac{\cot \theta}{2}}^{2} \int_{0}^{1} (x_1 x_2 - G(x))^2 x_2 dx_2 dx_1 = -\frac{\cot^3 \theta}{240} - \frac{\cot^2 \theta}{30} - \frac{\cot \theta}{12} + \frac{1}{3}. \]

Thus, (4.13) and (4.14) imply

**Figure 15.** Division of the integral region.
\[
\int_{\Omega_2,\sim,1} (x_1 x_2 - F(x))^2 dx + \int_{\Omega_2,\sim,2} (x_1 x_2 - G(x))^2 dx \\
= \frac{1}{3} + \frac{1}{6\tan\theta} - \frac{1}{20\tan^2\theta} + \frac{1}{360\tan^3\theta}.
\]

Combining (4.9), (4.12) and (4.15), we have

\[
B_{1,\sim} = \int_{\Omega_1,\sim} x_1^2 x_2^2 dx + \int_{\Omega_2,\sim,1} (x_1 x_2 - F(x))^2 dx + \int_{\Omega_2,\sim,2} (x_1 x_2 - G(x))^2 dx \\
= \frac{1}{12\tan\theta} - \frac{1}{60\tan^2\theta} - \frac{1}{720\tan^3\theta} + \frac{4}{9}.
\]

Third, we will compute \(\int_{\Omega_2,\sim,1} F^2(x) dx\) and \(\int_{\Omega_2,\sim,2} G^2(x) dx\) in the following.

In fact,

\[
\int_{\Omega_2,\sim,1} F^2(x) dx = \int_{1}^{1+c\tan\theta} \int_{0}^{1} F^2(x) dx_2 dx_1 \\
= \frac{1}{120\tan^3\theta},
\]

\[
\int_{\Omega_2,\sim,2} G^2(x) dx = \int_{1}^{2+c\tan\theta} \int_{0}^{1} G^2(x) dx_2 dx_1 \\
= \frac{1}{9} - \frac{1}{24\tan\theta} + \frac{1}{120\tan^2\theta} - \frac{11}{1440\tan^3\theta}.
\]

Combining (4.17) and (4.18), we have

\[
B_{2,\sim} = \int_{\Omega_2,\sim,1} F^2(x) dx + \int_{\Omega_2,\sim,2} G^2(x) dx \\
= \frac{1}{9} - \frac{1}{24\tan\theta} + \frac{1}{120\tan^2\theta} + \frac{1}{1440\tan^3\theta}.
\]

Thus,

\[
B_{1,\sim} + B_{2,\sim} = \frac{1}{24\tan\theta} - \frac{1}{120\tan^2\theta} - \frac{1}{1440\tan^3\theta} + \frac{5}{9}.
\]

Therefore,
\[ 8 \mathbb{E}(L_2^2(P_{\Omega_0,\sim})) = 1 - (B_{1,\sim} + B_{2,\sim}) \]

(4.21)

\[ = -\frac{\cot \theta}{24} + \frac{\cot^2 \theta}{120} + \frac{\cot^3 \theta}{1440} + \frac{4}{9}, \]

where \( \arctan \frac{1}{2} \leq \theta < \frac{\pi}{2} \).

For \( \theta = \frac{\pi}{2} \), by (4.8) we have

(4.22)

\[ 8 \mathbb{E}(L_2^2(P_{\Omega_0,\sim})) = \frac{4}{9}. \]

\[ \begin{align*}
\Omega'_{2,\sim} & \hspace{2cm} \Omega'_{1,\sim} \\
\Omega'_{2,\sim,1} & \hspace{2cm} \Omega'_{2,\sim,2}
\end{align*} \]

(a)

\[ \begin{align*}
\Omega'_{1,\sim} & \hspace{2cm} \Omega'_{2,\sim} \\
\Omega'_{1,\sim,1} & \hspace{2cm} \Omega'_{1,\sim,2}
\end{align*} \]

(b)

