

Minimizing point configurations for tensor product energies on the torus

Dmitriy Bilyk, Nicolas Nagel, Ian Ruohoniemi

Abstract

We study point configurations on the torus $\mathbb{T}^d$ that minimize interaction energies with tensor product structure which arise naturally in the context of discrepancy theory and quasi-Monte Carlo integration. Permutation sets on $\mathbb{T}^2$ and Latin hypercube sets in higher dimensions (i.e. sets whose projections onto coordinate axes are equispaced points) are natural candidates to be energy minimizers. We show that such point configurations that have only one distance in the vector sense minimize the energy for a wide range of potentials, in other words, such sets satisfy a tensor product version of universal optimality. This applies, in particular, to three- and five-point Fibonacci lattices. We also characterize all lattices with this property and exhibit some non-lattice sets of this type. In addition, we obtain several further structural results about global and local minimizers of tensor product energies.

1 Introduction

1.1 Energies with product structure

Denote by $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ the unit torus, where we will identify elements in $\mathbb{T}$ with those in the half-open interval $[0, 1)$. In the present paper we explore the properties of minimizers of interaction energies on $\mathbb{T}^d$ with kernels that have tensor product structure. To be more precise, consider the kernel $F : \mathbb{T}^d \rightarrow [0, \infty)$ of the form

$$F(t_1, \dots, t_d) = f(t_1) \dots f(t_d), \quad (1.1)$$

where $f : \mathbb{T} \rightarrow [0, \infty)$. Throughout the paper we will assume that f , viewed as a non-negative 1-periodic function on $\mathbb{R}$, is *even and continuous* on $(0, 1)$. In most situations f will actually be assumed to be continuous on the whole

real line, but some results also apply when f is singular at 0 in the sense of $\lim_{t \rightarrow 0} f(t) = +\infty$. We will refer to both F and f as *potentials* throughout the paper.

For a (multi)set of N points $X = \{\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\} \subset \mathbb{T}^d$ with components $\mathbf{x}^k = (x_1^k, \dots, x_d^k)$, we define the *tensor product energy* as

$$E_F(X) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} F(\mathbf{x}^k - \mathbf{x}^\ell) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} \prod_{i=1}^d f(x_i^k - x_i^\ell). \quad (1.2)$$

This quantity can be intuitively viewed as the total potential energy of N particles in positions $\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}$ which interact according to the potential given by F . One is then naturally interested in minimizing the energy (1.2) over all N -point configurations and finding minimizing configurations of points, i.e. equilibrium distributions. We shall call such configurations *energy minimizers*. While extensive literature exists on energy minimization in various domains in the case where the kernel $F(\mathbf{x}, \mathbf{y})$ depends on the Euclidean (or another) distance between $\mathbf{x}$ and $\mathbf{y}$ (see for example the book [1]), kernels with tensor product structures have not received the same attention [23, 29].

The condition that f is even, or equivalently $f(t) = f(1 - t)$, means that $f(x - y)$ depends only on the *geodesic distance*

$$\delta(x, y) = \min\{|x - y|, 1 - |x - y|\} \quad (1.3)$$

between x and y in $\mathbb{T}$. For most potentials of interest, f is decreasing on $(0, 1/2]$, as this corresponds to the interaction being repulsive in each coordinate.

We remark that the energy (1.2) of a configuration is invariant under symmetries of the torus (translations, reflections, and permutations of coordinates) and all uniqueness statements, e.g. in Definition 1.1 or Theorem 2.8, are to be understood modulo these symmetries. In particular, since the energy (1.2) is translation invariant, we may assume without loss of generality that $\mathbf{x}^0 = (0, \dots, 0)$.

Energies of the type (1.2) with tensor product kernels (1.1) naturally arise in several fields such as discrepancy theory and quasi-Monte Carlo (QMC) integration. Warnock's formula [22, 23, 33] states that minimizing the periodic L^2 -discrepancy of N points in $\mathbb{T}^d$ is equivalent to minimizing the energy (1.2) with the kernel (1.1), where $f(t) = \frac{1}{2} - t + t^2$ on $[0, 1)$. It is also known [23, 29] that minimizing the worst-case integration error of the quasi-Monte Carlo rule over a certain Sobolev space corresponds to the energy minimiza-

tion problem with the potential

$$f_p(t) = 1 + \frac{p}{2} \left(\frac{1}{6} - t + t^2 \right), \quad (1.4)$$

where $p > 0$ is a fixed parameter (notice that $f_6 = 3f$, where f is the aforementioned potential from Warnock's formula). Similar statements hold for other reproducing kernel Hilbert spaces (see the Appendix §A and [2, 3] or [30, §9]). The relation to discrepancy theory provides possible candidates for energy minimizers, namely low-discrepancy sets.

1.2 Dimension $d = 2$: Permutation sets and Fibonacci lattices

Some of the most well-known low-discrepancy point sets in $\mathbb{T}^2$ include the digit-reversing *van der Corput sets* when $N = 2^n$ (see e.g. [22] for a definition) and the *Fibonacci lattices* which will keep appearing in this paper: if $N = F_j$ is a Fibonacci number, the N -point Fibonacci lattice is defined by

$$\Phi_j := \left\{ \left(\frac{k}{F_j}, \left\{ \frac{kF_{j-1}}{F_j} \right\} \right) : k = 0, 1, \dots, F_j - 1 \right\}, \quad (1.5)$$

where $\{\cdot\}$ denotes the fractional part. This set can be thought of as a rational approximation to the famous Kronecker set $\{(k/N, \{k\varphi\}) : k \in \mathbb{N}_0\}$, where φ is the golden section. Fibonacci lattices have been studied extensively in the context of discrepancy theory and numerical integration and perform strongly in these settings [4, 5].

Both of the aforementioned examples, the van der Corput set and the Fibonacci lattice, have important structural properties. First, these are N -point configurations that are subsets of the $(N \times N)$ -grid

$$\left\{ \left(\frac{k}{N}, \frac{\ell}{N} \right) : k, \ell = 0, 1, \dots, N - 1 \right\}$$

in $[0, 1)^2 \simeq \mathbb{T}^2$. We shall call such sets *grid sets*. More importantly these examples are of the form

$$X_\sigma = \{\mathbf{x}^k\}_{k=0}^{N-1} = \left\{ \left(\frac{k}{N}, \frac{\sigma(k)}{N} \right) : k = 0, 1, \dots, N - 1 \right\}, \quad (1.6)$$

where σ is a permutation of the set $\{0, 1, \dots, N - 1\}$ (if we assume $\mathbf{x}^0 = (0, 0)$, that is we fix $\sigma(0) = 0$, we can view σ just as easily as a permutation of $\{1, \dots, N - 1\}$).

Definition 1.1. We shall call configurations $X \subset \mathbb{T}^2$ which (up to translations) have the form (1.6) *permutation sets*.

A special place among the permutation sets is taken by *integration lattices* (which we shall refer to simply as *lattices*), i.e. sets of the form

$$\left\{ \left(\frac{k}{N}, \left\{ \frac{hk}{N} \right\} \right) : k = 0, 1, \dots, N-1 \right\}, \quad (1.7)$$

where h and N are relatively prime integers. The condition that h and N are relatively prime guarantees that (1.7) is indeed a permutation set. Notice that such sets are restrictions to $[0, 1)^2$ of lattices in $\mathbb{R}^2$ generated by the vectors $(1/N, h/N)$ and $(0, 1)$, which are 1-periodic in both coordinate directions. Fibonacci lattices (1.5) are particular examples of such sets.

We note that permutation sets are those sets whose projections onto both x - and y -axes are N equispaced points, i.e. they are grid sets with exactly one point in each row and each column. This interpretation suggests that, since many energies in dimension $d = 1$ are minimized by equispaced points (see §2.1), it is natural to expect that permutation sets minimize tensor product energies on $\mathbb{T}^2$ for large classes of potentials. We then have the following question: *which conditions on the potential f guarantee that minimizers of the tensor product energy (1.2) are permutation sets?* This question admits a variety of versions and interpretations: e.g., when does this hold for all N , for specific values of N , or for all N large enough? When are energies minimized by lattices? Does it hold if sets are restricted to the $(N \times N)$ -grid?

Based on the particularly important role played by the Fibonacci lattices in low discrepancy constructions and on ample numerical evidence [23], we conjecture that whenever N is a Fibonacci number, *the Fibonacci lattice minimizes the energy for a broad class of potentials f* . This conjectured phenomenon (the same configuration minimizing a large variety of energies) is a natural analogue of the notion of *universal optimality* that has been studied on the sphere $\mathbb{S}^d$ [10] and in the plane [11], and is known to be very rare.

In the end, we remark that the condition that f is logarithmically convex is very natural as it is necessary and sufficient for the tensor product F to be convex as a function of d variables [31].

1.3 Higher dimensions $d \geq 3$

In dimensions $d \geq 3$ it becomes significantly harder to construct good point sets. To the knowledge of the authors, no study of global optimizers of L_2 -discrepancy or quasi-Monte Carlo worst case errors in dimensions $d \geq 3$

has been carried out so far and only asymptotic results are known [9, 19]. A straightforward extension of the concept of permutation sets in higher dimensions to sets on the N^d -grid $\{k/N : k = 0, 1, \dots, N-1\}^d$ leads to Latin (hyper)cubes.

Definition 1.2. A *Latin hypercube set* is defined by permutations $\sigma_2, \dots, \sigma_d$ of the set $\{0, 1, \dots, N-1\}$ via

$$X_{\sigma_2, \dots, \sigma_d} = \left\{ \left(\frac{k}{N}, \frac{\sigma_2(k)}{N}, \dots, \frac{\sigma_d(k)}{N} \right) : k = 0, 1, \dots, N-1 \right\}.$$

In other words, it is a set whose projections on all coordinate axes are equispaced points. Thus a *permutation set* is a Latin hypercube set with $d = 2$. Analogously to $\mathbb{T}^2$, lattices in $\mathbb{T}^d$ of the form

$$\left\{ \left(\frac{k}{N}, \left\{ \frac{h_2 k}{N} \right\}, \dots, \left\{ \frac{h_d k}{N} \right\} \right) : k = 0, 1, \dots, N-1 \right\},$$

which are generated by a vector $(\frac{1}{N}, \frac{h_2}{N}, \dots, \frac{h_d}{N})$ with $\gcd(h_2, N) = \dots = \gcd(h_d, N) = 1$, are common examples of Latin hypercube sets.

Just as in the case of permutation set in dimension $d = 2$, the same questions naturally arise in dimension $d \geq 3$: *For which potentials f do Latin hypercubes minimize the tensor product energy? Can one identify specific lattices or other Latin hypercube sets which minimize large classes of energies?*

1.4 Main results

Even though complete answers to the aforementioned questions appear to be out of reach for now, we provide a plethora of interesting results which open up a new direction in energy minimization. In particular, we prove the following:

- As a first indication that permutation and Latin hypercube sets might naturally be minimizers of tensor product energies in all dimensions $d \geq 2$, we show that points in (local) minimizers cannot coincide in any coordinate provided that the potential f generates sufficient repulsion at small scales (Theorem 3.1).
- In Theorem 2.3, which can be considered the main result of the paper, we prove that in all dimensions $d \geq 2$ *Latin hypercube sets with a single distance vector* (in other words, sets of maximal separation in which each interaction contributes the same amount to the energy) minimize

the tensor product energy with potential f whenever the energy with potential $g = \log f$ on $\mathbb{T}$ is minimized by equispaced points. Sufficient conditions for this property include the following:

- (a) f is logarithmically convex and decreasing on $(0, 1/2]$, or
- (b) $\log f(t)$ is absolutely monotone with respect to $s = \cos(2\pi t)$.

