

A Spectral Approach for the Design of Experiments: Design, Analysis and Algorithms

Bhavya Kailkhura

KAILKHURA1@LLNL.GOV

*Center for Applied Scientific Computing
Lawrence Livermore National Lab
Livermore, CA 94550, USA*

Jayaraman J. Thiagarajan

JJAYARAM@LLNL.GOV

*Center for Applied Scientific Computing
Lawrence Livermore National Lab
Livermore, CA 94550, USA*

Charvi Rastogi

CHARVIRASTOGI@IITB.AC.IN

*Department of EECS
Indian Institute of Technology
Bombay, MH 400076, India*

Pramod K. Varshney

VARSHNEY@SYR.EDU

*Departments of EECS
Syracuse University
Syracuse, NY 13244, USA*

Peer-Timo Bremer

BREMER5@LLNL.GOV

*Center for Applied Scientific Computing
Lawrence Livermore National Lab
Livermore, CA 94550, USA*

Editor: Animashree Anandkumar

Abstract

This paper proposes a new approach to construct high quality space-filling sample designs. First, we propose a novel technique to quantify the space-filling property and optimally trade-off uniformity and randomness in sample designs in arbitrary dimensions. Second, we connect the proposed metric (defined in the spatial domain) to the quality metric of the design performance (defined in the spectral domain). This connection serves as an analytic framework for evaluating the qualitative properties of space-filling designs in general. Using the theoretical insights provided by this spatial-spectral analysis, we derive the notion of optimal space-filling designs, which we refer to as space-filling spectral designs. Third, we propose an efficient estimator to evaluate the space-filling properties of sample designs in arbitrary dimensions and use it to develop an optimization framework for generating high quality space-filling designs. Finally, we carry out a detailed performance comparison on two different applications in varying dimensions: a) image reconstruction and b) surrogate modeling for several benchmark optimization functions and a physics simulation code for inertial confinement fusion (ICF). Our results clearly evidence the superiority of the proposed space-filling designs over existing approaches, particularly in high dimensions.

Keywords: design of experiments, space-filling, poisson-disk sampling, surrogate modeling, regression

1. Introduction

Exploratory analysis and inference in high dimensional parameter spaces is a ubiquitous problem in science and engineering. As a result, a wide-variety of machine learning tools and optimization techniques have been proposed to address this challenge. In its most generic formulation, one is interested in analyzing a high-dimensional function $f : \mathcal{D} \rightarrow \mathbb{R}$ defined on the d -dimensional domain \mathcal{D} . A typical approach for such an analysis is to first create an initial sampling $\mathcal{X} = \{\mathbf{x}_i \in \mathcal{D}\}_{i=1}^N$ of \mathcal{D} , evaluate f at all \mathbf{x}_i , and perform subsequent analysis and learning using only the resulting tuples $\{(\mathbf{x}_i, f(\mathbf{x}_i))\}_{i=1}^N$. Despite the widespread use of this approach, a critical question that still persists is: how should one obtain a high quality initial sampling \mathcal{X} for which the data $f(\mathcal{X})$ is acquired or generated? This challenge is typically referred to as Design of Experiments (DoE) and solutions have been proposed as early as (Fisher, 1935) that optimized agricultural experiments. Subsequently, DoE has received significant attention from researchers in different fields (Garud et al., 2017). It is also an important *building block* for a wide variety of machine learning applications, such as, supervised machine learning, neural network training, image reconstruction, reinforcement learning, etc. (for a detailed discussion see Section 10). In several scenarios, it has been shown that success crucially depends on the quality of the initial sampling \mathcal{X} . Currently, a plethora of sampling solutions exist in the literature with a wide-range of assumptions and statistical guarantees; see (Garud et al., 2017; Owen, 2009) for a detailed review of related methods. Conceptually, most approaches aim to cover the sampling domain as uniformly as possible, in order to generate the so called *space-filling* experimental designs (Joseph, 2016)¹. However, it is well known that uniformity alone does not necessarily lead to high performance. For example, optimal sphere packings lead to highly uniform designs, yet are well known to cause strong aliasing artifacts most easily perceived by the human visual system in many computer graphics applications. Instead, a common assumption is that a good design should balance uniformity and randomness². Unfortunately, an exact definition for what should be considered a good space-filling design remains elusive.

Most common approaches use various scalar metrics to encapsulate different notions of ideal sampling properties. One popular metric is the discrepancy of an experimental design, defined as an appropriate ℓ_p norm of the ratio of points within all (hyper-rectangular) sub-volumes of \mathcal{D} and the corresponding volume ratio. In other words, discrepancy quantifies the non-uniformity of a sample design. The most prominent examples of so called *discrepancy sequences* are Quasi-Monte Carlo (QMC) methods and their variants (Caffisch, 1998). In their classical form, discrepancy sequences are deterministic though extensions to incorporate randomness have been proposed, for example, using digital scrambling (Owen, 1995). Nevertheless, by optimizing for discrepancy these techniques focus almost exclusively on uniformity, and consequently even optimized QMC patterns can be quite structured and create

-
1. The term “space-filling” is used widely in the literature on design of experiments. However, in most cases, space-filling is meant in an intuitive sense and as a synonym for “evenly or uniformly spread”. Later in this paper, we will provide a technical definition of space-filling property and what should be considered a good sample design.
 2. In this paper, by randomness we mean that sample points are uniformly distributed over space. Here “uniform” is used in the sense that sample points follow a uniform probability distribution over the sampling region and that each location is equally likely to be selected as sample location, not in the sense that they are “evenly dispersed over the sampling region.”

aliasing artifacts. Furthermore, even the fastest known strategies for evaluating popular discrepancy measures require $O(N^2d)$ operations making evaluation, let alone optimization, for discrepancy difficult even for moderate dimensions. Finally, for most discrepancy measures, the optimal achievable values are not known. This makes it difficult to determine whether a poorly performing sample design (e.g., in terms of generalization (or test error) in a regression application and reconstruction error in an image reconstruction application) is due to the insufficiency of the chosen discrepancy measure or due to ineffective optimization.

Another class of metrics to describe sample designs are based on geometric distances. These can be used directly by, for example, optimizing the maximin or minimax distance of a sample design (Schlomer et al., 2011) or indirectly by enforcing empty disk conditions. The latter is the basis for the so-called Poisson disk samples (Lagae and Dutr, 2008), which aim to generate random points such that no two samples can be closer than a given minimal distance r_{min} , i.e. enforcing an empty disc of radius r_{min} around each sample. Typically, Poisson-type samples are characterized by the *relative radius*, ρ , defined as the ratio of the minimum disk radius r_{min} and the maximum possible disk radius r_{max} for N samples to cover the sampling domain. Similar to the discrepancy sequences, maximin and minimax designs exclusively consider uniformity, are difficult to optimize for especially in higher dimensions, and often lead to very regular patterns. Poisson disk samples use ρ to trade off randomness (lower ρ values) and uniformity (higher ρ values). A popular recommendation in 2-d is to aim for $0.65 \leq \rho \leq 0.85$ as a good compromise. However, there does not exist any theoretical guidance for choosing ρ and hence, optimal values for higher dimensions are not known. As discussed in more detail in Section 2, there also exist a wide variety of techniques that combine different metrics and heuristics. For example, Latin Hypercube sampling (LHS) aims to spread the sample points uniformly by stratification, and one can potentially optimize the resulting design using maximin or minimax techniques (Jin et al., 2005).

