ÖAW AUSTRIAN ACADEMY OF SCIENCES

RICAM – JOHANN RADON INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

Computable error bounds for quasi-Monte Carlo using points with non-negative local discrepancy

M. Gnewuch, P. Kritzer, A.B. Owen, Z. Pan

RICAM-Report 2023-25

Computable error bounds for quasi-Monte Carlo using points with non-negative local discrepancy

Michael Gnewuch
University of Osnabrück

Peter Kritzer
Austrian Academy of Sciences

Art B. Owen
Stanford University

Zexin Pan
Stanford University

August 2023

Abstract

Let $f : [0,1]^d \to \mathbb{R}$ be a completely monotone integrand as defined by Aistleitner and Dick (2015) and let points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1} \in [0,1]^d$ have a non-negative local discrepancy (NNLD) everywhere in $[0,1]^d$. We show how to use these properties to get a non-asymptotic and computable upper bound for the integral of f over $[0,1]^d$. An analogous non-positive local discrepancy (NPLD) property provides a computable lower bound. It has been known since Gabai (1967) that the two dimensional Hammersley points in any base $b \geqslant 2$ have non-negative local discrepancy. Using the probabilistic notion of associated random variables, we generalize Gabai's finding to digital nets in any base $b \geqslant 2$ and any dimension $d \geqslant 1$ when the generator matrices are permutation matrices. We show that permutation matrices cannot attain the best values of the digital net quality parameter when $d \geqslant 3$. As a consequence the computable absolutely sure bounds we provide come with less accurate estimates than the usual digital net estimates do in high dimensions. We are also able to construct high dimensional rank one lattice rules that are NNLD. We show that those lattices do not have good discrepancy properties: any lattice rule with the NNLD property in dimension $d \geqslant 2$ either fails to be projection regular or has all its points on the main diagonal.

Keywords: Associated random variables, Digital nets, Rank one lattices

1 Introduction

Quasi-Monte Carlo (QMC) sampling [7, 26] can have much better asymptotic accuracy than plain Monte Carlo (MC), but it does not come with the usual statistical error estimates that MC has. Those estimates can be recovered by randomized QMC (RQMC) [21, 29] based on independent replicates of QMC. In this paper we consider an alternative approach to uncertainty quantification for

QMC. For some special sampling points with a non-negative local discrepancy (NNLD) property described later and a suitably monotone integrand f, we can compute upper and lower bounds on the integral μ of f over the unit cube in d dimensions. Methods based on random replication can provide confidence intervals for μ that attain a desired level such as 95% or 99% asymptotically, as the number of replicates diverges. The method we consider attains 100% coverage for finite n.

Unlike the well-known bounds derived via the Koksma-Hlawka inequality [19], these bounds can be computed by practical algorithms. Convex optimization [2] has the notion of a certificate: a computable bound on the minimum value of the objective function. The methods we present here provide certificates for multidimensional integration of a completely monotone function.

This improved uncertainty quantification comes at some cost. Our versions of the method will be more accurate than MC for dimensions $d \leqslant 3$, as accurate as MC (apart from logarithmic factors) for $d = 4$ and less accurate than MC for $d \geqslant 5$. They also require some special knowledge of the integrand.

The problem is trivial and the solution is well known for $d = 1$. If $f : [0, 1] \to \mathbb{R}$ is nondecreasing then

$$\frac{1}{n}\sum_{i=0}^{n-1} f\Big(\frac{i}{n}\Big) \leqslant \int_0^1 f(x)\,\mathrm{d}x \leqslant \frac{1}{n}\sum_{i=1}^{n} f\Big(\frac{i}{n}\Big). \tag{1}$$

These bracketing inequalities hold even if some of the quantities in them are $\pm\infty$. This works because f is nondecreasing, the evaluation points in the left hand side are 'biased low' and those in the right hand side are 'biased high'.

To get a multivariate version of (1), we generalize the notion of points biased low to points biased towards the origin in terms of a non-negative local discrepancy (NNLD) property of the points. This property was shown to hold for two dimensional Hammersley points by Gabai [12] in 1967. We couple the NNLD property with a multivariate notion of monotonicity called complete monotonicity [1].

This paper is organized as follows. Section 2 gives some notation and then defines the properties of point sets and functions that we need. Theorem 1 there establishes the bracketing property we need. Section 3 gives fundamental properties of NNLD point sets with an emphasis on projection regular point sets. Only very trivial lattice rules, confined to the diagonal in $[0, 1]^d$, can be both projection regular and NNLD. Cartesian products preserve the NNLD property as well as an analogous non-positive local discrepancy property. Section 4 compares our bounds to those obtainable from the Koksma-Hlawka inequality. Section 5 shows that digital nets whose generator matrices are permutation matrices produce NNLD point sets. Section 6 gives a construction of rank one lattice rules that are NNLD. We conclude with a discussion and some additional references in Section 7.

2 Definitions and a bound

Here we define a non-negative local discrepancy (NNLD) property of the points we use as well as a complete monotonicity criterion for the integrand. We then establish bounds analogous to (1). First we introduce some notation.

2.1 Notation

For integer $b \geqslant 1$, let $\mathbb{Z}_b = \{0, 1, \dots, b-1\}$. The set $\{1, 2, \dots, d\}$ of variable indices is denoted by $[d]$. For $u \subseteq [d]$, we use $|u|$ for the cardinality of u and $-u$ for the complement $[d] \setminus u$, especially in subscripts and superscripts. The singleton $\{j\}$ may be abbreviated to just j and $-\{j\}$ to $-j$. For points $\boldsymbol{x}, \boldsymbol{z} \in [0,1]^d$ and a set $u \subseteq [d] = \{1, 2, \dots, d\}$ let $\boldsymbol{x}_u{:}\boldsymbol{z}_{-u}$ be the hybrid point with j'th component x_j for $j \in u$ and j'th component z_j for $j \notin u$.

The points with all coordinates 0 or all coordinates 1 are denoted by $\mathbf{0}$ and $\mathbf{1}$ respectively. When it is necessary to specify their dimension we use $\mathbf{0}_d$ and $\mathbf{1}_d$. The notation $\mathbb{1}\{A\}$ is for an indicator variable equal to 1 when A is true and 0 otherwise.

For integer $d \geqslant 1$ we will use the following precedence notion on $[0,1]^d$. For $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^d$ we say that $\boldsymbol{x} \leqslant \boldsymbol{z}$ when $x_j \leqslant z_j$ holds for all $j = 1, \dots, d$.

2.2 Non-negative local discrepancy

A QMC rule is given by a list of points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1} \in [0,1]^d$ and it yields the estimate

$$\hat{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} f(\boldsymbol{x}_i)$$

of μ. We refer to these points as a point set, P_n, though in any setting where some $\boldsymbol{x}_i$ are duplicated we actually treat P_n as a multiset, counting multiplicity of the points. The local discrepancy of P_n at $\boldsymbol{z} \in [0,1]^d$ is given by

$$\delta(\boldsymbol{z}) = \delta(\boldsymbol{z}; P_n) = \widehat{\mathrm{VOL}}([\mathbf{0}, \boldsymbol{z})) - \mathrm{VOL}([\mathbf{0}, \boldsymbol{z}))$$

where VOL is Lebesgue measure and $\widehat{\mathrm{VOL}}$ is the empirical measure with

$$\widehat{\mathrm{VOL}}([\mathbf{0}, \boldsymbol{z})) = \frac{1}{n} \sum_{i=0}^{n-1} 1_{\boldsymbol{x}_i \in [\mathbf{0}, \boldsymbol{z})}.$$

That is, VOL is $\mathbb{U}[0,1]^d$ while $\widehat{\mathrm{VOL}}$ is $\mathbb{U}(P_n)$. The quantity $D_n^* = \sup_{\boldsymbol{z} \in [0,1]^d} |\delta(\boldsymbol{z})|$ is called the star discrepancy of the point set P_n.

Definition 1. The point set P_n with points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}$ has non-negative local discrepancy (NNLD) if

$$\delta(\boldsymbol{z}) \geqslant 0 \tag{2}$$

for all $\boldsymbol{z} \in [0,1]^d$.

A distribution for $\boldsymbol{x} \in \mathbb{R}^d$ is positively lower orthant dependent [32] if

$$\Pr(\boldsymbol{x} \leqslant \boldsymbol{z}) \geqslant \prod_{j=1}^{d} \Pr(x_j \leqslant z_j)$$

for all $\boldsymbol{z} \in \mathbb{R}^d$. A sufficient condition for NNLD is that the $\mathbb{U}(P_n)$ distribution on $[0,1]^d$ is positively lower orthant dependent and that the marginal distributions $\mathbb{U}\{x_{0,j}, \dots, x_{n-1,j}\}$ for each $j = 1, \dots, d$ are stochastically smaller than $\mathbb{U}[0,1]$. The random variable X is stochastically smaller than the random variable Y if $\Pr(X \leqslant z) \geqslant \Pr(Y \leqslant z)$ for all $z \in \mathbb{R}$ and in that case we also say that the distribution of X is stochastically smaller than that of Y. There is a related notion of positive upper orthant dependence as well as two related notions of negative orthant dependence, both upper and lower.

In one dimension, the points $0, 1/n, \dots, (n-1)/n$ are NNLD. As mentioned earlier, $n = b^m$ Hammersley points in base $b \geqslant 2$ and dimension $d = 2$ are NNLD [12]. Those Hammersley points are constructed as follows. For $0 \leqslant i < n$ write $i = \sum_{k=1}^{m} a_i(k) b^{k-1}$ for digits $a_i(k) \in \{0, 1, \dots, b-1\}$ and set $i' = \sum_{k=1}^{m} a_i(m-k+1) b^{k-1}$. Then the i'th such Hammersley point is $\boldsymbol{x}_i = \big(i/n, i'/n\big)$ for $i = 0, 1, \dots, n-1$. Some further properties of the Hammersley points, related to the work of [12], are given by [3].