(See §2.1 for more details)

- We exhibit many sets in dimensions with a single distance vector in dimensions $d \geq 2$. In fact, in Theorem 2.8 we characterize all the lattices with a single distance vector. We also give many examples of non-lattice sets with this property. This provides a large collection of point configurations which satisfy this tensor product version of universal optimality.
- In particular, the 3- and 5-point Fibonacci lattices are easily seen satisfy the assumptions of Theorem 2.3. Thus the 5-point Fibonacci lattice Φ_5 minimizes the energy for potentials f as above, see Corollary 2.5, i.e. Φ_5 is universally optimal with respect to these classes of potentials (and a similar result holds for the 3-point Fibonacci lattice, Corollary 2.7).
- Moreover, the 5-point Fibonacci lattice Φ_5 minimizes the tensor product energy (1.2) among all 5-point permutation sets if f is only assumed to be decreasing on $(0, 1/2]$, Lemma 2.6.
- If f is decreasing and logarithmically convex on $(0, 1/2]$, lattices have the smallest energy among all sets of the same *permutation type* (see Definition 4.1), i.e. energy minimization under these restriction yields permutation sets (Theorem 4.2). As a corollary, under the same conditions every lattice is a local energy minimizer (Corollary 4.4).
- We also address energy of 4- and 6-point sets on $\mathbb{T}^2$. For $N = 4$, we find global minimizers of the energy when f is differentiable on $(0, 1)$, and decreasing and logarithmically convex on $(0, 1/2]$ (Proposition 5.2), which extends a numerical result from [23] to a wide range of potentials. Moreover, we observe that a 4-point permutation set almost never minimizes a tensor product energy and is never the unique minimizer (Remark 5.3).
- For $N = 6$ a similar construction yields a minimizer among all 6-point configurations restricted to a natural choice of permutation type (Proposition 5.4).

In the Appendix we provide a sampler of relevant potentials and discuss their properties.

2 Optimality of Latin hypercube sets with a single distance vector

In this section we will present the aforementioned collection of point configurations which minimize broad classes of tensor product energies (i.e. possess a certain version of universal optimality). To understand the relevant properties of the underlying potentials, we first discuss energy on the circle.

2.1 Equispaced points as energy minimizers on the circle

While the one-dimensional case $\mathbb{T}$ does not exhibit tensor product structure, it is an important starting point for our investigations. In this setting, if all points repel each other equally, it is natural to expect the energy

$$E_f(x^0, x^1, \dots, x^{N-1}) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} f(x^k - x^\ell)$$

to be minimized by *equispaced points*

$$\{x^0, x^1, \dots, x^{N-1}\} = \left\{ \frac{0}{N}, \frac{1}{N}, \dots, \frac{N-1}{N} \right\} \subset \mathbb{T}.$$

In the context of numerical integration, the optimality of equispaced points was proven in a broad sense in [34]. However, in the general energy setting, the precise characterizations of the potentials f for which this happens is still elusive. Some known sufficient conditions on f are formulated below. Recall that we always assume f (as a 1-periodic function on $\mathbb{R}$) to be even, i.e. $f(t) = f(1 - t)$.

Proposition 2.1. Let $N \in \mathbb{N}$. Then

$$E_f(x^0, x^1, \dots, x^{N-1}) \geq E_f\left(\frac{0}{N}, \frac{1}{N}, \dots, \frac{N-1}{N}\right) \quad (2.1)$$

for all $x^0, x^1, \dots, x^{N-1} \in [0, 1)$ provided that f satisfies any one of the following conditions:

- (i) f is convex and decreasing on $(0, 1/2]$;
- (ii) $f(t) = h(\cos(2\pi t))$ where $h : [-1, 1) \rightarrow [0, \infty)$ is absolutely monotone (up to an additive constant), i.e. h is a C^∞ -function satisfying $h^{(k)}(s) \geq 0$ for all $s \in (-1, 1)$ and all $k \in \mathbb{N}$;
- (iii) $f \in C(\mathbb{T})$ and $f(t) = \sum_{m=0}^{\infty} a_m \cos(2\pi mt)$ pointwise with $a_m \geq 0$ for all $m > 0$, and $a_m = 0$ for all $m > 0$ with $N|m$.

Condition (i) is classical and goes back to [17], see also [6]. Condition (ii) is a special case $d = 1$ of the celebrated theorem of Cohn and Kumar [10] on universal optimality on the sphere $\mathbb{S}^d$. It is not hard to see that these conditions overlap, but do not imply each other. Observe that these two conditions apply to all values of $N \in \mathbb{N}$ simultaneously. In fact, for a given value of N , the positivity of $h^{(k)}$ in condition (ii) only needs to be checked up to order $k \leq N - 1$. Condition (iii) is a simple calculus exercise (positivity of the coefficients even implies that the Fourier series converges absolutely and uniformly). Fejér kernels

$$f_\nu(t) = \frac{1}{\nu} \left(\frac{\sin(\nu\pi t)}{\sin(\pi t)} \right)^2 = 1 + 2 \sum_{m=1}^{\nu-1} \left(1 - \frac{m}{\nu} \right) \cos(2\pi mt)$$

are examples of potentials for which (iii) holds, but neither (i) nor (ii) applies for a fixed $N \geq \nu \geq 3$.

None of the conditions (i)–(iii) are necessary. Positive definiteness of f (equivalent to positivity of the coefficients in the cosine Fourier series) is a simple necessary condition for (2.1) to hold for all N large enough, but it is not sufficient.

If (2.1) holds, i.e. N equispaced points in $\mathbb{T}$ minimize the energy E_f , the potential f will be called *N -equispacing*. As an example, the aforementioned potential (1.4) arising in quasi-Monte Carlo integration is N -equispacing for all $N \in \mathbb{N}$ and for all $p > 0$ as it fulfills both (i) and (ii) (as follows from [27, equation (13)] above).

2.2 Sets with a single distance vector

Since tensor product energies depend on distances in individual coordinates independently, rather than on actual distances between points, it is natural to consider a vector-valued analogue of distances. For points $\mathbf{x}, \mathbf{y} \in \mathbb{T}^d$ define the *distance vector* between them as

$$\mathrm{dv}(\mathbf{x}, \mathbf{y}) = (\delta(x_{\tau(1)}, y_{\tau(1)}), \dots, \delta(x_{\tau(d)}, y_{\tau(d)})),$$

where δ denotes the geodesic distance on $\mathbb{T}$ as in (1.3) and τ is a permutation of $\{1, \dots, d\}$ such that

$$\delta(x_{\tau(1)}, y_{\tau(1)}) \leq \dots \leq \delta(x_{\tau(d)}, y_{\tau(d)}),$$

i.e. we take the geodesic distance in every component and sort them increasingly. For $X = \{\mathbf{x}^k\}_{k=0}^{N-1} \subseteq \mathbb{T}^d$ let

$$\text{dv}(X) = \{\text{dv}(\mathbf{x}^k, \mathbf{x}^\ell) : k \neq \ell\}$$

be the set of all distance vectors among points in X . We shall be interested in Latin hypercube sets $X \subseteq \mathbb{T}^d$ satisfying

$$\#\text{dv}(X) = 1,$$

i.e. having only *a single distance vector*. We remark that such sets are natural tensor product analogues of one-distance sets which have been well studied in other geometric contexts, e.g. on spheres (simplices with center of mass in $\mathbf{0}$) and projective spaces (equiangular tight frames), see [12]. However, this tensor product setting does not seem to have previously appeared in the literature.

Since for the tensor product energy E_F the interaction between two points only depends on the distance vector between them, we see that for a set $X \subseteq \mathbb{T}^d$, $\#X = N$ with a single distance vector $\{(\rho_1, \dots, \rho_d)\} = \text{dv}(X)$ we have

$$E_F(X) = (N^2 - N) \prod_{i=1}^d f(\rho_i). \quad (2.2)$$

The following lemma counts the frequency of different entries in the common distance vector in such a set. The main point of this computation is to show that the frequency distribution coincides with the distribution of distances in d copies of N equispaced points in $\mathbb{T}$.

Lemma 2.2. Let $X \subseteq \mathbb{T}^d$, $\#X = N$ be a Latin hypercube set with a single distance vector, whose common distance vector is given by $\{(\rho_1, \dots, \rho_d)\} = \text{dv}(X)$. Then for every $r = 1, \dots, \lfloor N/2 \rfloor$ it holds

$$\#\left\{i = 1, \dots, d : \rho_i = \frac{r}{N}\right\} = \begin{cases} \frac{2d}{N-1}, & r < \frac{N}{2}, \\ \frac{d}{N-1}, & r = \frac{N}{2}. \end{cases}$$

Proof. We will show, via double counting, that

$$(N^2 - N) \cdot \#\left\{i = 1, \dots, d : \rho_i = \frac{r}{N}\right\}$$

$$= d \cdot \# \left\{ (k, \ell) \in \{0, 1, \dots, N-1\}^2 : \delta \left(\frac{k}{N}, \frac{\ell}{N} \right) = \frac{r}{N} \right\},$$

from which the claim follows since

$$\# \left\{ (k, \ell) \in \{0, 1, \dots, N-1\}^2 : \delta \left(\frac{k}{N}, \frac{\ell}{N} \right) = \frac{r}{N} \right\} = \begin{cases} 2N, & r < \frac{N}{2}, \\ N, & r = \frac{N}{2}. \end{cases}$$

Since X is a Latin hypercube, there are permutations $\sigma_1, \dots, \sigma_d$ such that

$$X = \{\mathbf{x}^k\}_{k=0}^{N-1} = \left\{ \left(\frac{\sigma_1(k)}{N}, \dots, \frac{\sigma_d(k)}{N} \right) : k = 0, 1, \dots, N-1 \right\}$$

(with $\sigma_1 = \text{id}$ the identity on $\{0, 1, \dots, N-1\}$). Consider the array $A = [a_{(k,\ell),i}]$ with rows indexed by $(k, \ell) \in \{0, 1, \dots, N-1\}^2, k \neq \ell$ and columns indexed by $i = 1, \dots, d$, with entries

$$a_{(k,\ell),i} = \delta \left(\frac{\sigma_i(k)}{N}, \frac{\sigma_i(\ell)}{N} \right).$$

We will count for how many entries it holds $a_{(k,\ell),i} = \frac{r}{N}$. The row corresponding to (k, ℓ) consists, by definition, of the entries of $\text{dv}(\mathbf{x}^k, \mathbf{x}^\ell)$ in some order. By the single distance property, this row contains exactly

$$\# \left\{ i = 1, \dots, d : \rho_i = \frac{r}{N} \right\}$$

such entries, so that there are

$$(N^2 - N) \cdot \# \left\{ i = 1, \dots, d : \rho_i = \frac{r}{N} \right\}$$

in total of them in A . On the other hand, since σ_i is a permutation, in the i -th column we have

$$\begin{aligned} & \# \left\{ (k, \ell) \in \{0, 1, \dots, N-1\}^2 : \delta \left(\frac{\sigma_i(k)}{N}, \frac{\sigma_i(\ell)}{N} \right) = \frac{r}{N} \right\} \\ &= \# \left\{ (k, \ell) \in \{0, 1, \dots, N-1\}^2 : \delta \left(\frac{k}{N}, \frac{\ell}{N} \right) = \frac{r}{N} \right\}, \end{aligned}$$

so in all of A there are

$$d \cdot \# \left\{ (k, \ell) \in \{0, 1, \dots, N-1\}^2 : \delta \left(\frac{k}{N}, \frac{\ell}{N} \right) = \frac{r}{N} \right\}$$

such entries. This finishes the proof. $\square$

For an N -point Latin hypercube set with a single distance vector to exist in $\mathbb{T}^d$ one thus needs to fulfill the divisibility condition

$$\frac{N-1}{\gcd(2, N-1)} \Big| d, \quad (2.3)$$

in particular for a given dimension d they can only exist up to $N \leq 2d+1$ points.