**Figure 16.** Division of the integral region.

Considering the case \( 0 \leq \theta < \arctan \frac{1}{2} \), we denote the partition by \( \Omega'_\sim = \{\Omega'_{1,\sim}, \Omega'_{2,\sim}\} \), see Figure 16. Let

\[ q'_{1,\sim}(x) = \begin{cases} 
x_1x_2, x \in \Omega'_{1,\sim}, \\
x_1x_2 - H(x), x \in \Omega'_{2,\sim,1}, \\
x_1x_2 - J(x), x \in \Omega'_{2,\sim,2},
\end{cases} \]

and

\[ q'_{2,\sim}(x) = \begin{cases} 
0, x \in \Omega'_{1,\sim}, \\
H(x), x \in \Omega'_{2,\sim,1}, \\
J(x), x \in \Omega'_{2,\sim,2},
\end{cases} \]

where

(4.23)

\[ H(x) = \frac{1}{2} \cdot [x_2 - (1 - x_1) \tan \theta - \frac{1}{2}] \cdot [\cot \theta \cdot x_2 - 1 + x_1 - \frac{1}{2} \cot \theta], \]

and
\[ J(x) = \left[ x_2 - \tan \theta - \frac{1}{2} \right] \cdot x_1 + \frac{1}{2} x_1^2 \cdot \tan \theta. \]

Then we divide subsets \( \Omega_{1,\sim} = \{ \Omega_{1,\sim,1}, \Omega_{1,\sim,2} \} \) and \( \Omega_{2,\sim} = \{ \Omega_{2,\sim,1}, \Omega_{2,\sim,2} \} \) to facilitate calculation. See Figure 16.

So

\[ B_{1,\sim}' = \int_I q_{1,\sim}^2 \cdot \partial x = \int_{\Omega_{1,\sim}} x_1^2 \cdot \partial x + \int_{\Omega_{2,\sim,1}} (x_1 x_2 - H(x))^2 \cdot \partial x \]
\[ + \int_{\Omega_{2,\sim,2}} (x_1 x_2 - J(x))^2 \cdot \partial x, \]

and

\[ B_{2,\sim}' = \int_I q_{2,\sim}^2 \cdot \partial x = \int_{\Omega_{2,\sim,1}} H^2(x) \cdot \partial x + \int_{\Omega_{2,\sim,2}} J^2(x) \cdot \partial x. \]

If we follow the calculation process of (4.10)-(4.20), then we obtain

\[ B_{1,\sim}' = -\frac{4}{45} \tan^3 \theta - \frac{4}{15} \tan^2 \theta + \frac{\tan \theta}{3} + \frac{4}{9}, \]

and

\[ B_{2,\sim}' = \frac{2}{45} \tan^3 \theta + \frac{2}{15} \tan^2 \theta + \frac{\tan \theta}{6} + \frac{1}{9}. \]

Thus,

\[ B_{1,\sim}' + B_{2,\sim}' = -\frac{2}{45} \tan^3 \theta - \frac{2}{15} \tan^2 \theta + \frac{\tan \theta}{6} + \frac{5}{9}. \]

Hence,

\[ 8 \mathbb{E}(L_2^2(P_{\mathbb{W}_\sim})) = 1 - (B_{1,\sim}' + B_{2,\sim}') = \frac{4}{9} + \frac{2}{45} \tan^3 \theta + \frac{2}{15} \tan^2 \theta - \frac{\tan \theta}{6}, \]

where \( 0 \leq \theta < \arctan \frac{1}{2}. \)

Combining with (4.21) and considering the translation and stretch of the rectangle \( I = [0, 2] \times [0, 1] \) into

\[ I' = [a_1, a_1 + 2b] \times [a_2, a_2 + b], \]

we obtain
where $a_1 = \frac{m-2}{m}$, $a_2 = \frac{m-1}{m}$, $b = \frac{1}{m}$. $\Omega_\ast$ is the infinite family of equi-volume partitions defined in (2.5) and $\Omega_\ast^r$ is grid-based equi-volume partition defined in (2.6). The equal sign of (4.30) holds if and only if partition parameter $\theta = 0, \pi_2$. Noting that conclusion (4.30) is only for the two-dimensional case.