We will also make use of a complementary property: non-positive local discrepancy.

Definition 2. The point set P_n with points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}$ has non-positive local discrepancy (NPLD) if

$$\delta(\boldsymbol{z}) \leqslant 0 \tag{3}$$

for all $\boldsymbol{z} \in [0,1]^d$.

One of our techniques is to take NNLD points $\boldsymbol{x}_i$ and reflect them to $\mathbf{1} - \boldsymbol{x}_i$ to get points that oversample rectangular regions near $\mathbf{1}$. In doing so we will need to take care of two issues. One is that for $d \geqslant 2$, the complement of a hyperrectangle $[\mathbf{0}, \boldsymbol{a})$ under this transformation is not another hyperrectangle. The other is that even for $d = 1$, the complement of a half open interval $[0, a)$ is a closed interval $[a, 1]$.

To handle these issues we make two observations below. First, for an n-point set $P_n \subset [0,1]^d$ let us additionally define the local discrepancy with respect to closed boxes:

$$\overline{\delta}(\boldsymbol{z}) = \overline{\delta}(\boldsymbol{z}; P_n) = \widehat{\mathrm{VOL}}([\mathbf{0}, \boldsymbol{z}]) - \mathrm{VOL}([\mathbf{0}, \boldsymbol{z}]).$$

Observation 1. *The point set P_n has the NNLD property if and only if*

$$\overline{\delta}(\boldsymbol{z}) \geqslant 0 \quad \textit{for all } \boldsymbol{z} \in [0,1]^d. \tag{4}$$

This is due to the following reasoning: First, we always have $\overline{\delta}(\boldsymbol{z}) \geqslant \delta(\boldsymbol{z})$ for all $\boldsymbol{z} \in [0,1]^d$. Thus the NNLD property of P_n implies (4)*. For the converse, we*

assume that P_n satisfies (4) and consider two cases. If $z_j = 0$ for some $j \in [d]$ then $\delta(\boldsymbol{z}) = 0$. If instead $\min_{j\in[d]} z_j > 0$ then

$$\delta(\boldsymbol{z}) = \lim_{\varepsilon\downarrow 0} \overline{\delta}(\boldsymbol{z} - \varepsilon\mathbf{1}).$$

Either way, (2) holds, i.e., P_n is NNLD.

Observation 2. *The condition*

$$\overline{\delta}(\boldsymbol{z}) \leqslant 0 \quad \text{for all } \boldsymbol{z} \in [0,1]^d \tag{5}$$

implies that P_n has the NPLD property, since $\delta(\boldsymbol{z}) \leqslant \overline{\delta}(\boldsymbol{z})$ for all $\boldsymbol{z} \in [0,1]^d$. As a partial converse, if $P_n \subset [0,1)^d \cup \{\mathbf{1}\}$, then the NPLD property also implies condition (5). Indeed, in that case we have $\overline{\delta}(\mathbf{1}) = 0$ and

$$\overline{\delta}(\boldsymbol{z}) = \lim_{\varepsilon\downarrow 0} \delta(\boldsymbol{z} + \varepsilon\mathbf{1}) \leqslant 0 \quad \text{for all } \boldsymbol{z} \in [0,1)^d.$$

Now consider for any $\boldsymbol{z} \in [0,1)^d$ and any $\varnothing \neq u \subsetneq [d]$ the closed anchored box $[\mathbf{0}, (\boldsymbol{z}_u{:}\mathbf{1}_{-u})]$. Due to $P_n \subset [0,1)^d \cup \{\mathbf{1}\}$, it contains exactly the same number of points from P_n as the anchored box $[\mathbf{0}, (\boldsymbol{z}_u{:}\boldsymbol{z}^*_{-u})]$, where $\boldsymbol{z}^*$ is defined by $z^*_j := \max(\{x_{0,j}, \ldots, x_{n-1,j}\} \setminus \{1\})$ for $j = 1, \ldots, d$ taking $z^*_j = 0$ in case it is $\max(\varnothing)$. Consequently, we have

$$\overline{\delta}(\boldsymbol{z}_u{:}\mathbf{1}_{-u}) \leqslant \overline{\delta}(\boldsymbol{z}_u{:}\boldsymbol{z}^*_{-u}) \leqslant 0.$$

Hence for $d = 1$ we have equivalence of (5) and NPLD for all $P_n \subset [0,1]$. But if $d \geqslant 2$, then for arbitrary $P_n \subset [0,1]^d$ not contained in $[0,1)^d \cup \{\mathbf{1}\}$ the NPLD property does not necessarily imply condition (5), as a trivial example with $d = 2$, $n = 1$, $P_n = \{(1, 1/2)\}$ shows: $\delta(\boldsymbol{z}) = -\mathrm{VOL}([\mathbf{0}, \boldsymbol{z})) \leqslant 0$ for all $\boldsymbol{z} \in [0,1]^d$, but $\overline{\delta}((1, 1/2)) = 1 - 1/2 = 1/2 > 0$.

For $d = 1$ if the points in $\tilde{P}_n$ are $1 - x_i$ for the points x_i of P_n, then

$$\overline{\delta}(z; P_n) + \delta(1 - z; \tilde{P}_n) = 0,$$

i.e., $\overline{\delta}(z; P_n) = -\delta(1-z; \tilde{P}_n)$ for all $z \in [0,1]$. Then due to Observations 1 and 2, reflections of NNLD points are NPLD points and vice versa for $d = 1$.

In addition to reflection, we consider another useful transformation. Let $\tilde{\boldsymbol{x}}_i$ be the base b Hammersley points for $i = 0, \ldots, n-1$ where $n = b^m$ and $d = 2$. Then [4] show that

$$\boldsymbol{x}_i = (1/n + \tilde{x}_{i,1}, 1 - \tilde{x}_{i,2}) \tag{6}$$

are NPLD.

2.3 Completely monotone functions

Here we define completely monotone functions, describing them in words before giving the formal definition. If $\boldsymbol{x} \leqslant \boldsymbol{z}$, then a completely monotone function can increase but not decrease if any x_j is replaced by z_j. That is $f(\boldsymbol{x}_{-j}{:}\boldsymbol{z}_j) - f(\boldsymbol{x}) \geqslant 0$ always holds. Next, the size of this difference can only be increasing as some other component x_k is increased to z_k, so certain differences of differences must also be non-negative. This condition must hold for anywhere from 1 to d applications of differencing. The $|u|$-fold differences of differences are alternating sums of the form

$$\Delta_u(\boldsymbol{x}, \boldsymbol{z}) = \sum_{v \subseteq u} (-1)^{|u-v|} f(\boldsymbol{x}_{-v}{:}\boldsymbol{z}_v).$$

Note that the coefficient of $f(\boldsymbol{x}_{-u}{:}\boldsymbol{z}_u)$ in $\Delta_u(\boldsymbol{x}, \boldsymbol{z})$ is positive.

Definition 3. The function $f : [0,1]^d \to \mathbb{R}$ is completely monotone if $\Delta_u(\boldsymbol{x}, \boldsymbol{z}) \geqslant 0$ for all non-empty u and all $\boldsymbol{x}, \boldsymbol{z} \in [0,1]^d$ with $\boldsymbol{x}_u \leqslant \boldsymbol{z}_u$.

In [1], Aistleitner and Dick use completely monotone functions to analyze the total variation of f in the sense of Hardy and Krause, denoted by $V_{\mathrm{HK}}(f)$. See [28] for an account. From Theorem 2 of [1], if $V_{\mathrm{HK}}(f) < \infty$ then we can write

$$f(\boldsymbol{x}) = f(\boldsymbol{0}) + f^+(\boldsymbol{x}) - f^-(\boldsymbol{x})$$

where f^+ and f^- are completely monotone functions with $f^+(\boldsymbol{0}) = f^-(\boldsymbol{0}) = 0$. They call $f^+ - f^-$ the Jordan decomposition of f. The functions $f^\pm$ are uniquely determined.

If f is right-continuous and $V_{\mathrm{HK}}(f) < \infty$ then $f(\boldsymbol{x}) = \nu([\boldsymbol{0}, \boldsymbol{x}])$ for a uniquely determined signed Borel measure ν, by Theorem 3 of [1]. Let this signed measure have Jordan decomposition $\nu = \nu^+ - \nu^-$ for ordinary (unsigned) Borel measures $\nu^\pm$. Then $f^\pm(\boldsymbol{x}) = \nu^\pm([\boldsymbol{0}, \boldsymbol{x}] \setminus \{\boldsymbol{0}\})$.

The completely monotone functions that we study take the form

$$f(\boldsymbol{x}) = f(\boldsymbol{0}) + \lambda\, \nu([\boldsymbol{0}, \boldsymbol{x}]) \tag{7}$$

where ν is an arbitrary probability measure on $[0,1]^d$ (or, more precisely, on the Borel σ-algebra of $[0,1]^d$) and $\lambda \geqslant 0$. Note that every right-continuous completely monotone function f on $[0,1]^d$ can be represented in that way, see, e.g., [10, II.5.11 Korrespondenzsatz, p. 67].

If ν is absolutely continuous with respect to the Lebesgue measure, then we may represent f, due to the Radon-Nikodym theorem, as

$$f(\boldsymbol{x}) = f(\boldsymbol{0}) + \lambda \int_{[\boldsymbol{0}, \boldsymbol{x}]} g(\boldsymbol{z})\, \mathrm{d}\boldsymbol{z} \tag{8}$$

where g is a probability density on $[0,1]^d$, i.e., a non-negative Lebesgue integrable function on $[0,1]^d$ with integral equal to one.