In general, scalar metrics to evaluate the quality of a sample design tend not to be very descriptive. Especially in high dimensions different designs with, for example, the same ρ can exhibit widely different performance and for some discrepancy sequences the optimal designs converge to random samples in high dimensions (Morokoff and Caffisch, 1994; Wang and Sloan, 2008). Furthermore, one rarely knows the best achievable value of the metric, i.e. the lowest possible discrepancy, for a given problem which makes evaluating and comparing sampling designs difficult. Finally, most metrics are expensive to compute and not easily optimized. This makes it challenging in practice to create good designs in high dimensions and with large sample sizes.

To alleviate this problem, we propose a new technique to quantify the space-filling property, which enables us to systematically trade-off uniformity and randomness, consequently producing better quality sampling designs. More specifically, we use tools from statistical mechanics to connect the qualitative performance (in the spectral domain) of a sampling pattern with its spatial properties characterized by the pair correlation function (PCF). The PCF measures the distribution of point samples as a function of distances, thus, providing a holistic view of the space-filling property (See Figure 1(b)). Furthermore, we establish the connection between the PCF and the power spectral density (PSD) via the 1-D Hankel transform in arbitrary dimensions, thus providing a relation between the PCF and the quality metric of sampling quality to help subsequent design and analysis.

Using insights from the analysis of space-filling designs in the spectral domain, we provide design guidelines to systematically trade-off uniformity and randomness for a good sampling pattern. The analytical tractability of the PCF enables us to perform theoretical analysis in the spectral domain to derive the structure of optimal space-filling designs, referred to as *space-filling spectral design* in the rest of this paper. Next, we develop an edge corrected kernel density estimator based technique to measure the space-filling property via PCFs in arbitrary dimensions. In contrast to existing PCF estimation techniques, the proposed PCF estimator is both accurate and computationally efficient. Based on this estimator, we develop a systematic optimization framework and a novel algorithm to synthesize space-filling spectral designs. In particular, we propose to employ a weighted least-squares based gradient descent optimization, coupled with the proposed PCF estimator, to accurately match the optimal space-filling spectral design defined in terms of the PCF.

Note that there is a strong connection between the proposed space-filling spectral designs and coverage based designs such as Poisson Disk Sampling (PDS) (Gamito and Maddock, 2009). However, the major difference lies in the metric/criterion these techniques use to estimate and optimize the space-filling designs. Furthermore, existing works on PDS focus primarily on algorithmic issues, such as worst-case running times and numerical issues associated with providing high-quality implementations. However, different PDS methods often demonstrate widely different performances which raises the questions of how to evaluate the qualitative properties of different PDS patterns and how to define an optimal PDS pattern? We argue that, coverage (ρ) based metrics alone are insufficient for understanding the statistical aspects of PDS. This makes it difficult to generate high quality PDS patterns. As we will demonstrate below, existing PDS approaches largely ignore the randomness objective and instead concentrate exclusively on the coverage objective resulting in inferior sampling patterns compared to space-filling spectral designs, especially in high dimensions. Note that on the other hand, the proposed PCF based metric does not have these limitations and enables a comprehensive analysis of statistical properties of space-filling designs (including PDS), while producing higher quality sampling patterns compared to the state-of-the-art PDS approaches.

In (Kailkhura et al., 2016a), we use the PCF to understand the nature of PDS and provided theoretical bounds on the sample size of achievable PDS. Here we significantly extend our previous work and provide a more comprehensive analysis of the problem along with a novel space-filling spectral designs, an edge corrected PCF estimator, an optimization approach to synthesize the space-filling spectral designs and a detailed evaluation of the performance of the proposed sample design. The main contributions of this paper can be summarized as follows:

- We provide a novel technique to quantify the space-filling property of sample designs in arbitrary dimensions and systematically trade-off uniformity and randomness.
- We use tools from statistical mechanics to connect the qualitative performance (in the spectral domain) of a sample design with its spatial properties characterized by the PCF.
- We develop a computationally efficient edge corrected kernel density estimator based technique to estimate the space-filling property in arbitrary dimensions.

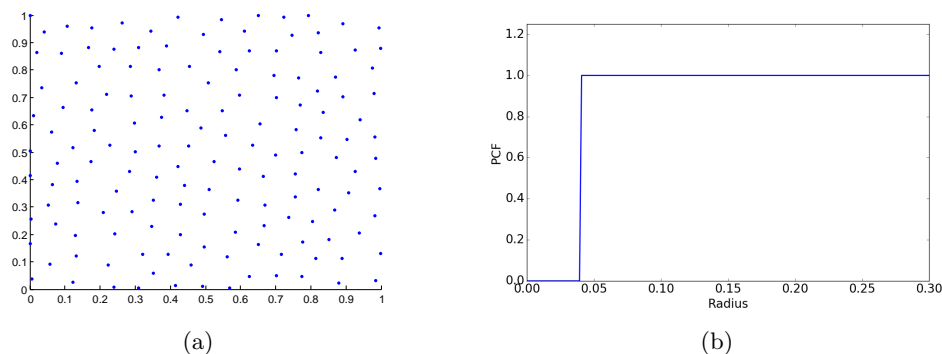


Figure 1: A sample design that balances *randomness* and *uniformity*. (a) Point distribution, and (b) Pair correlation function.

- Using theoretical insights obtained via spectral analysis of point distributions, we provide design guidelines for optimal space-filling designs.
- We devise a systematic optimization framework and a gradient descent optimization algorithm to generate high quality space-filling designs.
- We demonstrate the superiority of proposed space-filling spectral samples compared to existing space-filling approaches through rigorous empirical studies on two different applications: *a)* image reconstruction and *b)* surrogate modeling on several benchmark optimization functions and an inertial confinement fusion (ICF) simulation code.

2. Related Work

In this section, we provide a brief overview of existing approaches for creating space-filling sampling patterns. Note that the prior art for this long-studied research area is too extensive to cover in detail, and hence we recommend interested readers to refer to (Garud et al., 2017; Owen, 2009) for a more comprehensive review.

2.1 Latin Hypercube Sampling

Monte-Carlo methods form an important class of techniques for space-filling sample design. However, it is well known that Monte-Carlo methods are characterized by high variance in the resulting sample distributions. Consequently, variance reduction methods are employed in practice to improve the performance of simple Monte Carlo techniques. One example is stratified sampling using the popular Latin Hypercube Sampling (LHS) (McKay, 1992; Packham, 2015). Since its inception, several variants of LHS have been proposed with the goal of achieving better space-filling properties, in addition to reducing variance. A notable improvement in this regard are techniques that achieve LHS space filling not only in one-dimensional projections, but also in higher dimensions. For example, Tang (Tang, 1993; Leary et al., 2003) introduced orthogonal-array-based Latin hypercube sampling to improve space-filling in higher dimensional subspaces. Furthermore, a variety of space-filling criteria

such as entropy, integrated mean square error, minimax and maximin distances, have been utilized for optimizing LHS (Jin et al., 2005). A particularly effective and widely adopted metric is the maximin distance criterion, which maximizes the minimal distance between points to avoid designs with points too close to one another (Morris and Mitchell, 1995). A detailed study on LHS and its variants can be found in (Koehler and Owen, 1996).

2.2 Quasi Monte Carlo Sampling

Following the success of Monte-Carlo methods, Quasi-Monte Carlo (QMC) sampling was introduced in (Halton, 1964) and since then has become the de facto solution in a wide-range of applications (Caffisch, 1998; Wang and Sloan, 2008). The core idea of QMC methods is to replace the random or pseudo-random samples in Monte-Carlo methods with well-chosen deterministic points. These deterministic points are chosen such that they are highly uniform, which can be quantified using the measure of discrepancy. Low-discrepancy sequences along with bounds on their discrepancy were introduced in the 1960's by Halton (Halton, 1964) and Sobol (Sobol, 1967), and are still in use today. However, despite their effectiveness, a critical limitation of QMC methods is that error bounds and statistical confidence bounds of the resulting designs cannot be obtained due to the deterministic nature of low-discrepancy sequences. In order to alleviate this challenge, randomized quasi-Monte Carlo (RQMC) sampling has been proposed (LEcuyer and Lemieux, 2005), and in many cases shown to be provably better than the classical QMC techniques (Owen and Tribble, 2005). This has motivated the development of other randomized quasi-Monte Carlo techniques, for example, methods based on digital scrambling (Owen, 1995).