Next we will give a proof of (4.30) for $d-$dimensional case. We firstly prove the case $b = 1$ and $(a_1, a_2, \ldots, a_d) = (0, 0, \ldots, 0)$.

Let $I_d' = [0, 2] \times [0, 1] \times [0, 1]^{d-2}$. For $i = 1, 2$, we have

$$q_i' = q_i(x_1, x_2) \cdot \prod_{j=3}^{d} x_j,$$

where $q_i' = q_i(x_1, x_2)$ is defined as (4.5) for $\Omega''_\ast$.

Thus,

$$\int_{I_d'} q_i'^2 d\lambda = B_i' \cdot \int_{[0,1]^{d-2}} \prod_{j=3}^{d} x_j^2 dx_3 dx_4 \cdot \ldots dx_d = \frac{1}{3^{d-2}} \cdot B_i',$$

where $B_i' \ast , i = 1, 2$ have been calculated in (4.16) and (4.19) respectively.

As we have

$$\int_{I_d} \lambda([0, x]) d\lambda = \int_{[0,1]^{d-2}} \prod_{j=3}^{d} x_j dx_3 dx_4 \cdot \ldots dx_d = \frac{1}{2^{d-2}}.$$

Then we obtain,

$$\mathbb{E}(L_2^2(\Omega''_\ast)) = \frac{1}{2^{d-2}} = \frac{1}{3^{d-2}} \cdot (B_1' + B_2').$$

Now, for $I_d$ in (2.2), we define a vector

$$a = \{a_1, a_2, \ldots, a_d\}.$$

We then prove (4.2) is independent of $a$. In $I_d$, we choose $a = 0$, set

$$I_d^0 = [0, 2b] \times [0, b]^{d-1},$$
and

\[(4.34) \quad I^0_{d,m} = [0, \frac{2}{m}] \times [0, \frac{1}{m}]^{d-1}.
\]

It suffices to show that

\[(4.35) \quad \frac{1}{N^2 \lambda(I^0_d)} \sum_{i=1}^{N} \int_{I_d} q_i(x)(1-q_i(x))dx = \frac{1}{N^2 \lambda(I^0_0)} \sum_{i=1}^{N} \int_{I^0_d} q_i(x)(1-q_i(x))dx.
\]

We only consider \(N = 2\) in (4.35), this is because we choose \(K = I_d\) and \(K = I^0_d\) in (4.2) respectively. This means \(I_d, I^0_d\) are divided into two equal volume parts respectively.

Let

\[(4.36) \quad x_i - a_i = t_i, 1 \leq i \leq d.
\]

According to (4.3) and plugging (4.36) into the left side of (4.35), the desired result is obtained.

From (4.2) and let \(K = [0, 1]^d\), we have

\[(4.37) \quad \mathbb{E}L^2_2(P_{\Omega^*_{\sim}}) - \mathbb{E}L^2_2(P_{\Omega^*_i})
\]

\[= \frac{1}{N^2} \sum_{i=1}^{N} \int_{[0,1]^d} \tilde{q}_i(x)(1-\tilde{q}_i(x))dx - \frac{1}{N^2} \sum_{i=1}^{N} \int_{[0,1]^d} \tilde{q}_i(x)(1-\tilde{q}_i(x))dx,
\]

where

\[\tilde{q}_i(x) = \frac{\lambda(\Omega^*_{i,\sim} \cap [0,x])}{\lambda(\Omega^*_{i,\sim})}, \tilde{q}_i(x) = \frac{\lambda(\Omega^*_i \cap [0,x])}{\lambda(\Omega^*_i)}; \quad i = 1, 2,
\]

and

\[\tilde{q}_i(x) = \tilde{q}_i(x) = \frac{\lambda(Q_i \cap [0,x])}{\lambda(Q_i)}, \quad i = 3, 4, \ldots, N.
\]