2.4 Basic result

Here we present the basic integration bounds. To bracket μ we use up to $2n$ function evaluations using n each for the lower and upper limits. For some constructions it is possible that some function evaluations might be usable in both limits, reducing the cost of computation. For $d = 1$ we only need $n + 1$ evaluations.

Theorem 1. *Let f be a completely monotone function of the form (7). Let $P_n = \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}\} \subset [0,1]^d$, and put $\widetilde{P}_n = \{\mathbf{1} - \boldsymbol{x}_0, \dots, \mathbf{1} - \boldsymbol{x}_{n-1}\}$.*

(i) Let $\widetilde{P}_n$ have non-negative local discrepancy. Then

$$\overline{\mu} = \hat{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} f(\boldsymbol{x}_i) \geqslant \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{9}$$

(ii) Let P_n have non-positive local discrepancy. If additionally either $P_n \subset [0,1)^d \cup \{\mathbf{1}\}$ or ν is absolutely continuous with respect to the Lebesgue measure, then

$$\underline{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{1} - \boldsymbol{x}_i) \leqslant \int_{[0,1]^d} f(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x}. \tag{10}$$

Proof. Without loss of generality take $f(\mathbf{0}) = 0$ and $\lambda = 1$. Consequently, $f(\boldsymbol{x}) = \nu([\mathbf{0}, \boldsymbol{x}])$ for all $\boldsymbol{x} \in [0,1]^d$. We obtain

$$\mu = \int_{[0,1]^d} \nu([\mathbf{0}, \boldsymbol{x}]) \, \mathrm{d}\boldsymbol{x} = \int_{[0,1]^d} \int_{[0,1]^d} 1_{\boldsymbol{z} \leqslant \boldsymbol{x}} \, \mathrm{d}\nu(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{x}.$$

Reversing the order of integration,

$$\mu = \int_{[0,1]^d} \int_{[0,1]^d} 1_{\boldsymbol{z} \leqslant \boldsymbol{x}} \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\nu(\boldsymbol{z}) = \int_{[0,1]^d} \mathrm{VOL}([\boldsymbol{z}, \mathbf{1}]) \, \mathrm{d}\nu(\boldsymbol{z}). \tag{11}$$

Similarly,

$$\hat{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} \nu([\mathbf{0}, \boldsymbol{x}_i]) = \frac{1}{n} \sum_{i=0}^{n-1} \int_{[0,1]^d} 1_{\boldsymbol{z} \leqslant \boldsymbol{x}_i} \, \mathrm{d}\nu(\boldsymbol{z})$$

from which

$$\hat{\mu} = \int_{[0,1]^d} \frac{1}{n} \sum_{i=0}^{n-1} 1_{\boldsymbol{z} \leqslant \boldsymbol{x}_i} \, \mathrm{d}\nu(\boldsymbol{z}) = \int_{[0,1]^d} \widehat{\mathrm{VOL}}([\boldsymbol{z}, \mathbf{1}]) \, \mathrm{d}\nu(\boldsymbol{z}). \tag{12}$$

Combining (11) and (12) the integration error now satisfies

$$\begin{aligned} \hat{\mu} - \mu &= \int_{[0,1]^d} \Big(\widehat{\mathrm{VOL}}([\boldsymbol{z}, \mathbf{1}]) - \mathrm{VOL}([\boldsymbol{z}, \mathbf{1}] \Big) \, \mathrm{d}\nu(\boldsymbol{z}) \\ &= \int_{[0,1]^d} \overline{\delta}(\mathbf{1} - \boldsymbol{z}; \widetilde{P}_n) \, \mathrm{d}\nu(\boldsymbol{z}), \end{aligned} \tag{13}$$

where $\overline{\delta}(\mathbf{1}-\boldsymbol{z};\widetilde{P}_n)$ is the local discrepancy of $\widetilde{P}_n$ with respect to the anchored closed box $[\mathbf{0},\mathbf{1}-\boldsymbol{z}]$. Recall that ν is a positive measure.

For part (i), let $\widetilde{P}_n$ have the NNLD property. Due to Observation 1 we have $\overline{\delta}(\mathbf{1}-\boldsymbol{z};\widetilde{P}_n) \geqslant 0$ for all $\boldsymbol{z} \in [0,1]^d$. Hence $\hat{\mu} \geqslant \mu$, establishing (9).

For part (ii), let $\widetilde{P}_n$ have the NPLD property. If additionally $\widetilde{P}_n \subset [0,1)^d \cup \{\mathbf{1}\}$, then Observation 2 ensures that $\overline{\delta}(\mathbf{1}-\boldsymbol{z};\widetilde{P}_n) \leqslant 0$ for all $\boldsymbol{z} \in [0,1]^d$, establishing $\hat{\mu} \leqslant \mu$. If instead ν is absolutely continuous with respect to the Lebesgue measure, then we can replace $\overline{\delta}(\mathbf{1}-\boldsymbol{z};\widetilde{P}_n)$ in (13) by $\delta(\mathbf{1}-\boldsymbol{z};\widetilde{P}_n)$ without changing the integral. Hence we get again $\hat{\mu} \leqslant \mu$. In any case, exchanging the roles of P_n and $\widetilde{P}_n$ establishes (10). □

Theorem 1 provides an upper bound for μ when sampling from reflected NNLD points. This bound will approach μ as $n \to \infty$ if those points also satisfy $D_n^* \to 0$ as $n \to \infty$. To get a lower bound we can use reflected NPLD points, provided that either ν is absolutely continuous or those points all belong to $[0,1)^d \cup \{\mathbf{1}\}$. The NPLD points could be those given by equation (6). We find in Section 5 that NPLD points are not as simple to construct as NNLD points.

2.5 Example

Here is a simple example to illustrate these bounds. The integrand is known to be completely monotone because it is a multivariate cumulative distribution function (CDF). For $\boldsymbol{x} \in [0,1]^2$ we take

$$f(\boldsymbol{x}) = \Pr(X_1 \leqslant x_1, X_2 \leqslant x_2) \tag{14}$$

for $\boldsymbol{X} \sim \mathcal{N}(0,\Sigma)$ with $\Sigma = \left(\begin{smallmatrix} 1 & \rho \\ \rho & 1 \end{smallmatrix}\right)$ using $\rho = 0.7$. Due to (9), we can compute an upper bound for $\mu = \int_{[0,1]^2} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ by sampling at points $\mathbf{1}-\boldsymbol{x}_i$ where $\boldsymbol{x}_i \in [0,1]^2$ are the first $n = 2^m$ Hammersley points in any base $b \geqslant 2$. We can compute a lower bound for μ by first transforming Hammersley points via (6) to get NPLD points $\boldsymbol{x}_i$ and then sampling at $\mathbf{1}-\boldsymbol{x}_i$. Note that the point sets in these bounds are not extensible in that the points for $n = b^m$ are not necessarily reused for $n = b^{m+1}$.

Figure 1 shows the results for $n = 2^m$ and $1 \leqslant m \leqslant 13$. Over the given range, $n(\overline{\mu}-\underline{\mu})$ increases with n while $n(\overline{\mu}-\underline{\mu})/\log(n)$ decreases with n. The computed upper and lower bounds for $n = 2^{13}$ show that

$$0.5618735 \leqslant \mu \leqslant 0.5619890.$$

This function is so smooth and the dimension is so small that comparable accuracy could be attained by standard low dimensional integration methods with many fewer function evaluations. However, these computations took approximately five seconds in R on a MacBook Air M2 laptop, using the `mvtnorm` package [13, 14] to compute f. A more efficient integration could save only about five seconds and it would not come with guaranteed bounds.

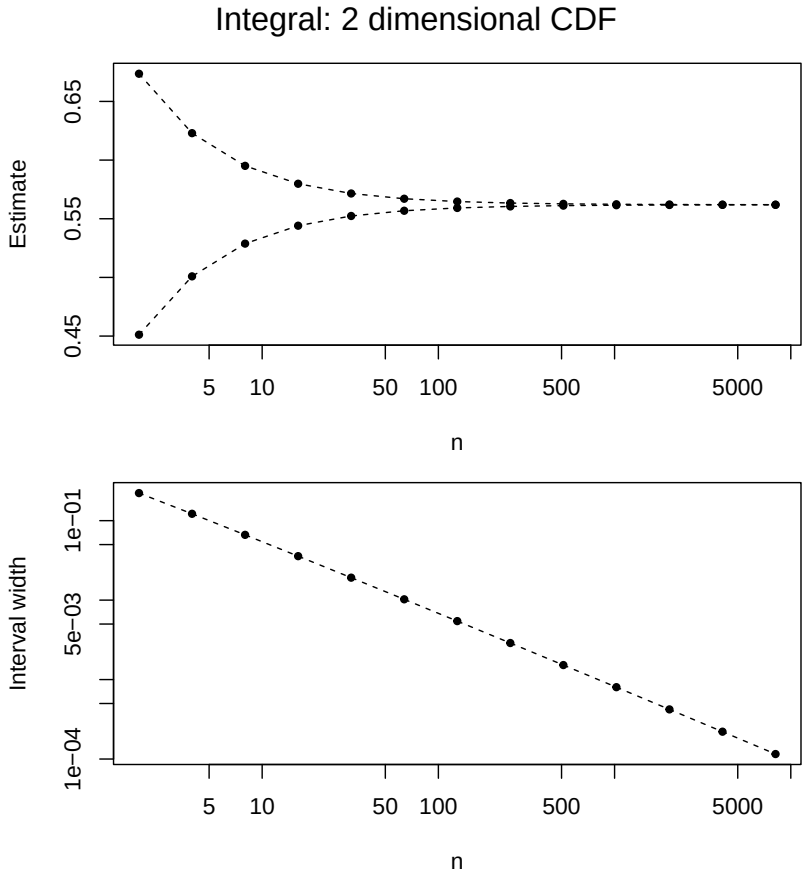

Figure 1: The top panel shows upper and lower bounds for $\mu = \int_{[0,1]^2} f(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}$ using transformations of the Hammersley points and $n = 2^m$ for $1 \leqslant m \leqslant 13$. The bottom panel plots the difference between those upper and lower bounds versus n, on a logarithmic scale.