2.3 Latin hypercube sets with a single distance vector as energy minimizers

The following is the main energy minimization result of this paper. It states that as long as $\log f$ is N -equispacing, then an N -point Latin hypercube set with a single distance vector $X \subset \mathbb{T}^d$ minimizes the corresponding tensor product energy E_F .

Theorem 2.3. Let $X \subseteq \mathbb{T}^d$, $\#X = N$ be a Latin hypercube set with a single distance vector. Let $f : \mathbb{T} \rightarrow (0, \infty)$ be a potential, such that the potential $g = \log f : \mathbb{T} \rightarrow \mathbb{R}$ is N -equispacing. Then

$$E_F(Y) \geq E_F(X)$$

for all $Y \subseteq \mathbb{T}^d$, $\#Y = N$.

According to Proposition 2.1, the result of Theorem 2.3 can be applied if $\log f$ is decreasing and convex (i.e. f is logarithmically convex) as a function of geodesic distance or if $\log f(t)$ is absolutely monotone when rewritten as a function of $s = \cos(2\pi t)$, see e.g. [20] for more information on logarithmic absolute monotonicity.

Thus, as mentioned earlier, this result can be viewed as an analogue of the classical universal optimality results on spheres and projective spaces [10, 12] in the sense that one configuration minimizes the energy for a large class of potentials. In fact, in [10] universal optimality was shown using the linear programming method for *sharp configurations*, i.e. sets which are spherical designs of high order and have few pairwise distances. The sets with a single distance vector have an extreme version of the latter property in the tensor product sense.

Proof of Theorem 2.3. Let $Y = \{\mathbf{y}^0, \mathbf{y}^1, \dots, \mathbf{y}^{N-1}\} \subseteq \mathbb{T}^d$, $\#Y = N$. Since the exponential function is convex, we can apply Jensen's inequality to obtain

$$E_F(Y) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} \prod_{i=1}^d f(y_i^k - y_i^\ell) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} \exp \left(\sum_{i=1}^d g(y_i^k - y_i^\ell) \right)$$

$$\geq (N^2 - N) \exp \left(\frac{1}{N^2 - N} \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} \sum_{i=1}^d g(y_i^k - y_i^\ell) \right).$$

By assumption, for all $i = 1, \dots, d$ we have

$$\begin{aligned} \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} g(y_i^k - y_i^\ell) &\geq \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} g\left(\frac{k}{N} - \frac{\ell}{N}\right) \\ &= \begin{cases} 2N \sum_{r=1}^{(N-1)/2} g\left(\frac{r}{N}\right), & N \text{ odd}, \\ 2N \sum_{r=1}^{(N-2)/2} g\left(\frac{r}{N}\right) + Ng\left(\frac{N}{2}\right), & N \text{ even} \end{cases} \end{aligned}$$

and by Lemma 2.2 thus

$$\begin{aligned} \frac{1}{N^2 - N} \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} \sum_{i=1}^d g(y_i^k - y_i^\ell) &\geq \frac{d}{N-1} \begin{cases} 2 \sum_{k=1}^{(N-1)/2} g\left(\frac{k}{N}\right), & N \text{ odd}, \\ 2 \sum_{k=1}^{(N-2)/2} g\left(\frac{k}{N}\right) + g\left(\frac{N}{2}\right), & N \text{ even} \end{cases} \\ &= \sum_{r=1}^{\lfloor N/2 \rfloor} \# \left\{ i = 1, \dots, d : \rho_i = \frac{r}{N} \right\} \cdot g\left(\frac{r}{N}\right), \end{aligned}$$

where $\{(\rho_1, \dots, \rho_d)\} = \text{dv}(X)$ is the unique distance vector in X . In total we get

$$E_F(Y) \geq (N^2 - N) \exp \left(\sum_{i=1}^d g(\rho_i) \right) = E_F(X)$$

by (2.2), finishing the proof. $\square$

We may readily apply this result to the potential (1.4) arising from QMC integration. As was already observed in [23], the potential f_p is logarithmically convex as long as $p \leq 6$. Combining condition (i) of Proposition 2.1 and Theorem 2.3 we obtain the following.

Corollary 2.4. Let $X \subseteq \mathbb{T}^d$, $\#X = N$ be a Latin hypercube set with a single distance vector. Then X is a minimizer for the energy given by the potential

$$f_p(t) = 1 + \frac{p}{2} \left(\frac{1}{6} - t + t^2 \right)$$

for all $0 < p \leq 6$.

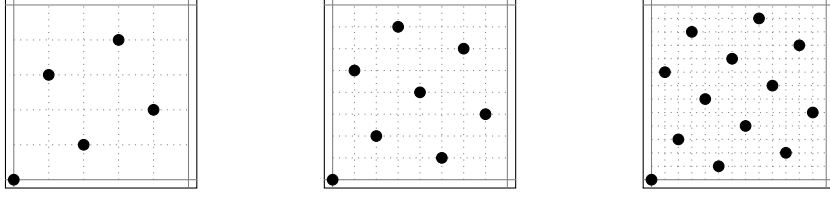


Figure 1: The 5-, 8- and 13-point Fibonacci lattices.

Of course, Theorem 2.3 is only valuable if we can demonstrate that sufficiently many sets of this form exist. We dedicate the next three subsections to constructing a large number of such sets, starting with the case $d = 2$ with particular emphasis on $N = 5$, the largest N for which there is such a set in $\mathbb{T}^2$ (see (2.3)).

2.4 Optimality of the 3- and 5-point Fibonacci lattices

We observe that the 5-point Fibonacci lattice

$$\Phi_5 = \left\{ \left(\frac{0}{5}, \frac{0}{5} \right), \left(\frac{1}{5}, \frac{3}{5} \right), \left(\frac{2}{5}, \frac{1}{5} \right), \left(\frac{3}{5}, \frac{4}{5} \right), \left(\frac{4}{5}, \frac{2}{5} \right) \right\}$$

has a single distance vector $(\frac{1}{5}, \frac{2}{5})$, see Figure 1, in particular $E_F(\Phi_5) = 20f(\frac{1}{5})f(\frac{2}{5})$ as in (2.2). Therefore, Theorem 2.3 applies in this case and we immediately obtain optimality of this configuration for a variety of potentials.

Corollary 2.5. If $\log f$ is 5-equispacing, then the 5-point Fibonacci lattice Φ_5 is an energy minimizer for the corresponding energy E_F . In particular, this holds if (see Proposition 2.1)

- f is decreasing and logarithmically convex on $(0, 1/2]$; or
- $f(t) = h(\cos(2\pi t))$ and $\log h$ is absolutely monotone on $[-1, 1)$.

We further observe that the 5-point Fibonacci lattice has the lowest energy among all other 5-point permutation sets merely under the assumption that f is decreasing on $(0, 1/2]$.

Lemma 2.6. Assume that f is decreasing on $(0, 1/2]$ (or just that $f(2/5) < f(1/5)$). Then the 5-point Fibonacci lattice Φ_5 minimizes the energy E_F among all 5-point permutation sets.

Proof. We note that permutation sets on $\mathbb{T}^2$ project down to equispaced points on $\mathbb{T}^1$ in both the x - and y -coordinates. The one-dimensional energy

of either of these projections is $5f(1/5) + 5f(2/5)$, so the energy of any permutation set in $\mathbb{T}^2$ must be some sum of the form $\sum_{i=1}^{10} a_i b_i$ where five of the a_i are $f(1/5)$, five of the a_i are $f(2/5)$, five of the b_i are $f(1/5)$ and five of the b_i are $f(2/5)$. By the rearrangement inequality [21] we note then that for any permutation set X ,

$$E_F(X) \geq \sum_{i=1}^{10} f(1/5)f(2/5) = 10f(1/5)f(2/5),$$

with equality only holding when $f(1/5) = f(2/5)$. Since $10f(1/5)f(2/5)$ is the energy of the 5-point Fibonacci lattice, we have the desired result. $\square$

We also notice that the 3-point Fibonacci lattice is equivalent, up to symmetries, to the diagonal lattice, consisting of points $(0, 0)$, $(\frac{1}{3}, \frac{1}{3})$, and $(\frac{2}{3}, \frac{2}{3})$, and has a single distance vector $(\frac{1}{3}, \frac{1}{3})$. Thus Theorem 2.3 also applies in this case and yields an identical corollary:

Corollary 2.7. If $\log f$ is 3-equispacing, then the 3-point Fibonacci lattice Φ_4 minimizes the tensor product energy E_F .

A statement similar to this has already been observed in [29] and a different proof of this fact in the case when f is logarithmically convex is given in §4.

Finally we would like to mention some other previously known results about the Fibonacci lattice and compare them to our results above. All of the prior results concerned specific potentials arising from QMC integration.

- It was proved in [29] that for the potential f_p in (1.4) with $0 < p \leq 9$ the 5-point Fibonacci lattice minimizes $E_F(X)$ among all 5-point configurations. For this particular potential, Corollary 2.4 yields a smaller range of $p \leq 6$. We do however want to point out that the method to prove Theorem 2.3 is completely elementary and does not rely on any computer assisted calculations.
- Let again f_p be the potential as in (1.4). Via an optimization algorithm it was shown in [23] that the N -point Fibonacci lattice for $N \in \{8, 13\}$ (which has multiple distance vectors, see Figure 1) is the energy minimizer for $p \in \{1, 6\}$ among all N -point configurations.

2.5 Classification of lattices with a single distance vector

In this section we will classify all lattices with a single distance vector. For small values of N this can be checked straightforwardly. Note that for each

$N = 2, 3, 4$ there is, up to torus symmetry, only one lattice in $\mathbb{T}^d$ with N points: the corresponding generating vector is

$$\left(\frac{1}{N}, \dots, \frac{1}{N}\right).$$

The lattices with $N = 2, 3$ points are easily seen to have a single distance vector, while the case of $N = 4$ defines precisely two distance vectors.

Theorem 2.8. A lattice $\Lambda \subseteq [0, 1)^d$, $\#\Lambda = N \geq 3$ has a single distance vector if and only if N is prime and the generating vector of Λ is of the form (up to torus symmetry)

$$\left(\underbrace{\frac{1}{N}, \dots, \frac{1}{N}}_{m \text{ times}}, \underbrace{\frac{2}{N}, \dots, \frac{2}{N}}_{m \text{ times}}, \dots, \underbrace{\frac{n}{N}, \dots, \frac{n}{N}}_{m \text{ times}}\right)$$

with $N = 2n + 1$ and $d = mn$.

