



Modification of the Maximin and ϕ_p (Phi) Criteria to Achieve Statistically Uniform Distribution of Sampling Points

Miroslav Vořechovský & Jan Eliáš

To cite this article: Miroslav Vořechovský & Jan Eliáš (2019): Modification of the Maximin and ϕ_p (Phi) Criteria to Achieve Statistically Uniform Distribution of Sampling Points, Technometrics, DOI: [10.1080/00401706.2019.1639550](https://doi.org/10.1080/00401706.2019.1639550)

To link to this article: <https://doi.org/10.1080/00401706.2019.1639550>

 View supplementary material 

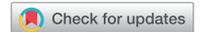
 Accepted author version posted online: 08 Jul 2019.
Published online: 28 Aug 2019.

 Submit your article to this journal 

 Article views: 39

 View related articles 

 View Crossmark data 



Modification of the Maximin and ϕ_p (Phi) Criteria to Achieve Statistically Uniform Distribution of Sampling Points

Miroslav Vořechovský and Jan Eliáš

Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic

ABSTRACT

This article proposes a sampling technique that delivers robust designs, that is, point sets selected from a design domain in the shape of a unit hypercube. The designs are guaranteed to provide a *statistically uniform* point distribution, meaning that every location has the same probability of being selected. Moreover, the designs are *sample uniform*, meaning that each individual design has its points spread evenly throughout the domain. The *sample uniformity* (often measured via a *discrepancy* criterion) is achieved using distance-based criteria (ϕ_p or Maximin), that is, criteria normally used in space-filling designs. We show that the standard intersite metrics employed in distance-based criteria (Maximin and ϕ_p (phi)) do *not* deliver statistically uniform designs. Similarly, designs optimized via centered L_2 discrepancy or support points are also not statistically uniform. When these designs (after optimization based on intersite distances) are used for Monte Carlo type of integration, their statistical nonuniformity is a serious problem as it may lead to a systematic bias. This article proposes using a periodic metric to guarantee the statistical uniformity of the family of distance-based designs. The presented designs used as benchmarks in the article are only taken from the class of Latin hypercube designs, which forces univariate projections to be uniform and improves accuracy in Monte Carlo integration of some functions. Supplementary materials for this article are available online.

ARTICLE HISTORY

Received January 2018
Accepted June 2019

KEYWORDS

Design of experiments for computer models; Latin hypercube sampling; Periodic metric; Space-filling designs; Uniform design; Variance reduction

1. Introduction

We consider numerical methods for the analysis of a computer program via experimentation (evaluations of the computer code). The computer model is denoted as g and it computes output quantities, y , based on the vector of the input variables, \mathbf{x} . Each realization of input vector \mathbf{x} is a point in a *design domain*, \mathcal{U} , which is considered to be a unit hypercube, $\mathcal{U} \equiv [0, 1]^s$. We assume that any transformations necessary to achieve the change from different domains and density functions to \mathcal{U} have already been carried out. The computer program is repeatable, that is, it returns exactly the same output for the same input. However, we assume that the computer model is an implementation of a complex deterministic function for which even a single evaluation at given point requires substantial computational effort.

There are various goals in computer experiments, such as finding a simple approximation of g , called the metamodel or surrogate model \hat{g} , that is sufficiently accurate over the region $A \subset \mathcal{U}$. Another task is the estimation of the size of the error $\hat{g}(\mathbf{x}) - g(\mathbf{x})$ for a given point $\mathbf{x} \in A$. Sometimes the target is to perform the sensitivity analysis of y with respect to changes in \mathbf{x} , or to find which x_i are the most important for each response in y , or to visualize the function g and find its extremes, or to uncover bugs in the implementation of g . We mainly focus our attention on estimating the s -dimensional integral of function

g over the whole design domain \mathcal{U} : $I(g) = \int_{[0,1]^s} g(\mathbf{x}) d\mathbf{x}$ via an average computed over n points selected from \mathcal{U} : $\hat{I}(g, \mathcal{D}) = \frac{1}{n} \sum_{i=1}^n g(\mathbf{x}_i)$. The set of integration points, $\{\mathbf{x}_i \in \mathcal{U}\}$, $i = 1, \dots, n$, is called the *design* or *sampling plan*. It is denoted by \mathcal{D} and is a rectangular matrix of n points with s coordinates. Since the evaluation of g is very expensive, n must be as low as possible. The selection of the design becomes an important step in the estimation of the integral because it can strongly influence the quality of the estimate. The quality of the estimate may be measured by the absolute difference between the exact value and the estimate

$$\epsilon(g, \mathcal{D}) = |I(g) - \hat{I}(g, \mathcal{D})|. \quad (1)$$

In this article, the ultimate goal is to find an optimality criterion for the selection of designs that favors *robust* designs with uniform coverage of the design domain. The designs may be used especially for the integration of black-box functions where little is known prior to experiments. The Koksma–Hlawka inequality (Hlawka 1961, Niederreiter 1992, Theorem 2.11) is perhaps the most well-known and popular multidimensional quadrature error bound (see also Koksma 1942/1943; Fang, Lin, et al. 2000; Fang and Ma 2001). The inequality imposes an upper bound on the absolute error as a product of two independent terms

$$\epsilon(g, \mathcal{D}) \leq V(g) D^*(\mathcal{D}), \quad (2)$$

CONTACT Miroslav Vořechovský  vorechovsky.m@vut.cz  Institute of Structural Mechanics, Faculty of Civil Engineering, Brno University of Technology, Brno, Czech Republic.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/TECH.

 Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/TECH.

© 2019 American Statistical Association and the American Society for Quality

where $V(g)$ is the total variation of the function g in the sense of Hardy and Krause and $D^*(\mathcal{D})$ is the star discrepancy of design \mathcal{D} , which does not depend on function g . For fixed $g(\cdot)$ this bound is minimized when \mathcal{D} has minimum discrepancy. We remark, however, that the inequality does not mean that a design with minimum discrepancy has minimum error $\epsilon(g, \mathcal{D})$. It does suggest that designs with a low discrepancy may be robust to the choice of $g(\cdot)$ because they have this property regardless of the value of $g(\cdot)$. Thus, the class of designs that minimize discrepancy has received considerable attention in the literature under the name *uniform designs*. The star discrepancy measures the *maximum deviation* between the empirical distribution function associated with the design $\mathcal{D} \in \mathcal{U}$ and the uniform distribution over \mathcal{U} . It is possible to generalize the star discrepancy by defining an L_p -star discrepancy (Hickernell 1998a), and related generalizations are also known for the bound in Equation (2). In the various algorithms that search for uniform designs, discrepancy is subject to minimization. There are various measures of discrepancy, depending on the definition (see, e.g., Hickernell 1998a, 1998b). These definitions also have corresponding formulas for their calculation, for example, modified L_2 discrepancy (Fang and Wang 1993), wrap-around L_2 discrepancy (Hickernell 1998b; Fang and Ma 2001), and centered L_2 discrepancy (Hickernell 1998b; Fang, Lin, et al. 2000) (see also Santner, Williams, and Notz 2003; Fang, Liu, et al. 2018).

An important class of designs for multidimensional integration is that of quasi Monte Carlo (QMC) and random QMC (RQMC) sequences, which were born in the 1950s and 1960s and are sometimes named *low-discrepancy sequences* or *digital nets/sequences*; see, for example, a review by Dick, Kuo, and Sloan (2013). Examples of QMC designs include the sequences by Halton (1960), Sobol' (1967, 1976), Niederreiter (1987, 1988, 1992), Faure (1981), the generalization of the Faure sequences by Tezuka (1995), and others. QMC methods are called “number-theoretic methods” (Fang and Wang 1993; Fang, Wang, and Bentler 1994) by some authors (e.g., “good lattice points,” etc.). Special types of digital higher order nets are known as “higher order polynomial lattice point sets” (Baldeaux et al. 2011). QMC sequences became well known for their ability to decrease the estimation variance of integrals in comparison with crude Monte Carlo integration, as well as for the faster convergence of the integral estimate (Owen 1999; L'Ecuyer and Lemieux 2005). The use of low-discrepancy sequences makes QMC designs a cheap alternative to algorithmic *discrepancy minimization* (Niederreiter 1992, chaps. 3 and 4). QMC sequences achieve optimal error rates for star discrepancy (Dick and Pillichshammer 2010). QMC methods may be thought of as de-randomized Monte Carlo samples. It has become apparent that QMC can be rerandomized (RQMC) to obtain sample-based error estimates. Various rerandomization methods have been developed (e.g., “random shift modulo one” by Cranley and Patterson (1976), “scramblings” by Owen (1995; 1998), and others). The random scrambling of QMC sequences eliminates their inherent bias while retaining their low-discrepancy properties. These randomizations are also known to have better projection properties of some digital nets (e.g., Sobol') to the first few dimensions (L'Ecuyer and Lemieux 2005), see below. Surveys of RQMC appear in Owen (1999) and

L'Ecuyer and Lemieux (2005). For smooth integrands, higher order scrambled digital nets have been shown to achieve optimal mean square error rates (Dick 2011).

In computer experiments, it is generally accepted that design points should be spread evenly throughout the entire design domain to provide information about all portions of the experimental region. These are called *space-filling* designs (see, e.g., Joseph 2016). Many criteria for space-fillingness are based on *distances* among points, for example, the Maximin and miniMax criteria (Johnson, Moore, and Ylvisaker 1990; Tan 2013; He 2017a; Mak and Joseph 2018a), the Audze–Eglājs criterion (Audze and Eglājs 1977), and the generalized ϕ_p criterion (Morris and Mitchell 1995). Space-filling designs may be obtained by exploiting the analogy with sphere packing problems (He 2017b). Recently, Mak and Joseph (2018b) suggested the use of a kind of space-filling design for Monte Carlo integration. Their “support points” are generated using the minimization of the *energy distance* between two distributions.

Designs optimized for model prediction should be non-collapsible: if only a subset of input variables is relevant for predicting the response, then the prediction error is related to the uniformity of the *projected* designs. Another frequent requirement placed on designs is for them to display *orthogonality*: this ensures that all specified parameters may be estimated independently of any others (estimation of main effects), and that interactions can also be estimated in a straightforward manner. There are numerous design types focused on orthogonality, such as various factorial designs, “orthogonal arrays,” “mutually orthogonal Latin squares,” etc. Some authors simply tend to decrease statistical correlations among vectors of samples of individual variables in designs (Morris and Mitchell 1995; Vořechovský and Novák 2009). As shown by Owen (1992), the bilinear part of the integrand is more accurately estimated if the sample correlations among input variables are negligible.

One way of obtaining a design with excellent projective properties is to select a design from the class of Latin hypercube designs (LHDs), a type of design suggested by Conover (1975). It has been shown that LHDs are especially suitable for evaluating the *expectation of functions* in computer experiments. There are a number of research articles concerning sampling efficiency for LHS and related sampling schemes (McKay, Conover, and Beckman 1979; Stein 1987; Tang 1993; Owen 1994; Shields and Zhang 2016). LHS never increases variance, in comparison to crude Monte Carlo sampling, but may decrease it. The combination of the standard ϕ_p criterion with LHS was suggested by Morris and Mitchell (1995). Their goal was to find designs which offer a compromise between the entropy/maximin criterion and the display of good projective properties in each dimension (as guaranteed by LHS). The combination of LHS with the standard Maximin criterion has gained considerable popularity (Liefvendahl and Stocki 2006; van Dam et al. 2007; Joseph and Hung 2008; van Dam, Rennen, and Husslage 2009; Grosso, Jamali, and Locatelli 2009; Rennen et al. 2010; Dette and Pepelyshev 2010; Husslage et al. 2011; Janssen 2013).

In this article, we propose a sampling technique that delivers designs suitable for Monte Carlo integration. The designs are *statistically uniform*, meaning that each location in \mathcal{U} has the

same probability of being selected. The standard intersite metrics used in distance-based criteria do *not* guarantee statistically uniform designs, which renders designs achieved using such criteria inapplicable for Monte Carlo integration. This article proposes a periodic metric for *distance-based criteria* which guarantees the statistical uniformity of the obtained designs. Furthermore, to reduce the variance of the numerical integration and also to maximize the amount of information about the analyzed function, we have developed designs with good *sample uniformity*, that is, each single point layout \mathcal{D} is uniform. *Sample uniformity* is often measured via a *discrepancy* criterion; here it is achieved by using a distance-based criterion (ϕ_p or Maximin) that is normally used in space-filling designs. The authors conjecture that an improvement in space-fillingness also leads to a decrease in discrepancy. To achieve uniform projections and also an additional improvement in integration accuracy, we have limited the designs investigated to the class of LHDs.