3 More about NNLD points

Here we collect some observations about properties that any $n \geqslant 1$ NNLD points in $[0,1]^d$ must necessarily have. Then we use those properties to describe constraints that the NNLD property imposes on customary QMC constructions (lattices and digital nets). Finally we show that the NNLD and NPLD properties are preserved by tensor products.

The first and most obvious property of NNLD points is that $\mathbf{0}$ must be one of those points or else there is a box $B = [\mathbf{0}, \boldsymbol{a})$ with $0 = \widehat{\mathrm{VOL}}(B) < \mathrm{VOL}(B)$ so that $\delta(\boldsymbol{a}) < 0$. Next it must be true that all n points belong to $[0, 1-1/n]^d$. Suppose to the contrary that $x_{i1} > 1 - 1/n$ for some $0 \leqslant i < n$. Then for some

$\epsilon > 0$ there exists $B = [0, 1 - 1/n + \epsilon) \times [0,1]^{d-1}$ with $\widehat{\mathrm{VOL}}(B) \leqslant (n-1)/n < \mathrm{VOL}(B)$ so that $\boldsymbol{x}_i$ are not NNLD. The same argument applies if $x_{ij} > 1 - 1/n$ for any i and any j.

Trivial constructions of NNLD points have $\boldsymbol{x}_i = (i/n)\mathbf{1} \in [0,1]^d$ for $0 \leqslant i < n$. We observe that these points as well as the Hammersley points for $d = 2$ have variables that are positively correlated. We will use a general positive dependence property in Sections 5 and 6 to construct more NNLD point sets. The NPLD construction in (6) creates a negative lower orthant dependence property for the components of $\boldsymbol{x}_i \in [0,1]^2$.

Many of the constructions P_n we consider are projection regular by which we mean that the projections of P_n onto each single coordinate are equal to the full set $\{0, 1/n, 2/n, \ldots, (n-1)/n\}$. Projection regularity is usually considered advantageous in QMC, as it guarantees a certain structure and even distribution of the integration node set, and simplifies the derivation of error bounds. However, combined with the NNLD property, it imposes a constraint on the point set that we will use to rule out certain constructions.

Proposition 1. *Let P_n be a point set with n points in $[0,1)^d$ that is projection regular. If P_n has the NNLD property, then P_n must contain the point*

$$\boldsymbol{x}_* = \left(\frac{n-1}{n}, \frac{n-1}{n}, \ldots, \frac{n-1}{n}\right).$$

Proof. Suppose that P_n is projection regular and does not contain $\boldsymbol{x}_*$. Then there must exist at least one two dimensional projection Q_n of P_n which does not contain the point $\boldsymbol{y}_* := (\frac{n-1}{n}, \frac{n-1}{n})$. Without loss of generality, assume that Q_n is the projection of P_n onto the first and second coordinates.

This implies, due to projection regularity, that at least two points of Q_n do not lie in the box $[\mathbf{0}, \boldsymbol{y}_*)$. Thus,

$$\delta(\boldsymbol{y}_*) = \widehat{\mathrm{VOL}}([\mathbf{0}, \boldsymbol{y}_*)) - \mathrm{VOL}([\mathbf{0}, \boldsymbol{y}_*)) \leqslant \frac{n-2}{n} - \frac{(n-1)^2}{n^2} = -\frac{1}{n^2}.$$

Therefore, P_n has negative local discrepancy for the box $[\mathbf{0}, \boldsymbol{y}_*) \times [0,1)^{d-2}$. □

Proposition 1 has some consequences for well known QMC points. We will consider digital nets and integration lattices. The most widely used and studied integration lattices are rank one lattices. Given a generating vector $\boldsymbol{g} = (g_1, \ldots, g_d) \in \mathbb{N}^d$ and a sample size $n \geqslant 1$, a rank one lattice uses points

$$\boldsymbol{x}_i = \Big(\frac{g_1 i}{n}, \frac{g_2 i}{n}, \ldots, \frac{g_d i}{n}\Big) \bmod 1$$

for $0 \leqslant i < n$ where the modulus operation above takes the fractional part of its argument. These n points form a group under addition modulo 1. More general integration lattices having ranks between 1 and d can also be constructed [6, 26, 33]. Lattice rules with ranks larger than 1 are seldom used. They also have the group structure.

Corollary 1. *For fixed $d, n \geqslant 1$ there is only one projection regular lattice point set in $[0,1)^d$ that consists of n points and has the NNLD property, namely the lattice point set*

$$\left\{\mathbf{0}, \frac{1}{n}\mathbf{1}, \frac{2}{n}\mathbf{1}, \ldots, \frac{n-1}{n}\mathbf{1}\right\},$$

whose points all lie on the main diagonal of the d-dimensional unit cube $[0,1)^d$.

Proof. Let P_n be a projection regular lattice point set, consisting of n points in $[0,1)^d$, that has NNLD. Due to Proposition 1, P_n has to contain the point $\boldsymbol{x}_* = \frac{n-1}{n}\mathbf{1}$. Due to the additive group structure of P_n, we have

$$k\boldsymbol{x}_* \bmod 1 = \frac{n-k}{n}\mathbf{1} \in P_n \quad \text{for } k = 0, 1, \ldots, n-1.$$

The set above has n distinct points, so they must be all of P_n. □

From Corollary 1 we see, in particular, that the only projection regular rank one lattices that are NNLD are trivial, and equivalent to taking all $g_j = 1$. If we also consider lattices that are not projection regular, then we can find constructions that are NNLD and do not only consist of points on the main diagonal of the unit cube $[0,1)^d$. See Theorem 3.

Now we look at (t, m, d)-nets [7, 26]. The most widely used (t, m, d)-nets are those of Sobol' in base $b = 2$. Sobol' points require one to choose parameters known as direction numbers, with those of [20] being especially prominent. By considering the point $\boldsymbol{x}_* = \mathbf{1}(1 - 1/n)$, we often find that such Sobol' points cannot be NNLD. The first and third components of $\boldsymbol{x}_i \in [0,1]^d$ for $d \geqslant 3$ are projection regular but, for $2 \leqslant m \leqslant 20$ they fail to contain $(1 - 1/n, 1 - 1/n)$. Therefore the projection of the Sobol' points onto those two dimensions fails to be NNLD and hence the d dimensional point set is not NNLD either.

Like lattice point sets, digital (t, m, d)-nets in base $b \geqslant 2$ have a group structure; this time it is based on the digitwise addition modulo b, which is performed in each component separately. Using this group structure and Proposition 1, we obtain a corollary with a similar flavor to Corollary 1, although with less dramatic consequences.

Corollary 2. *Let $d, m \geqslant 1$ and $b \geqslant 2$. Let*

$$\alpha_{b,m} = \sum_{\nu=1}^{m} b^{-\nu} = \frac{1 - b^{-m}}{b - 1}.$$

On the one hand, any digital (t, m, d)-net in base $b \geqslant 2$ that is projection regular and has the NNLD property contains the cyclic subgroup

$$\{\mathbf{0}, \alpha_{b,m}\mathbf{1}, 2\alpha_{b,m}\mathbf{1}, \ldots, (b-1)\alpha_{b,m}\mathbf{1}\},$$

which consists of b points on the main diagonal.

On the other hand, any (t, m, d)-net in base $b \geqslant 2$ has at most $b^{t+\lceil \frac{m-t}{d} \rceil}$ points on the main diagonal.

Proof. Let $n = b^m$, and let P_n be a projection regular digital (t, m, d)-net, consisting of n points in $[0,1)^d$, that has NNLD. Due to Proposition 1, P_n has to contain the point $\boldsymbol{x}_* = \frac{n-1}{n}\mathbf{1} = (b-1)\alpha_{b,m}\mathbf{1}$. Using the specific commutative group addition of P_n, we see that adding up $\boldsymbol{x}_*$ k times yields

$$k\boldsymbol{x}_* = (b-k)\alpha_{b,m}\mathbf{1} \in P_n$$

for $k = 0, 1, \ldots, b-1$.

Now let P_n be an arbitrary (t, m, d)-net in base b. Put $k := \lceil \frac{m-t}{d} \rceil$. We may partition the half-open unit cube $[0,1)^d$ into b^{m-t} half-open axis-parallel boxes (of the same shape and of volume b^{t-m}) with side length b^{-k} and, possibly, side length b^{1-k}. Due to the net property, each of these boxes contains exactly b^t points of P_n, and at most b^k of the boxes have a non-trivial intersection with the main diagonal. □

The next result shows that Cartesian products of finitely many NNLD (or NPLD) point sets are also NNLD (respectively NPLD).