Proof. Let $\Lambda \subseteq \mathbb{T}^d$ be a lattice with a single distance vector such that $\#\Lambda = N \geq 3$. By applying torus symmetries we may assume the generating vector to have the form

$$\mathbf{h} = \left(\frac{h_1}{N}, \dots, \frac{h_d}{N}\right)$$

with $1 = h_1 \leq \dots \leq h_d \leq \lfloor N/2 \rfloor$. In particular, the distance vector among points in Λ must be of the form

$$\text{dv}(\mathbf{0}, \mathbf{h}) = \left(\frac{h_1}{N}, \dots, \frac{h_d}{N}\right),$$

by Lemma 2.2 thus

$$\left(\frac{h_1}{N}, \dots, \frac{h_d}{N}\right) = \left(\underbrace{\frac{1}{N}, \dots, \frac{1}{N}}_{m \text{ times}}, \underbrace{\frac{2}{N}, \dots, \frac{2}{N}}_{m \text{ times}}, \dots, \underbrace{\frac{n}{N}, \dots, \frac{n}{N}}_{m \text{ times}}\right)$$

with $n = \frac{N-1}{2} \in \mathbb{N}$ and $m = \frac{2d}{N-1}$. Furthermore, N must be relatively prime to $1, \dots, \lfloor N/2 \rfloor$, so that it must already be prime itself.

It remains to show that any lattice with generating vector

$$\left(\underbrace{\frac{1}{N}, \dots, \frac{1}{N}}_{m \text{ times}}, \underbrace{\frac{2}{N}, \dots, \frac{2}{N}}_{m \text{ times}}, \dots, \underbrace{\frac{n}{N}, \dots, \frac{n}{N}}_{m \text{ times}}\right)$$

with $d = mn$ and $N = 2n + 1$ prime is a lattice with a single distance vector. Indeed, for an arbitrary $k \in \{1, \dots, N - 1\}$ consider the distance vector from $\mathbf{0} = (0, \dots, 0)$ to

$$\mathbf{x}^k = \left(\left\{ \frac{h_1 k}{N} \right\}, \dots, \left\{ \frac{h_d k}{N} \right\} \right)$$

with $h_1 = \dots = h_m = 1, h_{m+1} = \dots = h_{2m} = 2, \dots, h_{(n-1)m+1} = \dots = h_{nm} = n$. We will show that this distance vector does not depend on k . If there were $h, \tilde{h} \in \{1, \dots, n\}, h \neq \tilde{h}$ with

$$\left\{ \frac{hk}{N} \right\} = \left\{ \frac{\tilde{h}k}{N} \right\},$$

then we would either get

$$hk \equiv \tilde{h}k \pmod{N} \Rightarrow h = \tilde{h}$$

or

$$hk \equiv N - \tilde{h}k \pmod{N} \Rightarrow (h + \tilde{h})k \equiv 0 \pmod{N} \Rightarrow h + \tilde{h} = N,$$

both leading to contradictions since $h \neq \tilde{h}$ and $h + \tilde{h} \leq 2n < N$. Thus

$$\text{dv}(\mathbf{0}, \mathbf{x}^k) = \left(\underbrace{\frac{1}{N}, \dots, \frac{1}{N}}_{m \text{ times}}, \underbrace{\frac{2}{N}, \dots, \frac{2}{N}}_{m \text{ times}}, \dots, \underbrace{\frac{n}{N}, \dots, \frac{n}{N}}_{m \text{ times}} \right)$$

and by the lattice structure this is also the distance vector between any other two point from the lattice. $\square$

We see that if $2d + 1$ is prime, the bound on N mentioned after (2.3) can be attained by such a lattice construction.

2.6 Some non-lattice sets with a single distance vector

In addition to the lattices from the previous section, we can also construct the following 4-, 6-, 8-, and 9-point Latin hypercube sets with a single distance vector in $\mathbb{T}^3$, $\mathbb{T}^{10}$, $\mathbb{T}^{14}$, and $\mathbb{T}^{12}$ respectively:

$$X_4 = \left\{ \frac{1}{4}(0, 0, 0), \frac{1}{4}(1, 1, 2), \frac{1}{4}(2, 3, 1), \frac{1}{4}(3, 2, 3) \right\},$$

$$X_6 = \left\{ \frac{1}{6}(0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \frac{1}{6}(1, 1, 1, 1, 2, 2, 2, 2, 3, 3), \right.$$

$$\begin{aligned}
& \frac{1}{6}(2, 2, 2, 3, 3, 4, 5, 5, 1, 1), \frac{1}{6}(3, 3, 4, 2, 5, 1, 1, 4, 4, 5), \\
& \frac{1}{6}(4, 5, 3, 4, 1, 3, 4, 1, 5, 4), \frac{1}{6}(5, 4, 5, 5, 4, 5, 3, 3, 2, 2) \Big\}, \\
\\
X_8 = & \left\{ \frac{1}{8}(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \frac{1}{8}(1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4), \right. \\
& \frac{1}{8}(2, 2, 3, 5, 1, 4, 4, 5, 1, 6, 6, 7, 1, 3), \frac{1}{8}(3, 3, 2, 4, 3, 6, 6, 7, 6, 1, 1, 4, 5, 7), \\
& \frac{1}{8}(4, 5, 6, 3, 5, 1, 5, 1, 2, 4, 7, 1, 2, 6), \frac{1}{8}(5, 4, 7, 2, 7, 3, 7, 3, 5, 7, 4, 6, 6, 2), \\
& \left. \frac{1}{8}(6, 7, 5, 6, 4, 5, 1, 4, 7, 2, 5, 2, 3, 1), \frac{1}{8}(7, 6, 4, 7, 6, 7, 3, 6, 4, 5, 2, 5, 7, 5) \right\}, \\
\\
X_9 = & \left\{ \frac{1}{9}(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0), \frac{1}{9}(1, 1, 1, 2, 2, 2, 3, 3, 3, 4, 4, 4), \right. \\
& \frac{1}{9}(2, 2, 2, 4, 4, 4, 6, 6, 6, 8, 8, 8), \frac{1}{9}(3, 4, 5, 1, 6, 8, 1, 4, 7, 2, 3, 7), \\
& \frac{1}{9}(4, 3, 8, 7, 8, 6, 2, 8, 5, 5, 7, 3), \frac{1}{9}(5, 6, 4, 8, 1, 3, 4, 7, 1, 7, 2, 6), \\
& \frac{1}{9}(6, 5, 7, 5, 3, 1, 5, 2, 8, 1, 6, 2), \frac{1}{9}(7, 8, 3, 6, 5, 7, 7, 1, 4, 3, 1, 5), \\
& \left. \frac{1}{9}(8, 7, 6, 3, 7, 5, 8, 5, 2, 6, 5, 1) \right\}.
\end{aligned}$$

These point sets are again energy minimizers in the sense of Theorem 2.3. It can be observed that X_4 and X_8 are digital nets [14].

It is easy to see that for every $N \in \mathbb{N}$ there exists a Latin hypercube set $X \subseteq \mathbb{T}^d$, $\#X = N$ with a single distance vector in dimension $d = (N - 1)!$, namely where every coordinate of $\mathbb{T}^{(N-1)!}$ corresponds to a permutation σ of $\{0, 1, \dots, N - 1\}$ fixing $\sigma(0) = 0$. It is thus more interesting to find, for any given $N \in \mathbb{N}$, the smallest dimension $d = d(N)$ for which there exists an N -point Latin hypercube set with a single distance vector in $\mathbb{T}^d$. The condition (2.3) is necessary but not sufficient.

3 Energy minimizers have distinct coordinates

In the present section we demonstrate that, under very mild assumptions on the function f , both global and local energy minimizers must necessarily

have distinct coordinates. For example, in the two-dimensional case $\mathbb{T}^2$ this implies no two points in the minimizing configuration can lie on the same horizontal or vertical line. This brings us one step closer to understanding when permutation sets or Latin hypercube sets minimizer tensor product energies. The result is formalized in the following theorem:

Theorem 3.1. Let $f : \mathbb{R} \rightarrow (0, \infty)$ be continuous, 1-periodic, even and satisfy the following assumptions:

- (i) $f(t)$ is differentiable on $(0, 1)$,
- (ii) there are constants $C > 0$ and $\varepsilon > 0$ such that $f(t) \leq f(0) - Ct$ for all $0 \leq t < \varepsilon$.

Consider an N -point configuration $X = \{\mathbf{x}^0, \dots, \mathbf{x}^{N-1}\} \subseteq \mathbb{T}^d$ such that $x_1^0 = \dots = x_1^j$ for some $j \geq 1$ and $x_1^k \neq x_1^0$ for all $k > j$ (if such k exist). Define a perturbation of X by

$$X_r := \{\mathbf{x}^0(r), \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\},$$

where $\mathbf{x}^0(r) = (x_1^0 + r, x_2^0, \dots, x_d^0)$ for $r \in [-1/2, 1/2)$. Then the function

$$r \mapsto E_F(X_r)$$

does not have a local minimum at $r = 0$.

As a direct consequence we obtain the following.

Corollary 3.2. Let f satisfy the assumptions of Theorem 3.1 and let $X = \{\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\} \subseteq \mathbb{T}^d$ be a (local) minimizer of the corresponding energy E_F . Then

$$x_i^k \neq x_i^\ell$$

for all $i = 1, \dots, d$ and $k, \ell = 0, 1, \dots, N-1, k \neq \ell$.

That is, points in minimizing configurations must necessarily have distinct coordinates.

Remark. Assumption (ii) in Theorem 3.1 (which means that f has a sharp kink at $t = 0$) can be interpreted as saying that the repulsion induced by the interaction F is quite strong at small distances in each coordinate. We observe that if f is singular at zero, i.e. $f(0) = +\infty$, the statement of Corollary 3.2 would have been obvious. It is a known phenomenon [8] that when the local repulsion is weak, then minimizers of energies tend to cluster. Thus, this result can be viewed as complimentary: minimizers do not cluster when the local repulsion is strong.

Observe that assumption (ii) is satisfied by various natural kernels, in particular the aforementioned $f(t) = 1/2 - t + t^2$ for $t \in [0, 1]$ which corresponds to the periodic L_2 -discrepancy, or more generally the QMC potential (1.4) for $0 < p < 24$ (where it stays positive).

To prove Theorem 3.1, we begin with a simple observation: if a function has a downwards-opening kink, then it cannot have a local minimum there.

Lemma 3.3. Let $G : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ be a function with $G(r) \leq G(0) - C|r|$ in $-\varepsilon < r < \varepsilon$ for some $C > 0$. Let $H : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ be differentiable at $r = 0$. Then $r \mapsto G(r) + H(r)$ does not have a local minimum at $r = 0$.