2.3 Poisson Disk Sampling

While discrepancy-based designs have been popular among uncertainty quantification researchers, the computer graphics community has had long-standing success with coverage-based designs. In particular, Poisson disk sampling (PDS) is widely used in applications such as image/volume rendering. The authors in (Dippe and Wold, 1985; Cook, 1986) were the first to introduce PDS for turning regular aliasing patterns into featureless noise, which makes them perceptually less visible. Their work was inspired by the seminal work of Yellott *et.al.* (Yellott, 1983), who observed that the photo-receptors in the retina of monkeys and humans are distributed according to a Poisson disk distribution, thus explaining its effectiveness in imaging.

Due to the broad interest created by the initial work on PDS, a large number of approaches to generate Poisson disk distributions have been developed over the last decade (Gamito and Maddock, 2009; Ebeida et al., 2012, 2011; Ip et al., 2013; Bridson, 2007; Oztireli and Gross, 2012; Heck et al., 2013; Wei, 2008; Dunbar and Humphreys, 2006; Wei, 2010; Balzer et al., 2009; Geng et al., 2013; Yan and Wonka, 2012a, 2013; Ying et al., 2013b, 2014; Hou et al., 2013; Ying et al., 2013a; Guo et al., 2014; Wachtel et al., 2014; Xu et al., 2014; Ebeida et al., 2014; de Goes et al., 2012; Zhou et al., 2012). Most Poisson disk sample generation methods are based on dart throwing (Dippe and Wold, 1985; Cook, 1986), which attempts to generate as many darts as necessary to cover the sampling domain while not violating the Poisson disk criterion. Given the desired disk size r_{min} (or coverage ρ), dart throwing generates random samples and rejects or accepts each sample based on its distance

to the previously accepted samples. Despite its effectiveness, its primary shortcoming is the choice of termination condition, since the algorithm does not know whether or not the domain is fully covered. Hence, in practice, the algorithm has poor convergence, which in turn makes it computationally expensive. On the other hand, dart throwing is easy to implement and applicable to any sampling domain, even non-Euclidean. For example, Anirudh *et al.* use a dart throwing technique to generate Poisson disk samples on the Grassmannian manifold of low-dimensional linear subspaces (Anirudh *et al.*, 2017).

Reducing the computational complexity of PDS generation, particularly in low and moderate dimensions, has been the central focus of many existing efforts. To this end, approximate techniques that produce sample sets with characteristics similar to Poisson disk have been developed. Early examples (McCool and Fiume, 1992) are relatively simple and can be used for a wide range of sampling domains, but the gain in computational efficacy is limited. Other methods partition the space into grid cells in order to allow parallelization across the different cells and achieve linear time algorithms (Bridson, 2007). Another class of approaches, referred to as *tile-based* methods, have been developed for generating a large number of Poisson disk samples in 2-D. Broadly, these methods either start with a smaller set of samples, often obtained using other PDS techniques, and tile these samples (Wachtel *et al.*, 2014), or alternatively use a regular tile structure for placing each sample (Ostromoukhov *et al.*, 2004). With the aid of efficient data structures, these methods can generate a large number of samples efficiently. Unfortunately, these approximations can lead to low sample quality due to artifacts induced at tile boundaries and the inherent non-random nature of tilings. More recently, many researchers have explored the idea of partitioning the sampling space in order to avoid generating new samples that will be ultimately rejected by dart throwing. While some of these methods only work in 2-D (Dunbar and Humphreys, 2006; Ebeida *et al.*, 2011), the efficiency of other methods that are designed for higher dimensions (Gamito and Maddock, 2009; Ebeida *et al.*, 2012) drops exponentially with increasing dimensions. Finally, relaxation methods that iteratively increase the Poisson disk radius of a sample set (McCool and Fiume, 1992) by re-positioning the samples also exist. However, these methods have the risk of converging to a regular pattern with tight packing unless randomness is explicitly enforced (Balzer *et al.*, 2009; Schlomer *et al.*, 2011).

A popular variant of PDS is the maximal PDS (MPDS) distribution, where the *maximality* constraint requires that the sample disks overlap, in the sense that they cover the whole domain leaving no room to insert an additional point. In practice, maximal PDS tends to outperform traditional PDS due to better coverage. However, algorithmically guaranteeing maximality requires expensive checks causing the resulting algorithms to be slow in moderate (2-5) and practically unfeasible in higher (7 and above) dimensions. Though strategies to alleviate this limitation have been proposed in (Ebeida *et al.*, 2012), the inefficiency of MPDS algorithms in higher dimensions still persists. Interestingly, a common limitation of all existing MPDS approaches is that there is no direct control over the number of samples produced by the algorithm, which makes the use of these algorithms difficult in practice, since optimizing samples for a given sample budget is the most common approach.

As discussed in Section 1, the metrics used by the space-filling designs discussed above do not provide insights into how to systematically trade-off uniformity and randomness. Thereby, making the design and optimization of sampling pattern a cumbersome process. To alleviate this problem, we propose a novel metric for assessing the space-filling property

and connect the proposed metric (defined in the spatial domain) to the quality metric of design performance (defined in the spectral domain).

3. A Metric for Assessing Space-filling Property

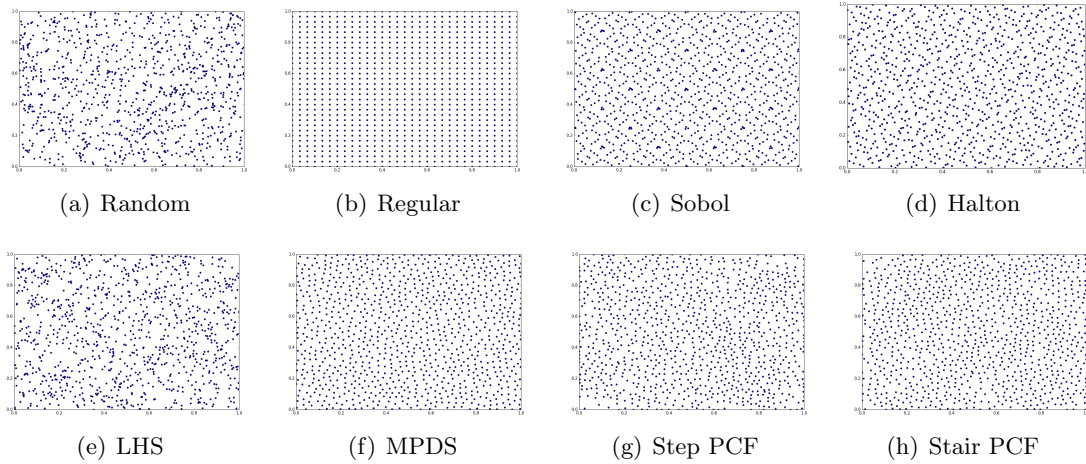


Figure 2: Visualization of 2-d point distributions obtained using different sample design techniques. In all cases, the number of samples N was fixed at 1000.