Lemma 1. *For positive integers d_1, d_2, n_1 and n_2, let $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{n_1-1} \in [0,1]^{d_1}$ and $\tilde{\boldsymbol{x}}_0, \ldots, \tilde{\boldsymbol{x}}_{n_2-1} \in [0,1]^{d_2}$ be NNLD point sets. Let $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{N-1} \in [0,1]^{d_1+d_2}$ for $N = n_1 n_2$ be the Cartesian product of those two point sets. Then $\boldsymbol{z}_0, \ldots, \boldsymbol{z}_{N-1}$ are NNLD points. If both $\boldsymbol{x}_i$ and $\tilde{\boldsymbol{x}}_i$ are NPLD then $\boldsymbol{z}_i$ are also NPLD.*

Proof. For any $\boldsymbol{z} \in [0,1]^{d_1+d_2}$ define $\boldsymbol{x} = \boldsymbol{z}_{[d_1]}$ and $\tilde{\boldsymbol{x}} = \boldsymbol{z}_{-[d_1]}$. Let VOL_1, VOL_2 and VOL denote Lebesgue measure on $[0,1]^{d_1}$, $[0,1]^{d_2}$ and $[0,1]^d$ for $d = d_1 + d_2$, respectively. Let $\widehat{\mathrm{VOL}}_1$, $\widehat{\mathrm{VOL}}_2$ and $\widehat{\mathrm{VOL}}$ be empirical measures for $\boldsymbol{x}_i$, $\tilde{\boldsymbol{x}}_i$ and $\boldsymbol{z}_i$ respectively. If $\boldsymbol{x}_i$ and $\tilde{\boldsymbol{x}}_i$ are NNLD then

$$\begin{aligned}
\widehat{\mathrm{VOL}}([\mathbf{0}_d, \boldsymbol{z})) &= \widehat{\mathrm{VOL}}_1([\mathbf{0}_{d_1}, \boldsymbol{x}))\widehat{\mathrm{VOL}}_2([\mathbf{0}_{d_2}, \tilde{\boldsymbol{x}})) \\
&\geqslant \mathrm{VOL}_1([\mathbf{0}_{d_1}, \boldsymbol{x}))\mathrm{VOL}_2([\mathbf{0}_{d_2}, \tilde{\boldsymbol{x}})) \\
&= \mathrm{VOL}([\mathbf{0}_d, \boldsymbol{z})).
\end{aligned}$$

Therefore $\delta(\boldsymbol{z}) \geqslant 0$ and $\boldsymbol{z}_i$ are NNLD. The same argument, with the inequalities reversed, applies to the NPLD case. □

4 Comparison to Koksma-Hlawka bounds

The Koksma-Hlawka inequality is

$$|\hat{\mu} - \mu| \leqslant D_n^* V_{\mathrm{HK}}(f) \tag{15}$$

where D_n^* denotes again the star discrepancy and $V_{\mathrm{HK}}(f)$ is the total variation of f in the sense of Hardy and Krause. We can be sure that

$$\hat{\mu} - D_n^* V_{\mathrm{HK}}(f) \leqslant \mu \leqslant \hat{\mu} + D_n^* V_{\mathrm{HK}}(f)$$

but the endpoints of this interval are in general far harder to compute than μ is. One difficulty is that $V_{\mathrm{HK}}(f)$ is a sum of $2^d - 1$ Vitali variations (see [28])

that in general are harder to compute than f itself is. However when $\tilde{f}$, defined by $\tilde{f}(\boldsymbol{x}) = f(\mathbf{1} - \boldsymbol{x})$ for every $\boldsymbol{x}$, is completely monotone then it is useful to work with an alternative definition of total variation $V_{\mathrm{HK}\mathbf{0}}$ (see [1]). For this definition, $V_{\mathrm{HK}\mathbf{0}}(\tilde{f}) = V_{\mathrm{HK}}(f)$, and $V_{\mathrm{HK}\mathbf{0}}(\tilde{f}) = \tilde{f}(\mathbf{1}) - \tilde{f}(\mathbf{0}) = f(\mathbf{0}) - f(\mathbf{1})$, see [1].

With an expression for total variation we still need a value or a bound for D_n^*. The computation of D_n^* is expensive, but in some instances it might be worth doing, and for a given set of points we could pre-compute D_n^*. It is possible to compute D_n^* exactly at cost $O(n^{d/2+1})$ for fixed d as $n \to \infty$, see [8]. The cost to compute D_n^* is exponential in the dimension d. If $n = d \to \infty$ together then computation of D_n^* is NP-complete, see [16, 15]. Nevertheless, there are algorithms known that provide either upper and lower bounds for D_n^* in moderate dimension, see [34], or lower bounds for D_n^* even in high dimensions, see [17]. For these and other facts about computing D_n^*, cf. [9].

Then, if we have computed a value $\varepsilon \geqslant D_n^*(P_n)$ we then get an interval

$$\hat{\mu} \pm \varepsilon(f(\mathbf{0}) - f(\mathbf{1}))$$

that is sure to contain μ, when $f(\mathbf{1} - \boldsymbol{x})$ is completely monotone, whether or not P_n is NNLD.

5 Digital net constructions

The NNLD points of [3, 12] are two dimensional Hammersley points which are a special kind of digital nets [7] in which the generator matrices are permutation matrices. In this section we show that digital nets constructed with permutation matrices can be used to get NNLD points with $n = b^m$ points for any integer base $b \geqslant 2$ in any dimension $d \geqslant 1$. This generalizes the result of [3, 12] which holds for $d = 2$. We obtain this generalization by a probabilistic argument using the notion of associated random variables from reliability theory [11]. We also show that there is a limit to how good digital nets can be when their generator matrices are permutation matrices.

5.1 Permutation digital nets

Here we describe how permutation digital nets are constructed. We won't need the more general definition of digital nets until we study them more closely in Section 5.3.

For a dimension $d \geqslant 1$, an integer base $b \geqslant 2$ and an integer $m \geqslant 1$ we choose d matrices $C^{(j)} \in \mathbb{Z}_b^{m\times m}$. For $n = b^m$ and indices $i = 0, 1, \ldots, n-1$, write $i = \sum_{k=1}^m a_{i,k} b^{k-1}$ for $a_{i,k} \in \mathbb{Z}_b$ and put $\vec{i} = (a_{i,1}, \ldots, a_{i,k})^{\mathsf{T}}$. Now let

$$\vec{x}_{ij} = C^{(j)}\vec{i} \mod b$$

have components $\vec{x}_{ij}(k) \in \mathbb{Z}_b$. Then $\boldsymbol{x}_i$ has j'th component

$$x_{ij} = \sum_{k=1}^m \vec{x}_{ij}(k) b^{-k} \in [0, 1).$$

Here we use arithmetic modulo b to define the digital nets. It is customary to only use arithmetic modulo b when b is a prime number and to use a generalization based on finite fields when $b = p^r$ for a prime number p and some power $r \geqslant 2$. Our proofs of NNLD properties exploit a monotonicity of integers modulo b whether or not b is a prime.

As an illustration, the first 16 Hammersley points in base $b \geqslant 2$ for $d = 2$ are constructed this way with

$$C^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad C^{(2)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{16}$$

Hammersley points for $d = 2$ and general $m \geqslant 1$ are constructed similarly, with $C^{(1)} = I_m$ and $C^{(2)}$ a 'reversed' identity matrix as in (16). The Hammersley points for $d \geqslant 3$ are constructed using different bases for different components [18].

5.2 Associated random variables

The settings with $d = 1$ or with $n = 1$ are trivial so we work with $d \geqslant 2$ and $n > 1$. The key ingredient in constructing a short proof of the NNLD property is the notion of associated random variables [11] that originated in reliability theory.

Definition 4. Random variables $T_1, \dots, T_m$ are associated if, for $\boldsymbol{T} = (T_1, \dots, T_m)$ we have $\mathrm{Cov}(g_1(\boldsymbol{T}), g_2(\boldsymbol{T})) \geqslant 0$ for all pairs of functions $g_1, g_2 : \mathbb{R}^m \to \mathbb{R}$ that are nondecreasing in each argument individually and for which $\mathbb{E}(g_1(\boldsymbol{T}))$, $\mathbb{E}(g_2(\boldsymbol{T}))$ and $\mathbb{E}(g_1(\boldsymbol{T})g_2(\boldsymbol{T}))$ all exist.

The next theorem uses points that are a digital net with permutation matrix generators, followed by shifting every component of each point to the right by a distance $1/n$. It shows that they oversample sets of the form $(\boldsymbol{z}, \boldsymbol{1}]$.

Theorem 2. *For integers $m \geqslant 1$, $b \geqslant 2$ and $d \geqslant 2$, let $\pi_1, \dots, \pi_d$ be permutations of $\{1, \dots, m\}$, not necessarily distinct. For $n = b^m$ and $i = 0, \dots, n-1$ and $k = 1, \dots, m$ define $a_i(k) \in \mathbb{Z}_b$ via $i = \sum_{k=1}^m a_i(k) b^{k-1}$. If $\boldsymbol{x}_i \in (0,1]^d$ has components*

$$x_{ij} = \frac{1}{n} + \sum_{k=1}^{m} b^{-k} a_i(\pi_j(k)), \quad j = 1, \dots, d \tag{17}$$

then for any $\boldsymbol{z} \in [0,1]^d$

$$\frac{1}{n} \sum_{i=0}^{n-1} \prod_{j=1}^{d} \mathbb{1}\{x_{ij} > 1 - z_j\} \geqslant \prod_{j=1}^{d} z_j. \tag{18}$$

Proof. We define a random index $i \sim \mathbb{U}\{0, 1, \dots, n-1\}$ which then implies that for each index j the digits $a_i(\pi_j(k)) \sim \mathbb{U}(\mathbb{Z}_b)$ independently for $k = 1, \dots, m$. For each $j = 1, \dots, d$ we have $x_{ij} \sim \mathbb{U}\{1/n, 2/n, \dots, 1\}$. Therefore for any $z_j \in [0, 1]$, $\Pr(x_{ij} > 1 - z_j) \geqslant z_j$.