2. Maximin and ϕ_p Criteria With Intersite and Periodic Metrics

Johnson, Moore, and Ylvisaker (1990) developed the idea of miniMax and Maximin designs for function estimation (e.g., using Kriging). A miniMax distance design is a set of points that minimizes the maximum distance from *any* point in \mathcal{U} to its nearest design point. It requires evaluating the supremum over an infinite set, which is computationally expensive to approximate. We focus on the Maximin criterion and its generalization to the ϕ_p criterion. The points in a design might be considered to be spread out when no two points in the design are “too” close together. One measure of the closeness of the points in set \mathcal{D} is the smallest distance between any two points in \mathcal{D} : $\min_{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}} d(\mathbf{x}_i, \mathbf{x}_j)$. A design that maximizes this measure is said to be a *Maximin* distance design (Johnson, Moore, and Ylvisaker 1990), abbreviated as Mm and denoted as \mathcal{D}_{Mm} . We now redefine this criterion by taking its inverse to obtain a *minimization* problem while the minimum distance is maximized

$$\phi_\infty(d; \mathcal{D}) = \left[\min_{\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{D}} d(\mathbf{x}_i, \mathbf{x}_j) \right]^{-1}. \quad (3)$$

The Maximin criterion can be viewed as a limiting case of a more general criterion proposed by Morris and Mitchell (1995), the ϕ_p criterion

$$\phi_p(d; \mathcal{D}) = \frac{1}{\binom{n}{2}} \sum_{\mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}} [d(\mathbf{x}_i, \mathbf{x}_j)]^{-p}, \quad (4)$$

where p is a positive integer power. For a fixed metric d and a power p , an n -point design is optimal if it minimizes the criterion in Equation (4). Taking the power $p = 2$ and the Euclidean intersite distance between all pairs of points leads to the Audze–Eglājs (ϕ_2) criterion (Audze and Eglājs 1977). Joseph and Hung (2008) use the power $p = 15$. Some authors advise using powers as high as $p = 50$ combined with the ℓ_1 metric d (Jin, Chen, and Sudjianto 2005; Viana, Venter, and Balabanov

2010; Pholdee and Bureerat 2015). The ϕ_p family of distance-based criteria has become very popular and is frequently used in research and in industry. Such criteria are also implemented in widely used software packages, both commercial and free (e.g., the built-in function “lhsdesign” in MATLAB or the “DiceDesign” package included as a part of R project (R Core Team 2013)). Note that while DiceDesign offers the possibility of using a combinatorial optimization algorithm, the “lhsdesign” function in MATLAB merely selects the best design from a user-defined number of randomly permuted LHDs, which results in very poorly optimized designs and thus the problem with their nonuniformity we focus at in Section 2.1 may not be detected. Also note that Equation (4) is sometimes used with an additional p th root of the result (a monotone transformation which does not affect the comparison of criteria for two designs).

The ϕ_p class of criteria (including Maximin and Audze–Eglājs) are dependent on a distance measure or a metric. Let $d(\cdot, \cdot)$ be a metric on $[0, 1]^s$. We focus on the Euclidean distance $d(\mathbf{x}_i, \mathbf{x}_j) = L_{ij}$ between points i and j . It is the length of the line segment connecting them and can be expressed as a function of their Cartesian coordinates

$$d(\mathbf{x}_i, \mathbf{x}_j) = L_{ij} \sqrt{\sum_{v=1}^s (\Delta_{ij,v})^2}, \quad \text{where } \Delta_{ij,v} = |x_{i,v} - x_{j,v}| \quad (5)$$

is the difference in their positions projected onto axis v . Such an *intersite* metric d is the standard way of measuring distances used in combination with distance-based criteria. Note that some authors (Johnson, Moore, and Ylvisaker 1990; Morris and Mitchell 1995; Ye, Li, and Sudjianto 2000; van Dam et al. 2007; van Dam, Rennen, and Husslage 2009; Viana, Venter, and Balabanov 2010) also propose using other metrics such as the ℓ_1 (rectangular or Manhattan) or ℓ_∞ (Chebyshev) distance.

Maximin designs are sometimes deemed *uniform*; see, for example, page 594 of the recent handbook Dean et al. (2015). This actually is not true, as designs with a lower ϕ_p criterion obviously tend to push the design points toward the boundaries of the hypercube. Indeed, Santner, Williams, and Notz (2003) noticed in their Example 5.4 that a combination of Euclidean distance with the Maximin criterion yields designs that “concentrate points on or near the boundary of \mathcal{U} .” Furthermore, they suggest that one should “remedy this by restricting the class of available designs to only include, say, LHDs.” Indeed, He (2017b) stated that most Maximin designs are unsuitable for integration purposes as the average of the outputs is a biased estimator for the mean response. However, in the same article it is also stated that maximin LHDs are suitable for integration purposes. However, we demonstrate in Section 2.1 that even LHDs optimized with the ϕ_p and Mm (ϕ_∞) criteria do *not* cover domains uniformly, so restricting the designs to LHDs does *not* provide a remedy to the problem. The problem is related to the boundaries of the design domain.

2.1. Statistical Nonuniformity of ϕ_p LHDs, CD_2 LHDs, and support points

In this section, we first demonstrate that designs optimized with the ϕ_p and Mm (ϕ_∞) criterion do *not* cover domains

uniformly, even when used with LHDs. We also show that the “centered L_2 discrepancy” (CD_2) criterion also leads to statistically *nonuniform* samples. Finally, the same behavior is demonstrated for support points (Mak and Joseph 2018b).

Statistical uniformity means that the probability that the i th experimental point will be located inside a chosen subset of the domain must be equal to V_S/V_D , with V_S being the subset volume and V_D the volume of the whole domain (for hypercube domain $V_D = 1^s = 1$). For the purpose of testing statistical uniformity, the whole unit volume of the hypercube is divided into bins of equal volume. In the article, we directly employ n^s LH-bins, which use the regular grid of coordinates along each dimension. In a statistically uniform design the probability that each of these bins will be filled must be identical. To perform a numerical test, N_{run} designs (sampling plans) have been simulated and optimized independently. After generating the N_{run} designs, the total number of sampled points is nN_{run} . The average number of points inside one bin should be $\bar{f} = nN_{\text{run}}/n^s = N_{\text{run}}/n^{s-1}$ for a *statistically uniform* design. For each bin, we now count the actual frequency of occurrence of the points inside that bin, f_a . Finally, we define a variable f (a normalized frequency) that can be calculated for

each bin

$$f = f_a \frac{1}{\bar{f}} = f_a \frac{n^{s-1}}{N_{\text{run}}}. \quad (6)$$

A statistically uniform design criterion should yield $f \rightarrow 1$ as $N_{\text{run}} \rightarrow \infty$ for all of the bins.

The results of the numerical study are shown in Figure 1 for various numbers of points, n , and dimensions, s , in two-dimensional images. The number of repetitive optimized designs used, $N_{\text{run}} = 10^4$, is high enough to reveal unwanted patterns. The gray shading represents the f value for individual LH bins. The first dimension (variable) is associated with the horizontal axis, the second variable with the vertical axis, and the third variable (if present) is expressed by repetitive two-dimensional images (slices) produced for different values of the third coordinate. Similarly, the fourth dimension (if present) is shown by repetitive views of three-dimensional plots made for different values of the fourth coordinate. The figures clearly show the *statistical nonuniformity* of the point density in the design domain for the original ϕ_∞ LHDs. In two-dimensional space, the corners are not sampled at all, but there is an area of highly probable points close to them followed again by an improbable region. A similar behavior is

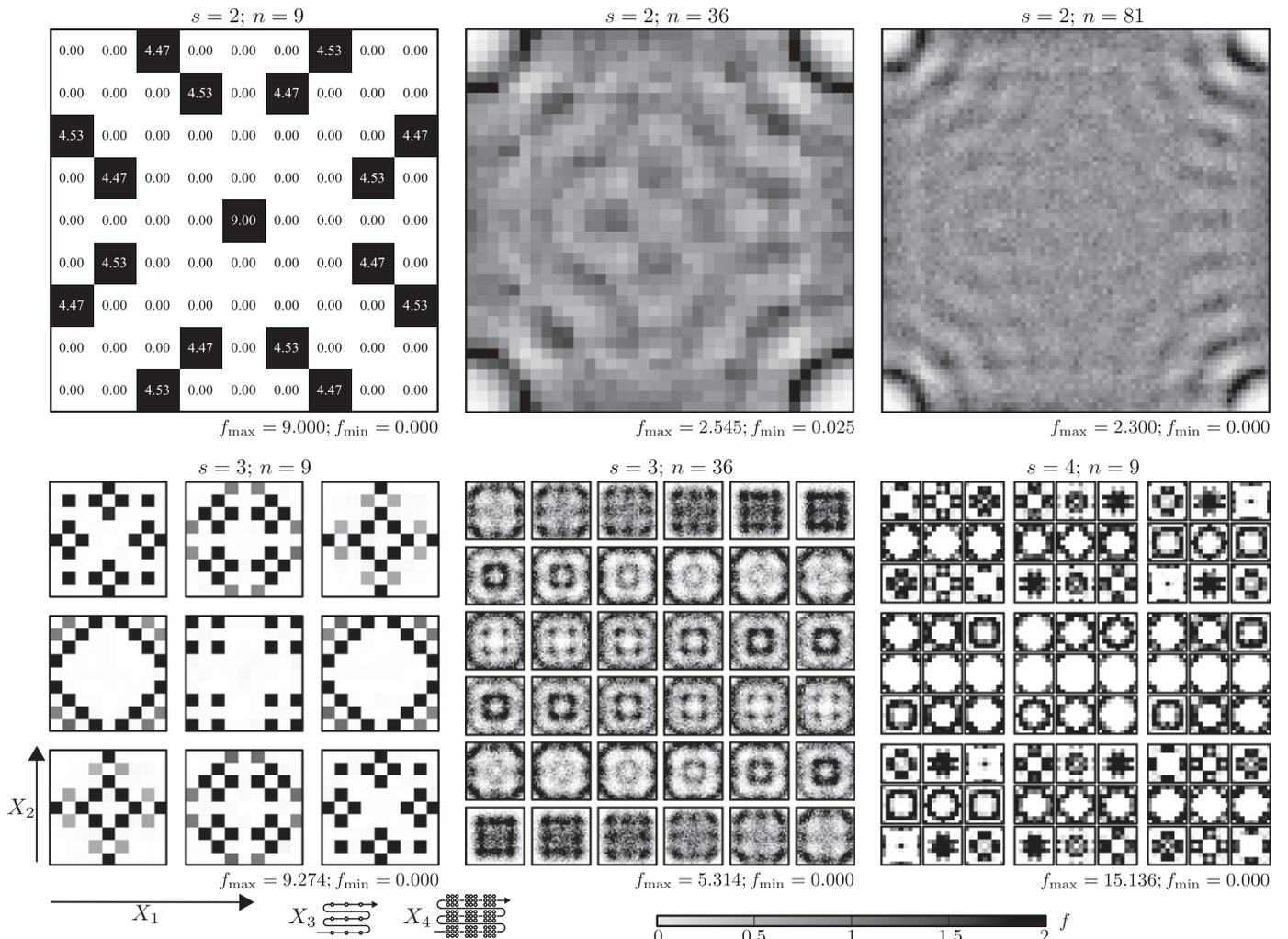


Figure 1. Relative frequencies f calculated for ϕ_∞ LHDs (intersite metric) using $N_{\text{run}} = 10^4$ designs. Top row: $s = 2$ dimensions. Bottom row: slices through a hypercube in $s = 3, 4$ dimensions.

observed in three-dimensional and four-dimensional hypercubes, where the corners of the domain are always sampled poorly.

Figure 2 shows the maps (histograms) for two-dimensional ϕ_p LHDs, where the exponents used in the optimization were varied: $p = \{2, 4, 8, 16\}$. Exponent $p = 2$ represents the Audze and Eglājs (1977) criterion and the highest exponent behaves in almost the same way as the Mm criterion; compare the map with Figure 1 top middle. Figure 2 documents that the increase

in p partly improves the *statistical uniformity* of LHDs but strong bias is still present even for large exponents.

The proof of a preference to remove samples from corner regions in ϕ_p LHDs is a simple extension to the proof made by Eliáš and Vořechovský (2016) for $p = 2$.

Similar behavior was detected in the case of CD_2 LHDs; see the examples in Figure 3. Additionally, Figure 3 also presents the relative frequencies for the support points recently proposed by Mak and Joseph (2018b). The support points can be generated

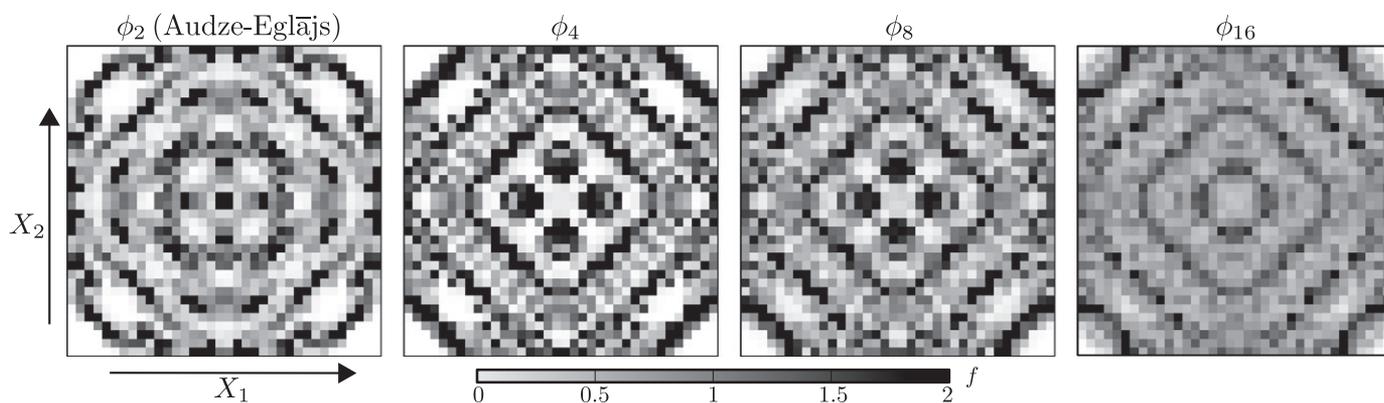


Figure 2. The role of power p in standard ϕ_p LHDs (intersite metric). The relative frequencies f were calculated from $N_{\text{run}} = 10^4$ designs with $s = 2$ variables and $n = 36$ points.