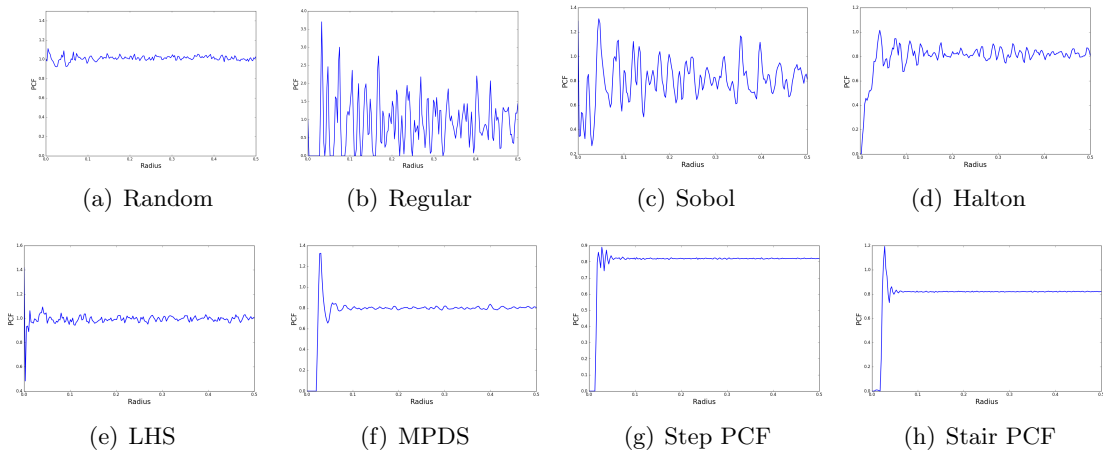


Figure 3: *Space-filling Metric*: Pair correlation functions, corresponding to the samples in Figure 2, characterize the coverage (and randomness) of point distributions obtained using different techniques.

In this section we first provide a definition of a space-filling design. Subsequently, we propose a metric to quantify space-filling properties of sample designs.

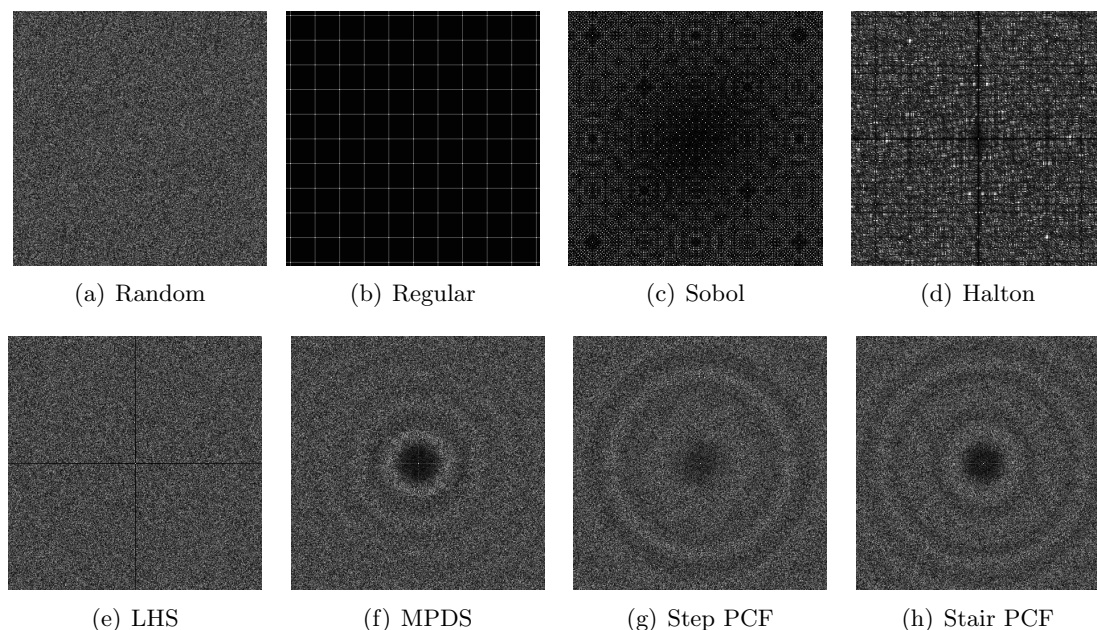


Figure 4: *Performance Quality Metric*: Power spectral density is used to characterize the effectiveness of sample designs, through the distribution of power in different frequencies.

3.1 Space-filling Designs

Without any prior knowledge of the function f of interest, a reasonable objective when creating \mathcal{X} is that the samples should be random to provide an equal chance of finding features of interest, e.g., local minima in an optimization problem, anywhere in \mathcal{D} . However, to avoid sampling only parts of the parameter space, a second objective is to cover the space in \mathcal{D} uniformly, in order to guarantee that all sufficiently large features are found. Therefore, a good space-filling design can be defined as follows:

Definition 1 *A space-filling design is a set of samples that are distributed according to a uniform probability distribution (Objective 1: Randomness) but no two samples are closer than a given minimum distance r_{min} (Objective 2: Coverage).*

Next, we describe the metric that we use to quantify the space-filling property of a sample design. The proposed metric is based on the spatial statistic, *pair correlation function* (PCF) and we will show that this metric is directly linked to the quality metric of design performance defined in the spectral domain.

3.2 Pair Correlation Function as a Space-filling Metric

In contrast to existing scalar space-filling metrics such as discrepancy, and coverage, the PCF characterizes the distribution of sample distances, thus providing a comprehensive description of the sample designs. A precise definition of the PCF can be given in terms

of the intensity λ and product density β of a point process (Illian et al., 2008; Oztireli and Gross, 2012).

Definition 2 *Let us denote the intensity of a point process \mathcal{X} as $\lambda(\mathcal{X})$, which is the average number of points in an infinitesimal volume around \mathcal{X} . For isotropic point processes, this is a constant value. To define the product density β , let $\{B_i\}$ denote the set of infinitesimal spheres around the points, and $\{dV_i\}$ indicate the volume measures of B_i . Then, we have³ $Pr(\mathbf{X}_1 = \mathbf{x}_1, \dots, \mathbf{X}_N = \mathbf{x}_N) = \beta(\mathbf{x}_1, \dots, \mathbf{x}_N)dV_1 \cdots dV_N$ which represents the probability of having points \mathbf{x}_i in the infinitesimal spheres $\{B_i\}$. In the isotropic case, for a pair of points, β depends only on the distance between the points, hence one can write $\beta(\mathbf{x}_i, \mathbf{x}_j) = \beta(\|\mathbf{x}_i - \mathbf{x}_j\|) = \beta(r)$ and $Pr(r) = \beta(r)dV_i dV_j$. The PCF is then defined as*

$$G(r) = \frac{\beta}{\lambda^2}. \tag{1}$$

Note that the PCF characterizes spatial properties of a sample design. However, in several cases, it is easier to link the quality metric of a sample design to its spectral properties. Therefore, we establish a connection between the spatial property of a sample design defined in PCF space to its spectral properties.

3.3 Connecting Spatial Properties and Spectral Properties of Space-filling Designs

Fourier analysis is a standard approach for understanding the qualitative properties of sampling patterns. Hence, we propose to analyze the spectral properties of sample designs, using tools such as the power spectral density, in order to assess their quality. For isotropic samples, a quality metric of interest is the radially-averaged power spectral density, which describes how the signal power is distributed over different frequencies.

Definition 3 *For a finite set of N points, $\{\mathbf{x}_j\}_{j=1}^N$, in a region with unit volume, the power spectral density of the sampling function $\sum_{j=1}^N \delta(\mathbf{x} - \mathbf{x}_j)$ is formally defined as*

$$P(\mathbf{k}) = \frac{1}{N} |S(\mathbf{k})|^2 = \frac{1}{N} \sum_{j,\ell} e^{-2\pi i \mathbf{k} \cdot (\mathbf{x}_\ell - \mathbf{x}_j)}, \tag{2}$$

where $|\cdot|$ denotes the l^2 -norm and $S(\mathbf{k})$ denotes the Fourier transform of the sampling function.