Let T_j be the value of the random variable x_{ij} where i is random and j is not. Letting γ_j be the inverse of the permutation π_j, we may write

$$T_j = x_{ij} = \frac{1}{n} + \sum_{k=1}^{m} b^{-\gamma_j(k)} a_i(k).$$

Independent random variables $a_i(k)$ are associated by Theorem 2.1 of [11]. Then $T_1, \dots, T_d$ are associated by result P4 of [11] because they are nondecreasing functions of $a_i(1), \dots, a_i(m)$.

For $d = 2$, let $g_1(\boldsymbol{T}) = \mathbb{1}\{x_{i1} > 1 - z_1\}$ and $g_2(\boldsymbol{T}) = \mathbb{1}\{x_{i2} > 1 - z_2\}$. These are nondecreasing functions of associated random variables and so by the definition of associated random variables

$$\Pr(x_{i1} > 1 - z_1, x_{i2} > 1 - z_2) \geqslant \Pr(x_{i1} > 1 - z_1) \Pr(x_{i2} > 1 - z_2).$$

Next, for $2 < r \leqslant d$ let $g_1(\boldsymbol{T}) = \prod_{j=1}^{r-1} \mathbb{1}\{x_{ij} > 1 - z_j\}$ and $g_2(\boldsymbol{T}) = \mathbb{1}\{x_{ir} > 1 - z_r\}$. Using induction we conclude that with our random i,

$$\Pr(x_{ij} > 1 - z_j,\ j = 1, \dots, d) \geqslant \prod_{j=1}^{d} \Pr(x_{ij} > 1 - z_j) \geqslant \prod_{j=1}^{d} z_j$$

which is equivalent to (18). □

Corollary 3. *For integer $b \geqslant 2$ and dimension $d \geqslant 2$ let $\tilde{\boldsymbol{x}}_0, \dots, \tilde{\boldsymbol{x}}_{n-1} \in [0, 1]^d$ be points of a digital net constructed in base b using permutation matrices as generators. Then the points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1} \in [0, 1]^d$ with $x_{ij} = 1 - (1/n + \tilde{x}_{ij})$ are NNLD.*

Proof. Pick $\boldsymbol{z} \in [0, 1]^d$. Now $\mathbb{1}\{x_{ij} < z_j\} = \mathbb{1}\{\tilde{x}_{ij} + 1/n > 1 - z_j\}$ and so

$$\widehat{\mathrm{VOL}}([\boldsymbol{0}, \boldsymbol{z})) = \frac{1}{n} \sum_{i=0}^{n-1} \prod_{j=1}^{d} \mathbb{1}\{x_{ij} < z_j\} = \frac{1}{n} \sum_{i=0}^{n-1} \prod_{j=1}^{d} \mathbb{1}\{\tilde{x}_{ij} + 1/n > 1 - z_j\} \geqslant \prod_{j=1}^{d} z_j$$

by Theorem 2. □

For $d = 2$ it was possible to turn an NNLD point set into an NPLD point set in (6) which includes a reflection $x_{i,2} = 1 - \tilde{x}_{i,2}$. If we were to reflect two or more components of an NNLD point set, then those components would take on a positive upper orthant dependence, which does not generally provide the negative lower orthant dependence we want for NPLD points. For projection regular NNLD points the reflection of $s \geqslant 2$ components will contain $\mathbf{1}_s/n$ and there will be a box $B = [\mathbf{0}_s, \mathbf{1}_s(1/n + \epsilon))$ with $\delta(B) = 1/n - (1/n + \epsilon)^s > 0$ for small enough $\epsilon > 0$.

5.3 Quality of permutation digital nets

It is clear on elementary grounds that a permutation digital net with two identical permutations among $\pi_1, \ldots, \pi_d$ would be very bad. The resulting points would satisfy $x_{ij} = x_{ij'}$ for $0 \leqslant i < n$ and some $1 \leqslant j < j' \leqslant d$. Here we show that our restriction to permutation digital nets rules out the best digital nets when $d \geqslant 3$. We begin with the definitions of these nets.

Definition 5. For integers $d \geqslant 1$, $b \geqslant 2$, and vectors $\boldsymbol{k}, \boldsymbol{a} \in \mathbb{N}^d$ with $a_j \in \mathbb{Z}_{b^{k_j}}$ for $j = 1, \ldots, d$ the Cartesian product

$$\mathcal{E}(\boldsymbol{k}, \boldsymbol{a}) = \prod_{j=1}^{d} \Big[\frac{a_j}{b^{k_j}}, \frac{a_j + 1}{b^{k_j}}\Big)$$

is an elementary interval in base b.

Definition 6. For integers $b \geqslant 2$, $d \geqslant 1$ and $0 \leqslant t \leqslant m$, the n points $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_{n-1}$ are a (t, m, d)-net in base b if

$$\widehat{\mathrm{VOL}}(\mathcal{E}(\boldsymbol{k}, \boldsymbol{a})) = \mathrm{VOL}(\mathcal{E}(\boldsymbol{k}, \boldsymbol{a}))$$

holds for all elementary intervals in base b for which $\sum_{j=1}^{d} k_j \leqslant m - t$.

Digital nets are (t, m, d)-nets. Other things being equal, smaller values of t denote better equidistribution of the points $\boldsymbol{x}_i$ which translates into a lower bound on D_n^* and hence a smaller upper bound in the Koksma-Hlawka inequality. From Theorem 4.10 of [26]

$$D_n^* = O\Big(\frac{b^t \log(n)^{d-1}}{n}\Big) + O\Big(\frac{\log(n)^{d-2}}{n}\Big) \tag{19}$$

where the implied constants depend only on d and b. The powers of $\log(n)$ are not negligible but they are also not seen in examples of integration errors [30].

The quality parameter of a permutation digital net can be very bad. For $d = 2$, taking the Hammersley construction yields $t = 0$ which is the best possible value. Here we show that for $d \geqslant 3$, the best available values of t are far from optimal.

The following definition and result are based on [24, Sect. 2.3].

Construction 1 (Digital Construction of (t, m, d)-Nets)**.** *For prime b, and $C^{(1)}, \ldots, C^{(d)} \in (\mathbb{F}_b)^{m \times m}$, let $\mathcal{C} = \{C^{(1)}, \ldots, C^{(d)}\}$. For $h \in \mathbb{F}_b^m$ define $p(h) \in [0, 1)^d$ componentwise by its b-adic digit expansion*

$$p(h)_j = \delta_1^{(j)}(h) b^{-1} + \delta_2^{(j)}(h) b^{-2} + \cdots + \delta_m^{(j)}(h) b^{-m} \in [0, 1), \quad j = 1, \ldots, d,$$

where $\delta^{(j)}(h) = (\delta_1^{(j)}(h), \ldots, \delta_m^{(j)}(h))$ is simply the vector $C^{(j)} h \in \mathbb{F}_b^m$. We define the point set

$$P(\mathcal{C}) = (p(h))_{h \in \mathbb{F}_b^m}. \tag{20}$$

Clearly, $|P(\mathcal{C})| = b^m$.

To assess the quality of $P(\mathcal{C})$*, we define the* quality criterion $\rho(\mathcal{C})$*: For* $\boldsymbol{m} = (m_1, m_2, \dots, m_d) \in \{0, 1, \dots, m\}^d$ *with* $|\boldsymbol{m}| = \sum_{j=1}^d m_j$ *let*

$$\mathcal{C}^{(\boldsymbol{m})} = \begin{pmatrix} C^{(1)}(1{:}m_1, \cdot) \\ C^{(2)}(1{:}m_2, \cdot) \\ \vdots \\ C^{(d)}(1{:}m_d, \cdot) \end{pmatrix} \in \mathbb{F}_b^{|\boldsymbol{m}| \times d}$$

where $C^{(j)}(1{:}m_j, \cdot) \in \mathbb{F}_b^{m_j \times d}$ *represents the first* m_j *rows of* $C^{(j)}$*. Now* $\rho(\mathcal{C})$ *is the maximum number* $\rho \in \{0, 1, \dots, m\}$ *such that for all* $\boldsymbol{m} \in \{0, 1, \dots, m\}^d$ *with* $|\boldsymbol{m}| = \rho$ *we have* $\operatorname{rank}(\mathcal{C}^{(\boldsymbol{m})}) = \rho$.

Proposition 2. *Let* $b, m, \mathcal{C}$*, and* $P(\mathcal{C})$ *be as in Construction 1. Then* $P(\mathcal{C})$ *is a* (t, m, d)*-net for* $t = m - \rho(\mathcal{C})$.

Observation 3. *The proposition shows that the best possible* t*-value* $t(\mathcal{C})$ *of* $P(\mathcal{C})$ *is at most* $m - \rho(\mathcal{C})$*. But similar arguments as in the corresponding proof of [24, Proposition 2.7] show that actually*

$$t(\mathcal{C}) = m - \rho(\mathcal{C}).$$

Proposition 3. *Let* $V := \{v_1, \dots, v_m\}$ *be a set of linearly independent vectors in* $\mathbb{F}_b^m$*. Let* $m = \ell d + r$*, where* $\ell \in \mathbb{N}_0$ *and* $0 \leqslant r < d$*. If the rows* $C_k^{(j)}$*,* $k = 1, \dots, m$*, of the matrices* $C^{(j)}$*,* $j = 1, \dots, d$*, are all contained in* V*, then* $\rho(\mathcal{C}) \leqslant 2\lfloor m/d \rfloor + 1$*. Therefore, the smallest* t*-value* $t(\mathcal{C})$ *of* $P(\mathcal{C})$ *satisfies*

$$t(\mathcal{C}) \geqslant (d-2)\lfloor m/d \rfloor + r - 1.$$

Proof. Consider the m row vectors

$$C_1^{(1)}, C_1^{(2)}, \dots, C_1^{(d)}, \quad C_2^{(1)}, C_2^{(2)}, \dots, C_2^{(d)}, \quad \dots \quad , C_{\ell+1}^{(1)}, C_{\ell+1}^{(2)}, \dots, C_{\ell+1}^{(r)}.$$

Case 1: Two of these row vectors are equal. Assume these rows are $C_k^{(j)}$ and $C_{k'}^{(j')}$. If $j = j'$, then we consider the matrix $C := \mathcal{C}^{(\boldsymbol{m})}$ with $m_j = \max\{k, k'\}$ and $m_\nu = 0$ for all $\nu \neq j$. Obviously, $\operatorname{rank}(C) \leqslant \max\{k, k'\} - 1$. Hence it follows that $\rho(\mathcal{C}) \leqslant \max\{k, k'\} - 1 \leqslant \lceil m/d \rceil - 1$. If $j \neq j'$, then we consider the matrix $C := \mathcal{C}^{(\boldsymbol{m})}$ with $m_j = k$, $m_{j'} = k'$, and $m_\nu = 0$ for all $\nu \notin \{j, j'\}$. Obviously, $\operatorname{rank}(C) \leqslant k + k' - 1$. Hence it follows that $\rho(\mathcal{C}) \leqslant k + k' - 1 \leqslant 2\lceil m/d \rceil - 1$.