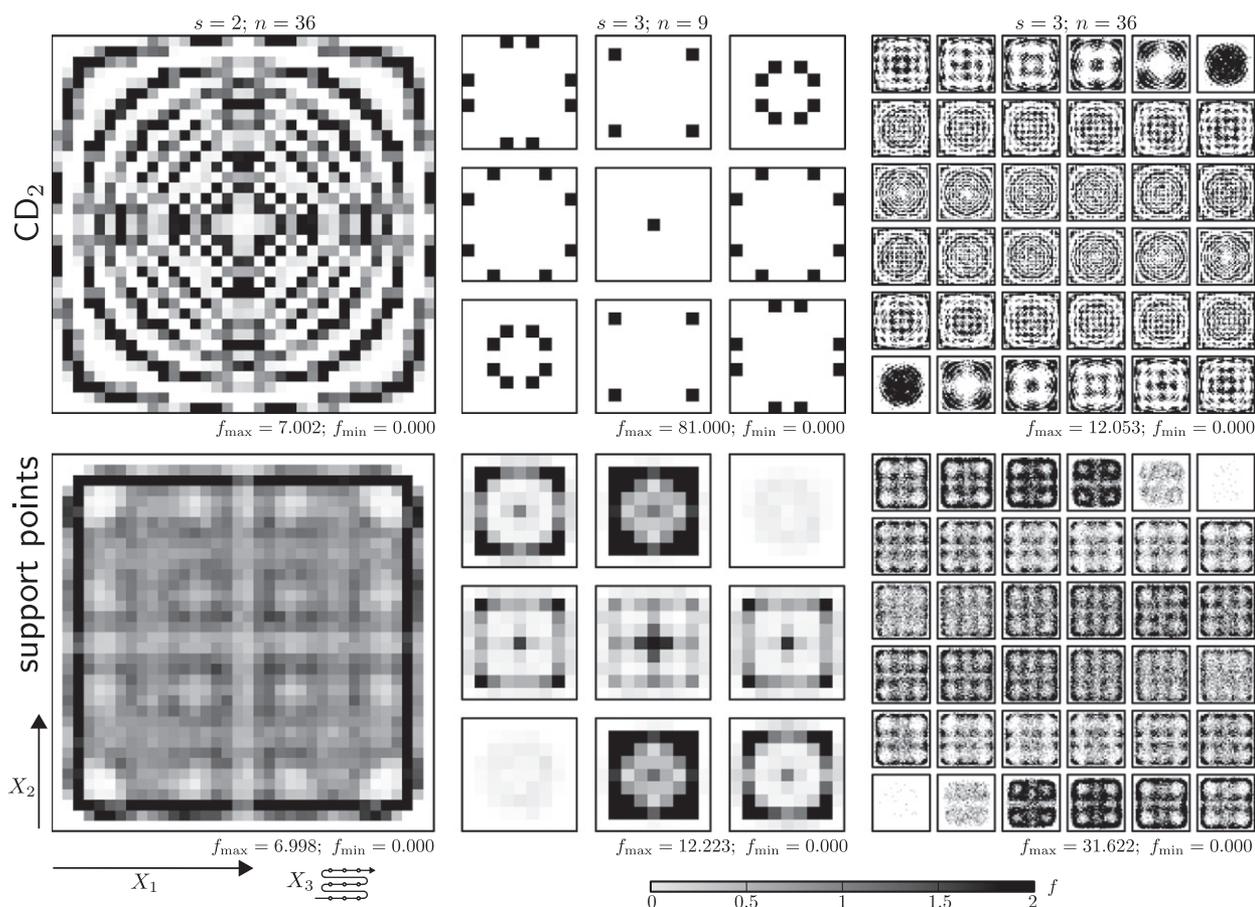


Figure 3. Relative frequencies f calculated for CD_2 LHDs (top row) and for support points (bottom row) with $n = \{9; 36\}$ points in $s = 2$ and 3 dimensions.

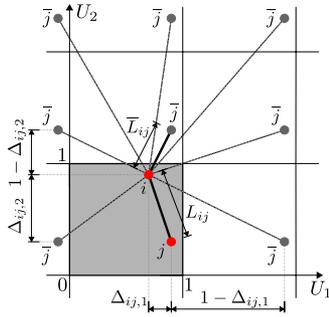


Figure 4. The Euclidean *intersite* distance, L_{ij} , between points i and j in a two-dimensional domain (light gray), and the distances between point i and periodic images of point j , which are denoted by \bar{j} . The shortest (periodic) distance is \bar{L}_{ij} . The periodic domain in two-dimensional is homeomorphic to the surface of a torus in three-dimensional (see Figure 5 right).

using the function $\text{sp}()$ in the R package “support” (Mak 2017). The histograms are processed using points generated from uniform distribution, but the results obtained with another distributions (transformed to uniform via the corresponding cumulative distribution function) produce similar results. support points are obtained via the minimization of the *energy distance* or “ \mathcal{E} -distance” between two distributions (Székely and Rizzo 2004, 2013). The “ \mathcal{E} -distance” uses the Euclidean (intersite) norm between points (all the intersite distances among the design points themselves and also their intersite distances to an underlying set of points in the domain), which is the reason why support points suffer from the same kind of *statistical nonuniformity* as the other distance-based criteria discussed in this section.

2.2. The Periodic Metric for Distance-Based Designs

We propose a very simple and computationally cheap remedy that provides *statistically uniform* designs while keeping the concept of the ϕ_p , Mm and other distance-based criteria unchanged. The proposal concerns modification of the metric. It is inspired by the modification of the Audze–Eglājs criterion recently proposed by the authors in Eliáš and Vořechovský (2016). In the present article, the idea is generalized to the whole family of distance-based ϕ_p criteria. The remedy is based on periodic repetition of the design domain, \mathcal{U} , together with the experimental points contained, \mathcal{D} , along all directions; see Figure 4 for an illustration for $s = 2$. The distance \bar{L}_{ij} between a pair of points i and the nearest image of point j is measured in a periodically extended space. This is sometimes referred to as the *minimum-image convention*.

In the periodic space, we now redefine the Euclidean distance between points \mathbf{x}_i and \mathbf{x}_j , formerly L_{ij} , to be the length of the line segment connecting point \mathbf{x}_i with the *closest* image of point \mathbf{x}_j . This is achieved by considering the *shortest* 1D *periodic projections* along each individual coordinate, $\Delta_{ij,v}$. The new periodic distance metric yields

$$\bar{d}(\mathbf{x}_i, \mathbf{x}_j) = \bar{L}_{ij} = \sqrt{\sum_{v=1}^s (\bar{\Delta}_{ij,v})^2}, \quad \text{where} \\ \bar{\Delta}_{ij,v} = \min(\Delta_{ij,v}, 1 - \Delta_{ij,v}) \quad (7)$$

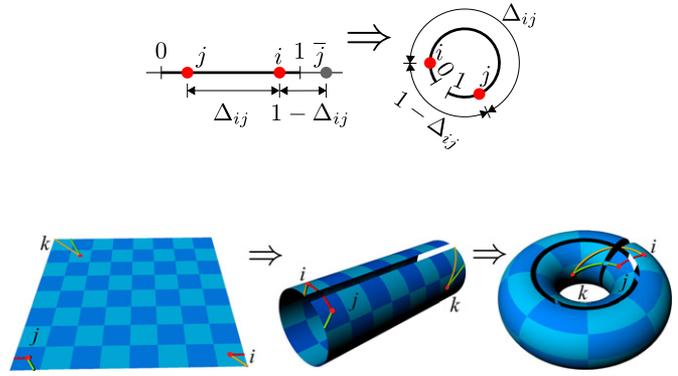


Figure 5. Illustration of space folding and the periodic interpoint distance \bar{d} . Top: $s = 1$. Bottom: $s = 2$ (the toroidal boundary conditions are meant for illustration only; the metric defined in Equation (7) does not measure geodesic distance on a curved surface).

with $\Delta_{ij,v}$ defined in Equation (5). One can easily show that such a function fulfills the conditions imposed on metrics: it provides a mapping $\mathcal{U} \times \mathcal{U} \rightarrow [0, \infty)$ which fulfills the conditions of (i) nonnegativity, (ii) the identity of indiscernibles, (iii) symmetry, and (iv) triangle inequality. Technically, this improvement is very easy to implement in computer programs and the additional amount of computer time necessary to perform the comparison and selection of the minima is negligible.

To distinguish between the original Mm (ϕ_∞) formulation (Johnson, Moore, and Ylvisaker 1990) and the proposed one based on periodic space, we call the new formulation the “*periodic* Maximin (pMm) criterion” and denote it by $\bar{\phi}_\infty(\mathcal{D})$. Similarly, the proposed “*periodic* ϕ_p criterion” is denoted by $\bar{\phi}_p(\mathcal{D})$. The definitions read simply

$$\bar{\phi}_\infty(\mathcal{D}) = \phi_\infty(\bar{d}; \mathcal{D}) \quad \text{and} \quad \bar{\phi}_p(\mathcal{D}) = \phi_p(\bar{d}; \mathcal{D}), \quad (8)$$

where the only difference from Equations (3) and (4) is in the metric used.

Taking only the shortest distance $\bar{d}(\mathbf{x}_i, \mathbf{x}_j)$ in Equation (8) results in the simplification of the infinite periodic extension of the design domain. This simplification is applied in many fields, for example, molecular dynamics. The difference in the design patterns when using the $\bar{\phi}_p$ criterion and when considering all the lengths in the periodic system (Figure 4) decreases as the exponent p increases. The analysis performed in Eliáš and Vořechovský (2016) for an exponent as low as $p = 2$ shows only an insignificant difference. It is, however, worse for a small number of points and a large number of dimensions, that is, in situations where the minimum periodic distance is comparable to other (longer) distances between images of the same points. As the exponent approaches ∞ and Maximin designs are obtained, the difference vanishes; the $\bar{\phi}_\infty$ criterion actually considers the full periodic space. Moreover, according to a recent article (Sadílek and Vořechovský 2018) the exponent p should be dependent on the dimension of the problem, s : the influence of long distances is sufficiently suppressed in favor of short distances when the exponent exceeds the critical value, which is the domain dimension. Therefore, $p > s$ and we recommend taking the exponent $p = s + 1$.

We claim that the periodic metric, \bar{d} , used in Equation (8) provides *statistically uniform* designs. The histograms studied

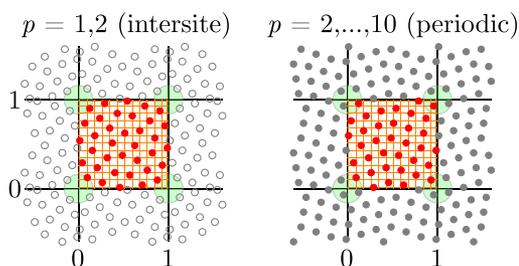


Figure 6. Examples of optimal LHDs with $n = 36$ points in $s = 2$ dimensions. Left: ϕ_p criterion with *intersite* distance. Right: $\bar{\phi}_p$ with *periodic* distance.

in Section 2.1 for the intersite metric (ϕ_p and Mm designs) become perfectly even for the periodic metric ($\bar{\phi}_p$ and pMm designs). The source of *statistical uniformity* actually lies in the *invariance* of $\bar{\phi}_\infty$ and $\bar{\phi}_p$ with respect to translation. If all the points in a periodic space are shifted by an arbitrary vector, the periodic Euclidean distances \bar{d} remain unchanged, and thus also the values of criteria based on them are not altered. Therefore, when generating the experimental points repetitively with an unbiased generator, even if there is a tendency to form a pattern, the pattern is always randomly shifted. Since this shift is unbiased, there is an *equal probability* of a point appearing in any location. As a result, *statistical uniformity* is guaranteed.