The radially-averaged power spectral density (PSD) is denoted using $P(k)$. Next, we show that the connection between spectral properties of a d -dimensional isotropic sample design and its corresponding pair correlation function can be obtained via the d -dimensional Fourier transform or more efficiently using the 1-d Hankel transform.

3. We denote a realization of random variables $\mathbf{X}_1, \dots, \mathbf{X}_N$ by $\mathbf{x}_1, \dots, \mathbf{x}_N$.

Proposition 4 For an isotropic sample design with N points, $\{\mathbf{x}_j\}_{j=1}^N$, in a d -dimensional region with unit volume, the pair correlation function $G(r)$ and radially averaged power spectral density $P(k)$ are related as follows:

$$G(r) = 1 + \frac{V}{2\pi N} H [P(k) - 1] \quad (3)$$

where V is the volume of the sampling region and $H[\cdot]$ denotes the 1-d Hankel transform, defined as

$$H(f(k)) = \int_0^\infty k J_0(kr) f(k) dk,$$

with $J_0(\cdot)$ denoting the Bessel function of order zero.

Proof Note that PSD and PCF of a sample design are related via the d -dimensional Fourier transform as follows:

$$\begin{aligned} P(\mathbf{k}) &= 1 + \frac{N}{V} F(G(\mathbf{r}) - 1) \\ &= 1 + \frac{N}{V} \int_{\mathbb{R}^d} (G(\mathbf{r}) - 1) \exp(-i\mathbf{k}\cdot\mathbf{r}) d\mathbf{r}. \end{aligned}$$

It can be shown that, for radially symmetric or isotropic functions, the above relationship simplifies to

$$P(k) = 1 + 2\pi \frac{N}{V} H [G(r) - 1].$$

Next, using the inverse property of the Hankel transform, i.e.,

$$H_0^{-1}(f(r)) = \int_0^\infty r J_0(kr) f(r) dr,$$

we have

$$G(r) = 1 + \frac{V}{2\pi N} H [P(k) - 1]. \quad (4)$$

■

Proposition 4 is important as it enables us to qualitatively understand space-filling designs by first mapping them into the PCF space constructed based on spatial distances between points and, then, evaluating and understanding spectral properties of sample designs.

In Figure 3, we show the PCF⁴ of some commonly used 2-d sample designs ($N = 1000$) illustrated in Figure 2. As can be observed, both regular grid samples and QMC sequences have significant oscillations in their PCFs, which can be attributed to their structured nature. Regular grid sample design demonstrates a large disk radius r_{min} ($G(r) = 0$ for $0 \leq r \leq r_{min}$) as every sample is at least r_{min} apart from the rest of the samples, which in turn implies a better coverage. However, in practice, they perform poorly (e.g., in terms of generalization (or test error) in regression application and reconstruction error in image reconstruction application) compared to randomized sample designs and this can

4. Note that for non-isotropic sample designs, d -dimensional PCF (Illian et al., 2008) can be more descriptive.

be understood by studying their spectral properties. In contrast, random sample (Monte-Carlo) designs have a constant PCF with nearly no oscillations, since point samples are uncorrelated, thus, $P(r) = \lambda dx \lambda dy$ and theoretically have $G(r) = 1, \forall r$. Furthermore, the LHS design has a similar PCF as random designs with the exception of a small, yet non-zero, r_{min} .

Other variants of PDS like MPDS, Step PCF and Stair PCF designs attempt to trade-off between coverage ($G(r) = 0$ for $0 \leq r \leq r_{min}$) and randomness $G(r) = 1$, for $r > r_{min}$. Note that the Step and the Stair PCF methods are space-filling spectral designs proposed later in this paper. However, upon a careful comparison, it can be seen that MPDS has a larger peak and more oscillations in its PCF compared to the proposed designs. In fact, our empirical studies show that the amount of oscillations in the PCF of the MPDS design significantly increases with dimensions.

Next, in Figure 4, we show the corresponding PSDs of the different sample designs. It can be seen that, oscillations in PCF directly correspond to oscillations in PSDs. For example, the oscillatory behavior of the PCF for regular and QMC sequences cause a non-uniform distribution of power in their corresponding PSDs. Furthermore, the larger peak height in the PCF of MPDS implies that a large amount of power is concentrated in a small frequency band instead of power being distributed over all frequencies. In Section 5, we will analyze the effect of the shape of PCF on the performance of a sample design in detail.

It is important to note that, not every PCF (or PSD) is physically realizable by a sample design. In fact, there are two necessary mathematical conditions⁵ that a sample design must satisfy to be realizable.

Definition 5 (Realizability) *A PCF can be defined to be potentially realizable through a sample design, if it satisfies the following conditions:*

- *its PCF must be non-negative, i.e., $G(r) \geq 0, \forall r$, and*
- *its corresponding PSD must be non-negative, i.e., $P(k) \geq 0, \forall k$.*

As both the PSD and the PCF characteristics are strongly tied to each other (as shown in Proposition 4), these two conditions limit the space of realizable space-filling spectral designs. The results from this section will serve as tools for qualitatively understanding and, thus, designing optimal space-filling spectral designs in the following sections.

4. Space-filling Spectral Designs

In this section, we first formalize desired characteristics of a good space-filling design, as given in Definition 1. Next, we will describe the proposed framework for creating space-filling spectral designs.

Definition 6 *A set \mathcal{X} of N random samples $\{\mathbf{X}_i\}_{i=1}^N$ in a sampling domain \mathcal{D} can be characterized as a space-filling design, if $\mathcal{X} = \{\mathbf{X}_i = \mathbf{x}_i \in \mathcal{D}; i = 1, \dots, N\}$ satisfy the following two objectives:*

5. Whether or not these two conditions are not only necessary but also sufficient is still an open question (however, no counterexamples are known).

- $\forall \mathbf{X}_i \in \mathcal{X}, \forall \Delta \mathcal{D} \subseteq \mathcal{D} : Pr(\mathbf{X}_i = \mathbf{x}_i \in \Delta \mathcal{D}) = \int_{\Delta \mathcal{D}} d\mathbf{X}$
- $\forall \mathbf{x}_i, \mathbf{x}_j \in \mathcal{X} : \|\mathbf{x}_i - \mathbf{x}_j\| \geq r_{min}$

where r_{min} is referred to as the coverage radius.

In the above definition, the first objective states that the probability of a random sample $\mathbf{X}_i \in \mathcal{X}$ falling inside a subset $\Delta \mathcal{D}$ of \mathcal{D} is equal to the hyper-volume of $\Delta \mathcal{D}$ (uniform distribution). The second condition enforces the minimum distance constraint between point sample pairs for improving coverage.

A Poisson design enforces the first condition alone, in which case the number of samples that fall inside any subset $\Delta \mathcal{D} \subseteq \mathcal{D}$ obeys a discrete Poisson distribution. Though easier to implement, Poisson sampling often produces distributions where the samples are grouped into clusters and leaves holes in possibly the regions of interest. In other words, this increases the risk of missing important features, when only the samples are used for analysis. Consequently, a sample design that distributes random samples in a uniform manner across \mathcal{D} is preferred, so that clustering patterns are not observed. The coverage condition explicitly eliminates the clustering behavior by preventing samples from being closer than r_{min} . A space-filling design can be defined conveniently in the PCF domain and we refer to this as the space-filling spectral design, due to its direct connection to the spectral domain properties.

