sciendo

uniform distribution theory

DOI: 10.2478/UDT-2023-0005
Unif. Distrib. Theory **18** (2023), no.1, 65–82

ON THE EXPECTED $\mathcal{L}_2$–DISCREPANCY OF JITTERED SAMPLING

Nathan Kirk — Florian Pausinger

School of Mathematics and Physics, Queen's University Belfast, Belfast, UNITED KINGDOM

ABSTRACT. For $m, d \in \mathbb{N}$, a jittered sample of $N = m^d$ points can be constructed by partitioning $[0,1]^d$ into m^d axis-aligned equivolume boxes and placing one point independently and uniformly at random inside each box. We utilise a formula for the expected $\mathcal{L}_2$–discrepancy of stratified samples stemming from general equivolume partitions of $[0,1]^d$ which recently appeared, to derive a closed form expression for the expected $\mathcal{L}_2$–discrepancy of a jittered point set for any $m, d \in \mathbb{N}$. As a second main result we derive a similar formula for the expected Hickernell $\mathcal{L}_2$–discrepancy of a jittered point set which also takes all projections of the point set to lower dimensional faces of the unit cube into account.

Communicated by Friedrich Pillichshammer

1. Introduction

1.1. Jittered sampling and discrepancy

For $m, d \in \mathbb{N}$, classical *jittered sampling* for $N = m^d$ combines the simplicity of grids with uniform random sampling by partitioning $[0,1]^d$ into m^d axis-aligned congruent cubes and placing a random point inside each of them; see Figure 1. Jittered sampling is sometimes referred to as 'stratified sampling' in the literature, but we will use the term 'stratified sampling' in a broader sense as in [15, 16]. Thus, in the following *stratified sampling* means that $[0,1]^d$

2020 Mathematics Subject Classification: 11K38, 05A18, 60C05.
Keywords: jittered sampling; star-discrepancy; $\mathcal{L}_2$-discrepancy; Hickernell $\mathcal{L}_2$-discrepancy.
The first author is supported by a UKRI Engineering and Physical Sciences Council PhD Studentship, grant number EP/N509541/1.

is partitioned into N subsets $\Omega_1, \ldots, \Omega_N$ and the ith point in $\mathcal{P}_{\mathbf{\Omega}}$ is chosen uniformly in the ith set of the partition (and stochastically independent of the other points), $i = 1, \ldots, N$. If all sets Ω_i have the same volume, the partition is called *equivolume*. If $N = m^d$ and the partition consists of the above mentioned axis-aligned congruent cubes, we obtain *jittered sampling* as a special case, i.e., jittered sampling is an example of an equivolume, stratified sampling scheme.

In the classical setting, discrepancy theory concerns itself with the study of the irregularity of distribution of point sets contained in the unit cube. There are numerous notions and variants of discrepancy with wide ranging applications in numerical integration, computer graphics, machine learning and option pricing in financial mathematics to name just a few [2, 18, 19, 22].

One such discrepancy measure is the so-called $\mathcal{L}_p$*−discrepancy* for $1 \leq p < \infty$; in this paper we will focus our attention primarily on the case $p = 2$ while briefly mentioning the star discrepancy in order to state the motivation for this note. For a point set $\mathcal{P}_N = \{\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}\}$ contained in $[0,1]^d$, we define the $\mathcal{L}_p$−discrepancy as

$$\mathcal{L}_p(\mathcal{P}_N) := \left(\int_{[0,1]^d} \left| \frac{\#\left(\mathcal{P}_N \cap [0, \mathbf{x})\right)}{N} - |[0, \mathbf{x})| \right|^p d\mathbf{x} \right)^{1/p}$$

in which $1 \leq p < \infty$, $\#\big(\mathcal{P}_N \cap [0, \mathbf{x})\big)$ counts the number of indices $0 \leq i < N$ such that $\mathbf{x}_i \in [0, \mathbf{x})$ and $|[0, \mathbf{x})|$ is the standard Lebesque measure of the subset $[0, \mathbf{x}) = \prod_{j=1}^{d} [0, x_j)$ with $\mathbf{x} = (x_1, \ldots, x_d)$. For an infinite sequence $\mathcal{X}$, we calculate the discrepancy of its first N elements; i.e., the $\mathcal{L}_p$−discrepancy of $\mathcal{X}$ is defined to be the $\mathcal{L}_p$−discrepancy of the first N terms of $\mathcal{X}$ for an arbitrary, but fixed N. Note that for ease of notation, we simply write $\mathcal{L}_p(\cdot)$ instead of $\mathcal{L}_{p,N}(\cdot)$ since we only work with finite point sets in this note.

As will be the case throughout this paper, when $\mathcal{P}_N$ is a set of random samples the *mean* p^{th} *power* $\mathcal{L}_p$*−discrepancy* $\mathbb{E}\mathcal{L}_p^p(\mathcal{P}_N)$ is often utilised as the discrepancy measure, where $\mathbb{E}$ denotes the probabilistic expectation. Usually we simply refer to this as the mean (or expected) $\mathcal{L}_p$−discrepancy. The $\mathcal{L}_p$−discrepancy as defined above is simply taking the $\mathcal{L}_p$ norm of the discrepancy function (the deviation of the measure of a test set $[0, \mathbf{x})$ from the fraction of points lying inside $[0, \mathbf{x})$). One can also take the $\mathcal{L}_\infty$ norm of the discrepancy function to create another measure of irregularity of distribution called the *star-discrepancy*, defined as

$$D_N^*(\mathcal{P}) = \sup_{\mathbf{x} \in [0,1]^d} \left| \frac{\#\left(\mathcal{P}_N \cap [0, \mathbf{x})\right)}{N} - |[0, \mathbf{x})| \right|.$$

The *Hickernell* $\mathcal{L}_2$*-discrepancy* was introduced by Hickernell in [8,9]; note that some ideas can be traced back to [26], while, to the best of our knowledge, the explicit definition appeared for the first time in [8, 9]. It considers not only the ordinary $\mathcal{L}_2-$discrepancy of a point set in the d-dimensional unit cube, but also the $\mathcal{L}_2-$discrepancy of all projections of the point set onto lower-dimensional faces of the d-dimensional unit cube.