To compare LHDs optimized with respect to ϕ_p (standard intersite) and the proposed $\bar{\phi}_p$ (periodic distances), Figure 6 provides examples for 36 points in two dimensions. The nonuniformity of point coverage (visible here via the empty corners emphasized by green circles) in the case of the standard intersite distance (left) is remedied in the case of periodic distances (right). The exponents p for which these designs represent optimal LHDs are listed above the scatterplots. Each single design with intersite metric will have empty corners as demonstrated statistically in Figure 2. The periodic repetition of the design helps with revealing the statistical (non)uniformity in the vicinity of boundaries. Using this visualization technique for individual designs also reveals empty corners for ϕ_∞ LHDs that can be downloaded from the “space-filling designs” database by van Dam et al. (2017).

The proposed $\bar{\phi}_p$ -optimal designs are not only *statistically uniform*, they also retain the original desirable space-filling property of the distance-based family of criteria, which is that the points in each individual optimized design tend to be evenly spread over \mathcal{U} (*sample uniformity*); see, for example, Figure 6 right.

Figure 7 shows individual LHDs obtained with the proposed periodic criterion, $\bar{\phi}_p$, in the case of 14 points in two dimensions. The plots on the left and in the middle show two designs selected from the same pattern, that is, designs that feature an identical set of periodic interpoint distances. However, the layout in the middle is a rotated, reflected and shifted version of the left-hand one and therefore looks like a different design when plotted individually.

To document the role of the exponent, p , the $\bar{\phi}_\infty$ LHD (pMmLHD) design on the right-hand side of Figure 7 is plotted. One can clearly see that a design minimizing $\bar{\phi}_\infty$ may lead to somewhat less *sample uniformity* within a single layout compared to $\bar{\phi}_p$ as it only focuses on the extreme (minimum) distance. The rest of the points may find various locations

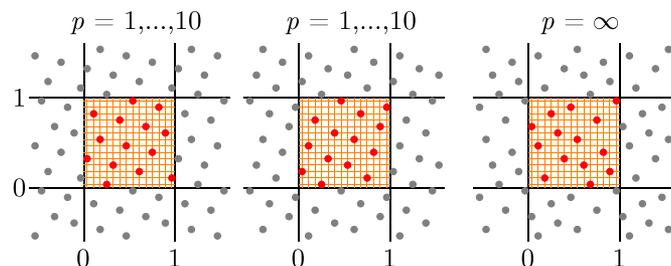


Figure 7. Examples of optimal $\bar{\phi}_p$ LHDs with 14 points in two dimensions (periodic), with various exponents ($p = \infty$ implies the $\bar{\phi}_\infty$ criterion).

providing they spread over a distance greater than or equal to \bar{L}_{\min} . Indeed, all three designs in Figure 7 have exactly the same minimum periodic distance between the closest pairs of points: $\bar{L}_{\min} \approx 0.2259$. The vast majority of $\bar{\phi}_\infty$ LHDs feature the minimum distance many more times than $\bar{\phi}_p$ with a high (yet finite) exponent, p . This makes $\bar{\phi}_\infty$ designs look “more random” or “less spread-out” than when a $\bar{\phi}_p$ criterion is used with a finite exponent, p (compare the degree of contrast of Figure 1 middle top with Figure 2 right), which may not be desirable in some applications.

2.3. Optimization of LHDs

In this section, we describe how we construct optimal or near-optimal LHDs. The design, \mathcal{D} , sometimes called the “sampling plan,” is a matrix of size $n \times s$. When combining sampling strategies with fixed n coordinates for each separate variable (as in the case of our LHDs with centers of n equidistant intervals), the only way to optimize the sample with respect to a particular criterion (correlation, Audze–Eglājs, ϕ_p , Mm, etc.) is to change the mutual ordering of these coordinates. This is achieved via combinatorial optimization used in conjunction with the optimization criteria.

To minimize the desired criterion, one can search among all $(n!)^{s-1}$ possible mutual orderings of these fixed coordinates (or $(n!)^{s-1} / (s-1)!$ configurations if we consider designs equivalent if they differ by a permutation of coordinates). An exhaustive search thus quickly becomes prohibitive with respect to time even for moderate values of n and s . We use an exchange algorithm based on simulated annealing optimization (Vořechovský and Novák 2009) to search for a good solution. The performance of the optimization is governed by a *cooling schedule*, that is, a set of parameters determining the initial and final temperature, the number of temperature steps over time, the number of swaps at a given temperature, etc. This “shuffling” algorithm does not guarantee that the best possible arrangement (the minimum of the criterion) will be found. Details regarding this heuristic optimization algorithm can be found in Vořechovský and Novák (2009).

Another option would be based on the “latinization” of points that are obtained by exploiting the physical analogy between a design and a set of free points with repulsive forces that are derived from the potential represented by the optimization criterion. Such a dissipative dynamical system is used in Vořechovský, Mašek, and Eliáš (2017), Mašek and Vořechovský (2018), and Vořechovský, Mašek, and Eliáš (2019), together with

a formulation of the equations of motion, their computer implementation using GPU and a study of the designs' performance in numerical integration.

3. Uniform Designs Obtained via Direct Discrepancy Minimization

This article focuses mainly on the performance of the proposed designs in multidimensional integration. It is well known that the integration accuracy of designs is closely related to their extreme discrepancy (Niederreiter 1992). It is thus important to compare the proposed designs not only to standard QMC and RQMC sequences, but to also to LHDs obtained by direct discrepancy minimization. Hickernell (1998a; 1998b) proposed a family of generalized L_2 discrepancies, and from among them the *centered* L_2 -discrepancy (CD_2) and *wrap-around* L_2 -discrepancy (WD_2) have been widely used in theoretical studies and practical applications. From here on, we focus on CD_2 and WD_2 .

Perhaps the most widely used is the centered L_2 discrepancy (Hickernell 1998b)

$$[CD_2(\mathcal{D})]^2 = \frac{13^s}{12} - \frac{2}{n} \sum_{i=1}^n \prod_{v=1}^s \left[1 + \frac{1}{2} \left| x_{i,v} - \frac{1}{2} \right| - \frac{1}{2} \left(x_{i,v} - \frac{1}{2} \right)^2 \right] + \frac{1}{n^2} \sum_{i=1}^n \sum_{i=j}^n \prod_{v=1}^s \left[1 + \frac{1}{2} \left| x_{i,v} - \frac{1}{2} \right| + \frac{1}{2} \left| x_{j,v} - \frac{1}{2} \right| - \frac{1}{2} |x_{i,v} - x_{j,v}| \right]. \quad (9)$$

The expected values and variances of $[CD_2]^2$ for simple random sampling and LHDs are given in Fang, Ma, and Winker (2000). CD_2 is a rotation and reflection-invariant discrepancy (replacing coordinate $x_{i,v}$ by $1 - x_{i,v}$ leaves the centered discrepancy unchanged). Instead of anchoring the discrepancy to the origin (as many discrepancies do), it refers to the center of the hypercube. It is invariant under permuting variables and/or simulations (points), and it measures projection uniformity. We claim that finite samples optimized via CD_2 discrepancy do *not* enjoy the property of statistical uniformity, and we document this behavior on frequency maps in Section 2.1.

Another of the various discrepancies, the wrap-around L_2 discrepancy (WD_2 for short), proposed by Hickernell (1998b), has nice properties. WD_2 is invariant when reordering simulations and relabeling coordinates (rotation-invariant), and it is also reflection-invariant. It does not involve the corner or center points, so the discrepancy is truly “unanchored.” In the context of the present article, the important property is that WD_2 is *invariant under coordinate shift*. That is why the studied WD_2 LHDs are *statistically uniform*. In this way, it enjoys the same property as the proposed combination of the distance-based $\bar{\phi}_p$ criterion enhanced by the shift-invariant metric. Hickernell (1998b) also pointed out that the wrap-around discrepancy satisfies the Koksma–Hlawka inequality ($[WD_2]^2$ may serve as a measure of the upper bound of the *squared* error of the estimator $\hat{I}(g, \mathcal{D})$), see Fang and Ma (2001).

The analytical expression of WD_2 can be written as

$$[WD_2(\mathcal{D})]^2 = - \left(\frac{4}{3} \right)^s + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \prod_{v=1}^s \left[\frac{3}{2} - \Delta_{ij,v} (1 - \Delta_{ij,v}) \right], \quad (10)$$

where $\Delta_{ij,v} = |x_{i,v} - x_{j,v}|$ is the intersite distance projection. The expected values and variances of $[WD_2]^2$ for simple random sampling and LHDs are given in Fang and Ma (2001); these values can be used to provide information about the performance of the estimator $\hat{I}(g, \mathcal{D})$. WD_2 -optimal designs have perfectly uniform projections on each single dimension (they are inherently LHDs).

The proposed $\bar{\phi}_p$ designs are deemed to have a high degree of *sample uniformity* even though *discrepancy* is not directly subject to minimization; only the distance-based *periodic* criterion is. Figures 6 and 7, which were obtained for $\bar{\phi}_p$ LHDs, visually support our conjecture that a decrease in the $\bar{\phi}_p$ criterion also decreases discrepancy. Figure 8 compares the actual $[CD_2]^2$ and $[WD_2]^2$ values and also the proposed $\bar{\phi}_p$ and $\bar{\phi}_\infty$ criteria for various designs. The designs under comparison are:

- LHDs with random ordering,
- QMC and RQMC sequences,
- CD_2 -LHDs and WD_2 -LHDs, both obtained via combinatorial optimization,
- the proposed $\bar{\phi}_\infty$ LHDs and $\bar{\phi}_p$ LHDs (with exponent $p = s + 1$), both obtained via combinatorial optimization.
- support points.

In this article, the QMC and RQMC sequences are taken from the R package “fOptions” (Wuertz, Setz, and Chalabi 2017; R Core Team 2013). Note that the same sequences can also be obtained with the “randtoolbox” package in R (Dutang and Savický 2019). For the QMC sequences we have selected Sobol', as it exhibited the best overall performance among all of the QMCs studied within this article. The group of RQMC sequences is represented by scrambled Sobol' sequences; the scrambling option is set to three in the “runif.sobol” function of the “fOptions” package (or the “sobol” function of the “randtoolbox” package), meaning that two types of scrambling are applied to the Sobol' sequence: Owen scrambling (Owen 1995, 1998, 1999) and also the Faure–Tezuka (Tezuka and Faure 2003) type of scrambling; see also Hong and Hickernell (2003).

It can be seen that the best designs in each criterion are always those that were subject to optimization with respect to that criterion. This is untrue only for large sample sizes when the combinatorial algorithm loses its ability to find the global minimum, or at least a good local one. Also, in $s = 10$ dimensions the algorithm could not outperform $\bar{\phi}_p$ LHDs when searching for $\bar{\phi}_\infty$ LHDs (see also the discussion in Section 2). Generally, optimizing in $s = 10$ dimensions seems to be a hard task for a heuristic algorithm.

The best overall performance is exhibited by the proposed $\bar{\phi}_p$ LHDs. They naturally have the best $\bar{\phi}_p$ criterion values, but they also show very low discrepancies. The opposite is not true: QMC, RQMC, and discrepancy-optimized designs are low-discrepancy designs but their space-filling ability