4.1 Defining a Space-filling Spectral Design in Spatial Domain

For Poisson design, point locations are not correlated and, therefore, $P(r) = \lambda dx \lambda dy$. This implies that for Poisson designs $G(r) = 1$. Similarly, for space-filling designs, due to the minimum distance constraint between the point sample pairs, we do not have any point samples in the region $0 \leq r < r_{min}$. Consequently, space-filling spectral designs are defined as a step pair correlation function in the spatial domain (*Step PCF*).

Proposition 7 *Given the desired coverage radius r_{min} , a space-filling spectral design is defined in the spatial domain as*

$$G(r - r_{min}) = \begin{cases} 0 & \text{if } r < r_{min} \\ 1 & \text{if } r \geq r_{min}. \end{cases}$$

As a consequence of Proposition 4, space-filling spectral designs can equivalently be defined in the spectral domain.

4.2 Defining a Space-filling Spectral Design in Spectral Domain

We derive the power spectral density of the space-filling spectral design using the connection established in Section 3. Following our earlier notation, we denote the d -dimensional power spectral density by $P(\mathbf{k})$ and d -dimensional PCF by $G(\mathbf{r})$.

Proposition 8 (*Kailkhura et al., 2016a*) *Given the desired coverage radius r_{min} , a d -dimensional space-filling spectral design \mathcal{X} , with N sample points in a sampling domain*

\mathcal{D} of volume V , can be defined in the PSD domain as

$$P(k) = 1 - \frac{N}{V} \left(\frac{2\pi r_{\min}}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_{\min})$$

where $J_{\frac{d}{2}}(\cdot)$ is the Bessel function of order $d/2$.

Proof We know that,

$$P(\mathbf{k}) = 1 + \frac{N}{V} F(G(\mathbf{r}) - 1), \tag{5}$$

$$= 1 + \frac{N}{V} \int_{\mathbb{R}^d} (G(\mathbf{r}) - 1) \exp(-i\mathbf{k}\cdot\mathbf{r}) d\mathbf{r}, \tag{6}$$

where $F(\cdot)$ denotes the d -dimensional Fourier transform. Note that for the radially symmetric or isotropic functions, i.e., $G(r)$ where $r = \|\mathbf{r}\|$, the above relationship simplifies to

$$P(k) = 1 + \frac{N}{V} (2\pi)^{\frac{d}{2}} k^{1-\frac{d}{2}} H_{\frac{d}{2}-1} \left(r^{\frac{d}{2}-1} (G(r) - 1) \right), \tag{7}$$

where

$$H_v(f(r))(k) = \int_0^\infty r J_v(kr) f(r) dr$$

is the 1-d Hankel transform of order v with J being the Bessel function. To derive the PSD of a step function, we first evaluate the Hankel transform of $f(r) = r^{\frac{d}{2}-1}(G(r) - 1)$ where $G(r)$ is a step function.

$$\begin{aligned} H_{\frac{d}{2}-1} \left(r^{\frac{d}{2}-1} (G(r) - 1) \right) &= \int_0^\infty r^{\frac{d}{2}} J_{\frac{d}{2}-1}(kr) (G(r) - 1) dr \\ &= - \int_0^{r_{\min}} r^{\frac{d}{2}} J_{\frac{d}{2}-1}(kr) dr \\ &= - \frac{r_{\min}^{\frac{d}{2}}}{k} J_{\frac{d}{2}}(kr_{\min}) \end{aligned}$$

Using this expression in (7), we obtain

$$P(k) = 1 - \frac{N}{V} \left(\frac{2\pi r_{\min}}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_{\min}). \tag{8}$$

■

This proposition connects the spatial properties of a space-filling spectral design, defined via the PCF, to its spectral properties. The motivation for this is the fact that in several cases, it is easier to link the qualitative performance of a sample design to its spectral properties. In the next section, we will develop the relation between spectral properties and an qualitative measure of the performance, which in turn provides us guidelines for designing better space-filling spectral sampling patterns.

5. Qualitative Analysis of Space-filling Spectral Designs

In this section, we derive insights regarding the qualitative performance of space-filling spectral designs. To this end, we analyze the impact of the shape of the PCF on the reconstruction performance. Further, for a tractable analysis, we consider the task of recovering the class of periodic functions using space-filling spectral designs and analyze the reconstruction error as a function of their spectral properties. The analysis presented in this section will clarify how the shape of the PCF of a sample design directly impacts its reconstruction performance.

5.1 Analysis of Reconstruction Error for Periodic Functions

Let us denote the Fourier transform of the sample design \mathcal{X} by \mathcal{S} . The function to be sampled and its corresponding Fourier representation are denoted by \mathcal{I} and $\hat{\mathcal{I}}(k)$ respectively. Now, the spectrum of the sampled function is given by $\hat{\mathcal{I}}_s(k) = \mathcal{S} * \hat{\mathcal{I}}(k)$. Note that a sampling pattern with a finite number of points is comprised of two components, a DC peak at the origin and a noisy remainder $\bar{\mathcal{S}}$. Thus, equivalently, we have $\hat{\mathcal{I}}_s(k) = \{N\delta(k) + \bar{\mathcal{S}}\} * \hat{\mathcal{I}}(k)$. The error introduced in the process of function reconstruction is the difference between the reconstructed and the original functions:

$$\mathcal{E}(k) = |\hat{\mathcal{I}}_s(k)/N - \mathcal{I}(k)|^2 = |\bar{\mathcal{S}} * \hat{\mathcal{I}}(k)/N|^2$$

where we have divided the R.H.S. by N to normalize the energy of \mathcal{I}_s . For error analysis, we focus on the low frequency content of the error term, since the high frequency components are removed during the reconstruction process.

Denoting the power spectrum without the DC component by $\bar{\mathcal{P}}(k)$, for a constant function the error simplifies to

$$\mathcal{E}(k) \propto |\bar{\mathcal{S}}(k)|^2 \propto \bar{\mathcal{P}}(k). \quad (9)$$

This, as stated above, allows for the characterization of the error in terms of the spectral properties of the sampling pattern used.

Next, we consider an important class of functions, the family of periodic functions, for further analysis. All periodic functions with a finite period can be expressed as a Fourier series⁶, which is a summation of sine and cosine terms

$$\mathcal{I}(x) = a_0 + \sum_{m=1}^M a_m \cos(2\pi mx) + \sum_{m=1}^M b_m \sin(2\pi mx).$$

The Fourier transform of this function is equivalently a summation of pulses:

$$\hat{\mathcal{I}}(k) = a_0 \delta(k) + \sum_{m=1}^M a_m \left(\frac{1}{2} (\delta(k-m) + \delta(k+m)) \right) + \sum_{m=1}^M b_m \left(\frac{1}{2} (\delta(k+m) - \delta(k-m)) \right).$$

6. In the subsequent analysis, the number of terms in the Fourier series, M , is an arbitrary value which can be replaced by infinity for non-differentiable/discontinuous functions. Note that the Fourier series of periodic functions that are smooth (no discontinuity and no sharp corners) is finite.