Proof of Theorem 3.1. Without loss of generality let $x_1^0 = \dots = x_1^j = 0$ be the common coordinate. Consider the perturbation $\mathbf{x}^0(r) = (r, x_2^0, \dots, x_d^0)$ while keeping all other coordinates fixed. We separate the terms of $E_F(X_r)$ that depend on r and treat the rest as a constant $K_1 > 0$

$$E_F(X_r) = f(r) \cdot 2 \sum_{k=1}^j \prod_{i=2}^d f(x_i^0 - x_i^k) + 2 \sum_{k=j+1}^{N-1} f(r - x_1^k) \prod_{i=2}^d f(x_i^0 - x_i^k) + K_1.$$

Let

$$G(r) = f(r) \cdot 2 \sum_{k=1}^j \prod_{i=2}^d f(x_i^0 - x_i^k)$$

and

$$H(r) = 2 \sum_{k=j+1}^{N-1} f(r - x_1^k) \prod_{i=2}^d f(x_i^0 - x_i^k) + K_1.$$

By (i) and since $x_1^k \neq 0$ for $k > j$ we have that $H(r)$ is differentiable at $r = 0$. Also note that

$$K_2 = 2 \sum_{k=1}^j \prod_{i=2}^d f(x_i^0 - x_i^k) > 0$$

is a positive constant since f only takes on positive values and $j \geq 1$ so that the sum is nonempty. Hence (ii) implies that

$$G(r) \leq K_2(f(0) - C|r|) = G(0) - CK_2|r|$$

in a neighborhood around $r = 0$. Thus G and H satisfy the assumptions of Lemma 3.3, hence $E_F(X_r) = G(r) + H(r)$ does not have a local minimum in $r = 0$. $\square$

4 Permutation type and winding number

In this section we show that lattices minimize tensor product energies with logarithmically convex potential f among large classes of sets which share structural similarities with the lattice. In particular, lattices are local minimizers of such energies. To this end, we prove a lower bound on the energy in terms of the permutation type of the configuration.

Given a point set in $\mathbb{T}^d \simeq [0, 1)^d$, we can always assume that the x_1 -coordinates are ordered, i.e. $x_1^k \leq x_1^\ell$ for $k < \ell$. However the same need not hold for the other coordinates. Instead when $d \geq 2$ we can order the points by one of the other coordinates and obtain a permutation that relates this ordering to our original ordering. We obtain the following definition:

Definition 4.1. For a given point set $X = \{\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\} \subset \mathbb{T}^d \simeq [0, 1)^d$, suppose $S = (\sigma_1, \sigma_2, \dots, \sigma_d)$ is a d -tuple of permutations of $\{0, 1, \dots, N-1\}$ satisfying

$$x_i^{\sigma_i^{-1}(0)} \leq x_i^{\sigma_i^{-1}(1)} \leq \dots \leq x_i^{\sigma_i^{-1}(N-1)}$$

for all i . Then we say S is a *permutation type* of X .

For example, we note that the Latin hypercube set $X_{\sigma_2, \dots, \sigma_d}$ given in Definition 1.2 has permutation type $(\text{id}, \sigma_2, \dots, \sigma_d)$. We will generally assume that the first coordinates are ordered, i.e. σ_1 is the identity permutation, and by translating so that $\mathbf{x}^0 = (0, 0, \dots, 0)$ we will also assume that $\sigma_i(0) = 0$ for all i . The permutation type is uniquely defined if and only if all points have distinct coordinates in the sense of §3. For any permutation σ we will define the k -th *winding number* associated with σ by

$$\omega_\sigma(k) := \#\{j : \sigma((j+k) \bmod N) < \sigma(j)\}.$$

For notational simplicity, input values outside of the domain of σ will be implicitly assumed to be taken modulo N . Conceptually, this quantity counts how many times the permutation winds around the torus, moving along it with step size k . With this definition we are now ready to state the lower bound for the energy.

Theorem 4.2. Assume the configuration $X = \{\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\} \subset \mathbb{T}^d$ has unique permutation type $(\sigma_1, \sigma_2, \dots, \sigma_d)$. Then the following inequality holds for all potentials f decreasing and logarithmically convex on $(0, 1/2]$:

$$E_F(\mathbf{x}^0, \dots, \mathbf{x}^{N-1}) \geq N \sum_{k=1}^{N-1} \prod_{i=1}^d f\left(\frac{\omega_{\sigma_i}(k)}{N}\right).$$

If f is strictly decreasing, equality is obtained if and only if the point set $X = \{\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{N-1}\}$ is a lattice.

Proof. By regrouping the terms, the energy can be written in the form

$$E_F(\mathbf{x}^0, \dots, \mathbf{x}^{N-1}) = \sum_{\ell=1}^{N-1} \sum_{k=0}^{N-1} F(\mathbf{x}^k - \mathbf{x}^{k+\ell}), \quad (4.1)$$

if we identify $\mathbf{x}^{k+\ell} = \mathbf{x}^{k+\ell-N}$ in the case $k + \ell > N - 1$. For $x, y \in [0, 1) \simeq \mathbb{T}$ we define

$$\delta_{\circlearrowleft}(x, y) = \begin{cases} y - x, & y \geq x, \\ 1 + y - x, & y < x, \end{cases}$$

which can be thought of as the *counterclockwise* geodesic distance. Note that this need not always equal the geodesic distance δ defined in (1.3), but the periodicity and evenness of f implies that the value of f is independent of the choice between these two distances as its argument. Thus we rewrite

$$F(\mathbf{x}^k - \mathbf{x}^{k+\ell}) = \prod_{i=1}^d f(x_i^k - x_i^{k+\ell}) = \prod_{i=1}^d f(\delta_{\circlearrowleft}(x_i^k, x_i^{k+\ell})).$$

Since f is logarithmically convex and decreasing on $(0, 1/2]$, it is log-convex on $(0, 1)$, and hence the function F is convex as a function of d variables on $(0, 1)^d$. Therefore, we can apply Jensen's inequality to the inner sum in the right-hand side of (4.1) and we obtain

$$\begin{aligned} E_F(\mathbf{x}^0, \dots, \mathbf{x}^{N-1}) &= \sum_{\ell=1}^{N-1} \sum_{k=0}^{N-1} F(\delta_{\circlearrowleft}(x_1^k, x_1^{k+\ell}), \dots, \delta_{\circlearrowleft}(x_d^k, x_d^{k+\ell})) \\ &\geq N \sum_{\ell=1}^{N-1} F\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{\circlearrowleft}(x_1^k, x_1^{k+\ell}), \dots, \frac{1}{N} \sum_{k=0}^{N-1} \delta_{\circlearrowleft}(x_d^k, x_d^{k+\ell})\right) \\ &= N \sum_{\ell=1}^{N-1} \prod_{i=1}^d f\left(\frac{1}{N} \sum_{k=0}^{N-1} \delta_{\circlearrowleft}(x_i^k, x_i^{k+\ell})\right). \end{aligned} \quad (4.2)$$

Observe that $\delta_{\circlearrowleft}(x_i^k, x_i^{k+\ell}) = 1 + x_i^{k+\ell} - x_i^k$ if and only if $x_i^{k+\ell} < x_i^k$, i.e. $x_i^{\sigma_i^{-1}(\sigma_i(k+\ell))} < x_i^{\sigma_i^{-1}(\sigma_i(k))}$, which is equivalent to $\sigma_i(k + \ell) < \sigma_i(k)$. This occurs for $\omega_{\sigma_i}(\ell)$ values of k (note that unique permutation type means that we can ignore the case of equality between coordinates). Otherwise it equals $x_i^{k+\ell} - x_i^k$. Telescoping the sum, we conclude that

$$\sum_{\ell=0}^{N-1} \delta_{\circlearrowleft}(x_i^k, x_i^{k+\ell}) = \omega_{\sigma_i}(k),$$

and thus, as desired,

$$E_F(\mathbf{x}^0, \dots, \mathbf{x}^{N-1}) \geq N \sum_{k=1}^{N-1} \prod_{i=1}^d f\left(\frac{\omega_{\sigma_i}(k)}{N}\right).$$

Finally, we note that there are two cases for equality in Jensen's inequality in (4.2): either f is linear on the relevant domain (which cannot be the case if f is strictly decreasing and logarithmically convex on $(0, 1/2]$) or the argument of F is independent of k , i.e.

$$\delta_{\circlearrowleft}(x_i^k - x_i^{k+\ell}) = \delta_{\circlearrowleft}(x_i^{k'} - x_i^{k'+\ell}) \text{ for all } i, \ell, k, k',$$

which by their periodic structure is fulfilled by lattices. Conversely, if this condition is satisfied we note that since a unique permutation type means none of the coordinates coincide, the $\ell = 1$ case implies that our configuration is a lattice. This proves the desired equivalence. $\square$

We immediately obtain the following corollary:

Corollary 4.3. Assume that the potential f is decreasing and logarithmically convex on $(0, 1/2]$. Let $X \subset \mathbb{T}^d$ be an N -point integration lattice, and let $Y \subset \mathbb{T}^d$ be an N -point configuration with the same permutation type as X . Then

$$E_F(Y) \geq E_F(X).$$

In other words, lattices minimize the energy among all sets with the same permutation type. Observing that a small perturbation of a lattice does not change the permutation type, we arrive directly at local optimality.

Corollary 4.4. If a point set X is a lattice and f is decreasing and logarithmically convex on $(0, 1/2]$, then X is a local minimum of E_F .

The last two corollaries exhibit an example of the behavior discussed in the introduction: certain permutation sets (namely, lattices) minimize the energy among large classes of more general (non-permutation) sets.

Finally we note that the case $N = 3$ can also be addressed by these methods because this case has only two permutation types in each coordinate, and all permutation sets in this case are lattices (up to the symmetries of the torus). Thus a special case of Corollary 2.7 when f is logarithmically convex, can be obtained as a consequence of Corollary 4.3.

5 The cases $N = 4$ and $N = 6$ in $\mathbb{T}^2$

In the setting of $\mathbb{T}^2$, the cases $N = 4$ and $N = 6$ are unique: the only possible lattices are the identity lattice that places all points on $y = x$, and its reflection that places all points on $y = 1 - x$. However, unlike in the case $N = 3$ (which like $N = 5$ is covered by Corollary 2.7) there are other permutation sets that are not one of these two cases.

We first observe that the permutation set whose permutation type is the identity cannot strictly minimize the energy, even among permutation sets.

Proposition 5.1. Let $N \in \mathbb{N}$ and let $f : \mathbb{T} \rightarrow [0, \infty)$ be a potential. Then, for all permutations σ of $\{0, 1, \dots, N-1\}$, we have

$$E_F(X_{\text{id}}) \geq E_F(X_\sigma),$$

where id denotes the identity permutation and X_σ is a permutation set with permutation type σ (see (1.6)).

Proof. The energy of X_σ takes the form

$$E_F(X_\sigma) = \sum_{\substack{k, \ell=0 \\ k \neq \ell}}^{N-1} f\left(\frac{k-\ell}{N}\right) f\left(\frac{\sigma(k)-\sigma(\ell)}{N}\right).$$

As multisets we have

$$\begin{aligned} & \left\{ f\left(\frac{k-\ell}{N}\right) : k, \ell \in \{0, 1, \dots, N-1\}, k \neq \ell \right\} \\ &= \left\{ f\left(\frac{\sigma(k)-\sigma(\ell)}{N}\right) : k, \ell \in \{0, 1, \dots, N-1\}, k \neq \ell \right\}, \end{aligned}$$

by the rearrangement inequality [21] the energy E_F is thus maximized if

$$f\left(\frac{k-\ell}{N}\right) = f\left(\frac{\sigma(k)-\sigma(\ell)}{N}\right)$$

for all k, ℓ , which is the case for the identity permutation. $\square$

Thus, for $N = 4$ and 6 , candidates for energy minimizers must be something other than a lattice. We shall consider these cases in detail.