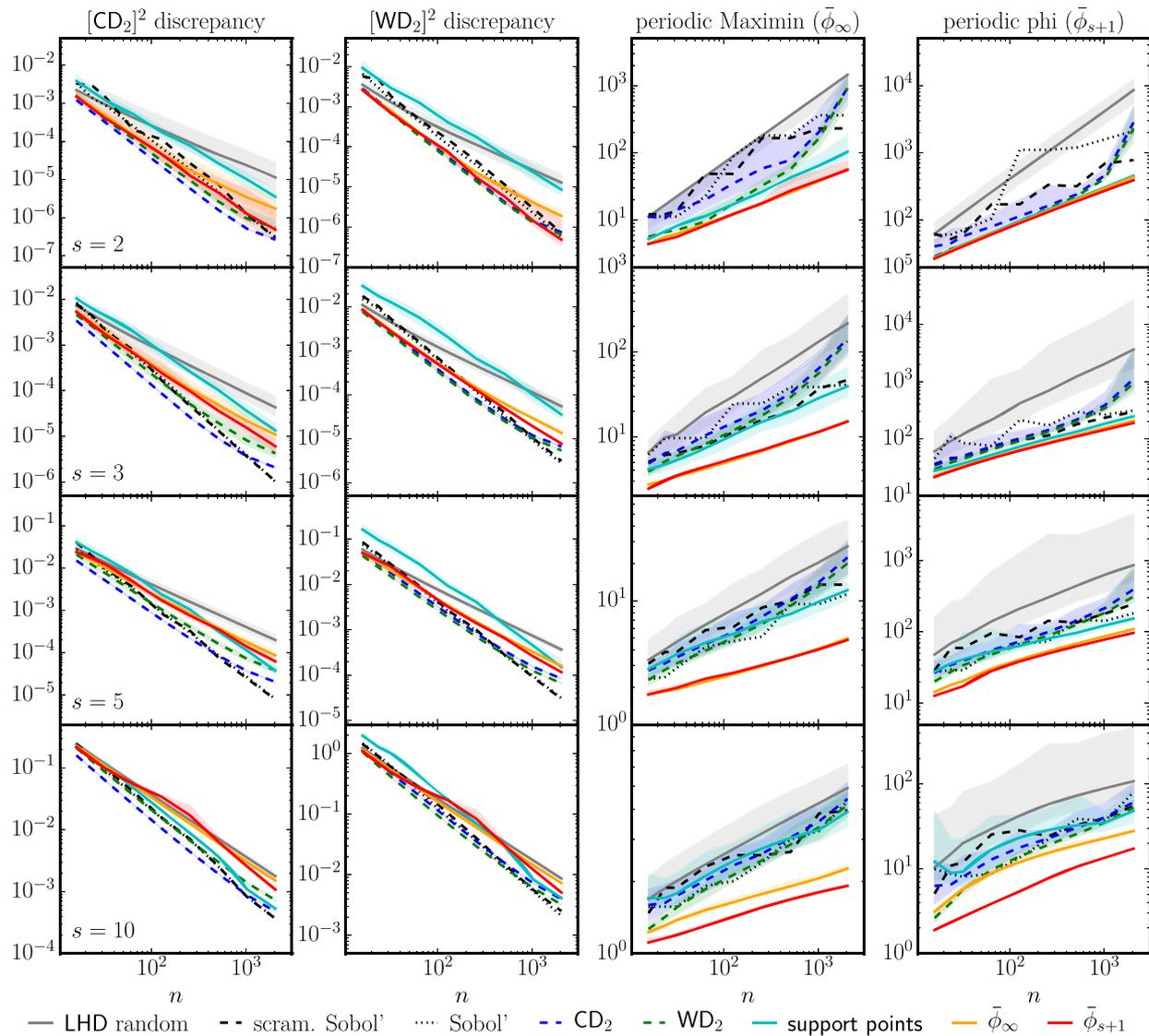


Figure 8. Mutual comparison of various criteria values for designs optimized via various criteria in dimensions $s = 2, 3, 5$, and 10 (from top to bottom). In the two left columns: discrepancies (CD_2 and WD_2). In the two right columns: distance-based criteria ($\bar{\phi}_\infty$ and $\bar{\phi}_p$). The scatterbands correspond to the 5th and 95th percentiles.

(measured via $\bar{\phi}_p$ and ϕ_∞) is quite poor. The $\bar{\phi}_\infty$ criterion reveals clusters of points that are in WD_2 LHDs, CD_2 LHDs, and QMC and RQMC sequences.

Finally, we should remark that the designs obtained with the proposed $\bar{\phi}_p$ criterion possess a high degree of orthogonality (low correlations between variables).

4. Numerical Examples

In this section, we compare the integration errors for four groups of designs: (i) CD_2 LHDs and WD_2 LHDs, (ii) QMC and RQMC sequences, (iii) support points, and finally (iv) the proposed $\bar{\phi}_p$ LHDs and $\bar{\phi}_\infty$ LHDs. We have purposely omitted studies featuring ϕ_p LHDs, as they were not developed for numerical integration, provide statistically nonuniform designs and are only rarely recommended for this purpose in the literature. A special class of ϕ_p LHDs, namely ϕ_2 LHDs (Audze-Eglājs), were studied in (Eliáš and Vořechovský 2016).

The error, denoted as $\epsilon(g, \mathcal{D})$, which is the absolute difference between the exact solution and its numerical estimate, depends on the problem at hand and is a function of a particular design, \mathcal{D} . We have selected three functions (products of Gaussian and uniform variables, and an indicator function) and shall analyze the errors in a subspace of smooth functions by exploiting Chebyshev polynomials.

Since the designs are not unique for most of the techniques under comparison, we consider N_{run} realizations of the designs and thus also obtain N_{run} values of $\epsilon(g, \mathcal{D})$. Therefore, the error can be treated as a random variable (we use a pseudo-random number generator when preparing the designs) and we report the median surrounded by a colored band between the 5th and 95th percentiles of $\epsilon(g, \mathcal{D})$ to assess the variability of the error. In an ideal situation, the error is zero with zero variance. The dependence of the error on the sample size is reported in double logarithmic scale to reveal power laws.

There is only one design obtained with QMC and RQMC sequences for a particular sample size and domain dimension.

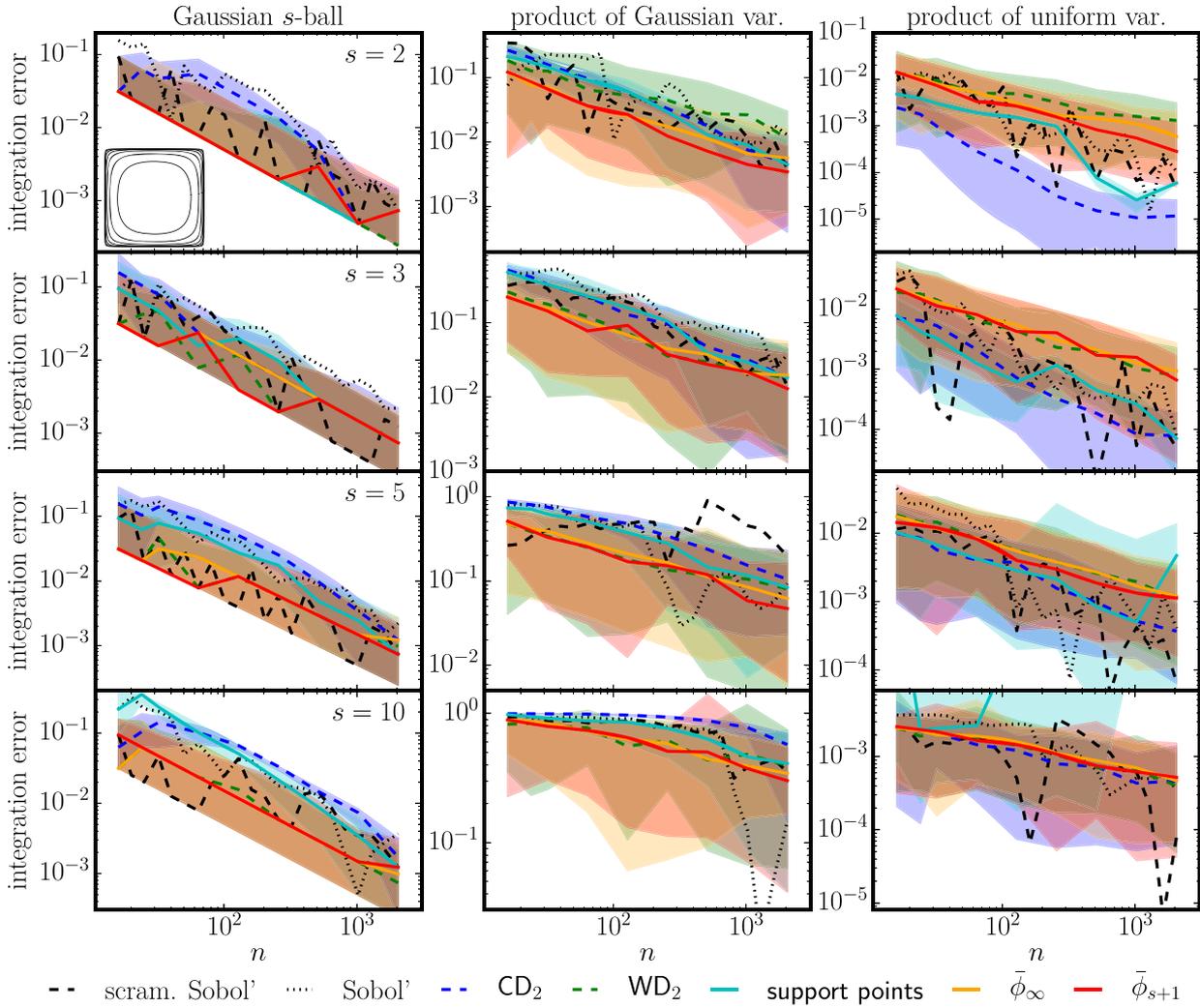


Figure 9. Convergence of integration errors for three simple analytical functions. The inset for $s = 2$ dimensions shows the evolution of the failure region boundary (Gaussian s -ball) for sample sizes $n = 16, 32, 64, 128, 256, 512, 1024$, and 2048 .

4.1. Determination of Failure Probability Using Gaussian s -Ball

The first example is motivated by the estimation of the failure probability of an engineering system. We consider a computational model that signals failure via an indicator function, $I_f(\mathbf{x})$. This function returns one for failure and zero otherwise. The probability of failure is then given as $p_f = \int_{\mathcal{U}} I_f(\mathbf{x}) d\mathbf{x}$. Monte Carlo estimation of this integral can be obtained as an average of the indicator function: $\hat{p}_f = \frac{1}{n} \sum_{i=1}^n I_f(\mathbf{x}_i)$. This estimator is unbiased if all points selected from \mathcal{U} have the same probability of being selected.

We define the model in terms of the jointly Gaussian random vector \mathbf{y} with standard and independent marginals, y_v . The probability density function of vector \mathbf{y} reads: $\varphi(\mathbf{y}) = \frac{1}{(2\pi)^{s/2}} \exp(-\frac{1}{2}\mathbf{y}^T\mathbf{y})$. Failure is assumed to occur when the combination of input parameters is too high. In our case, it occurs when the distance of a point \mathbf{y} from the origin is greater than a predefined value r , that is, $r^2 < \rho^2 = \mathbf{y}^T\mathbf{y}$. The failure region is therefore outside the s -ball $S_r \in \mathbb{R}^s$ and the probability of failure reads $p_f(r, s) = 1 - \int_{S_r} \varphi(\mathbf{y}) d\mathbf{y}$. This integral can be simplified using the hyper-spherical coordinate system with the radial coordinate ρ

$$\begin{aligned} p_f(r, s) &= 1 - \frac{\int_0^r \exp(-\rho^2/2) \rho^{s-1} d\rho}{\int_0^\infty \exp(-\rho^2/2) \rho^{s-1} d\rho} \\ &= 1 - \frac{2^{(1-\frac{s}{2})}}{\Gamma(\frac{s}{2})} \int_0^r \exp(-\rho^2/2) \rho^{s-1} d\rho. \end{aligned} \quad (11)$$

The Gaussian variables, y_v , can be transformed into coordinates, x_v , in the unit hypercube via a component-wise transformation involving the standard Gaussian cumulative distribution function: $x_v = \Phi(y_v)$, $v = 1, \dots, s$. We purposely vary the selected radius r depending on the number of integration points, n , so that the probability of failure $p_f = 8.5/n$ (on average only 8.5 points should fall into the failure domain). This is a tough test which measures whether the sampling scheme has the correct proportion of points close to the boundaries of \mathcal{U} for any sample size n . The test reveals insufficient coverage of corners by points in the cases of both CD_2 LHDs and Gaussian support points as their p_f estimates are severely inaccurate.

The failure probability estimation error is defined as $\epsilon_f = |\hat{p}_f(\mathcal{D}) - p_f|$. The left column in Figure 9 shows the evolution of error in various dimensions. CD_2 LHDs strongly deviate from p_f as they do not have statistically uniform coverage of \mathcal{U} (in $s = 10$ dimensions they often estimate $\hat{p}_f = 0$ and therefore the error

is equal to p_f). The lowest average errors for this function are obtained with the proposed $\bar{\phi}_\infty$ LHDs and $\bar{\phi}_p$ LHDs, and also with WD_2 LHDs, that is, the techniques that have, on average, the correct proportion of points in the corner regions. QMC and RQMC sequences provide, on average, the same rate of convergence, but the trend is not as smooth as with $\bar{\phi}_p$ LHDs. Scrambling applied to Sobol' sequences improved the accuracy considerably.

4.2. Product of Independent Variables

In the second example, we wish to estimate the standard deviation of a product of independent random variables. Two alternatives are studied. In the first alternative we study variables y_ν , again following the standard Gaussian joint probability density

$$g_G(\mathbf{y}) = \prod_{\nu=1}^s y_\nu = \prod_{\nu=1}^s \Phi^{-1}(x_\nu) \quad (12)$$

and we again use component-wise mapping: $y_\nu = \Phi^{-1}(x_\nu)$; x_ν are independent uniform variables with uniform density in \mathcal{U} . The exact solution $\sigma_G = 1$, see Eliáš and Vořechovský (2016). The estimation reads

$$\hat{\sigma}_G(\mathcal{D}) = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (g_G(\mathbf{y}_i) - \hat{\mu}_G)^2}, \quad (13)$$

where $\hat{\mu}_G(\mathcal{D}) = \frac{1}{n} \sum_{i=1}^n g_G(\mathbf{y}_i)$. The integration error is defined as $\epsilon_G = |\hat{\sigma}_G(\mathcal{D}) - \sigma_G|$. Function $g_G(\mathbf{y})$ is smooth with many symmetries and is easily integrable, but it is particularly sensitive to the regions in the vicinity of the *corners* of \mathcal{U} , where the product of Gaussian variables tends to infinity. For example, the placement of just one point into a corner of the hypercube ($x_{i,\nu}$ equals either $0.5/n$ or $1 - 0.5/n$) results in an extremely high or low result for the estimated product. That is why the standard and CD_2 LHDs (and also ϕ_p LHDs, ϕ_∞ LHDs and Gaussian support points) provide strongly underestimated values of $\hat{\sigma}_G$: these criteria suppress the corner regions completely (see Section 2.1). Figure 9 middle shows again that the lowest errors are obtained with the proposed $\bar{\phi}_\infty$ LHDs and $\bar{\phi}_p$ LHDs. WD_2 LHDs provide comparable medians but the variance of the estimator tends to be slightly higher. The convergence plots of QMC and RQMC sequences are quite serrated, but for selected sample sizes very accurate estimates are obtained.