Making substitutions, $a_m + b_m = A_m, a_m - b_m = B_m$, we obtain

$$\mathcal{E}(k) = \frac{1}{4N^2} \left| 4a_0 \bar{\mathcal{S}}(k) + \sum_{m=1}^M (A_m \bar{\mathcal{S}}(k+m) + B_m \bar{\mathcal{S}}(k-m)) \right|^2.$$

The reconstruction error can then be upper bounded as follows:

$$\mathcal{E}(k) \leq \frac{1}{4N} \left[4a_0^2 \bar{\mathcal{P}}(k) + \sum_{m=1}^M (A_m^2 \bar{\mathcal{P}}(k+m) + B_m^2 \bar{\mathcal{P}}(k-m)) \right]. \quad (10)$$

In the case of a single sinusoidal function, $\cos(2\pi fx)$, using triangle inequality, this becomes (Heck et al., 2013)

$$\mathcal{E}(k) \leq \frac{1}{4N} \left[\bar{\mathcal{P}}(k+f) + \bar{\mathcal{P}}(k-f) + 2\sqrt{\bar{\mathcal{P}}(k+f)\bar{\mathcal{P}}(k-f)} \right]. \quad (11)$$

Even though this is only an upper bound and the theoretic analysis is restricted to periodic functions, we have empirically found that it accurately predicts the characteristics of the sampling error for a broad range of complex functions and provides useful guidelines (more details are provided in Section 9).

The above analysis implies that to assess the quality of the sample designs, one can analyze their spectral behavior. More specifically, the above analysis suggests that to minimize the reconstruction error (Eq. (10) and (11)): (a) the power spectra of the sample design should be close to zero, and (b) for errors to be broadband white noise (uniform over frequencies), the power spectra should be a constant. Note that in several applications, e.g., image reconstruction, most relevant information is predominantly at low frequencies. In such scenarios, this naturally leads to the following criteria for sample designs: (a) the spectrum should be close to zero for low frequencies which indicates the range of frequencies that can be represented with almost no error, (b) the spectrum should be a constant for high frequencies or contain minimal amount of oscillations in the power spectrum. However, as we will see next, there exist a trade-off between low frequency power and high frequency oscillations in power spectra.

5.2 Effect of PCF Characteristics on Sampling Performance

Based on the two criteria discussed above, we assess the effect of the shape of the PCF on the quality of space-filling designs in the spectral domain. Note that PCFs of the samples constructed in practice (Figure 2) often demonstrates the following characteristics: (a) presence of a zero-region characterized by r_{min} , (b) a large peak around r_{min} , and (c) damped oscillations. To model and analyze these characteristics, we consider the following parametric PCF family⁷

$$G(r; r_{min}, \delta, a, c) = G(r - r_{min}) + (a - 1) (G(r - r_{min}) - G(r - r_{min} - \delta)) + \frac{a - 1}{4r} \exp(-r/2) \sin(c \times r - c) G(r - r_{min}) \quad (12)$$

7. Note that there exist a broad range of parametric space-filling spectral designs. However, finding PCFs that are realizable is a nontrivial problem because the space of functions that obey the realizability conditions is not easy to parametrize.

where $G(r - r_{\min})$ is the Step function, peak width $\delta \geq 0$ and the peak height $a \geq 1$ and last term in (12) corresponds to damped oscillations. This family is a generalization of Step PCF, with additional parameterization of peak height and oscillations in the PCF.

5.2.1 EFFECT OF PEAK HEIGHT ON SPECTRAL PROPERTIES

In order to study the impact of increasing peak height in the PCF on the PSD characteristics, we conduct an empirical study. We compute the PSD of a sample design with the following parameters: $N = 195, r_{\min} = 0.02, \delta = 0.005$. Note that we vary the PCF peak height a , which actually reflects the behavior of existing coverage based PDS algorithms. As shown in Figure 5(a), increasing a results in both significantly *higher* low frequency power and *larger* high frequency oscillations. As expected, the PSD of the Step PCF (or $a = 1$) performs the best, i.e., the spectrum is close to zero for low frequencies and constant for high frequencies.

5.2.2 EFFECT OF DISK RADIUS ON SPECTRAL PROPERTIES

Next, we study the importance of choosing an appropriate r_{\min} (or coverage ρ) while generating sample distributions. In Figure 5(b), we show the PSD for $N = 195$ and $a = 1$, with varying disk radius values r_{\min} . For a fixed sample budget, as we increase the radius, we observe two contrasting changes in the PSD: (i) the spectrum tends to be close to zero at low frequencies and (ii) an increase in oscillations for high frequencies. Consequently, there is a trade-off between low frequency power and high frequency oscillations in power spectra which can be controlled by varying r_{\min} . However, the increase in oscillations are less significant compared to the gain in the zero-region. Furthermore, in several applications, low frequency content is more informative, and hence one may still attempt to maximize r_{\min} or coverage.

5.2.3 EFFECT OF OSCILLATIONS ON SPECTRAL PROPERTIES

Finally, we study the effect of oscillations in the PCF on the power distribution in the spectral domain. In Figure 5(c), we plot the PSD for $a = 1$ with varying amounts of oscillations controlled via the parameter c . It can be seen that introducing oscillations in the PCF results in significantly *higher* low frequency power and *larger* high frequency oscillations. As expected, the PSD of the Step PCF (or $c = 0$) behaves the best.

In summary, the discussion in this section suggests that the PCF of an ideal space-filling spectral design should have the following three properties: (a) large r_{\min} , (b) small peak height, and (c) low oscillations. Since, the Step PCF satisfies these three properties, it is expected to be a good space-filling spectral design. Next, we consider the problem of optimizing the parameter of the Step PCF design, i.e. r_{\min} .

6. Optimization of Step PCF based Space-filling Spectral Designs

The proposed space-filling metric enjoys mathematical tractability and is supported by theoretical results as defined in Section 4. This enables us to obtain new insights for optimizing Step PCF based space-filling spectral designs. In particular, (a) For a fixed r_{\min} , we obtain the maximum number of point samples in any arbitrary dimension d , (b) For a

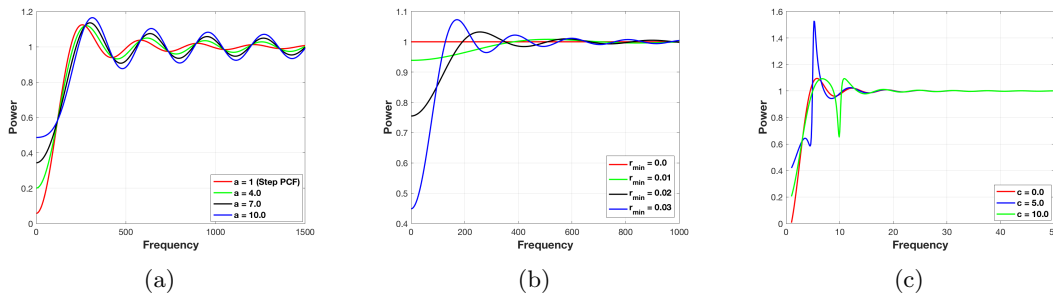


Figure 5: (a) Effect of peak height in the PCF on power spectra, (b) Effect of disk radius in the PCF on power spectra, (c) Effect of oscillations in the PCF on power spectra.

fixed sampling budget N , we derive the maximum achievable r_{\min} in arbitrary dimension d .

6.1 Case 1: Fixed r_{\min}

The problem of finding the maximum number of point samples in a Step PCF based space-filling spectral design with a given disk radius r_{\min} can be formalized as follows:

$$\begin{aligned}
 & \text{maximize} && N \\
 & \text{subject to} && P(k) \geq 0, \forall k \\
 & && G(r - r_{\min}) \geq 0, \forall r,
 \end{aligned} \tag{13}$$

where $P(k) = 1 - \frac{N}{V} \left(\frac{2\pi r_{\min}}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_{\min})$. Note that a space-filling spectral design has to satisfy realizability constraints as defined in Definition 5.