Case 2: All of these row vectors are different. Consider $C_{\ell+1}^{(d)}$. Then there exist $1 \leqslant j < d$ and $1 \leqslant h \leqslant \ell + 1$ or $j = d$ and $1 \leqslant h \leqslant \ell$ such that $C_{\ell+1}^{(d)} = C_h^{(j)}$.

Now we argue similarly as in case 1: If $j = d$, then it is easy to see that $\rho(\mathcal{C}) \leqslant \ell = \lfloor m/d \rfloor$. If $j \neq j'$, then $\rho(\mathcal{C}) \leqslant h + \ell \leqslant 2\ell + 1 \leqslant 2\lfloor m/d \rfloor + 1$.

In any case, we have shown that $\rho(\mathcal{C}) \leqslant 2\lfloor m/d \rfloor + 1$. □

Corollary 4. *Let $m = \ell d + r$, where $\ell \in \mathbb{N}$ and $0 \leqslant r < d$. If $C^{(1)}, \dots, C^{(d)} \in \mathbb{F}_b^{m\times m}$ are all permutation matrices, then the smallest t-value $t(\mathcal{C})$ of $P(\mathcal{C})$ satisfies*

$$t(\mathcal{C}) \geqslant (d-2)\lfloor m/d \rfloor + r - 1.$$

Proof. This follows directly from Proposition 3, since the rows of the matrices $C^{(1)}, \dots, C^{(d)}$ are all in $\{e_1, \dots, e_m\}$, where e_i denotes the i-th standard unit vector of $\mathbb{F}_b^m$. □

Let us represent the permutation matrix where row k has a one in column $\pi(k)$ as simply the column vector with entries $\pi(k)$. Then we can represent our permutation nets with an $m \times d$ matrix Π with j'th column π_j. For example the Hammersley points with generator matrices I_m and reversed I_m are represented this way by

$$\Pi = \begin{pmatrix} 1 & m \\ 2 & m-1 \\ \vdots & \vdots \\ m & 1 \end{pmatrix}. \tag{21}$$

For $d = 3$ we want $\Pi \in \{1, \dots, m\}^{m\times 3}$ with the largest possible value of

$$\rho = \min\bigl\{k + k' \mid \Pi_{k,j} = \Pi_{k',j'}, 1 \leqslant j < j' \leqslant 3\bigr\} - 1.$$

Then we get quality parameter $t = m - \rho$. If we simply adjoin a third column to Π in (21) the best ρ we can get is $m/2$ if m is even and $(m+1)/2$ if m is odd. These lead to $t \geqslant m/2$ if m is even and $t \geqslant (m-1)/2$ if m is odd, which is much worse than the bound in Corollary 4. For $t = m/2$ the first term in (19) is $O(b^{m/2}\log(n)^2/n) = O(\log(n)^2/\sqrt{n})$ because $b = n^{1/m}$.

If $m = 3\ell$, then we can choose the first ℓ rows of Π to be

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \vdots & \vdots & \vdots \\ 3\ell-2 & 3\ell-1 & 3\ell \end{pmatrix}.$$

Let us label these first ℓ rows of Π by $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_\ell \in \mathbb{N}^3$. Now, for $\boldsymbol{r} = (a, b, c)$ let $\boldsymbol{r}' = (b, c, a)$ and $\boldsymbol{r}'' = (c, a, b)$ be one and two rotations of the elements of $\boldsymbol{r}$ to the left with wraparound. By taking the rows of Π in this order

$$\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_\ell,\ \ \boldsymbol{r}'_\ell, \boldsymbol{r}'_{\ell-1}, \dots, \boldsymbol{r}'_1,\ \ \boldsymbol{r}''_\ell, \boldsymbol{r}''_{\ell-1}, \dots, \boldsymbol{r}''_1$$

we get $\rho = 2\ell$ and hence $t = m/3$. This is very close to the bound $\lfloor m/d \rfloor + 0 - 1 = m/3 - 1$ from Corollary 4. We prefer the ordering

$$\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_\ell,\ \ \boldsymbol{r}'_\ell, \boldsymbol{r}''_\ell,\ \ \boldsymbol{r}'_{\ell-1}, \boldsymbol{r}''_{\ell-1},\ \ \boldsymbol{r}'_{\ell-2}, \boldsymbol{r}''_{\ell-2},\quad \dots\ \ \boldsymbol{r}'_2, \boldsymbol{r}''_2,\ \ \boldsymbol{r}'_1, \boldsymbol{r}''_1$$

because while it attains the same value of t it has fewer pairs of columns for which $k + k' = 2\ell + 1$. With $t = m/3$ for $d = 3$ the first term in (19) is $O(b^t \log(n)^2/n) = O(n^{-2/3} \log(n)^2)$.

Using the same method for $d = 4$ and $m = 4\ell$ we can get $\rho = 2\ell = m/2$, implying that $t = m/2$, and yielding a rate of $O(b^t \log(n)^3/n) = O(n^{-1/2} \log(n)^3)$. This result for $d = 4$ matches the rate for plain MC apart from the power of $\log(n)$. So the 100% error bounds available from NNLD sampling come with a logarithmic accuracy penalty in comparison to plain MC.

A second choice for $d = 4$ is to use a Cartesian product of two Hammersley point sets with $\sqrt{n}$ points each. The error of such a Cartesian product would ordinarily be the same as that of the individual Hammersley rules in two dimensions with their reduced sample sizes. That is $O(n^{-1/2} \log(n))$ which is then a better logarithmic factor than the 4 dimensional permutation nets attain.

For $d = 3$ we could also use a Cartesian product of Hammersley points with $n = b^2$ points and a one dimensional grid $\{0, 1/n, \dots, 1 - 1/n\}$. This then uses $N = n^2$ points and we expect an error of $O(\log(n)/n) = O(\log(N)/N^{1/2})$ which is a worse rate than we can get with the permutation net in $[0, 1]^3$.

5.4 Other generator matrices

Permutation matrices are not the only generator matrices that can produce points with the NNLD property. For digital nets in base 2, we know from Proposition 1 that if $C^{(1)} = I_m$ then we must have $C^{(j)}\mathbf{1}_m = \mathbf{1}_m \bmod 2$. This in turn implies that every row of $C^{(j)}$ must have an odd number of 1s in it. A numerical search shows there are 221 choice of nonsingular $C^{(2)}$ when $m = 4$ and $C^{(1)} = I_4$. Below are some examples:

$$C^{(2)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Nevertheless, it is hard to find an example where non-permutation matrices perform better than permutation matrices with respect to the t-value. When $d = 3$, one can verify, either by lengthy reasoning or brute-force enumeration, that NNLD digital nets constructed by non-permutation matrices cannot attain a better t-value than those constructed by permutation matrices for $m \leqslant 7$ and $b = 2$.

6 Non-trivial Rank 1 lattices that are NNLD

Here we consider special cases of rank-1 lattice rules that are suboptimal in terms of discrepancy, but produce NNLD points. While they can be defined in any dimension $d \geqslant 2$ it is only for dimension 1 that they are projection regular. Therefore the conclusions from Proposition 1 and Corollary 1 do not hold for them when $d > 1$.

Theorem 3. *For integers $m \geqslant d$ and $b \geqslant 2$ and $0 \leqslant i < n = b^m$, let*

$$\boldsymbol{x}_i = \Big(\frac{i}{n}, \frac{ib}{n}, \dots, \frac{ib^{j-1}}{n}, \dots, \frac{ib^{d-1}}{n}\Big) \mod 1.$$

Then points $\boldsymbol{x}_0, \dots, \boldsymbol{x}_{n-1}$ are NNLD.

Before proving this theorem we note that these points are quite poor for integration; however, the structure of the points can be useful for showing good integration bounds in suitably weighted spaces, see [5]. There are only b^{d-j+1} unique values of x_{ij}. Further, when $|j - j'|$ is small the points $(x_{ij}, x_{ij'})$ lie within at most $b^{|j-j'|}$ lines in $[0,1)^2$ and have a large discrepancy.