For any nonempty subset $s \subseteq \{1 : d\}$ of coordinate indices, let $[0,1]^s$ denote the $|s|-$dimensional unit cube spanned by the coordinate axes in s and let $\mathbb{N}_m^s$ be the set of $|s|-$dimensional vectors with coordinates in s consisting only of entries from $\{1 : m\}$. Likewise, if $\mathbf{\Omega} = \{\Omega_1, \ldots, \Omega_N\}$ is a jittered partition of $[0,1]^d$ and $\mathcal{P}_{\mathbf{\Omega}}$ is the jittered $N-$element point set obtained from the partition $\mathbf{\Omega}$, then $\mathbf{\Omega}^s = \{\Omega_1^s, \ldots, \Omega_N^s\}$ and $\mathcal{P}_{\mathbf{\Omega}}^s$ denote the projection of the partition and jittered point set onto $[0,1]^s$ respectively. Similarly, $\mathbf{x}^s$ denotes the projection of a general vector $\mathbf{x} \in [0,1]^d$ into $[0,1]^s$. Then, the *Hickernell* $\mathcal{L}_2$*-discrepancy* of a point set $\mathcal{P} \subset [0,1]^d$ is given by

$$D_{H,2}(\mathcal{P}) := \left(\sum_{\emptyset \neq s \subseteq \{1:d\}} \int_{[0,1]^s} \left| \frac{\#(\mathcal{P}^s \cap [0,\mathbf{x}^s)}{N} - |[0,\mathbf{x}^s)| \right|^2 d\mathbf{x}^s \right)^{1/2}$$

or more concisely using the notation already established,

$$D_{H,2}(\mathcal{P}) := \left(\sum_{\emptyset \neq s \subseteq \{1:d\}} \mathcal{L}_2^2(\mathcal{P}^s) \right)^{1/2}, \tag{1}$$

where $\mathcal{P}^s$ denotes the projection of the point set $\mathcal{P}$ into the cube $[0,1]^s$.

In fact, Hickernell defined this discrepancy for general p: For $1 \leq p < \infty$, the *Hickernell* $\mathcal{L}_p$*−discrepancy* is defined as

$$D_{H,p}(\mathcal{P}) := \left(\sum_{\emptyset \neq s \subseteq \{1:d\}} \mathcal{L}_p^p(\mathcal{P}^s) \right)^{1/p}.$$

Constructions of deterministic point sets are widely used in the context of numerical integration via quasi-Monte Carlo (QMC) methods due to the low discrepancy value leading to improved order of the integration error. We refer the reader to the Koksma-Hlawka inequality [1, 2, 8–10, 13, 24]. The optimal order for the $\mathcal{L}_2-$discrepancy of a finite point set contained inside $[0,1]^d$ for $d \geq 2$ is

$$\mathcal{O}\left((\log N)^{\frac{d-1}{2}} / N \right),$$

conversely, for a set of i.i.d uniform random points (a Monte Carlo point set) the expected discrepancy has order $\mathcal{O}(N^{-1/2})$.

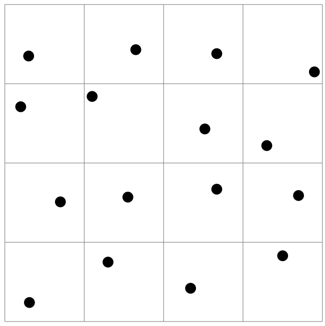

FIGURE 1. Jittered sampling for $d = 2$ and $m = 4$.

An element of randomness is often desirable or even necessary in real world simulations. Therefore to achieve the best of both worlds, it is of interest to select a deterministic point set and utilise a randomisation technique. The output of this process is aptly called, a *randomised quasi-Monte Carlo* (RQMC) point set; RQMC point sets often have the benefit of possessing better distribution properties than MC samples while also featuring an element of randomness, useful for simulation of real world phenomena. Jittered sampling is an example of an RQMC point set.

1.2. Exact formulas for the $\mathcal{L}_2$-discrepancy

Despite recurring criticism (see for example [20, 21]) the $\mathcal{L}_2$-discrepancy is a very popular notion of discrepancy mostly due to its simplicity. In contrast to other notions of discrepancy, we not only have an explicit formula for the $\mathcal{L}_2$-discrepancy of an arbitrary point set, known as Warnock's formula [25], but we also know the optimal order of the $\mathcal{L}_2$-discrepancy, i.e., we can look for point sets that are optimal with respect to $\mathcal{L}_2$-discrepancy; see [2, Section 3.2] and references therein. We remark that Warnock's formula can actually be traced back to an older paper of Koksma [12] in which a version for $d = 1$ is stated using the same idea as Warnock used later. In [5,7] a fast implementation of Warnock's formula for low dimensional point sets was presented and later thoroughly discussed in [4, 20]. The algorithms outperform a straightforward calculation, but the running time of the algorithms still grows exponentially with the dimension d. In addition, it should be noted that the direct calculation has to be done with sufficiently high numerical precision to avoid errors caused by annihilation [4, 7, 20].

Warnock's formula as presented in [2, Proposition 2.15] to calculate the $\mathcal{L}_2$-discrepancy of a given point set holds for any point set

$$\mathcal{P} = \{\mathbf{x}_0, \dots, \mathbf{x}_{N-1}\} \subseteq [0,1]^d.$$

We have

$$\mathcal{L}_2(\mathcal{P})^2 = \frac{1}{3^d} - \frac{2}{N}\sum_{n=0}^{N-1}\prod_{i=0}^{d}\frac{1-x_{n,i}^2}{2} + \frac{1}{N^2}\sum_{m,n=0}^{N-1}\prod_{i=0}^{d}\min(1-x_{m,i}, 1-x_{n,i}), \quad (2)$$

in which $x_{n,i}$ is the i-th component of the $\mathbf{x}_n$. For even powers p, there exists a generalisation of Warnock's formula for the $\mathcal{L}_p$-discrepancy; see [17].

Apart from this general formula, it is often possible to give explicit, closed formulas for the $\mathcal{L}_2$-discrepancy of particular deterministic sequences. As an example we mention the exact formula for the symmetrised Hammersley point set as derived in [14].

1.3. Main aim and outline

To put our results in a wider context, we recall that $(0, m, d)$-nets in base b denote an important type of deterministic d-dimensional point sets with b^m points which are widely studied and used [2, 18]. In [9] Hickernell derived an explicit formula for the expected $\mathcal{L}_2$-discrepancy of *randomised* $(0, m, d)$-nets. Moreover, B. Doerr [3] recently proved a tight star discrepancy estimate for jittered sampling. That is, the order of magnitude for the star discrepancy of a jittered point set is

$$\mathbb{E}D_N^* \in \Theta\left(\frac{\sqrt{d}\sqrt{1+\log(N/d)}}{N^{\frac{1}{2}+\frac{1}{2d}}}\right).$$

for all m and d with $m \geq d$. However, to the best of our knowledge, there is no such statement regarding an estimate or exact formula for the expected $\mathcal{L}_2-$discrepancy of a jittered sample for arbitrary dimension d and number of points $N = m^d$. Our main result is as follows:

Theorem 1.1. *Let* $\mathbf{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^d\}$ *be a jittered partition of* $[0,1]^d$ *for* $m \geq 2$. *Then*

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{m^{2d}}\left[\left(\frac{m}{2}\right)^d - \left(\frac{m}{2} - \frac{1}{6}\right)^d\right]. \quad (3)$$

Remark 1. In Proposition 5.1 we derive a similar formula for the $\mathcal{L}_2$-discrepancy of projections of jittered point samples in $[0,1]^d$ to lower dimensional faces of the unit cube.