Let \(I^0_{d,m} = \{\Omega^*_{1,\sim}, \Omega^*_{2,\sim}\}, I^0_{d,m} = \{\Omega^*_{1}, \Omega^*_{2}\}\) denote two different partitions of \(I^0_{d,m}\). It can easily be seen only \(I^0_{d,m}\) contributes to the difference between two expected \(L_2\)-discrepancies, thus
\[ \mathbb{E}L_2^2(P_{\Omega^*_{\sim}}) - \mathbb{E}L_2^2(P_{\Omega^*}) \]
\[ = \frac{1}{N^2} \sum_{i=1}^{2} \int_{I_{0,d,m}^i} (\tilde{q}_i(x) - \bar{q}_i(x))dx + \frac{1}{N^2} \sum_{i=1}^{2} \int_{I_{0,d,m}^i} (\tilde{q}_i^2(x) - \bar{q}_i^2(x))dx \]
\[ = \frac{1}{N} \sum_{i=1}^{2} \int_{I_{0,d,m}^i} (\lambda(\tilde{\Omega}_i \cap [0, x])) - \lambda(\Omega_i \cap [0, x]))dx \]
\[ + \frac{1}{N^3} \sum_{i=1}^{2} \int_{I_d}^\prime (\lambda^2(\tilde{\Omega}_i \cap [0, x])) - \lambda^2(\Omega_i \cap [0, x]))dx \]
\[ = \frac{1}{N} \sum_{i=1}^{2} \int_{I_{0,d,m}^i} (\lambda(\tilde{\Omega}_i \cap [0, x])) - \lambda(\Omega_i \cap [0, x]))dx \]
\[ + \frac{1}{N^3} \sum_{i=1}^{2} \int_{I_d}^\prime (\lambda^2(\tilde{\Omega}_i \cap [0, x])) - \lambda^2(\Omega_i \cap [0, x]))dx. \]

Furthermore, employing (4.2) again, we have
\[ \mathbb{E}(L_2^2(P_{\Omega^*_{\sim}})) - \mathbb{E}(L_2^2(P_{\Omega^*})) \]
\[ = \frac{1}{8} \sum_{i=1}^{2} \int_{I_d} q_i^\prime_{i,\sim}(x)(1 - q_i^\prime_{i,\sim}(x))dx - \frac{1}{8} \sum_{i=1}^{2} \int_{I_d} q_i^\prime_{i,\sim}(x)(1 - q_i^\prime_{i,\sim}(x))dx \]
\[ = \frac{1}{8} \sum_{i=1}^{2} \int_{I_d} (\lambda(\tilde{\Omega}_i^\prime_{i,\sim} \cap [0, x])) - \lambda(\Omega_i^\prime_{i,\sim} \cap [0, x]))dx \]
\[ + \frac{1}{8} \sum_{i=1}^{2} \int_{I_d} (\lambda^2(\tilde{\Omega}_i^\prime_{i,\sim} \cap [0, x])) - \lambda^2(\Omega_i^\prime_{i,\sim} \cap [0, x]))dx. \]

Combining with (4.38) and (4.39), we obtain
\[ (4.40) \quad \mathbb{E}L_2^2(P_{\Omega^*_{\sim}}) - \mathbb{E}L_2^2(P_{\Omega^*}) = \frac{1}{N^3} \cdot \left[ 8\mathbb{E}(L_2^2(P_{\Omega^*_{\sim}})) - 8\mathbb{E}(L_2^2(P_{\Omega^*})) \right]. \]

Combining with (4.20), (4.28) and (4.31), the proof is completed.

4.2. Proof of Theorem 3.1 Following the proof process of Lemma 4.1, we obtain Theorem 3.1.
4.3. **Proof of Theorem 3.5**. We only consider the case \( \arctan \frac{1}{2} \leq \theta \leq \frac{\pi}{2} \), the calculation of case \( 0 \leq \theta < \arctan \frac{1}{2} \) is similar to it. First, we have

\[
P_2(\theta) = -\frac{1}{24\tan \theta} + \frac{1}{120\tan^2 \theta} + \frac{1}{1440\tan^3 \theta}.
\]

Then from Lemma 4.1, we obtain

\[
(4.41) \quad \mathbb{E}L_2^2(P_{\Omega^*}) - \mathbb{E}L_2^2(P_{\Omega^*}) = \frac{1}{N^3} \cdot \frac{1}{3d-2} \cdot P_2(\theta),
\]

where

\[
P_{\Omega^*} = \{U_1, U_2, \ldots, U_N\},
\]

and

\[
P_{\Omega^*} = \{W_1, W_2, \ldots, W_N\},
\]
denote stratified samples under different partition models \( \Omega^* \) and \( \Omega^*_i \) respectively.