5.1 The $N = 4$ case

In the case $N = 4$, up to symmetries of the torus, there is only one other permutation set besides the identity permutation, namely

$$X^{(4)} := \{(0, 0), (1/4, 1/2), (1/2, 1/4), (3/4, 3/4)\}, \quad (5.1)$$

which by Proposition 5.1 minimizes the energy among all 4-point permutation sets. If f is differentiable on $(0, 1)$, one can perform gradient descent to approach a local minimizer. Following off of an observation in [23] we claim that with appropriate convexity conditions this will actually yield a global minimizer, and we present a simple condition how to determine it.

Proposition 5.2. Let f be differentiable on $(0, 1)$, as well as decreasing and logarithmically convex on $(0, 1/2]$. Then there exists an $a \in (0, 1/2)$ such that the 4-point configuration

$$X_a^{(4)} := \{(0, 0), (a, 1/2), (1/2, a), (1/2 + a, 1/2 + a)\}$$

is a minimizer of E_F . This value of a satisfies

$$f'(a)f(1/2) = f'(1/2 - a)f(1/2 - a). \quad (5.2)$$

If f is strictly logarithmically convex, this value of a is unique.

We would like to point out that in [23] such a set was obtained as the energy minimizer of E_F with $f(t) = \frac{1}{2} - t + t^2$ (Fig. 2) using a proof assisted with numerics. Proposition 5.2 provides a simple way to find such minimizers for a wide range of potentials and admits a transparent proof, which we now present.

Proof. We note that f being logarithmically convex results in the tensor product potential F being convex on $(0, 1) \times (0, 1)$, see [31]. This means that as a function of $\mathbf{x}^i$ and $\mathbf{x}^j$, the expression $F(\mathbf{x}^i - \mathbf{x}^j)$ is convex when the order of the x - and y -coordinates of the two points is fixed. Since the sum of convex functions is convex, we conclude that E_F is convex as a function of the coordinates of our three unfixed points (we assume $\mathbf{x}^0 = (0, 0)$) when we restrict it to point configurations of a specific permutation type. Viewing the coordinates of our three unfixed points as an element of $[0, 1]^6$, we observe that a convex combination of two element in this space that are of the same permutation type will result in another element of the same permutation type. Therefore the set of all configurations of the same permutation type is a convex set and thus any local minimizer of E_F in this set is a minimizer among all point sets of the same permutation type.

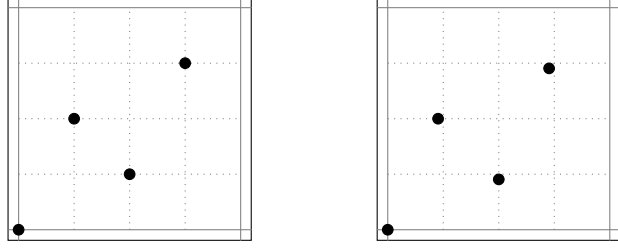


Figure 2: The best 4-point permutation set and the energy minimizer $X_a^{(4)}$ for the potential $f(t) = \frac{1}{2} - t + t^2$. Here, $a = 0.22669\dots$ is the solution to the cubic equation $2a^3 + a - \frac{1}{4} = 0$.

From Corollary 4.3 we know that the identity lattice minimizes the energy among sets of the same permutation type, however by Proposition 5.1 it will not be a (unique) minimizer among all 4-point configurations in $\mathbb{T}^2$. Considering torus symmetries, we need to check only one other permutation type, namely the one of $X_a^{(4)}$. Thus, we only need to show that, for a satisfying (5.2), $X_a^{(4)}$ is stationary with respect to E_F in order to demonstrate that it is an energy minimizer. For this, we consider the gradient of the energy functional $E_F(\mathbf{0}, \mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3)$ in terms of the 6 variables $\mathbf{x}^1 = (x_1^1, x_2^1)$, $\mathbf{x}^2 = (x_1^2, x_2^2)$, $\mathbf{x}^3 = (x_1^3, x_2^3)$. Taking the partial derivative with respect to any of these 6 coordinates, we see that it is 0 if and only if (5.2) holds. We thus conclude that $X_a^{(4)}$ is a minimizer. We know such an a exists because

$$E_F(X_a^{(4)}) = 4f(a)f(1/2) + 2f(1/2 - a)^2$$

is a convex function in a and thus must have a minimum (the cases of $a = 0$ and $a = 1/2$ are not minimizers by Corollary 3.2). If f is strictly logarithmically convex then $E_F(X_a^{(4)})$ is a strictly convex function of a and has a unique minimizer, so (5.2) is satisfied by only one value of a . $\square$

With this result we reduce the complexity of the problem down to a single variable, making it much more approachable for potentials satisfying the given condition. If we consider the potential given in (1.4), for example, we find that the energy minimizer is $X_a^{(4)}$ with $a = a(p)$ satisfying the cubic

$$\frac{p^2}{2}a^3 + 2p\left(1 - \frac{p}{24}\right)a - \frac{p}{2}\left(1 - \frac{p}{24}\right) = 0.$$

It can be noted that for $0 < p \leq 6$, where (1.4) is logarithmically convex, we have $|a(p) - \frac{1}{4}| < 0.024$. This phenomenon, that global optimizers do not stray too far from permutation sets, was first observed in [23]. Figure 2 shows

the global minimizer for the concrete potential of the periodic L_2 -discrepancy, equivalently $p = 6$ above.

Remark 5.3. We finish this subsection by noting that the permutation set (5.1), i.e. $X_{1/4}^{(4)}$, will only satisfy the condition given in (5.2) if $f(1/2) = f(1/4)$, in which case it will also yield the same energy as the identity permutation. Thus, for potentials that satisfy the conditions of Proposition 5.2, there cannot be a unique energy minimizer that is a permutation set.

5.2 The $N = 6$ case

In the $N = 6$ case there are many more possible permutation sets, with a total of 9 cases when taking symmetries into account, see Figure 3. However, assuming that f is decreasing on $(0, 1/2]$ (so that $f(1/6) \geq f(1/3) \geq f(1/2)$) and convex (so that $2f(1/3) \leq f(1/6) + f(1/2)$) it can be shown with a direct calculation that, up to symmetries, permutation set with the lowest energy is

$$X^{(6)} = \{(0, 0), (1/6, 1/3), (1/3, 2/3), (1/2, 1/6), (2/3, 5/6), (5/6, 1/2)\},$$

which up to symmetries is the only permutation set that does not have $(1/6, 1/6)$ as a distance vector.

We cannot deduce the global optimality of $X_a^{(6)}$ as in Proposition 5.2 because unlike the $N = 4$ case, we have multiple non-lattice permutation types that remain distinct under reflection, rotation, and translation. However, we can construct a configuration which is optimal among all sets with the same permutation type as $X^{(6)}$.

Proposition 5.4. Assume f is differentiable on $(0, 1)$, as well as decreasing and logarithmically convex on $(0, 1/2]$. Then there exists $a \in (0, 1/4)$ such that the point configuration

$$X_a^{(6)} := \{(0, 0), (a, 1/2 - a), (2a, 1 - 2a), (1/2, 1/2 - 2a), \\ (a + 1/2, 1 - a), (1/2 + 2a, 1/2)\}$$

is a minimizer of E_F among all 6-point configurations with the same permutation type. This value of a satisfies

$$f'(a)f(1/2 - a) + f'(2a)f(2a) = f'(1/2 - a)f(a) + f'(1/2 - 2a)f(1/2). \quad (5.3)$$

If f is strictly logarithmically convex, this value of a is unique.

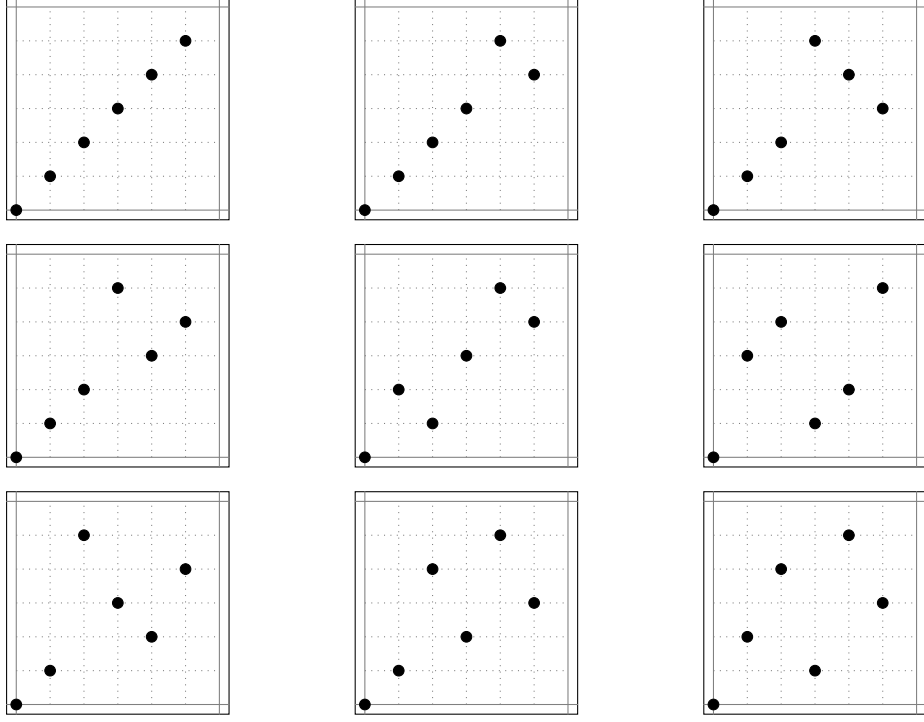


Figure 3: The 9 possible distinct permutation sets in approximate order of decreasing energy, with the best possible configuration last.

Proof. Just as in the $N = 4$ case we note that since f is logarithmically convex and differentiable, we can find a minimizer given a specific permutation type by finding a configuration of the permutation type that is stationary with respect to the energy E_F in the x - and y -coordinates of every point. It is straightforward to calculate the 10 partial derivatives with respect to the coordinates of $\mathbf{x}^1 = (x_1^1, x_2^1), \dots, \mathbf{x}^5 = (x_1^5, x_2^5)$. Setting these partial derivatives to 0, one obtains condition (5.3) for $X_a^{(6)}$. We use the same argument as in the proof of Proposition 5.2 to claim that a minimizer exists, and that it is unique in the case that f is strictly logarithmically convex. $\square$

Unlike the case $N = 4$, condition (5.3) can be conceivably satisfied by the permutation set $X_{1/6}^{(6)}$ and in this case the condition becomes

$$f'(1/6)(f(1/3) - f(1/2)) = f'(1/3)(f(1/6) - f(1/3)).$$

Nevertheless, for the majority of potentials this equality will not be satisfied and thus in general the energy minimizer for $N = 6$ will not be a permutation set, see Figure 4.