As a second main result, we derive a closed formula for the Hickernell $\mathcal{L}_2$-discrepancy of jittered sampling:

Theorem 1.2. *Let* $\boldsymbol{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^d\}$ *be a jittered partition of* $[0,1]^d$ *for* $m \geq 2$. *Then*

$$\mathbb{E}D_{H,2}^2(\mathcal{P}_{\boldsymbol{\Omega}}) = \sum_{j=1}^{d} \frac{1}{m^{d+j}} \binom{d}{j} \left[\left(\frac{m}{2}\right)^j - \left(\frac{m}{2} - \frac{1}{6}\right)^j \right].$$

Based on Theorem 1.1 we can derive the following corollary. Note that in contrast to the result of Doerr, the implicit constants hidden in the Θ notation actually depend on d, hence this result is not as strong as the result of Doerr.

Corollary 1.3. *For a jittered partition* $\boldsymbol{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^d\}$ *of* $[0,1]^d$ *containing* $N = m^d$ *points where* $m \geq 2$ *and* d *is fixed, we have*

$$\begin{aligned} \left(\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\boldsymbol{\Omega}})\right)^{1/2} &= \Theta\left(\frac{1}{m^{\frac{d}{2}+\frac{1}{2}}}\right) \\ &= \Theta\left(\frac{1}{N^{\frac{1}{2}+\frac{1}{2d}}}\right), \\ \left(\mathbb{E}D_{H,2}^2(\mathcal{P}_{\boldsymbol{\Omega}})\right)^{1/2} &= \Theta\left(\frac{1}{N^{\frac{1}{2}+\frac{1}{2d}}}\right). \end{aligned}$$

Proof. To show the first assertion, we rewrite the difference in brackets in Theorem 1.1 as

$$\left(\frac{m-1}{2} + \frac{1}{2}\right)^d - \left(\frac{m-1}{2} + \frac{1}{3}\right)^d$$

and expand it using the binomial theorem:

$$\sum_{k=0}^{d} \binom{d}{k} \left(\frac{m-1}{2}\right)^k \left[\left(\frac{1}{2}\right)^{d-k} - \left(\frac{1}{3}\right)^{d-k} \right]$$

Each summand in this sum is positive. To get a lower bound, it is sufficient to truncate the sum and only consider the terms for $k = d$ and $k = d-1$, i.e.,

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\boldsymbol{\Omega}}) \geq \frac{1}{m^{2d}} \frac{d}{6} \left(\frac{m-1}{2}\right)^{d-1}.$$

Similarly, to obtain an upper bound, we simply notice that the leading terms cancel in the difference, such that we get a polynomial in m of degree $d-1$. Taking the square root and substituting $m = N^{1/d}$ leads to the desired result.

The second assertion can be derived in a similar fashion. □

Remark 2. By Jensen's inequality [11] we get that

$$\mathbb{E}\big(\mathcal{L}_2(\mathcal{P}_{\mathbf{\Omega}})\big) = O\left(\mathbb{E}\big(\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}})\big)^{1/2}\right)$$

and hence $\mathbb{E}\big(\mathcal{L}_2(\mathcal{P}_{\mathbf{\Omega}})\big) = O\left(m^{-\frac{d}{2}-\frac{1}{2}}\right)$.

Our proof utilises a proposition from [16] (see also [15]) regarding the expected discrepancy of stratified samples obtained from an equivolume partition of the cube.

Proposition 1.4 (Proposition 3, [16])**.** *If $\mathbf{\Omega} = \{\Omega_1, \ldots, \Omega_N\}$ is an equivolume partition of a compact convex set $K \subset \mathbb{R}^d$ with $|K| > 0$, then*

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{N^2|K|} \sum_{i=1}^{N} \int_K q_i(\mathbf{x})\big(1 - q_i(\mathbf{x})\big) d\mathbf{x} \tag{4}$$

with $q_i(\mathbf{x}) = \frac{|\Omega_i \cap [0,\mathbf{x})|}{|\Omega_i|}$.

The structure of the paper is as follows. In Section 2, we derive the formula for the discrepancy of jittered sets in dimension 2. This calculation is much simpler than for arbitrary $d \geq 3$ and is used to highlight some difficulties in the general case. In Section 3 we introduce several lemmas needed to derive the proof of Theorem 1.1 in Section 4. Section 5 contains the proof of Theorem 1.2.

2. The special case $d = 2$

For $m \geq 2$, let $\mathbb{N}_m^2$ denote the set of all ordered pairs with entries from $\{1 : m\}$ and let $\mathbf{\Omega} = \{\Omega_{(i,j)} : (i,j) \in \mathbb{N}_m^2\}$ be a jittered partition of $[0,1]^2$. We define two subsets of the unit square which will be of importance to us in the following derivation.

For a given vector $(i,j) \in \mathbb{N}_m^2$, define

$$I_v := \left\{\mathbf{x} = (x_1, x_2) \in [0,1]^2 : \frac{i-1}{m} \leq x_1 \leq \frac{i}{m} \text{ and } \frac{j}{m} \leq x_2 \leq 1\right\}$$

and similarly,

$$I_h := \left\{\mathbf{x} = (x_1, x_2) \in [0,1]^2 : \frac{i}{m} \leq x_1 \leq 1 \text{ and } \frac{j-1}{m} \leq x_2 \leq \frac{j}{m}\right\}.$$

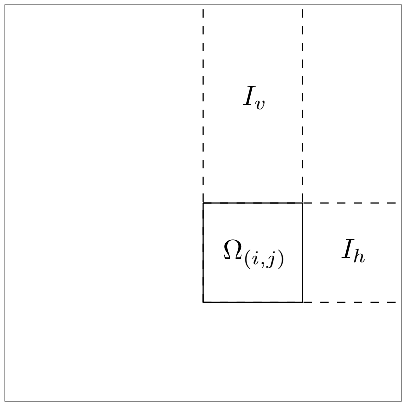

FIGURE 2. The regions I_h and I_v for a given partitioning set $\Omega_{(i,j)}$.

LEMMA 2.1. *Let* $\mathbf{\Omega} = \{\Omega_{(i,j)} : (i,j) \in \mathbb{N}_m^2\}$ *be a jittered partition of* $[0,1]^2$ *for* $m \geq 2$ *with* $q_{(i,j)}(\mathbf{x}) = \frac{|\Omega_{(i,j)} \cap [0,\mathbf{x})|}{|\Omega_{(i,j)}|}$. *Then for each* $(i,j) \in \mathbb{N}_m^2$,

$$q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right) \neq 0 \quad \textit{if and only if } \mathbf{x} \in \Omega_{(i,j)} \cup I_h \cup I_v.$$

P r o o f. We prove the contrapositive statement. If $\Omega_{(i,j)} \subset [0,\mathbf{x})$, then $\left|\Omega_{(i,j)} \cap [0,\mathbf{x})\right| = |\Omega_{(i,j)}|$. Hence,

$$q_{(i,j)} = 1 \quad \text{and} \quad q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right) = 0.$$

Similarly, if $\Omega_{(i,j)} \cap [0,\mathbf{x}) = \emptyset$, then $\left|\Omega_{(i,j)} \cap [0,\mathbf{x})\right| = 0$. Hence,

$$q_{(i,j)} = 0 \quad \text{and} \quad q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right) = 0.$$

Therefore, the integrand is non-zero if and only if $\Omega_{(i,j)}$ and $[0,\mathbf{x})$ have a non-trivial intersection and $\Omega_{(i,j)} \not\subseteq [0,\mathbf{x})$, i.e., when $\mathbf{x} \in \Omega_{(i,j)} \cup I_h \cup I_v$. □

We illustrate the regions for which $q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right) \neq 0$ in Figure 2.