Now, for arbitrary test set \( R = [0, x) \subset [0, 1]^d \), we consider the following discrepancy function,

\[
(4.42) \quad \Delta_{\mathscr{D}}(x) = \frac{1}{N} \sum_{n=1}^{N} 1_R(W_n) - \lambda(R).
\]

For an equivolume partition \( \Omega = \{\Omega_1, \Omega_2, \ldots, \Omega_N\} \), we divide the test set \( R \) into two parts, one is the disjoint union of \( \Omega_i \) entirely contained by \( R \) and another is the union of remaining pieces which are the intersections of some \( \Omega_j \) and \( R \), i.e.,

\[
(4.43) \quad R = \bigcup_{i \in I_0} \Omega_i \cup \bigcup_{j \in J_0} (\Omega_j \cap R),
\]

where \( I_0, J_0 \) are two index-sets.

Let

\[
T = \bigcup_{j \in J_0} (\Omega_j \cap R),
\]

then from (4.42), we have

\[
(4.44) \quad \Delta_{\mathscr{D}}(x) = \frac{1}{N} \sum_{n=1}^{N} 1_R(W_n) - \lambda(R) = \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) - \lambda(T),
\]

where (4.44) is based on the fact discrepancy function equals 0 on \( \bigcup_{i \in I_0} \Omega_i \).

According to the definition of \( L_2 \)-discrepancy and (4.44), it follows that
\[ (4.45) \quad \mathbb{E}(L_2^2(D_N, P_{\Omega}^*)) = \mathbb{E}\left( \int_{[0,1]^d} \left| \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) - \lambda(T) \right|^2 dx \right). \]

Consider the whole sum in (4.45) as a random variable which is defined on a region \( P_{\Omega} \). Besides we set the probability measure be \( w \), then we have

\[ \mathbb{E}(L_2^2(D_N, P_{\Omega}^*)) = \int_{P_{\Omega}} \int_{[0,1]^d} \left| \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) - \lambda(T) \right|^2 dxdw \]

(4.46)

\[ = \int_{[0,1]^d} \int_{P_{\Omega}} \left| \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) - \lambda(T) \right|^2 dw dx. \]

It can easily be checked that,

\[ \mathbb{E}\left( \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) \right) = \int_{P_{\Omega}} \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) dw = \lambda(T). \]

Hence,

\[ (4.47) \quad \int_{P_{\Omega}} \left| \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) - \lambda(T) \right|^2 dw = \text{Var}\left( \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) \right). \]

Let \( \sigma_j^2 = \text{Var}(1_T(W_j)) \), then we have

\[ (4.48) \quad \Sigma^2 = \text{Var}\left( \sum_{n=1}^{N} 1_T(W_n) \right) = \sum_{n=1}^{N} \text{Var}(1_T(W_n)) = \sum_{j \in J_0} \sigma_j^2. \]

Hence, from (2.1), we get

\[ (4.49) \quad \Sigma^2 \leq d \cdot N^{1 - \frac{1}{d}}. \]

Therefore,

\[ (4.50) \quad \mathbb{E}L_2^2(P_{\Omega}^*) = \int_{[0,1]^d} \text{Var}\left( \frac{1}{N} \sum_{n=1}^{N} 1_T(W_n) \right) dx \leq \frac{d}{N^{1 + \frac{1}{d}}}. \]
Combining with (4.41) and (4.50), the desired result is proved.