Proof. We write $i = \sum_{k=1}^m a_i(k) b^{k-1}$ and then

$$nx_{ij} = b^{j-1} \sum_{k=1}^{m} a_i(k) b^{k-1} \mod b^m = \sum_{k=1}^{m+1-j} a_i(k) b^{j+k-2}.$$

For $i \sim \mathbb{U}\{0, 1, \dots, n-1\}$ the digits $a_i(1), \dots, a_i(m)$ are independent $\mathbb{U}(\mathbb{Z}_b)$ random variables. Hence they are associated random variables which makes $nx_{i1}, \dots, nx_{id}$ and hence $x_{i1}, \dots, x_{id}$ into associated random variables. Finally, x_{ij} has the uniform distribution on $\{0, 1/n_j, 2/n_j, \dots, 1 - 1/n_j\}$ where $n_j = n/b^{j-1}$. This distribution is stochastically smaller than $\mathbb{U}[0,1]$ and so $\boldsymbol{x}_i$ are NNLD. □

The values x_{ij} for $0 \leqslant i < b^m$ in these lattices take $n_j = b^{d-j+1}$ distinct values ℓ/n_j for $0 \leqslant \ell < n_j$ with each of those values appearing n/n_j times. As such they constitute a left endpoint integration rule on n_j points and so for nonperiodic smooth integrands we anticipate an error rate of $O(n_j^{-1})$. For this to be better than plain MC we require $n_j \geqslant \sqrt{n}$ or $j \leqslant m/2$. While a better rate is available for periodic integrands, those cannot be completely monotone unless they are constant.

7 Discussion and further references

We find that it is possible to get computable bounds on some integrals by using points with a suitable bias property (non-negative local discrepancy (NNLD)) on integrands with a suitable monotonicity property (complete monotonicity). The method of associated random variables is useful for showing that a given point set is NNLD.

There are several generalizations of multivariate monotonicity in [25]. They include the complete monotonicity discussed here as well as the more commonly considered monotonicity in each of the d inputs one at a time. The complexity of integrating coordinate-wise monotone functions has been studied by [27, 31]. Scrambled (t, m, d)-nets have been shown to be negatively orthant dependent if and only if $t = 0$ [35]. Similarly, it was shown in [36] that randomly shifted and

jittered (RSJ) rank-1 lattices based on a random generator are also negatively orthant dependent and that, in some sense, one cannot achieve this result by employing less randomness. Using the NLOD property of the distribution of these RQMC points, it follows from [23] that for functions which are monotone in each variable scrambled nets and RSJ rank-1 lattices cannot increase variance over plain Monte Carlo in any dimension d.

While complete monotonicity is a very special property, its applicability can be widened by the method of control variates. If $h(\cdot)$ is completely monotone with known integral θ, we will in some settings be able to find $\lambda_+ > 0$ for which $f + \lambda_+ h$ is a completely monotone function of $\boldsymbol{x}$. Then by Theorem 1 we can compute an upper bound $B_+ \geqslant \mu + \lambda_+\theta$ and conclude that $\mu \leqslant B_+ - \lambda_+\theta$. Similarly a lower bound can be found by choosing λ_- such that $\lambda_- h - f$ is a completely monotone function of $\boldsymbol{x}$, using Theorem 1 to get an upper bound $\lambda_-\theta - \mu \leqslant B_-$ and then concluding that $\mu \geqslant \lambda_-\theta - B_-$. Details on how to choose h and find $\lambda_\pm$ are beyond the scope of this article.

The customary way to quantify uncertainty in QMC is to use RQMC replicates with statistically derived asymptotic confidence intervals. For a recent thorough empirical evaluation of RQMC, see [22], who found the usual confidence intervals based on the central limit theorem to be even more reliable than sophisticated bootstrap methods. Here we have found an alternative computable non-asymptotic approach with 100% coverage, but so far it does not give very good accuracy for high dimensions.

Acknowledgments

We thank Josef Dick, David Krieg, Frances Kuo, Dirk Nuyens and Ian Sloan for discussions. Much of this work took place at the MATRIX Institute's location in Creswick Australia as part of their research program on 'Computational Mathematics for High-Dimensional Data in Statistical Learning', in February 2023, and the paper was finalized during the Dagstuhl Seminar 23351 'Algorithms and Complexity for Continuous Problems', in Schloss Dagstuhl, Wadern, Germany, in August 2023. We are grateful to MATRIX and to the Leibniz Center Schloss Dagstuhl. The contributions of ABO and ZP were supported by the U.S. National Science Foundation under grant DMS-2152780. Peter Kritzer is supported by the Austrian Science Fund (FWF) Project P34808. For the purpose of open access, the authors have applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission.

References

[1] Ch. Aistleitner and J. Dick. Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arithmetica*, 167(2):143–171, 2015.

[2] S. Boyd and L. Vandeberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[3] L. De Clerck. A method for exact calculation of the stardiscrepancy of plane sets applied to the sequences of Hammersley. *Monatsh. Math.*, 101:261–278, 1986.

[4] J. Dick and P. Kritzer. A best possible upper bound on the star discrepancy of (t, m, 2)-nets. *Monte Carlo Methods and Applications*, 12(1):1–17, 2006.

[5] J. Dick, P. Kritzer, G. Leobacher, and F. Pillichshammer. A reduced fast component-by-component construction of lattice points for integration in weighted spaces with fast decreasing weights. *Journal of Computational and Applied Mathematics*, 276:1–15, 2015.

[6] J. Dick, P. Kritzer, and F. Pillichshammer. *Lattice Rules: Numerical Integration, Approximation, and Discrepancy*. Springer Nature, 2022.

[7] J. Dick and F. Pillichshammer. *Digital sequences, discrepancy and quasi-Monte Carlo integration*. Cambridge University Press, Cambridge, 2010.

[8] D.P. Dobkin, D. Eppstein, and D.P. Mitchell. Computing the discrepancy with applications to supersampling patterns. *ACM Trans. Graph.*, 15:354–376, 1996.

[9] C. Doerr, M. Gnewuch, and M. Wahlström. Calculation of discrepancy measures and applications. In Chen W., Srivastav A., and Travaglini G., editors, *A Panorama of Discrepancy Theory*, pages 621–678. Springer, 2014.

[10] J. Elstrodt. *Maß und Integrationstheorie*. Springer Spektrum, Berlin, 2018. 8th edition.

[11] J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.

[12] H. Gabai. On the discrepancy of certain sequences mod 1. *Illinois Journal of Mathematics*, 11(1):1–12, 1967.

[13] A. Genz and F. Bretz. *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics. Springer-Verlag, Heidelberg, 2009.

[14] A. Genz, F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Scheipl, and T. Hothorn. *mvtnorm: Multivariate Normal and t Distributions*, 2021. R package version 1.1-3.

[15] P. Giannopoulos, C. Knauer, M. Wahlströhm, and D. Werner. Hardness of discrepancy computation and ε-net verification in high dimension. *Journal of Complexity*, 28:162–176, 2012.

[16] M. Gnewuch, A. Srivastav, and C. Winzen. Finding optimal volume subintervals with k points and computing the star discrepancy are NP-hard. *Journal of Complexity*, 24:154–172, 2008.

[17] M. Gnewuch, M. Wahlström, and C. Winzen. A new randomized algorithm to approximate the star discrepancy based on Threshold Accepting. *SIAM J. Numer. Anal.*, 50:781–807, 2012.

[18] J. M Hammersley. Monte Carlo methods for solving multivariable problems. *Annals of the New York Academy of Sciences*, 86(3):844–874, 1960.

[19] F. J. Hickernell. Koksma-Hlawka inequality. *Wiley StatsRef: Statistics Reference Online*, 2014.

[20] S. Joe and F. Y. Kuo. Constructing Sobol' sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing*, 30(5):2635–2654, 2008.

[21] P. L'Ecuyer and C. Lemieux. A survey of randomized quasi-Monte Carlo methods. In M. Dror, P. L'Ecuyer, and F. Szidarovszki, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic Publishers, 2002.

[22] P. L'Ecuyer, M. K. Nakayama, A. B. Owen, and B. Tuffin. Confidence intervals for randomized quasi-Monte Carlo estimators. Technical Report `https://inria.hal.science/hal-04088085/`, 2023.

[23] C. Lemieux. Negative dependence, scrambled nets, and variance bounds. *Mathematics of Operations Research*, 43(1):228–251, 2018.

[24] J. Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Springer-Verlag, Heidelberg, 1998.

[25] K. Mosler and M. Scarsini. Some theory of stochastic dominance. *Lecture Notes-Monograph Series*, pages 261–284, 1991.

[26] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. S.I.A.M., Philadelphia, PA, 1992.

[27] E. Novak. Quadrature formulas for monotone functions. *Proc. of the AMS.*, 115:59–68, 1992.

[28] A. B. Owen. Multidimensional variation for quasi-Monte Carlo. In J. Fan and G. Li, editors, *International Conference on Statistics in honour of Professor Kai-Tai Fang's 65th birthday*, 2005.

[29] A. B. Owen. *Practical Quasi-Monte Carlo*. Draft available at `https://artowen.su.domains/mc/practicalqmc.pdf`, 2023.

[30] A. B. Owen and Z. Pan. Where are the logs? In *Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer*. Springer, Cham, 2022.

[31] A. Papageorgiou. Integration of monotone functions of several variables. *Journal of Complexity*, 9(2):252–268, 1993.

[32] M. Shaked. A general theory of some positive dependence notions. *Journal of Multivariate Analysis*, 12(2):199–218, 1982.

[33] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Oxford Science Publications, Oxford, 1994.

[34] E. Thiémard. Optimal volume subintervals with k points and star discrepancy via integer programming. *Math. Methods Oper. Res.*, 54:21–45, 2001.

[35] J. Wiart, C. Lemieux, and G. Y. Dong. On the dependence structure and quality of scrambled (t, m, s)-nets. *Monte Carlo Methods and Applications*, 27(1):1–26, 2021.

[36] M. Wnuk and M. Gnewuch. Note on pairwise negative dependence of random rank-1 lattices. *Operations Research Letters*, 48:410–414, 2020.