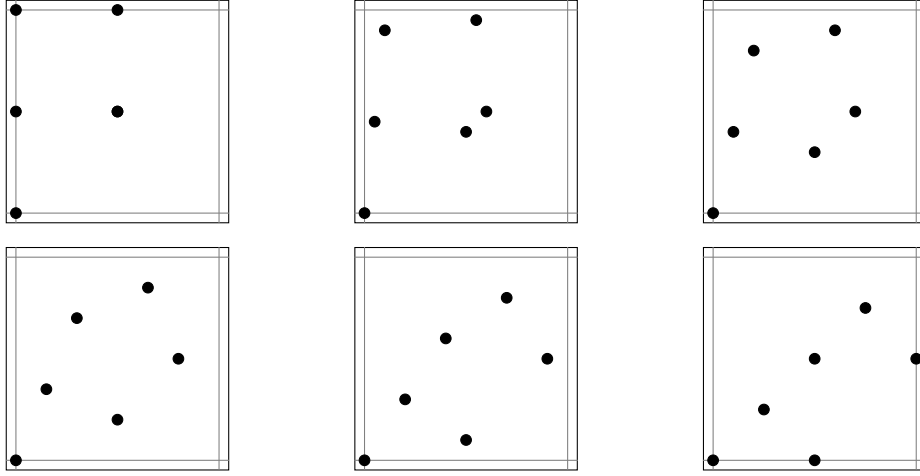


Figure 4: $X_a^{(6)}$ for $a \in \{0, 0.05, 0.1, 0.15, 0.2, 0.25\}$ (with the double points $(0.5, 0.5)$ and $(0, 0) \simeq (0, 1)$ for $a = 0$). For the potential $f(t) = \frac{1}{2} - t + t^2$ the minimizer is attained for $a = 0.16375 \dots$ solving $20a^3 - 15a^2 + \frac{13}{2}a - \frac{3}{4} = 0$.

Acknowledgment

DB was supported by the Simons Travel Support for Mathematicians and the Fulbright-NAWI Graz Visiting Professor fellowship. NN was supported by the ESF Plus Young Researchers Group ReSIDA-H2 (SAB, 10064931). We would also like to thank Alexey Glazyrin for fruitful discussions.

Appendix

A A collection of potentials

In this section we collect some potentials which are relevant in the context of tensor product energy. Natural examples arise from discrepancy theory, quasi-Monte Carlo integration and harmonic analysis. We also give other examples which might shed light on optimal point configurations.

Quasi-Monte Carlo integration. Let $s \in \mathbb{N}$ and let $a_{-1}, a_0, a_1, \dots, a_s \geq 0$ be parameters with $a_s > 0$ and $a_{-1} + a_0 > 0$. Let $H_{\mathbf{a}}^s(\mathbb{T})$ be the Sobolev space of periodic functions $f : \mathbb{T} \rightarrow \mathbb{R}$ such that $f^{(k)} \in L_2(\mathbb{T})$ for all weak derivatives of order $k = 0, 1, \dots, s$, equipped with the norm (which is a norm

by the assumptions on the parameters a_i)

$$\|f\|_{H_{\mathbf{a}}^s(\mathbb{T})}^2 := a_{-1} \left(\int_{\mathbb{T}} f(x) dx \right)^2 + \sum_{k=0}^s a_k \int_{\mathbb{T}} f^{(k)}(x)^2 dx.$$

This turns $H_{\mathbf{a}}^s(\mathbb{T})$ into a reproducing kernel Hilbert space with kernel given by

$$K_{\mathbf{a}}(x, y) = \sum_{m \in \mathbb{Z}} \frac{\exp(2\pi i m(x - y))}{a_{-1} \delta_{m,0} + \sum_{k=0}^s a_k (2\pi m)^{2k}}.$$

For d dimensions we simply take a d -fold tensor product $H_{\mathbf{a}}^s(\mathbb{T}^d) = H_{\mathbf{a}}^s(\mathbb{T}) \otimes \cdots \otimes H_{\mathbf{a}}^s(\mathbb{T})$, whose kernel can be given by

$$K_{\mathbf{a}}^d(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^d K_{\mathbf{a}}(x_i, y_i).$$

Spaces like these are sometimes also called *Korobov spaces* [13, 15, 18, 26], although note that some authors use this term for different spaces [16, 24, 25]. For $X \subseteq \mathbb{T}^d$, $\#X = N$ the worst case error of the quasi-Monte Carlo rule can be determined to be [30]

$$\sup_{\|f\|_{H_{\mathbf{a}}^s(\mathbb{T}^d)} \leq 1} \left| \frac{1}{N} \sum_{\mathbf{x} \in X} f(\mathbf{x}) - \int_{\mathbb{T}^d} f(\mathbf{x}) d\mathbf{x} \right|^2 = -\frac{1}{(a_{-1} + a_0)^d} + \frac{1}{N^2} \sum_{\mathbf{x}, \mathbf{y} \in X} K_{\mathbf{a}}^d(\mathbf{x}, \mathbf{y}).$$

Minimizing the worst case error is thus equivalent to minimizing

$$\sum_{\substack{\mathbf{x}, \mathbf{y} \in X \\ \mathbf{x} \neq \mathbf{y}}} K_{\mathbf{a}}^d(\mathbf{x}, \mathbf{y}),$$

which, since the kernel is of the form $K_{\mathbf{a}}^d(\mathbf{x}, \mathbf{y}) = \prod_{i=1}^d f_{\mathbf{a}}(x_i - y_i)$, can be interpreted as a tensor product energy. A particular example would be, for $p > 0$ a parameter, $a_{-1} = 1$, $a_s = p^{-1}$ and $a_0 = a_1 = \cdots = a_{s-1} = 0$, for which the kernel is of the form

$$K_{\mathbf{a}}(x, y) = 1 + (-1)^{s-1} \frac{p}{(2s)!} B_{2s}(x - y),$$

where B_{2s} is a Bernoulli polynomial of degree $2s$, given by

$$B_{2s}(t) = -(2s)! \sum_{m \neq 0} \frac{\exp(2\pi i m t)}{(2\pi i m)^{2s}}. \quad (\text{A.1})$$

The case of $s = 1$ and $p \in \{1, 6\}$ was considered in [23]. More generally, these potentials were studied for $s = 1$ in [29]. Similar but slightly different potentials were considered in [7].

(Smooth) Discrepancy. Discrepancy quantifies how uniform a set $X \subseteq \mathbb{T}^d$ is in a geometric way. We describe the definition here and show its relation to the quasi-Monte Carlo method. Let $s \in \mathbb{N}$ and for $x \in \mathbb{R}$ define the B-splines

$$B_{0,1}^1(x) := \chi_{[0,1)}(x)$$

and recursively

$$B_{0,1}^s(x) := \int_{x-1}^x B_{0,1}^{s-1}(u) \, du = \int_0^1 B_{0,1}^{s-1}(x-u) \, du.$$

Note that $\text{supp } B_{0,1}^s = [0, s]$. With $(x)_+ := \max\{0, x\}$ we even have explicitly

$$B_{0,1}^s(x) = \frac{1}{(s-1)!} \sum_{k=0}^s (-1)^k \binom{s}{k} (x-k)_+^{s-1}.$$

For $0 \leq \alpha < 1$ and $0 < \delta \leq 1$ set

$$B_{\alpha,\delta}^s(x) := B_{0,1}^s\left(\frac{x-\alpha}{\delta}\right) = \delta^{-1} \int_0^\delta B_{\alpha,\delta}^{s-1}(x-u) \, du$$

with $\text{supp } B_{\alpha,\delta}^s = [\alpha, \alpha + s\delta]$. For $x \in [0, 1) \simeq \mathbb{T}$ introduce the periodization

$$\tilde{B}_{\alpha,\delta}^s(x) := \sum_{k \in \mathbb{Z}} B_{\alpha,\delta}^s(k+x) = \delta^{-1} \int_0^\delta \tilde{B}_{\alpha,\delta}^{s-1}(x-u) \, du,$$

where $x-u \in \mathbb{T}$ with wrap-around on the torus. Finally, let

$$\varphi_{\alpha,\delta}^s(x) := \delta^{s-1} \tilde{B}_{\alpha,\delta}^s(x) = \int_0^\delta \varphi_{\alpha,\delta}^{s-1}(x-u) \, du. \quad (\text{A.2})$$

It is easy to verify that

- $\varphi_{\alpha,\delta}^1(x) = \chi_{[\alpha, \{\alpha+\delta\})}(x)$ in the sense of periodic intervals [22], where $\{\cdot\}$ denotes the fractional part,
- $\varphi_{\alpha,1}^s(x) \equiv 1$ and
- $\int_0^1 \varphi_{\alpha,\delta}^s(x) \, dx = \delta^s$.

Let us collect some facts about these functions $\varphi_{\alpha,\delta}^s$.

Lemma A.1. We have the Fourier representation (with convergence in L_2)

$$\varphi_{\alpha,\delta}^s(x) = \delta^s + \sum_{m \neq 0} \left(\frac{1 - \exp(-2\pi i m \delta)}{2\pi i m} \right)^s \exp(2\pi i m(x - \alpha)). \quad (\text{A.3})$$

Furthermore, for all $x, y \in [0, 1)$ it holds

$$\int_0^1 \int_0^1 \varphi_{\alpha,\delta}^s(x) \varphi_{\alpha,\delta}^s(y) d\delta d\alpha = \frac{1}{2s+1} + \frac{(-1)^{s-1}}{(s!)^2} B_{2s}(|x - y|). \quad (\text{A.4})$$

Proof. We prove (A.3) via induction on s . By periodicity we may assume $\alpha = 0$. For $s = 1$ we have $\varphi_{0,\delta}^1(x) = \chi_{[0,\delta)}(x)$ with Fourier coefficients

$$\int_0^1 \varphi_{0,\delta}^1(x) \exp(-2\pi i m x) dx = \begin{cases} \delta, & m = 0, \\ \frac{1 - \exp(-2\pi i m \delta)}{2\pi i m}, & m \neq 0. \end{cases}$$

Assuming the formula holds for $s - 1$, we have (using periodicity)

$$\begin{aligned} \int_0^1 \varphi_{0,\delta}^s(x) \exp(-2\pi i m x) dx &= \int_0^\delta \int_0^1 \varphi_{0,\delta}^{s-1}(x - u) \exp(-2\pi i m x) dx du \\ &= \int_0^\delta \exp(-2\pi i m u) du \int_0^1 \varphi_{0,\delta}^{s-1}(x) \exp(-2\pi i m x) dx \\ &= \begin{cases} \delta^s, & m = 0, \\ \left(\frac{1 - \exp(-2\pi i m \delta)}{2\pi i m} \right)^s, & m \neq 0. \end{cases} \end{aligned}$$

This shows (A.3). We use this Fourier representation to show the second formula (A.4). Indeed, we can then write

$$\begin{aligned} &\int_0^1 \int_0^1 \varphi_{\alpha,\delta}^s(x) \varphi_{\alpha,\delta}^s(y) d\delta d\alpha \\ &= \frac{1}{2s+1} + \sum_{m \neq 0} \left(\int_0^1 \left| \frac{1 - \exp(-2\pi i m \delta)}{2\pi i m} \right|^{2s} d\delta \right) \exp(2\pi i m(x - y)). \end{aligned}$$