A question may arise about the behavior of the integration error when the variables are simply independent standard uniformly distributed random variables x_ν . In this second alternative, variables y_ν become directly equal to x_ν , that is, without the additional inverse Gaussian transformation. The equation for the model, Equation (12), simply reads $g_U(\mathbf{x}) = \prod_{\nu=1}^s x_\nu$. The exact mean value $\mu_U = 2^{-s}$ and the exact standard deviation $\sigma_U = \sqrt{(4^s - 3^s)/(12^s)}$. The standard deviation is estimated as $\hat{\sigma}_U$ using Equation (13), in which the correct model $g_U(\mathbf{x})$ is used and the estimation of the mean value reads $\hat{\mu}_U(\mathcal{D}) = \frac{1}{n} \sum_{i=1}^n g_U(\mathbf{x}_i)$.

The right column in Figure 9 shows errors defined as $\epsilon_U = |\hat{\sigma}_U(\mathcal{D}) - \sigma_U|$. In this case, CD_2 LHDs seem to provide

excellent results, along with support points, QMC and RQMC sequences. QMC and RQMC sequences generally perform well for smooth integrands (Dick, Kuo, and Sloan 2013). The integrand employing $g_U(\mathbf{x})$ is not sensitive to errors in the “corners” of domain \mathcal{U} . Therefore, CD_2 LHDs also perform very well.

Recalling our desire to achieve design *robustness*, we shall drop the CD_2 LHDs and support points from the next numerical example with the argument that the designs are fundamentally wrong due to their statistical *nonuniformity*. The following numerical study is performed only with statistically uniform sampling techniques.

4.3. Integration of the Subspace Formed by Chebyshev Polynomials

After employing three simple functions, we now measure the performance of integration schemes on a class of smooth functions using Chebyshev polynomials. Chebyshev polynomials of the first kind form a sequence of functions and when they are mapped onto our interval of interest, $\langle 0, 1 \rangle$, the first two polynomials read: $T_0(x) = 1$ and $T_1(x) = 2x - 1$. Higher terms can be constructed using the recurrence relation: $T_k(x) = 2T_1(x)T_{k-1}(x) - T_{k-2}(x)$. These polynomials form a sequence of orthogonal functions on $\langle 0, 1 \rangle$. Any arbitrary function, $f(\mathbf{x})$, can be written as an infinite linear combination of T_k s and also approximated by a function, $g(\mathbf{x})$, that is written as a finite sum exploiting Chebyshev polynomials as

$$f(\mathbf{x}) = \sum_{k=0}^{\infty} a_k T_k(\mathbf{x}) \approx g(\mathbf{x}) = \sum_{k=0}^K a_k T_k(\mathbf{x}), \quad (14)$$

where K is the truncation threshold.

The extension of this concept into an s -dimensional unit hypercube is straightforward. Our basis functions are now s -dimensional products of unidimensional polynomials of arbitrary order

$$\begin{aligned} f(\mathbf{x}) \approx g(\mathbf{x}) &\equiv \sum_{k_1=0}^K \sum_{k_2=0}^K \dots \sum_{k_s=0}^K a_{k_1, k_2, \dots, k_s} T_{k_1}(x_1) T_{k_2}(x_2) \dots T_{k_s}(x_s) \\ &= \sum_{\mathbf{k}} a_{\mathbf{k}} T_{\mathbf{k}}(\mathbf{x}). \end{aligned} \quad (15)$$

We keep the truncation thresholds K identical for all dimensions. The last expression is a symbolic simplification of the previous one where the indexing vector \mathbf{k} runs over all s -dimensional vectors with integer elements in the range $\langle 0, K \rangle$.

Studying the integration of an arbitrary function, $I_f = \int_{\mathcal{U}} f(\mathbf{x}) d\mathbf{x}$ will now be limited to a subspace of functions g , denoted as $\mathcal{G}(\mathbb{R}^s)$, that can be represented by the truncated series with a finite level K

$$I_g = \int_{[0,1]^s} g(\mathbf{x}) d\mathbf{x} = \sum_{\mathbf{k}} a_{\mathbf{k}} \int_0^1 \int_0^1 \dots \int_0^1 T_{\mathbf{k}}(\mathbf{x}) dx_1 dx_2 \dots dx_s. \quad (16)$$

A particular function from the subset $\mathcal{G}(\mathbb{R}^s)$, is determined by a selection of vector of coefficients $a_{\mathbf{k}}$.

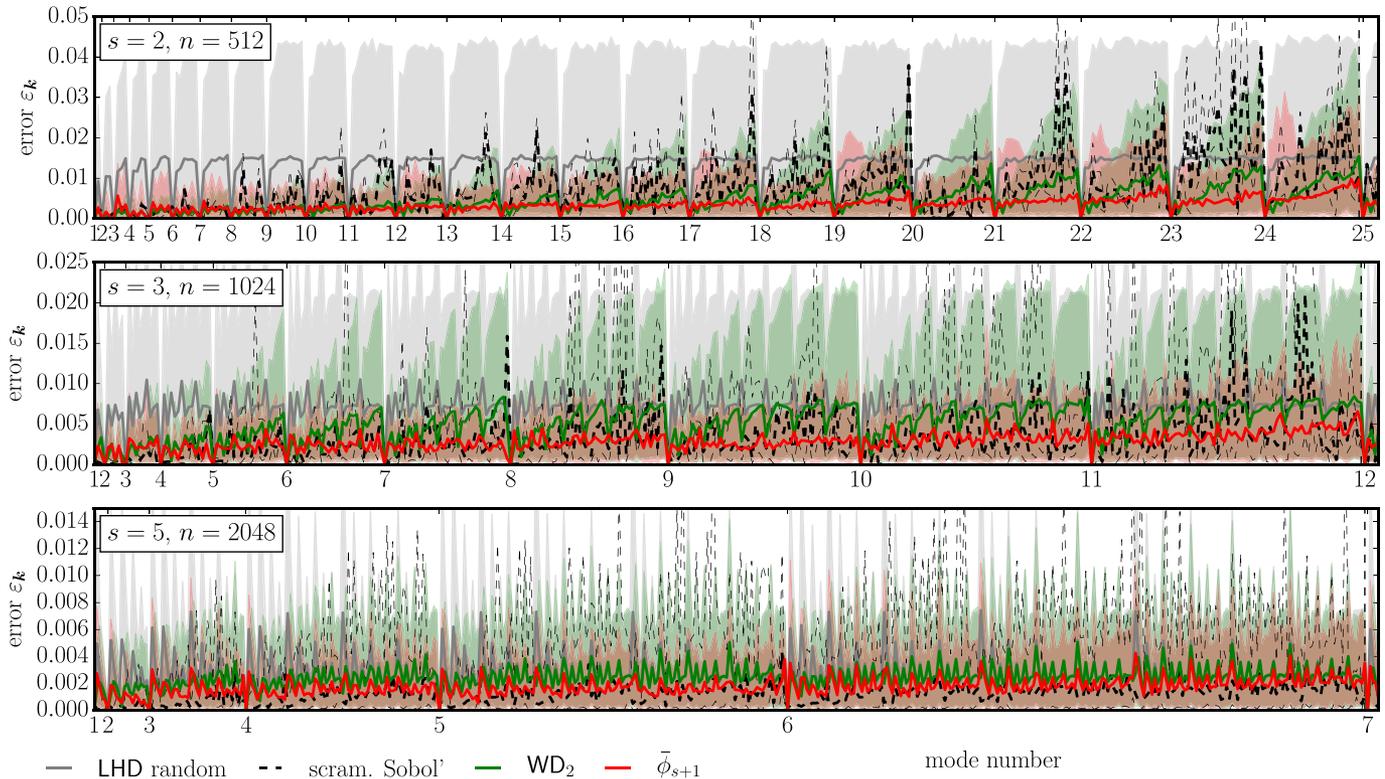


Figure 10. Integration errors for modes of Chebyshev polynomials with increasing mode number. The dimensions are (from the top): $s = 2, 3$, and 5 . The horizontal axis shows the maximum mode index in vector \mathbf{k} . The medians are surrounded by the 5th and 95th percentiles.

The integral in Equation (16) is estimated using a finite sample of points by an average using the truncated Chebyshev polynomial representation

$$\hat{I}_g(\mathcal{D}) = \sum_k^K a_k \left(\frac{1}{n} \sum_{i=1}^n T_k(x_i) \right). \quad (17)$$

For a given vector \mathbf{k} , the absolute error of such a numerical integration reads

$$\epsilon_{\mathbf{k}} = \left| \int_0^1 \int_0^1 \dots \int_0^1 T_{\mathbf{k}}(\mathbf{x}) dx_1 dx_2 \dots dx_s - \frac{1}{n} \sum_{i=1}^n T_{\mathbf{k}}(\mathbf{x}_i) \right|. \quad (18)$$

The total error due to the numerical integration of a particular function $g(\mathbf{x})$ becomes a linear combination of errors for each vector \mathbf{k} : $\epsilon = \sum_{\mathbf{k}} a_{\mathbf{k}} \epsilon_{\mathbf{k}}$. The error is dependent (i) on the vector of coefficients $a_{\mathbf{k}}$ (the selection of a particular function from $\mathcal{G}(\mathbb{R}^s)$) and, (ii) on the ability to integrate individual Chebyshev polynomials $T_{\mathbf{k}}(\mathbf{x})$. Any design that integrates these functions well can also effectively integrate any g function from $\mathcal{G}(\mathbb{R}^s)$. The exact integrals of $T_{\mathbf{k}}$ needed in Equation (18) are simply multiplications of exact integrals of unidimensional polynomials T_n . These have the analytical expression

$$\int_0^1 T_k(x) dx = \begin{cases} 0 & \text{for odd } k, \\ \frac{1}{1-k^2} & \text{for even } k. \end{cases} \quad (19)$$

The unidimensional Chebyshev polynomials T_k are wavy functions with k roots. The number of integration points that attempts to integrate such a function with reasonable confidence

should definitely not be lower than k . The same concept can be applied in a multidimensional space where, assuming an orthogonal grid of points, there are $\sqrt[s]{n}$ points in each dimension. Therefore, an s -dimensional design with n points cannot be expected to exhibit a good performance with Chebyshev terms where \mathbf{k} contains modes higher than $\lceil \sqrt[s]{n} \rceil$. Also, the performance of a design should not depend on the order of modes as we would like to see the same performance for any arbitrary permutation of \mathbf{k} . For that reason, we will always integrate all permutations of \mathbf{k} and report the statistical data. Therefore, only sorted vectors \mathbf{k} will be used, that is, $\mathbf{k} = \{1, 1, 2\}$ actually represents all three different permutations $\{1, 1, 2\}$, $\{2, 1, 1\}$, and $\{1, 2, 1\}$. We can also order all the sorted \mathbf{k} that have no items greater than $\lceil \sqrt[s]{n} \rceil$ in the following way: $\mathbf{k}_a \leq \mathbf{k}_b \Leftrightarrow \forall v: k_v^a \leq k_v^b$.