LEMMA 2.2. *For given* $(i,j) \in \mathbb{N}_m^2$, *we have the following*

$$q_{(i,j)}(\mathbf{x}) = \begin{cases} m^2\left(x_1 - \frac{i-1}{m}\right)\left(x_2 - \frac{j-1}{m}\right) & \textit{for } \mathbf{x} \in \Omega_{(i,j)}, \\ m\left(x_2 - \frac{j-1}{m}\right) & \textit{for } \mathbf{x} \in I_h, \\ m\left(x_1 - \frac{i-1}{m}\right) & \textit{for } \mathbf{x} \in I_v. \end{cases} \tag{5}$$

P r o o f. We discuss the case when $\mathbf{x} \in \Omega_{(i,j)}$ and note that the other cases can be derived in a similar manner. If $\mathbf{x} \in \Omega_{(i,j)}$, then by definition $\frac{i-1}{m} \leq x_1 \leq \frac{i}{m}$ and $\frac{j-1}{m} \leq x_2 \leq \frac{j}{m}$. Then recall the definition of $q_{(i,j)}(\mathbf{x})$ from (4), and therefore

$$\begin{aligned} q_{(i,j)}(\mathbf{x}) &= \frac{\left|\Omega_{(i,j)} \cap [0,\mathbf{x})\right|}{\left|\Omega_{(i,j)}\right|} \\ &= \frac{\left(x_1 - \frac{i-1}{m}\right)\left(x_2 - \frac{j-1}{m}\right)}{1/m^2} \\ &= m^2\left(x_1 - \frac{i-1}{m}\right)\left(x_2 - \frac{j-1}{m}\right) \end{aligned}$$

as required. □

Theorem 2.3. *Let* $\mathbf{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^2\}$ *be a jittered partition of* $[0,1]^2$ *for* $m \geq 2$, *then*

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{6m-1}{36m^4}.$$

P r o o f. To calculate the mean $\mathcal{L}_2$−discrepancy of two dimensional jittered sampling, we start from equation (4). Let $K = [0,1]^2$, $N = m^2$ and assign a position vector (i,j) with $1 \leq i,j \leq m$ to the set of the partition with lower left vertex (i,j). Using this notation and Lemma 2.1 we obtain

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{m^{2d}} \sum_{(i,j)\in\mathbb{N}_m^2} \int_J q_{(i,j)}(\mathbf{x})\big(1 - q_{(i,j)}(\mathbf{x})\big)d\mathbf{x} \tag{6}$$

with

$$q_{(i,j)}(\mathbf{x}) = \frac{\left|\Omega_{(i,j)} \cap [0,\mathbf{x})\right|}{\left|\Omega_{(i,j)}\right|} \quad \text{and} \quad J = \Omega_{(i,j)} \cup I_h \cup I_v.$$

For a given set in position (i,j) we can evaluate the integral:

$$\begin{aligned} \int_J q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right)d\mathbf{x} \quad = \quad & \int_{\frac{i-1}{m}}^{\frac{i}{m}} \int_{\frac{j-1}{m}}^{\frac{j}{m}} q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right)dx_2x_1 \\ & + \int_{\frac{i}{m}}^{1} \int_{\frac{j-1}{m}}^{\frac{j}{m}} q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right)dx_2x_1 \\ & + \int_{\frac{i-1}{m}}^{\frac{i}{m}} \int_{\frac{j}{m}}^{1} q_{(i,j)}(\mathbf{x})\left(1 - q_{(i,j)}(\mathbf{x})\right)dx_2x_1. \end{aligned}$$

Evaluating the integrals and using (6) gives,

$$\begin{aligned}\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\boldsymbol{\Omega}}) &= \frac{1}{m^4}\sum_{\boldsymbol{i}\in\mathbb{N}_m^2}\frac{5}{36m^2}+\frac{m-i}{6m^2}+\frac{m-j}{6m^2}\\ &= \frac{1}{m^4}\sum_{i,j=1}^{m}\frac{5}{36m^2}+\frac{m-i}{6m^2}+\frac{m-j}{6m^2}\\ &= \frac{6m-1}{36m^4}.\end{aligned}$$

as required. □

Corollary 2.4. *For a jittered partition* $\boldsymbol{\Omega}$ *of* $[0,1]^2$*,*

$$\left(\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\boldsymbol{\Omega}})\right)^{1/2} = \Theta\left(N^{-3/4}\right).$$

Proof. This follows from $N = m^2$. □

Remark 3. Again, from Corollary 2.4 and Jensen's inequality, for a two dimensional jittered sample $\mathcal{P}_{\boldsymbol{\Omega}}$ we have

$$\mathbb{E}\mathcal{L}_2(\mathcal{P}_{\boldsymbol{\Omega}}) = \mathcal{O}\left(N^{-3/4}\right).$$

3. Several key lemmas

For $m \geq 2$, let $\mathbb{N}_m^d$ denote the set of all $d-$tuples with entries from $\{1:m\}$. Let $\boldsymbol{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} = (i_1,\ldots,i_d) \in \mathbb{N}_m^d\}$ denote a jittered partition of $[0,1]^d$ with $\boldsymbol{i} \in \mathbb{N}_m^d$ acting as the position vector of the jittered set $\Omega_{\boldsymbol{i}}$. If u is a nonempty and *strict* subset of $\{1:d\}$, then for every $\boldsymbol{i} \in \mathbb{N}_m^d$ we can define the set

$$\mathcal{I}_{\boldsymbol{i}}^u = \left\{\mathbf{x} \in [0,1]^d : \frac{i_j-1}{m} \leq x_j \leq \frac{i_j}{m} \text{ for } j \notin u, \frac{i_k}{m} \leq x_k \leq 1 \text{ for } k \in u\right\}. \quad (7)$$

Note that there are 2^d-2 such subsets for each $\boldsymbol{i} \in \mathbb{N}_m^d$. There are $d-1$ different types of subsets $\mathcal{I}_{\boldsymbol{i}}^u$ which can be classified by the cardinality of the set u, denoted throughout by the usual $|u|$. It is therefore natural to say that a set is of *type*$-|u|$. See Figure 2 for visual aid and examples of the regions $\mathcal{I}_{\boldsymbol{i}}^u$ for $d=2$. Specifically, I_v can be identified with $\mathcal{I}_{\boldsymbol{i}}^{\{2\}}$ and similarly I_h with $\mathcal{I}_{\boldsymbol{i}}^{\{1\}}$ in the two dimensional case.