Since for $t \in \mathbb{R}$ it holds

$$|1 - \exp(-it)|^2 = -\exp(-it)(1 - \exp(it))^2,$$

for $t = 2\pi m \delta$ we have

$$\int_0^1 |1 - \exp(-2\pi i m \delta)|^{2s} d\delta = (-1)^s \int_0^1 \exp(-2\pi i s m \delta) (1 - \exp(2\pi i m \delta))^{2s} d\delta$$

$$\begin{aligned}
&= \sum_{k=0}^{2s} (-1)^{s+k} \binom{2s}{k} \int_0^1 \exp(2\pi i(k-s)m\delta) d\delta \\
&= \binom{2s}{s}.
\end{aligned}$$

This finally leads to the desired result (using (A.1))

$$\begin{aligned}
\int_0^1 \int_0^1 \varphi_{\alpha,\delta}^s(x) \varphi_{\alpha,\delta}^s(y) d\delta d\alpha &= \frac{1}{2s+1} + \binom{2s}{s} \sum_{m \neq 0} \frac{\exp(2\pi i m(x-y))}{(2\pi m)^{2s}} \\
&= \frac{1}{2s+1} + \frac{(-1)^{s-1}}{(s!)^2} B_{2s}(|x-y|).
\end{aligned}$$

□

The *s-smooth periodic L_2 -discrepancy* of a point set $X \subseteq \mathbb{T}^d$, $\#X = N$ will be defined as

$$L_{2,s}^{\text{per}}(X)^2 := \int_{[0,1]^d} \int_{(0,1]^d} \left| \frac{1}{N} \sum_{\mathbf{x} \in X} \prod_{i=1}^d \varphi_{\alpha_i, \delta_i}^s(x_i) - \int_{\mathbb{T}^d} \prod_{i=1}^d \varphi_{\alpha_i, \delta_i}^s(x_i) d\mathbf{x} \right|^2 d\delta d\alpha,$$

where $\underline{\alpha} = (\alpha_1, \dots, \alpha_d)$ and $\underline{\delta} = (\delta_1, \dots, \delta_d)$. From Lemma A.1 we conclude the following *Warnock-type formula* for periodic discrepancy (the case for $s = 1$ already appearing in [22, 23]).

Theorem A.2. It holds

$$\begin{aligned}
&L_{2,s}^{\text{per}}(X)^2 \\
&= -\frac{1}{(2s+1)^d} + \frac{1}{N^2} \sum_{\mathbf{x}, \mathbf{y} \in X} \prod_{i=1}^d \left(\frac{1}{2s+1} + \frac{(-1)^{s-1}}{(s!)^2} B_{2s}(|x_i - y_i|) \right).
\end{aligned}$$

While the general equivalence of QMC and (smooth) discrepancy is well-known [32], this explicit formula does not seem to have appeared in the literature so far. Comparing with $K_{\mathbf{a}}^d$ from the previous part, we see that this is equivalent to the QMC worst case error over the space $H_{\mathbf{a}}^s(\mathbb{T}^d)$ with parameters

$$a_{-1} = 2s+1, a_0 = a_1 = \dots = a_{s-1} = 0, a_s = \binom{2s}{s}^{-1}.$$

Diaphony. Diaphony [28, 35] of $X \subseteq \mathbb{T}^d$, $\#X = N$ is a quantity given by (without the square root)

$$\sum_{\mathbf{m} \in \mathbb{Z}} \left| \frac{1}{N} \sum_{\mathbf{x} \in X} \exp(2\pi i \mathbf{m}^\top \mathbf{x}) \right|^2 \prod_{i=1}^d \begin{cases} 1, & m_i = 0, \\ \frac{1}{2\pi m_i}, & m_i \neq 0. \end{cases}$$

This turns out to be the worst case error for the quasi-Monte Carlo rule over the space $H_{\mathbf{a}}^1(\mathbb{T}^d)$ with parameters $a_{-1} = 1, a_0 = 0, a_1 = 1$.

Tensorized Poisson kernel. The *Poisson kernel* with parameter $r \in (0, 1)$ is given by

$$P_r(t) = \sum_{m \neq 0} r^{|m|} \exp(2\pi i m t) = \frac{1 - r^2}{1 - 2r \cos(2\pi t) + r^2}$$

for $t \in \mathbb{T}$. Tensorization leads to the kernel

$$K_{P_r}^d(\mathbf{x}, \mathbf{y}) := \prod_{i=1}^d P_r(x_i - y_i),$$

For fixed r this can be seen as a kernel over a Sobolev space of infinite dominating mixed smoothness (with certain exponential decay of the Fourier coefficients).

Logarithmically convex potentials. We finish by mentioning some logarithmically convex potentials. Here k can be seen as a shape parameter and c is a parameter that controls the logarithmic convexity for a certain range.

- Polynomial potential

$$f(t) = 1 + c(t^k + (1-t)^k), \quad k > 1, 0 < c \leq k - 1$$

- Another polynomial potential

$$f(t) = 1 + c|2t - 1|^k, \quad k > 1, 0 < c \leq k - 1$$

- Potential via the entropy function

$$f(t) = 1 + c(t \ln t + (1-t) \ln(1-t)), \quad 0 < c \leq c_0 = 1.23325675 \dots$$

- Ellipse potential

$$f(t) = 1 - c\sqrt{t(1-t)}, \quad 0 < c \leq (3/2)^{3/2} = 1.8371173 \dots$$

References

- [1] Sergiy V. B., D. P. Hardin, and E. B. Saff. *Discrete Energy on Rectifiable Sets*. Springer, 2019.
- [2] J. Baldeaux and J. Dick. QMC rules of arbitrary high order: reproducing kernel Hilbert space approach. *Constr. Approx.*, 30(3):495–527, 2009.
- [3] A. Berlinet and C. Thomas-Agnan. *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Springer New York, NY, First edition, 2003.
- [4] D. Bilyk, V. N. Temlyakov, and R. Yu. Fibonacci sets and symmetrization in discrepancy theory. *Journal of Complexity*, 28(1):18–36, 2012.
- [5] D. Bilyk, V. N. Temlyakov, and R. Yu. The L_2 discrepancy of two-dimensional lattices. In *Recent Advances in Harmonic Analysis and Applications: in Honor of Konstantin Oskolkov*, pages 63–77. Springer, 2012.
- [6] J. Brauchart, D. P. Hardin, and E. B. Saff. Discrete energy asymptotics on a Riemannian circle. *Uniform Distribution Theory*, 7(2):77–108, 2012.
- [7] J. S. Brauchart and J. A. Dick. A characterization of Sobolev spaces on the sphere and an extension of Stolarsky’s invariance principle to arbitrary smoothness. *Constr Approx*, 38:397–445, 2013.
- [8] J. A. Carrillo, A. Figalli, and F. S. Patacchini. Geometry of minimizers for the interaction energy with mildly repulsive potentials. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 34(5):1299–1308, 2017.
- [9] W. W. L. Chen and M. M. Skriganov. Explicit constructions in the classical mean squares problem in irregularities of point distribution. *Journal für die reine und angewandte Mathematik*, 2002(545):67–95, 2002.
- [10] H. Cohn and A. Kumar. Universally optimal distribution of points on spheres. *J Amer Math Soc*, 20:99–148, 2007.
- [11] H. Cohn, A. Kumar, S. D. Miller, D. Radchenko, and M. Viazovska. Universal optimality of the E_8 and Leech lattices and interpolation formulas. *Ann. of Math. (2)*, 196(3):983–1082, 2022.
- [12] H. Cohn, A. Kumar, and G. Minton. Optimal simplices and codes in projective spaces. *Geom. Topol.*, 20(3):1289–1357, 2016.

- [13] J. Dick. Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM Journal on Numerical Analysis*, 46(3):1519–1553, 2008.
- [14] J. Dick and F. Pillichshammer. *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge University Press, 2010.
- [15] J. Dick, I. H. Sloan, X. Wang, and H. Woźniakowski. Good lattice rules in weighted Korobov spaces with general weights. *Numerische Mathematik*, 103:63–97, 2006.
- [16] D. Dũng, V. Temlyakov, and T. Ullrich. *Hyperbolic cross approximation*. Springer, 2018.
- [17] L. Fejes Tóth. On the sum of distances determined by a pointset. *Acta Math. Acad. Sci. Hungar.*, 7:397–401, 1956.
- [18] T. Goda and P. L’Ecuyer. Construction-free median quasi-Monte Carlo rules for function spaces with unspecified smoothness and general weights. *SIAM Journal on Scientific Computing*, 44(4):A2765–A2788, 2022.
- [19] T. Goda, K. Suzuki, and T. Yoshiki. An explicit construction of optimal order quasi-Monte Carlo rules for smooth integrands. *SIAM Journal on Numerical Analysis*, 54(4):2664–2683, 2016.
- [20] B.-N. Guo and F. Qi. A property of logarithmically absolutely monotonic functions and the logarithmically complete monotonicity of a power-exponential function. *Politehn. Univ. Bucharest Sci. Bull. Ser. A Appl. Math. Phys.*, 72(2):21–30, 2010.
- [21] G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988.
- [22] A. Hinrichs, R. Kritzing, and F. Pillichshammer. Extreme and periodic L_2 discrepancy of plane point sets. *Acta Arithmetica*, 199:163–198, 2021.
- [23] A. Hinrichs and J. Oettershagen. Optimal point sets for quasi-Monte Carlo integration of bivariate periodic functions with bounded mixed derivatives. In Ronald Cools and Dirk Nuyens, editors, *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC, Leuven, Belgium, April 2014*, pages 385–405. Springer International Publishing, 2016.

- [24] B. Kåsemodel, N. Nagel, and T. Ullrich. Tent transformed order 2 nets and quasi-Monte Carlo rules with quadratic error decay. *arXiv:2505.10955*, 2025.
- [25] N. M. Korobov. Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk SSSR*, 124:1207–1210, 1959.
- [26] F. Y. Kuo. Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces. *Journal of Complexity*, 19(3):301–320, 2003. Oberwolfach Special Issue.
- [27] D. H. Lehmer. Interesting series involving the central binomial coefficient. *The American Mathematical Monthly*, 92(7):449–457, 1985.
- [28] V. F. Lev. On two versions of L_2 -discrepancy and geometrical interpretation of diaphony. *Acta Mathematica Hungarica*, 64(4):281–300, 1995.
- [29] N. Nagel. On the global optimality of Fibonacci lattices in the torus. *arXiv:2502.17082*, 2025.
- [30] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.
- [31] V. Shikhman and O. Stein. On jet-convex functions and their tensor products. *Optimization*, 61(6):717–731, 2012.
- [32] V. Temlyakov. Connections between numerical integration, discrepancy, dispersion, and universal discretization. *The SMAI Journal of computational mathematics*, 5:185–209, 2019.
- [33] T. T. Warnock. Computational investigations of low-discrepancy point sets. In S. K. Zaremba, editor, *Applications of Number Theory to Numerical Analysis*, pages 319–343. Academic Press, 1972.
- [34] A. A. Zhensykbayev. The best quadrature formula for some classes of periodic differentiable functions. *Izv. Akad. Nauk SSSR Ser. Mat.*, 41(5):1110–1124, 1200, 1977.
- [35] P. Zinterhof. Über einige Abschätzungen bei der Approximation von Funktionen mit Gleichverteilungsmethoden. *Österr Akad Wiss Math-Naturwiss Kl*, 185(2):121–132, 1976.