Figure 10 shows the integration errors $\epsilon_{\mathbf{k}}$ for the individual mode vectors \mathbf{k} . Randomly permuted LHDs (gray lines) exhibit the highest median errors and also display the widest bands between the 5th and 95th percentiles. The scrambled Sobol' sequences (RQMC) exhibit excellent performance for the first modes when the functions are simple, but as the polynomials become very wavy, they show quite serrated profiles with occasionally large 95th percentile errors. LHDs obtained by direct WD_2 discrepancy minimization seem to perform better in two-dimensional but are outperformed by RQMC in higher dimensions. The best overall performance is achieved with the proposed $\bar{\phi}_p$ LHDs: with the exception of the very smooth functions (the first modes), they provide very robust results (low errors with low variability). In five dimensions their average efficiency seems to decrease. This is caused by the fact

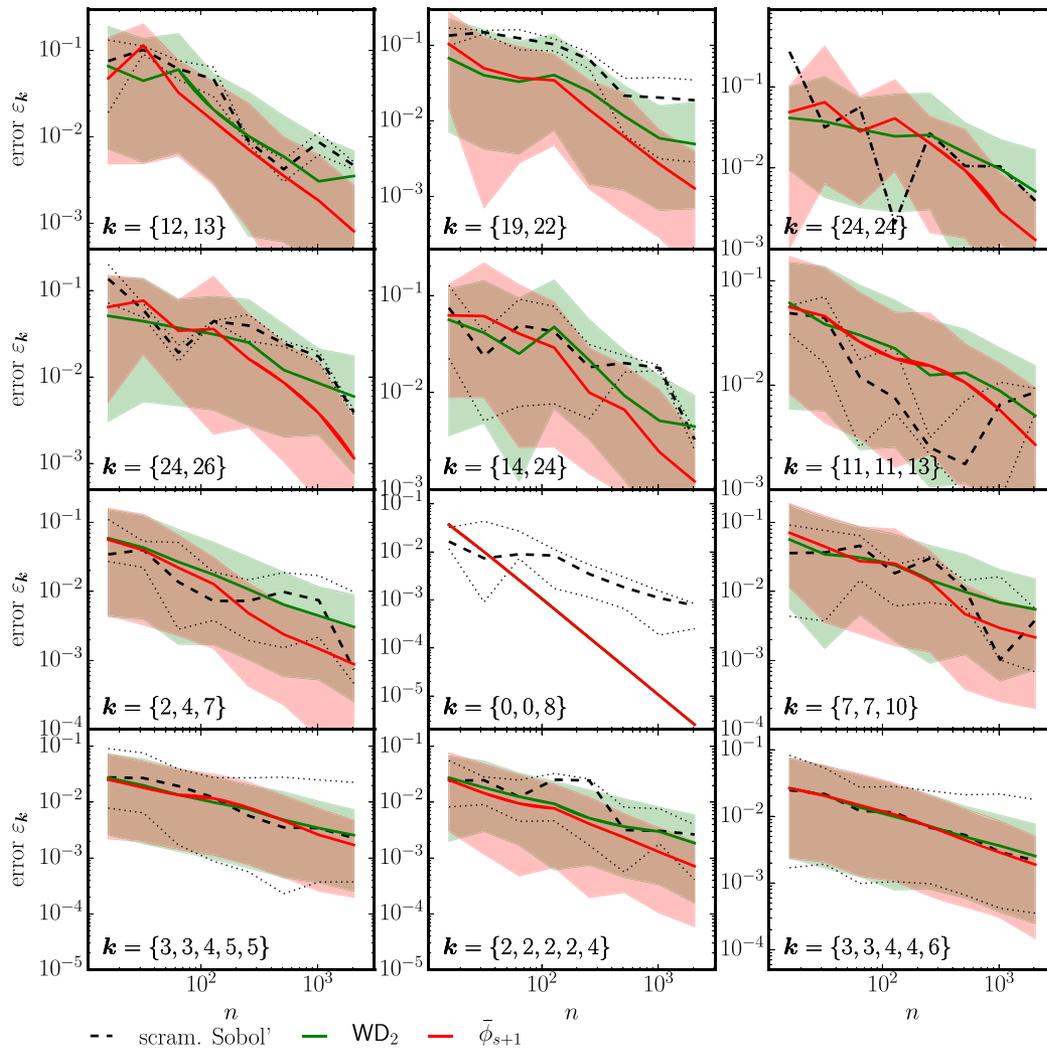


Figure 11. Evolution of integration error with n for different Chebyshev polynomials in various dimensions.

that for $s = 5$, the number of combinations of low modes is so high that the whole spectrum shown is in a regime where the advantage of perfectly equispaced integration points is still not fully utilized. The number of points, n , would have to be much higher to study higher modes in five dimensions. But still, the $\bar{\phi}_p$ LHDs have the lowest 95th percentiles even in five dimensions.

Another way of looking at the results is to fix a particular vector, \mathbf{k} , from Figure 10, and plot the evolution of the integration error as a function of sample size, n . These results are plotted in Figure 11 for twelve randomly selected vectors \mathbf{k} in different dimensions. The same lines and colors are used.

Special cases occur when all components of vector \mathbf{k} are identical (such as $\{24, 24\}$). Due to the symmetries present, permuting variables in the RQMC sequence provides identical error values as the permutations of coordinates are identical.

Another special case occurs when the vectors are formed by all zeros except for one mode (such as $\mathbf{k} = \{0, 0, 8\}$). In these cases, all of the LHDs exhibit no variance as this is essentially the integration of a univariate function using an equidistant grid of points, and therefore all our LHDs integrate it identically.

The average convergence rate of uniform designs (WD_2 LHDs) and RQMC seems to be about the same as the convergence rate achieved by the proposed $\bar{\phi}_p$ LHDs. However, for selected vectors \mathbf{k} , the variance obtained with $\bar{\phi}_p$ LHDs (the band between the 5th and 95th percentiles) seems to be somewhat lower for $\bar{\phi}_p$ LHDs, and also the average trend is smoother. Finally, when the function happens to be univariate, the $\bar{\phi}_p$ LHD scheme fully exploits the perfect projection regularity. Due to the certain degree of robustness they exhibit, we see the proposed $\bar{\phi}_p$ LHDs as being slightly superior to the other types of design under comparison.

5. Conclusions

The Maximin (ϕ_∞) criterion was originally designed to deliver samples that possess a certain kind of optimality when building a metamodel (Kriging). Later, a generalization in the form of the ϕ_p criterion was proposed; it also covered the Audze–Eglājs criterion. Despite the original goal concerning metamodeling, some authors in the past were tempted to use these criteria for other purposes, such as the generation of optimal samples for Monte Carlo type integration, especially when restricted to LHDs.

- The ϕ_∞ and ϕ_p criteria (with an intersite metric) favor a nonuniform probability of point selection within \mathcal{U} . Limiting the designs to LHDs does not fix the problem. Instead, such designs provide erroneous estimates of integrals (with almost no variability). The integration errors become pronounced for higher s and also for low exponents p in the ϕ_p criterion.
- Systematically nonuniform probabilities of point selection are also obtained when employing support points or the CD_2 criterion for direct discrepancy minimization. CD_2 LHDs provide erroneous estimates for integrals that are sensitive especially to points near the corners of the design domain. WD_2 discrepancy seems to be a correct figure of merit for direct discrepancy minimization when the target is a design for numerical integration; it yields *statistically uniform* designs with low discrepancy.
- The proposed $\bar{\phi}_\infty$ and $\bar{\phi}_p$ criteria (with a periodic metric) yield designs that are *statistically uniform*. Consideration of the periodicity of the design space is a simple remedy that comes with no additional effort—the standard and periodic versions have the same computational complexity. The proposed criteria based on a periodic metric are invariant under translation in any direction, but they are not rotation-invariant.
- Apart from the required *statistical uniformity*, the samples obtained with the proposed $\bar{\phi}_p$ criterion also possess excellent space-fillingness.
- The simultaneous fulfillment of *statistical uniformity* and *sample uniformity* by $\bar{\phi}_p$ designs makes them very *robust*: the integrals have stable convergence rates with a small degree of variability for a broad range of integrands. The study performed on Chebyshev polynomials showed that higher modes (wavy functions) are integrated using the proposed $\bar{\phi}_p$ LHDs more accurately and with smaller error variance than the other schemes under comparison.

The integration errors obtained with $\bar{\phi}_p$ LHDs are, on average, as good as the errors obtained with the studied scrambled Sobol' sequences. However, the convergence of average error to zero is smoother for $\bar{\phi}_p$ LHDs when compared to the sequences.

- $\bar{\phi}_p$ -optimal designs have excellent discrepancy (almost as good as that of designs optimized via discrepancy). However, WD_2 -optimal designs and low discrepancy (RQMC) sequences do *not* reach the same level of distance-based criteria as $\bar{\phi}_p$ -optimal designs do.

Additionally, we would like to mention several other points.

Optimal and near-optimal LHDs. The LHDs presented in this article are constructed using a heuristic combinatorial optimization technique that does not guarantee that optimal designs will be obtained. We have found that for large designs, optimization finds sets with a similar objective function value, but they correspond to a local minimum. Escaping from these easily discovered minima in favor of the global minimum is hard. This slightly impairs the results obtained for the proposed $\bar{\phi}_p$ LHDs and ϕ_∞ LHDs (in terms of integration errors and also WD_2 discrepancy), and also the CD_2 LHDs and WD_2 LHDs. Estimations obtained with designs found in reasonable time provide slightly

higher errors and also higher error variance than the optimal designs.

Applicability to non-LHDs. The numerical studies presented in this article were performed with LHDs. However, the proposed periodic distance-based criteria can equally be used to optimize designs that are not limited to LHDs.

Applicability to other metrics. The article is limited to the Euclidean distance measure. However, the proposed extension to periodic space can also be applied to (nonisotropic) metrics (such as the ℓ_1 or ℓ_∞ distance) in the very same way.

Subspaces. The proposed ϕ_p and $\bar{\phi}_p$ criteria optimize s -dimensional periodic distances. Restriction to LHDs additionally guarantees that one-dimensional projections are perfectly spaced. To guarantee the space-filling properties of designs in various possible projections into subspaces of dimensions from 2 to $s - 1$, the criterion would have to be extended to consider all of these projections. Such a generalization is straightforward and can be defined in a manner similar to that used by Joseph, Gul, and Ba (2015) or Mu and Xiong (2016, eqs. (11) and (12)).

Function prediction. Maximin designs were originally proposed for Kriging-type prediction (Johnson, Moore, and Ylvisaker 1990). This article presents periodic versions of Maximin and general ϕ_p criteria that we propose for use in Monte Carlo integration. The suitability of the proposed criteria for function prediction remains an open question.

Sample size extension. Comparison with QMC and RQMC is not completely fair as they are able to add points one-by-one (open designs), while in the proposed $\bar{\phi}_p$ LHDs the sample size n is known a-priori (closed designs). However, the LHDs obtained by direct WD_2 minimization were also closed designs and therefore the findings obtained regarding the effect of discrepancy on integration error can be compared to those for the $\bar{\phi}_p$ criterion.

Designs with good space-filling properties are constructed mainly for one-stage computer experiments; sequential sample extension can be carried out as well (see, e.g., Kong, Ai, and Tsui 2016). We remark that using the method proposed in (Vořechovský 2015), existing LHDs can be extended (by sets of points that are multiples of the current sample sizes) and optimized via any criterion. This technique is implemented in FReET software (Novák, Vořechovský, and Teplý 2014).

$\bar{\phi}_p$ criterion and integration error. The relationship between discrepancy and integration error has received considerable attention in the literature. What remains a challenge is the theoretical analysis of the relationship between the *periodic*-distance-based criteria proposed in the article and (i) the integration error, and also (ii) the discrepancy measures.

Supplementary Materials

The supplementary materials contain $N_{\text{run}} = 50$ realizations of the proposed $\bar{\phi}_p$ LHDs with the distance exponent $p = s + 1$ for samples sizes $n = \{16, 32, 64, 128, 256, 512, 1024, 2048\}$ and dimensions $s = \{2, 3, 5, 10\}$.

The designs were obtained with the combinatorial optimization algorithm by Vořechovský and Novák (2009).

Funding

The authors acknowledge financial support provided by the Czech Science Foundation under project no. 16-22230S.