Lemma 3.1. *Let $\mathbf{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^d\}$ be a jittered partition of $[0,1]^d$ for $m \geq 2$ with $q_{\boldsymbol{i}}(\mathbf{x}) = \frac{|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})|}{|\Omega_{\boldsymbol{i}}|}$. Then for each $\boldsymbol{i} \in \mathbb{N}_m^d$,*

$$q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big) \neq 0 \quad \textit{if and only if } \mathbf{x} \in \Omega_{\boldsymbol{i}} \cup \bigcup \mathcal{I}_{\boldsymbol{i}}^u$$

in which the union is over all subsets $\mathcal{I}_{\boldsymbol{i}}^u$ as defined in (7) for $\emptyset \neq u \subset \{1:d\}$.

P r o o f. It is more convenient to prove this via the contrapositive statement. That is,

$$q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big) = 0 \quad \text{if and only if } \mathbf{x} \notin \Omega_{\boldsymbol{i}} \cup \bigcup \mathcal{I}_{\boldsymbol{i}}^u.$$

Hence, given $\boldsymbol{i} \in \mathbb{N}_m^d$ suppose that $\mathbf{x} \notin \Omega_{\boldsymbol{i}} \cup \bigcup \mathcal{I}_{\boldsymbol{i}}^u$ for any nonempty choice of $u \subset \{1:d\}$. Then by definition, we are either in the situation that there exists at least one $j \in \{1:d\}$ such that $x_j \leq \frac{i_j - 1}{m}$ which implies $\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x}) = \emptyset$. Hence,

$$q_{\boldsymbol{i}}(\mathbf{x}) = \frac{|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})|}{|\Omega_{\boldsymbol{i}}|} = 0.$$

Or we have that $x_j \geq \frac{i_j}{m}$ for all $j \in \{1:d\}$ which implies $\Omega_{\boldsymbol{i}} \subseteq [0,\mathbf{x})$. In this case

$$q_{\boldsymbol{i}}(\mathbf{x}) = \frac{|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})|}{|\Omega_{\boldsymbol{i}}|} = \frac{|\Omega_{\boldsymbol{i}}|}{|\Omega_{\boldsymbol{i}}|} = 1.$$

In both cases, $q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big) = 0$ as required. In the other direction, suppose, $q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big) = 0$. Then either

$$q_{\boldsymbol{i}}(\mathbf{x}) = 0 \Rightarrow |\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})| = 0 \Rightarrow \Omega_{\boldsymbol{i}} \cap [0,\mathbf{x}) = \emptyset,$$

and therefore $\mathbf{x} \notin \Omega_{\boldsymbol{i}} \cup \bigcup \mathcal{I}_{\boldsymbol{i}}^u$. Alternatively if $q_{\boldsymbol{i}}(\mathbf{x}) = 1$, then

$$|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})| = |\Omega_{\boldsymbol{i}}| \Rightarrow \Omega_{\boldsymbol{i}} \subset [0,\mathbf{x}) \Rightarrow \mathbf{x} \notin \Omega_{\boldsymbol{i}} \cup \mathcal{I}_{\boldsymbol{i}}^u$$

for any nonempty $u \subset \{1:d\}$ since $\frac{i_j}{m} \leq x_j \leq 1$ for all $j \in \{1:d\}$. The statement of the Lemma is now proved. □

We can now adapt equation (4). Let $K = [0,1]^d$ and $N = m^d$, then

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{(m^d)^2} \sum_{\boldsymbol{i} \in \mathbb{N}_m^d} \int_{[0,1]^d} q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big) d\mathbf{x}, \tag{8}$$

where

$$q_{\boldsymbol{i}}(\mathbf{x}) = \frac{|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})|}{|\Omega_{\boldsymbol{i}}|}.$$

Next, we discard those regions of $[0,1]^d$ where the integrand vanishes as shown in Lemma 3.1. Therefore, we get,

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{m^{2d}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \left[\int_{\Omega_i} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} + \sum_{\emptyset\neq u\subset\{1:d\}} \int_{\mathcal{I}_i^u} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} \right]. \tag{9}$$

Lemma 3.2. *For a given* $\boldsymbol{i} \in \mathbb{N}_m^d$

$$q_{\boldsymbol{i}}(\mathbf{x}) = \begin{cases} m^d \prod_{j=1}^d \left(x_j - \frac{i_j-1}{m}\right), & \text{for} \quad \mathbf{x} \in \Omega_{\boldsymbol{i}} \\ m^{d-|u|} \prod_{j\notin u} \left(x_j - \frac{i_j-1}{m}\right), & \text{for} \quad \mathbf{x} \in \mathcal{I}_{\boldsymbol{i}}^u. \end{cases} \tag{10}$$

Proof. Fix $\boldsymbol{i} = (i_1,\dots,i_d) \in \mathbb{N}_m^d$. Both cases can be proven simultaneously by noting that $\mathcal{I}_{\boldsymbol{i}}^{\emptyset} = \Omega_{\boldsymbol{i}}$. Suppose that $\mathbf{x} \in \mathcal{I}_{\boldsymbol{i}}^u$ for any $u \subset \{1:d\}$, then by definition

$$\frac{i_j-1}{m} \le x_j \le \frac{i_j}{m} \quad \text{for } j\notin u \quad \text{and} \quad \frac{i_k}{m} \le x_k \le 1 \quad \text{for } k \in u.$$

Now to determine the desired expression for the function $q_{\boldsymbol{i}}(\mathbf{x})$, we are interested in the intersection between the jittered set $\Omega_{\boldsymbol{i}}$ and the test box $[0,\mathbf{x})$. From the observation above, we can conclude that

$$|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})| = \frac{1}{m^{|u|}} \prod_{j\notin u} \left(x_j - \frac{i_j-1}{m}\right).$$

Hence,

$$q_{\boldsymbol{i}}(\mathbf{x}) = \frac{|\Omega_{\boldsymbol{i}} \cap [0,\mathbf{x})|}{|\Omega_{\boldsymbol{i}}|} = \frac{\frac{1}{m^{|u|}} \prod_{j\notin u}\left(x_j - \frac{i_j-1}{m}\right)}{1/m^d} = m^{d-|u|} \prod_{j\notin u} \left(x_j - \frac{i_j-1}{m}\right)$$

as required. For completeness, when $u=\emptyset$ (i.e., $I_{\boldsymbol{i}}^{\emptyset} = \Omega_{\boldsymbol{i}}$) the expression becomes

$$q_{\boldsymbol{i}}(\mathbf{x}) = m^d \prod_{j=1}^d \left(x_j - \frac{i_j-1}{m}\right)$$

as required. □

Lemma 3.3. *Let* $\boldsymbol{i} \in \mathbb{N}_m^d$ *and let* u *be a nonempty, strict subset of* $\{1:d\}$*. Then*

$$\int_{I_i^u} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} = \left(\frac{3^{d-|u|}-2^{d-|u|}}{(6m)^{d-|u|}}\right) \prod_{j\in u}\left(1-\frac{i_j}{m}\right).$$