5. Conclusion

We study expected $L_2$–discrepancy under a class of new convex equivolume partitions. First, the expected $L_2$–discrepancy under two partition models are compared. Second, the explicit expected $L_2$–discrepancy upper bounds under the new partition models are obtained. So the optimal partition model that minimizes expected $L_2$–discrepancy is found and an optimal expected $L_2$–discrepancy upper bound is given explicitly under a class of new convex equivolume partitions. In future, star discrepancy will be studied under a class of convex equal volume partitions, which will have more corresponding applications.

References

[1] A. G. M. Ahmed, H. Perrier and D. Coeurjolly, et al, Low-discrepancy blue noise sampling, ACM Trans. Graph., 35(2016), 1-13.
[2] C. Aistleitner and M. Hofer, Probabilistic discrepancy bound for Monte Carlo point sets, Math. Comp., 83(2014), 1373-1381.
[3] R. F. Bass and K. Gröchenig, Random sampling of multivariate trigonometric polynomials, SIAM J. Math. Anal., 36(2004), 773-795.
[4] R. F. Bass and K. Gröchenig, Random sampling of bandlimited functions, Israel J. Math., 177(2010), 1-28.
[5] J. Beck, Some upper bounds in the theory of irregularities of distribution, Acta Arith., 43(1984), 115-130.
[6] J. Beck, Irregularities of distribution. I, Acta Math., 159(1987), 1-49.
[7] G. Bianchi, R. J. Gardner and M. Kiderlen, Phase retrieval for characteristic functions of convex bodies and reconstruction from covariograms, J. Amer. Math. Soc., 24(2011), 293-343.
[8] E. J. Candès and T. Tao, Near-optimal signal recovery from random projections: universal encoding strategies? IEEE Trans. Inform. Theory, 52(2006), 5406–5425.
[9] C. Cervellera and M. Muselli, Deterministic design for neural network learning: An approach based on discrepancy, IEEE Trans. Neural Netw., 15(2004), 533-544.
[10] H. S. Chan, T. Zickler and Y. M. Lu, Monte Carlo non-local means: random sampling for large-scale image filtering, IEEE Trans. Image Process., 23(2014), 3711-3725.
[11] K. Chiu, P. Shirley and C. Wang, Multi-jittered sampling, Graphics Gems, 4(1994), 370-374.
[12] F. Cucker and S. Smale, On the mathematical foundations of learning, Bull. Amer. Math. Soc., 39(2002), 1-49.
[13] F. Cucker and D. X. Zhou, Learning theory: an approximation theory viewpoint, Cambridge University Press., 2007.
[14] J. Dick and F. Pillichshammer, Digital Nets and Sequences. Discrepancy Theory and quasi-Monte Carlo Integration, Cambridge University Press., 2010.
[15] J. Dick and F. Pillichshammer, On the mean square weighted $L_2$ discrepancy of randomized digital $(t, m, s)$-nets over $\mathbb{Z}_2$, Acta Arith., 117(2005), 371-403.
[16] L. L. Cristea, J. Dick and F. Pillichshammer, On the mean square weighted $L_2$ discrepancy of randomized digital nets in prime base, J. Complexity, 22(2006), 605-629.
[17] J. Dick and F. Pillichshammer, Discrepancy theory and quasi-Monte Carlo integration. In: W. W. L. Chen, A. Srivastav, G. Travaglini(eds.), Panorama in Discrepancy Theory, Springer Verlag, Cham, 2014, 539-619.
[18] J. Dick, A. Hinrichs and F. Pillichshammer, A note on the periodic \( L_2 \)-discrepancy of Korobov’s \( p \)-sets, Arch. Math. (Basel), 115(2020), 67-78.
[19] B. Doerr, M. Gnewuch and A. Srivastav, Bounds and constructions for the star-discrepancy via \( \delta \)-covers, J. Complexity, 21(2005), 691-709.
[20] B. Doerr, A sharp discrepancy bound for jittered sampling, Math. Comp. (2022), http://doi.org/10.1090/mcom/3727.