References

- Audze, P., and Eglājs, V. (1977), “New Approach for Planning Out of Experiments,” *Problems of Dynamics and Strengths*, 35, 104–107 (in Russian). [2,3,5]
- Baldeaux, J., Dick, J., Greslehner, J., and Pillichshammer, F. (2011), “Construction Algorithms for Higher Order Polynomial Lattice Rules,” *Journal of Complexity*, 27, 281–299. [2]
- Conover, W. (1975), “On a Better Method for Selecting Input Variables,” unpublished Los Alamos National Laboratories manuscript, reproduced as Appendix A of “Latin Hypercube Sampling and the Propagation of Uncertainty in Analyses of Complex Systems” by J.C. Helton and F.J. Davis, Sandia National Laboratories Report SAND2001-0417. [2]
- Cranley, R., and Patterson, T. N. L. (1976), “Randomization of Number Theoretic Methods for Multiple Integration,” *SIAM Journal on Numerical Analysis*, 13, 904–914. [2]
- Dean, A., Morris, M., Stufken, J., and Bingham, D. (2015), *Handbook of Design and Analysis of Experiments*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, Boca Raton, FL: CRC Press. [3]
- Detle, H., and Pepelyshev, A. (2010), “Generalized Latin Hypercube Design for Computer Experiments,” *Technometrics*, 52, 421–429. [2]
- Dick, J. (2011), “Higher Order Scrambled Digital Nets Achieve the Optimal Rate of the Root Mean Square Error for Smooth Integrands,” *The Annals of Statistics*, 39, 1372–1398. [2]
- Dick, J., Kuo, F. Y., and Sloan, I. H. (2013), “High-Dimensional Integration: The Quasi-Monte Carlo Way,” *Acta Numerica*, 22, 133–288. [2,11]
- Dick, J., and Pillichshammer, F. (2010), *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*, New York: Cambridge University Press. [2]
- Dutang, C., and Savický, P. (2019), “randtoolbox: Generating and Testing Random Numbers,” R Package Version 1.30.0. [8]
- Eliáš, J., and Vořechovský, M. (2016), “Modification of the Audze-Eglājs Criterion to Achieve a Uniform Distribution of Sampling Points,” *Advances in Engineering Software*, 100, 82–96. [5,6,9,11]
- Fang, K.-T., Lin, D. K. J., Winker, P., and Zhang, Y. (2000), “Uniform Design: Theory and Application,” *Technometrics*, 42, 237–248. [1,2]
- Fang, K.-T., Liu, M.-Q., Qin, H., and Zhou, Y.-D. (2018), *Theory and Application of Uniform Experimental Designs*, Lecture Notes in Statistics (Vol. 221, 1st ed.), Singapore: Springer. [2]
- Fang, K.-T., and Ma, C.-X. (2001), “Wrap-Around L_2 -Discrepancy of Random Sampling, Latin Hypercube and Uniform Designs,” *Journal of Complexity*, 17, 608–624. [1,2,8]
- Fang, K.-T., Ma, C.-X., and Winker, P. (2000), “Centered L_2 -discrepancy of Random Sampling and Latin Hypercube Design, and Construction of Uniform Designs,” *Mathematics of Computation*, 71, 275–296. [8]
- Fang, K.-T., and Wang, Y. (1993), *Number-Theoretic Methods in Statistics* (1st ed.), London; New York: Chapman and Hall/CRC. [2]
- Fang, K.-T., Wang, Y., and Bentler, P. M. (1994), “Some Applications of Number-Theoretic Methods in Statistics,” *Statistical Science*, 9, 416–428. [2]
- Faure, H. (1981), “Discrépances de Suites Associées à un Système de Numération (en Dimension un) [Discrepancy of Sequences Associated With a Number System (in Dimension One)],” *Bulletin de la Société Mathématique de France*, 109, 143–182. [2]
- Grosso, A., Jamali, A., and Locatelli, M. (2009), “Finding Maximin Latin Hypercube Designs by Iterated Local Search Heuristics,” *European Journal of Operational Research*, 197, 541–547. [2]
- Halton, J. (1960), “On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals,” *Numerische Mathematik*, 2, 84–90. [2]
- He, X. (2017a), “Interleaved Lattice-Based Minimax Distance Designs,” *Biometrika*, 104, 713–725. [2]
- (2017b), “Rotated Sphere Packing Designs,” *Journal of the American Statistical Association*, 112, 1612–1622. [2,3]
- Hickernell, F. J. (1998a), “A Generalized Discrepancy and Quadrature Error Bound,” *Mathematics of Computation*, 67, 299–322. [2,8]
- (1998b), “Lattice Rules: How Well Do They Measure Up?” in *Random and Quasi-Random Point Sets*, eds. P. Hellekalek and G. Larcher, New York, NY: Springer New York, pp. 109–166. [2,8]
- Hlawka, E. (1961), “Funktionen von beschränkter Variation in der Theorie der Gleichverteilung,” *Annali di Matematica Pura ed Applicata*, 54, 325–333. [1]
- Hong, H. S., and Hickernell, F. J. (2003), “Algorithm 823: Implementing Scrambled Digital Sequences,” *ACM Transactions on Mathematical Software*, 29, 95–109. [8]
- Husslage, B. G. M., Rennen, G., van Dam, E. R., and den Hertog, D. (2011), “Space-Filling Latin Hypercube Designs for Computer Experiments,” *Optimization and Engineering*, 12, 611–630. [2]
- Janssen, H. (2013), “Monte-Carlo Based Uncertainty Analysis: Sampling Efficiency and Sampling Convergence,” *Reliability Engineering & System Safety*, 109, 123–132. [2]
- Jin, R., Chen, W., and Sudjianto, A. (2005), “An Efficient Algorithm for Constructing Optimal Design of Computer Experiments,” *Journal of Statistical Planning and Inference*, 134, 268–287. [3]
- Johnson, M., Moore, L., and Ylvisaker, D. (1990), “Minimax and Maximin Distance Designs,” *Journal of Statistical Planning and Inference*, 2, 131–148. [2,3,6,14]
- Joseph, V. R. (2016), “Space-Filling Designs for Computer Experiments: A Review,” *Quality Engineering*, 28, 28–35. [2]
- Joseph, V. R., Gul, E., and Ba, S. (2015), “Maximum Projection Designs for Computer Experiments,” *Biometrika*, 102, 371–380. [14]
- Joseph, V. R., and Hung, Y. (2008), “Orthogonal-Maximin Latin Hypercube Designs,” *Statistica Sinica*, 18, 171–186. [2,3]
- Koksma, J. F. (1942/1943), “Een algemeene stelling uit de theorie der gelijkmatige verdeling modulo 1,” *Mathematica B*, 11, 7–11. [1]
- Kong, X., Ai, M., and Tsui, K. L. (2016), “Design for Sequential Follow-Up Experiments in Computer Emulations,” *Technometrics*, 60, 61–69. [14]
- L’Ecuyer, P., and Lemieux, C. (2005), “Recent Advances in Randomized Quasi-Monte Carlo Methods,” in *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, International Series in Operations Research & Management Science (Vol. 46, Part V), eds. M. Dror, P. L’Ecuyer, and F. Szidarovszky, New York, NY: Springer, pp. 419–474. [2]
- Liefvendahl, M., and Stocki, R. (2006), “A Study on Algorithms for Optimization of Latin Hypercubes,” *Journal of Statistical Planning and Inference*, 136, 3231–3247. [2]
- Mak, S. (2017), “support: Support Points,” R Package Version 0.1.2. [6]
- Mak, S., and Joseph, V. R. (2018a), “Minimax and Minimax Projection Designs Using Clustering,” *Journal of Computational and Graphical Statistics*, 27, 166–178. [2]
- (2018b), “Support Points,” *The Annals of Statistics*, 46, 2562–2592. [2,4,5]
- Mašek, J., and Vořechovský, M. (2018), “Parallel Implementation of Hyper-Dimensional Dynamical Particle System on CUDA,” *Advances in Engineering Software*, 125, 178–187. [7]
- McKay, M. D., Conover, W. J., and Beckman, R. J. (1979), “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output From a Computer Code,” *Technometrics*, 21, 239–245. [2]
- Morris, M. D., and Mitchell, T. J. (1995), “Exploratory Designs for Computational Experiments,” *Journal of Statistical Planning and Inference*, 43, 381–402. [2,3]
- Mu, W., and Xiong, S. (2016), “On Algorithmic Construction of Maximin Distance Designs,” *Communications in Statistics—Simulation and Computation*, 46, 1–14. [14]
- Niederreiter, H. (1987), “Point Sets and Sequences With Small Discrepancy,” *Monatshefte für Mathematik*, 104, 273–337. [2]

- (1988), “Low-Discrepancy and Low-Dispersion Sequences,” *Journal of Number Theory*, 30, 51–70. [2]
- (1992), *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM). [1,2,8]
- Novák, D., Vořechovský, M., and Teplý, B. (2014), “FReET: Software for the Statistical and Reliability Analysis of Engineering Problems and FReET-D: Degradation Module,” *Advances in Engineering Software*, 72, 179–192. [14]
- Owen, A. B. (1992), “A Central Limit Theorem for Latin Hypercube Sampling,” *Journal of the Royal Statistical Society, Series B*, 54, 541–551. [2]
- (1994), “Controlling Correlations in Latin Hypercube Samples,” *Journal of the American Statistical Association*, 89, 1517–1522. [2]
- (1995), “Randomly Permuted $(t; m; s)$ -Nets and $(t; s)$ -Sequences,” in *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing* (Proceedings of a Conference at the University of Nevada, Las Vegas, Nevada, USA, June 23–25, 1994), Lecture Notes in Statistics (Vol. 106), eds. R. E. Caflisch, B. Moskowitz, H. Niederreiter, and P. J.-S. Shiue, New York: Springer-Verlag, pp. 299–317. [2,8]
- (1998), “Scrambling Sobol’ and Niederreiter-Xing Points,” *Journal of Complexity*, 14, 466–489. [2,8]
- (1999), “Monte Carlo Quasi-Monte Carlo and Randomized Quasi-Monte Carlo,” in *Monte-Carlo and Quasi-Monte Carlo Methods 1998* (Proceedings of a Conference Held at the Claremont Graduate University, Claremont, California, USA, June 22–26, 1998), Lecture Notes in Statistics, eds. H. Niederreiter and J. Spanier, Berlin, Heidelberg: Springer-Verlag, pp. 86–97. [2,8]
- Pholdee, N., and Bureerat, S. (2015), “An Efficient Optimum Latin Hypercube Sampling Technique Based on Sequencing Optimisation Using Simulated Annealing,” *International Journal of Systems Science*, 46, 1780–1789. [3]
- R Core Team (2013), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. [3,8]
- Rennen, G., Husslage, B., Van Dam, E. R., and Den Hertog, D. (2010), “Nested Maximin Latin Hypercube Designs,” *Structural and Multidisciplinary Optimization*, 41, 371–395. [2]
- Sadilek, V., and Vořechovský, M. (2018), “Evaluation of Pairwise Distances Among Points Forming a Regular Orthogonal Grid in a Hypercube,” *Journal of Civil Engineering and Management*, 24, 410–523. [6]
- Santner, T. J., Williams, B. J., and Notz, W. I. (2003), *The Design and Analysis of Computer Experiments*, Springer Series in Statistics, New York: Springer-Verlag. [2,3]
- Shields, M. D., and Zhang, J. (2016), “The Generalization of Latin Hypercube Sampling,” *Reliability Engineering & System Safety*, 148, 96–108. [2]
- Sobol’, I. (1967), “On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals,” *USSR Computational Mathematics and Mathematical Physics*, 7, 784–802. [2]
- (1976), “Uniformly Distributed Sequences With an Additional Uniform Property,” *USSR Computational Mathematics and Mathematical Physics*, 16, 236–242 (Short Communication). [2]
- Stein, M. (1987), “Large Sample Properties of Simulations Using Latin Hypercube Sampling,” *Technometrics*, 29, 143–151. [2]
- Székeley, G. J., and Rizzo, M. L. (2004), “Testing for Equal Distributions in High Dimensions,” *InterStat*, 5, 1–16. [6]
- (2013), “Energy Statistics: A Class of Statistics Based on Distances,” *Journal of Statistical Planning and Inference*, 143, 1249–1272. [6]
- Tan, M. H. Y. (2013), “Minimax Designs for Finite Design Regions,” *Technometrics*, 55, 346–358. [2]
- Tang, B. (1993), “Orthogonal Array-Based Latin Hypercubes,” *Journal of the American Statistical Association*, 88, 1392–1397. [2]
- Tezuka, S. (1995), *Uniform Random Numbers: Theory and Practice*, The Springer International Series in Engineering and Computer Science (Vol. 315), Boston, MA: Springer. [2]
- Tezuka, S., and Faure, H. (2003), “ t -Binomial Scrambling of Digital Nets and Sequences,” *Journal of Complexity*, 19, 744–757. [8]
- van Dam, E., den Hertog, D., Husslage, B., and Rennen, G. (2017), “Space-Filling Designs: A Free Online Database of Space-Filling Designs for (Computer) Experiments,” available at <https://spacefillingdesigns.nl/>. [7]
- van Dam, E. R., Husslage, B., den Hertog, D., and Melissen, H. (2007), “Maximin Latin Hypercube Designs in Two Dimensions,” *Operations Research*, 55, 158–169. [2,3]
- van Dam, E. R., Rennen, G., and Husslage, B. (2009), “Bounds for Maximin Latin Hypercube Designs,” *Operations Research*, 57, 595–608. [2,3]
- Viana, F. A. C., Venter, G., and Balabanov, V. (2010), “An Algorithm for Fast Optimal Latin Hypercube Design of Experiments,” *International Journal for Numerical Methods in Engineering*, 82, 135–156. [3]
- Vořechovský, M. (2015), “Hierarchical Refinement of Latin Hypercube Samples,” *Computer-Aided Civil and Infrastructure Engineering*, 30, 394–411. [14]
- Vořechovský, M., Mašek, J., and Eliáš, J. (2017), “Dynamical Model of Interacting Particles for the Construction of Audze-Eglajs Designs in a Periodic Space,” in *Safety, Reliability, Risk, Resilience and Sustainability of Structures and Infrastructure—12th Int. Conf. on Structural Safety and Reliability, Vienna, Austria*, eds. C. Bucher, B. R. Ellingwood, and D. M. Frangopol, Vienna: TU-Verlag, pp. 1374–1383. [7]
- (2019), “Distance-Based Optimal Sampling in a Hypercube: Analogies to N-Body Systems,” *Advances in Engineering Software* (in review). [7]
- Vořechovský, M., and Novák, D. (2009), “Correlation Control in Small Sample Monte Carlo Type Simulations I: A Simulated Annealing Approach,” *Probabilistic Engineering Mechanics*, 24, 452–462. [2,7,15]
- Wuertz, D., Setz, T., and Chalabi, Y. (2017), “fOptions’: Rmetrics—Pricing and Evaluating Basic Options,” R Package Version 3.0.1. [8]
- Ye, K. Q., Li, W., and Sudjianto, A. (2000), “Algorithmic Construction of Optimal Symmetric Latin Hypercube Designs,” *Journal of Statistical Planning and Inference*, 90, 145–159. [3]