Moreover,

$$\int_{\Omega_i} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} = \frac{3^d-2^d}{(6m)^d}.$$

P r o o f. For a given $\boldsymbol{i} \in \mathbb{N}_m^d$, consider a type$-|u|$ subset of $[0,1]^d$. We can suppose that $u = \{1, 2, \ldots, |u|\} \subset \{1 : d\}$ without loss of generality. Then using the general form of $q_{\boldsymbol{i}}(\mathbf{x})$, the second integral in (9) has the form

$$\int_{I_i^u} q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} = \int_{I_i^u} q_{\boldsymbol{i}}(\mathbf{x}) - q_{\boldsymbol{i}}^2(\mathbf{x})d\mathbf{x}$$

$$= m^{d-|u|} \int_{I_i^u} \prod_{j \notin u} \left(x_j - \frac{i_j - 1}{m} \right) d\mathbf{x} - m^{2(d-|u|)} \int_{I_i^u} \prod_{j \notin u} \left(x_j - \frac{i_j - 1}{m} \right)^2 d\mathbf{x}$$

$$= m^{d-|u|} \int_{\frac{i_1}{m}}^1 \cdots \int_{\frac{i_{|u|}}{m}}^1 \prod_{j=|u|+1}^{d} \int_{\frac{i_j - 1}{m}}^{\frac{i_j}{m}} \left(x_j - \frac{i_j - 1}{m} \right) dx_j dx_{|u|} \ldots x_1$$

$$- m^{2(d-|u|)} \int_{\frac{i_1}{m}}^1 \cdots \int_{\frac{i_{|u|}}{m}}^1 \prod_{j=|u|+1}^{d} \int_{\frac{i_j - 1}{m}}^{\frac{i_j}{m}} \left(x_j - \frac{i_j - 1}{m} \right)^2 dx_j dx_{|u|} \ldots x_1$$

$$= m^{d-|u|} \left(\frac{1}{2m^2} \right)^{d-|u|} \left(1 - \frac{i_{|u|}}{m} \right) \cdots \left(1 - \frac{i_1}{m} \right)$$

$$- m^{2(d-|u|)} \left(\frac{1}{3m^3} \right)^{d-|u|} \left(1 - \frac{i_{|u|}}{m} \right) \cdots \left(1 - \frac{i_1}{m} \right)$$

$$= \frac{1}{(2m)^{d-|u|}} \prod_{j \in u} \left(1 - \frac{i_j}{m} \right) - \frac{1}{(3m)^{d-|u|}} \prod_{j \in u} \left(1 - \frac{i_j}{m} \right)$$

$$= \left(\frac{1}{(2m)^{d-|u|}} - \frac{1}{(3m)^{d-|u|}} \right) \prod_{j \in u} \left(1 - \frac{i_j}{m} \right)$$

$$= \left(\frac{3^{d-|u|} - 2^{d-|u|}}{(6m)^{d-|u|}} \right) \prod_{j \in u} \left(1 - \frac{i_j}{m} \right)$$

as required.

To show the second statement in the lemma we use the formula of Lemma 3.2 for $q_{\boldsymbol{i}}(\mathbf{x})$ and $\mathbf{x} \in \Omega_{\boldsymbol{i}}$. We notice that the calculation is similar to the above resulting in

$$\int_{\Omega_i} q_{\boldsymbol{i}}(\mathbf{x})\big(1 - q_{\boldsymbol{i}}(\mathbf{x})\big)d\mathbf{x} = \frac{3^d - 2^d}{(6m)^d}. \qquad \square$$

4. Proof of Theorem 1.1

Starting from (9), we use Lemma 3.3 to rewrite

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{m^{2d}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \left[\int_{\Omega_{\boldsymbol{i}}} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big) d\mathbf{x} + \sum_{\emptyset\neq u\subset\{1:d\}} \int_{\mathcal{I}_{\boldsymbol{i}}^u} q_{\boldsymbol{i}}(\mathbf{x})\big(1-q_{\boldsymbol{i}}(\mathbf{x})\big) d\mathbf{x} \right]$$

as

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) = \frac{1}{m^{2d}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \left[\frac{3^d-2^d}{(6m)^d} + \sum_{\emptyset\neq u\subset\{1:d\}} \left(\frac{3^{d-|u|}-2^{d-|u|}}{(6m)^{d-|u|}} \right) \prod_{j\in u} \left(1-\frac{i_j}{m}\right) \right].$$

We gather the terms in the last line by including the case $u = \emptyset$ in the summation over subsets of $\{1 : d\}$.

$$\begin{aligned} \mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) &= \frac{1}{m^{2d}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \left[\sum_{u\subset\{1:d\}} \left(\frac{3^{d-|u|}-2^{d-|u|}}{(6m)^{d-|u|}} \right) \prod_{j\in u} \left(1-\frac{i_j}{m}\right) \right] \\ &= \frac{1}{m^{2d}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \sum_{k=0}^{d} \sum_{\substack{u\subset\{1:d\}\\|u|=k}} \frac{3^{d-k}-2^{d-k}}{(6m)^{d-k}} \prod_{j\in u} \left(1-\frac{i_j}{m}\right) \\ &= \frac{1}{m^{2d}} \sum_{k=0}^{d} \frac{3^{d-k}-2^{d-k}}{(6m)^{d-k}} \sum_{\substack{u\subset\{1:d\}\\|u|=k}} \underbrace{\sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \prod_{j\in u} \left(1-\frac{i_j}{m}\right)}_{(*)}. \qquad (11) \end{aligned}$$

The quantity $(*)$ is equal to the same value for all subsets u with $|u| = k > 0$. Hence we calculate the double sum in (11) by letting $u = \{1, 2, \dots, k\}$, setting $j_l = m - i_l$ for $1 \le l \le k$ and noting that there are $\binom{d}{k}$ subsets of $\{1 : d\}$ of size k. Thus,

$$\begin{aligned} &\sum_{\substack{u\subset\{1:d\}\\|u|=k}} \sum_{\boldsymbol{i}\in\mathbb{N}_m^d} \prod_{j\in u} \left(1-\frac{i_j}{m}\right) = \sum_{i_{k+1},\dots,i_d=1}^{m} \binom{d}{k} \sum_{i_1,\dots,i_k=1}^{m} \prod_{j=1}^{k} \left(1-\frac{i_j}{m}\right) \\ &= m^{d-k} \binom{d}{k} \sum_{j_1,\dots,j_k=0}^{m-1} \left(\frac{j_1 j_2 \cdots j_k}{m^k} \right) \\ &= m^{d-k} \binom{d}{k} \cdot m^{-k} \left(\frac{m(m-1)}{2} \right)^k = m^{d-k} \binom{d}{k} \left(\frac{m-1}{2} \right)^k. \end{aligned}$$

Incorporating the last derivation into (11), we obtain

$$\begin{aligned}\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}) &= \frac{1}{m^{2d}}\sum_{k=0}^{d}\binom{d}{k}\left[\left(\frac{1}{2}\right)^{d-k} - \left(\frac{1}{3}\right)^{d-k}\right]\left(\frac{m-1}{2}\right)^{k} \\ &= \frac{1}{m^{2d}}\left[\left(\frac{m-1}{2}+\frac{1}{2}\right)^{d} - \left(\frac{m-1}{2}+\frac{1}{3}\right)^{d}\right]\end{aligned}$$

as required.