[21] C. Doerr, M. Gnewuch and M. Wahlström, Calculation of Discrepancy Measures and Applications, In: W. Chen, A. Srivastav, G. Travaglini(eds), A Panorama of Discrepancy Theory, Springer, Heidelberg, 2107(2014), 621-678.
[22] H. Edelsbrunner and F. Pausinger, Approximation and convergence of the intrinsic volume, Adv. Math., 287(2016), 674-703.
[23] K. T. Fang, C. X. Ma and P. Winker, Centered \( L_2 \)-discrepancy of random sampling and Latin hypercube design, and construction of uniform designs, Math. Comp., 71(2002), 275-296.
[24] H. Führ and J. Xian, Relevant sampling in finitely generated shift-invariant spaces, J. Approx. Theory, 240(2019), 1-15.
[25] D. Frenkel, K. J. Schrenk and S. Martiniani, Monte Carlo sampling for stochastic weight functions, Proc. Natl. Acad. Sci., 114(2017), 6924-6929.
[26] R. J. Gardner, M. Kiderlen and P. Milanfar, Convergence of algorithms for reconstructing convex bodies and directional measures, Ann. Statist., 34(2006), 1331-1374.
[27] M. Gnewuch, Bracketing numbers for axis-parallel boxes and applications to geometric discrepancy, J. Complexity, 24(2008), 154-172.
[28] M. Gnewuch and N. Hebbinghaus, Discrepancy bounds for a class of negatively dependent random points including Latin hypercube samples, Ann. Appl. Probab., 31(2021), 1944-1965.
[29] M. Kiderlen and F. Pausinger, On a partition with a lower expected \( L_2 \)-discrepancy than classical jittered sampling, J. Complexity, 70(2022), https://doi.org/10.1016/j.jco.2021.101616.
[30] M. Kiderlen and F. Pausinger, Discrepancy of stratified samples from partitions of the unit cube, Monatsch. Math., 195(2021), 267-306.
[31] D. Krieg, Optimal Monte Carlo methods for \( L_2 \)-approximation, Constr. Appr., 49(2019), 385–403.
[32] Y. Lai, Monte Carlo and Quasi-Monte carlo methods and their applications, Ph.D. Dissertation, Department of Mathematics, Claremont Graduate University, California, USA, 1999.
[33] Y. Lai, Intermediate rank lattice rules and applications to finance, Appl. Numer. Math., 59(2009), 1-20.
[34] F. Liang, Dynamically weighted importance sampling in Monte Carlo computation, J. Am. Stat. Assoc., 97(2002), 807-821.
[35] M. D. McKay, W. J. Conover and R. J. Beckman, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21(1979), 239-245.
[36] H. Niederreiter, Random number generation and Quasi-Monte Carlo methods, SIAM, Philadelphia, 1992.
[37] F. Pausinger and A. M. Svane, A Koksma-Hlawka inequality for general discrepancy systems, J. Complexity, 31(2015), 773-797.
F. Pausinger and S. Steinerberger, On the discrepancy of jittered sampling, *J. Complexity*, 33(2016), 199-216.

M. Stein, Large sample properties of simulations using Latin hypercube sampling, *Technometrics*, 29(1987), 143-151.

A. W. van der Vaart and J. A. Wellner, Weak convergence and empirical processes, Springer Series in Statistics, Springer-Verlag, New York, With applications to statistics, 1996.

W. H. Wong and F. Liang, Dynamic weighting in Monte Carlo and optimization, *Proc. Natl. Acad. Sci.*, 94(1997), 14220–14224.

H. Zhu and J. Dick, Discrepancy estimates for acceptance-rejection samplers using stratified inputs, In: Cools R., Nuyens D. (eds) Monte Carlo and quasi-Monte Carlo Methods. Springer Proceedings in Mathematics and Statistics, vol 163, Springer, Cham, 2016.

**J. Xian, Department of Mathematics and Guangdong Province Key Laboratory of Computational Science, Sun Yat-sen University, 510275 Guangzhou, China.**

*Email address: xianjun@mail.sysu.edu.cn*

**X. Xu, Department of Mathematics, Sun Yat-sen University, 510275 Guangzhou, China.**

*Email address: xuxd26@mail2.sysu.edu.cn*