5. Proof of Theorem 1.2

For the Hickernell discrepancy, we are required to calculate the discrepancy of all projections of the point set in addition to that of the original set. Therefore, we first derive a formula for the expected $\mathcal{L}_2-$discrepancy of a $d-$dimensional jittered point set projected onto a lower dimensional face of the unit cube.

Proposition 5.1. *Let $\mathbf{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}_m^d\}$ be a jittered partition of $[0,1]^d$ for $m \geq 2$. For a nonempty subset $s \subseteq \{1 : d\}$, we have*

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}_{\mathbf{\Omega}}^s) = \frac{1}{m^{d+|s|}}\left[\left(\frac{m-1}{2}+\frac{1}{2}\right)^{|s|} - \left(\frac{m-1}{2}+\frac{1}{3}\right)^{|s|}\right]. \tag{12}$$

Proof. Suppose $\mathcal{P}_{\mathbf{\Omega}}$ is a jittered sample contained in $[0,1]^d$ with $N = m^d$ points. We follow the notation and general method as presented in [15] and consider the random variable

$$Z_{\mathbf{x}^s}(\mathcal{P}_{\mathbf{\Omega}}^s) = \frac{\#(\mathcal{P}_{\mathbf{\Omega}}^s \cap [0,\mathbf{x}^s))}{N},$$

where $\mathcal{P}_{\mathbf{\Omega}}^s$ denotes the projected jittered point set into the unit cube $[0,1]^s$. It is easy to observe that for some nonempty $s \subseteq \{1 : d\}$, the projection of a $d-$dimensional jittered partition onto $[0,1]^s$ is again a jittered partition (but with more than one point in each set) and hence equivolume. By Proposition 1 from [15], we can conclude that

$$\mathbb{E}(Z_{\mathbf{x}^s}(\mathcal{P}_{\mathbf{\Omega}}^s)) = |[0,\mathbf{x}^s)|. \tag{13}$$

From this unbiasedness of the variable $Z_{\mathbf{x}^s} = Z_{\mathbf{x}^s}(\mathcal{P}^s_{\mathbf{\Omega}})$ and applying Tonelli's theorem with $p = 2$, we have

$$\begin{aligned} \mathbb{E}\mathcal{L}_2^2(\mathcal{P}^s_{\mathbf{\Omega}}) &= \int_{[0,1]^s} \mathbb{E}\left|Z_{\mathbf{x}^s} - |[0,\mathbf{x}^s)|\right|^2 d\mathbf{x}^s \\ &= \int_{[0,1]^s} \mathbb{E}\left|Z_{\mathbf{x}^s} - \mathbb{E}Z_{\mathbf{x}^s}\right|^2 d\mathbf{x}^s = \int_{[0,1]^s} \mathbb{V}\mathrm{ar}(Z_{\mathbf{x}^s}) d\mathbf{x}^s \end{aligned} \tag{14}$$

with the last equality due to the expression $\mathbb{E}|Z_{\mathbf{x}^s} - \mathbb{E}Z_{\mathbf{x}^s}|^2$ which denotes the second central moment of $Z_{\mathbf{x}^s}$, i.e., the variance.

Now we take note that for a given projection of $\mathcal{P}_{\mathbf{\Omega}}$ into dimension $d-1$, the number of jittered sets is reduced by a factor of m with each set containing m points. Continuing in a similar manner, when projecting onto a dimension $1 \leq |s| < d$, there are $\frac{N}{m^{d-|s|}}$ jittered sets each containing $m^{d-|s|}$ points. Therefore, the random variable $N \cdot Z_{\mathbf{x}^s} = \#(\mathcal{P}^s_{\mathbf{\Omega}} \cap [0,\mathbf{x}^s))$ can be treated as the sum of $\frac{N}{m^{d-|s|}}$ groups of $m^{d-|s|}$ identical Bernoulli variables with success probabilities $\tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)$ for $\boldsymbol{i} \in \mathbb{N}^s_m$, where

$$\tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s) = \frac{|\Omega^s_{\boldsymbol{i}} \cap [0,\mathbf{x}^s)|}{|\Omega^s_{\boldsymbol{i}}|}.$$

Therefore,

$$N^2 \mathbb{V}\mathrm{ar}\big(Z_{\mathbf{x}^s}(\mathcal{P}^s_{\mathbf{\Omega}})\big) = \mathbb{V}\mathrm{ar}\big(N \cdot Z_{\mathbf{x}^s}(\mathcal{P}^s_{\mathbf{\Omega}})\big) = \sum_{\boldsymbol{i} \in \mathbb{N}^s_m} m^{d-|s|} \tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)\big(1 - \tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)\big).$$

Entering the last into (14) with $N = m^d$ yields

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}^s_{\mathbf{\Omega}}) = \frac{1}{m^{d+|s|}} \sum_{\boldsymbol{i} \in \mathbb{N}^s_m} \int_{[0,1]^s} \tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)\big(1 - \tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)\big) d\mathbf{x}^s. \tag{15}$$

For a given projection onto $[0,1]^s$ with $\emptyset \neq s \subseteq \{1:d\}$, the probabilities $\tilde{q}_{\boldsymbol{i}}(\mathbf{x}^s)$ have the same form as stated in Lemma 9 in the new appropriate dimension $|s|$. Hence (15) can be simplified in a similar manner to the standard $\mathcal{L}_2-$discrepancy calculation and can be written as in the statement of the proposition, i.e.,

$$\mathbb{E}\mathcal{L}_2^2(\mathcal{P}^s_{\mathbf{\Omega}}) = \frac{1}{m^{d+|s|}} \left[\left(\frac{m-1}{2} + \frac{1}{2}\right)^{|s|} - \left(\frac{m-1}{2} + \frac{1}{3}\right)^{|s|} \right]. \tag{16}$$

□

Proposition 5.1 is our main tool in the proof of Theorem 1.2.

P r o o f o f T h e o r e m 1.2. Let $\mathcal{P}_{\mathbf{\Omega}}$ be a jittered point set contained in $[0,1]^d$. We have that

$$\mathbb{E}D^2_{H,2}(\mathcal{P}_{\mathbf{\Omega}}) := \sum_{\emptyset \neq s \subseteq \{1:d\}} \mathbb{E}\mathcal{L}_2^2(\mathcal{P}^s_{\mathbf{\Omega}}).$$

From Proposition 5.1 we obtain,

$$\begin{aligned}\mathbb{E}D^2_{H,2}(\mathcal{P}_{\mathbf{\Omega}}) &= \sum_{\emptyset\neq s\subseteq\{1:d\}} \frac{1}{m^{d+|s|}}\left[\left(\frac{m-1}{2}+\frac{1}{2}\right)^{|s|}-\left(\frac{m-1}{2}+\frac{1}{3}\right)^{|s|}\right]\\ &= \sum_{j=1}^{d}\frac{1}{m^{d+j}}\binom{d}{j}\left[\left(\frac{m-1}{2}+\frac{1}{2}\right)^{j}-\left(\frac{m-1}{2}+\frac{1}{3}\right)^{j}\right].\end{aligned}$$

where the last equality is due to the fact that we have the uniform distribution inside each original jittered set, hence the projection onto any $|s|-$dimensional cube will have the same distribution. The binomial coefficient takes care of the fact that we have $\binom{d}{j}$ many projections for a given dimension $1 \leq j \leq d$. □

Remark 4. Note that in analogy to Theorem 2.3 and for $\mathbf{\Omega} = \{\Omega_{\boldsymbol{i}} : \boldsymbol{i} \in \mathbb{N}^2_m\}$ being a jittered partition of $[0,1]^2$ and $m \geq 2$, we have that

$$\mathbb{E}D^2_{H,2}(\mathcal{P}_{\mathbf{\Omega}}) = \frac{18m-1}{36m^4}.$$

Acknowledgement. We would like to thank Markus Kiderlen for fruitful discussions and help with the simplification of (11). In addition, we would like to thank the anonymous referee for their constructive feedback and for bringing several references to our attention.

REFERENCES

[1] AISTLEITNER C.—PAUSINGER, F.—TICHY, R. F.—SVANE, A. M.: *On functions of bounded variation*, Math. Proc. Cambridge Philos. Soc. **162** (2017), 405–418.

[2] DICK, J.—PILLICHSHAMMER, F.: *Digital Nets and Sequences*. Cambridge University Press, Cambridge, 2010.

[3] DOERR, B.: *A sharp discrepancy bound for jittered sampling*,Math. Comp. **91** (2022), 1871–1892.

[4] DOERR, C.—GNEWUCH, M.—WAHLSTRÖM, M.: *Calculation of Discrepancy Measures and Applications* In: *Panorama of Discrepancy Theory*, (W. W. L. Chen, A. Srivastav, G. Travaglini, eds.), Lecture Notes in Mathematics Vol. 2107, Springer-Verlag (2014). pp. 621–678.

[5] FRANK, K.—HEINRICH, S.: *Computing discrepancies of Smolyak quadrature rules,* J. Complexity **12** (1996), 287–314.

[6] GNEWUCH, M.—SRIVASTAV, A.—WINZEN C.: *Finding optimal volume subintervals with k points and calculating the star discrepancy are NP-hard problems*, J. Complexity **25** (2009), 115–127.

[7] HEINRICH, S.: *Efficient algorithms for computing the L2 discrepancy*, Math. Comput. **65** (1996), 1621–1633.

[8] HICKERNELL, F. J.: *A generalized discrepancy and quadrature error bound*, Math. Comp. **67** (1998), 299–322.

[9] HICKERNELL, F. J.: *The mean square discrepancy of randomized nets*, ACM Trans. on Modeling and Computer Simulation. **6** (1996), no. 4, 274–296.
[10] HLAWKA, E.: *Funkionen von Beschränkter Variation in der Theorie der Gleichverteilung*, Ann. Mat Pura Appl. **54** (1961), no. 4, 325–333.
[11] JENSEN, J. L. W. V.: *Sur les fonctions convexes et les inégalités entre les valeurs moyennes*, Acta Math. **30** (1906), no. 1, 175–193.
[12] KOKSMA, J. F.: *Some integrals in the theory of uniform distribution modulo 1*, Mathematica, B (Zutphen) **11** (1942), 49–52. (In Dutch)
[13] KOKSMA, J. F.: *A general theorem from the theory of uniform distribution modulo 1*, Mathematica, Zutphen B. **11** (1942), 7–11. (In Dutch)
[14] KRITZINGER, R.: *An exact formula for the $\mathcal{L}_2$-discrepancy of the symmetrized Hammersley point set*, Math. Comput. Simulation **143** (2018), 3–13.
[15] KIDERLEN, M.—PAUSINGER, F.: *Discrepancy of stratified samples from partitions of the unit cube*, Monatsh. Math. **195** (2022), 267–306.
[16] KIDERLEN, M., PAUSINGER, F.: *On a partition with a lower expected $\mathcal{L}_2-$discrepancy than classical jittered sampling*, J. Complexity **70** (2022), Article ID 101616, 13 pp.
[17] LEOBACHER, G.—PILLICHSHAMMER, F.: *Bounds for the weighted L_p-discrepancy and tractability of integration*, J. Complexity, **19** (2003), 539–547 .
[18] LEOBACHER, G.—PILLICHSHAMMER, F.: *Introduction to Quasi-Monte Carlo Integration and Applications*. Birkhäuser/springer, Cham, 2014.
[19] LEMIEUX, C.: *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Science + Business Media, Springer, New York, 2009.
[20] MATOUŠEK, J.: *On the L_2-discrepancy for anchored boxes*, J. Complexity **14** (1998), 527–556.
[21] MATOUŠEK, J.: *Geometric Discrepancy*. Springer-Verlag, Berlin, 1999.
[22] NIEDERREITER, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM, Philadelphia, 1992.
[23] PAUSINGER, F.—STEINERBERGER, S.: *On the discrepancy of jittered sampling*, J. Complexity **33** (2016), 199–216.
[24] PAUSINGER, F.—SVANE, A. M.: *A Koksma-Hlawka inequality for general discrepancy systems*, J. Complexity **31** (2015), 773–797.
[25] WARNOCK, T. T.: *Computational investigations of low discrepancy point sets*, In: Analysis, Proc. Sympos. Univ. Montreal 1971 (S. K. Zaremba, ed.), *Appl. Number Theory numer.* 1972, pp. 319–343.
[26] ZAREMBA, S. K.: *Some applications of multidimensional integration by parts*, Ann. Polon. Math. **21** (1968), 8–96.

Received September 16, 2022
Accepted February 27, 2023

Nathan Kirk
Florian Pausinger
School of Mathematics and Physics
Queen's University Belfast
University Road
BT7 1NN
UNITED KINGDOM
E-mail: nkirk09@qub.ac.uk
f.pausinger@qub.ac.uk