Proposition 9 (Kailkhura et al., 2016a) *For a fixed disk radius r_{\min} , the maximum number of point samples possible for a realizable Step PCF based space-filling spectral design in the sampling region with volume V can be approximated as*

$$N \approx \frac{V\Gamma\left(\frac{d}{2} + 1\right)}{\pi^{\frac{d}{2}} r_{\min}^d}.$$

Proof Using the definition of the Step PCF function, the constraint $G(r - r_{\min})$ is trivially satisfied. Note that the constraint $P(k) \geq 0, \forall k$ is equivalent to $\min_k P(k) \geq 0$. In other

words,

$$\begin{aligned}
& \min_k 1 - \rho \left(\frac{2\pi r_{\min}}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_{\min}) \geq 0 \\
\Leftrightarrow & \max_k \rho \left(\frac{2\pi r_{\min}}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_{\min}) \leq 1 \\
\Leftrightarrow & \rho (2\pi)^{\frac{d}{2}} r_{\min}^d \max_k \left(\frac{J_{\frac{d}{2}}(kr_{\min})}{(kr_{\min})^{\frac{d}{2}}} \right) \leq 1 \\
\Leftrightarrow & \rho (2\pi)^{\frac{d}{2}} r_{\min}^d \frac{1}{2^{\frac{d}{2}} \Gamma(\frac{d}{2} + 1)} \lesssim 1 \tag{14} \\
\Leftrightarrow & N \lesssim \frac{V \Gamma(\frac{d}{2} + 1)}{(\pi)^{\frac{d}{2}} r_{\min}^d}
\end{aligned}$$

where, in (14) we have used the fact that $J_v(x) \approx (x/2)^v / \Gamma(v + 1)$ and $\rho = N/V$. \blacksquare

Note that for the 2-dimensional case, we have $\frac{J_1(kr_{\min})}{kr_{\min}} = \text{jinc}(kr_{\min})$ where $\text{jinc}(\cdot)$ is the sombrero function (sometimes called besinc function or jinc function). Now using the fact that $\text{jinc}(x)$ has the maximum value equal to $1/2$, for a fixed disk radius r_{\min} , the maximum number of point samples possible in a 2-d Step PCF based space-filling spectral design is given by

$$N = V / \pi (r_{\min})^2,$$

which again corroborates our bound in Proposition 9.

6.2 Case 2: Fixed N

Alternately, we can also derive the bound for the disk radius of Step PCF with a fixed sampling budget N as follows:

$$\begin{aligned}
& \text{maximize } r_{\min} \\
& \text{subject to } P(k) \geq 0, \forall k \\
& \quad G(r - r_{\min}) \geq 0, \forall r
\end{aligned} \tag{15}$$

Proposition 10 (*Kaikhura et al., 2016a*) *For a fixed sampling budget N , the maximum possible disk radius r_{\min} for a realizable Step PCF based space-filling spectral design in the sampling region with volume V can be approximated as*

$$r_{\min} \approx \sqrt[d]{\frac{V \Gamma(\frac{d}{2} + 1)}{\pi^{\frac{d}{2}} N}}.$$

Proof The proof is similar to the one in Proposition 9. \blacksquare

6.3 Relative Radius of Step PCF

As mentioned before, the current literature characterizes coverage by the fraction ρ of the maximum possible radius r_{max} for N samples to cover the sampling domain, such that $r_{min} = \rho r_{max}$. The maximum possible disk radius is achieved by the deterministic hexagonal lattice (Schreiber, 1943) and can be approximated in a d dimensional sampling region as $r_{max} \approx \sqrt[d]{\frac{A_d}{C_d N}}$. Here, A_d is the hypervolume of the sampling domain and $C_d = V_d/r^d$ with V_d being the hypervolume of a hypersphere with radius r . Note that a uniformly distributed point set can have a relative radius of 0, and the relative radius of a hexagonal lattice equals 1 (in 2-d). Next, we derive a closed-form expression for the relative radius of Step PCF based design.

Proposition 11 *For a fixed sampling budget N , the maximum relative radius ρ for Step PCF based space-filling spectral design in the sampling region with volume V is given by $\rho = \frac{1}{2 \sqrt[d]{\eta_d}}$ where η_d is maximal density of a sphere packing in d -dimensions.*

Proof Let us denote by $r_{max} = \arg \min_r \eta_d$, then, the maximal density of a sphere packing with N samples in d -dimensions is given by

$$\eta_d = \frac{N \pi^{d/2}}{\Gamma(1 + \frac{d}{2})} \frac{r_{max}^d}{V} \tag{16}$$

$$\Leftrightarrow \eta_d = \left(\frac{r_{max}}{r_{min}} \right)^d \tag{17}$$

$$\Leftrightarrow \rho = \frac{1}{2 \sqrt[d]{\eta_d}} \tag{18}$$

where equality in (17) uses Proposition (10). ■

For $d = 2$ and 3, the relative radius simplifies to:

$$\rho = 0.5 \sqrt[2]{\frac{\pi \sqrt{3}}{6}}, \text{ for } d = 2, \text{ and}$$

$$\rho = 0.5 \sqrt[3]{\frac{\pi \sqrt{2}}{6}}, \text{ for } d = 3.$$

Note that finding the maximal density of a sphere packing for an arbitrary high dimension (except in $d = 2, 3$ and recently in 8, 24 (viazovska, 2017; Cohn et al., 2017)) is an open problem. Note that best known packings are often lattices, thus, we use the best known lattices to be an approximation of r_{max} in our analysis⁸.

In Figure 6, we plot the relative radius $\rho = r_{min}/r_{max}$ of Step PCF for different dimensions d . It is interesting to notice that the relative radius of Step PCF based designs

8. We use relative radius as a metric only for analysis and not for design optimization.

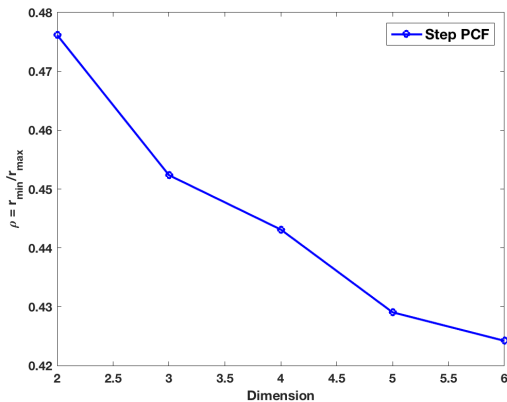


Figure 6: Relative radius $\rho = r_{min}/r_{max}$ of Step PCF based space-filling spectral design for different dimensions d .

increases as the dimension d increases, i.e., Step PCF based designs approach a more regular pattern. Further, note that for a fixed sampling budget both r_{min} and r_{max} increase as the number of dimensions increases. The Step PCF based designs maintain randomness by keeping the PCF flat, but this comes at a cost: the disk radius r_{min} of these patterns is very small (as can be seen from Figure 6). For several applications, covering the space better (by trading-off randomness) is more important. In the next section, we will propose a new class of space-filling spectral designs that can achieve a much higher r_{min} at the small cost of compromising randomness by introducing a single peak into an otherwise flat PCF.

7. Space-filling Spectral Designs with Improved Coverage

To improve the coverage of Step PCF base space-filling spectral design, in this section, we propose a novel space-filling spectral design which systematically trades-off randomness with coverage of the resulting samples. Note that the randomness property can be relaxed either by increasing the peak height of the PCF, or by increasing the amounts oscillations in the PCF (as discussed in Section 5.2). For simplicity⁹, we adopt the former strategy and use only the peak height parameter. More specifically, as an alternative to Step PCF, we design the following generalization which we refer as the Stair PCF design.

7.1 Stair PCF based Space-filling Spectral Design

Now, we define the proposed *Stair* PCF based space-filling design and quantify the gains achieved in the coverage characteristics (i.e. r_{min}).

Stair PCF in the Spatial Domain: The Stair PCF construction is defined as follows:

9. In our initial experiments, we found that increasing the peak height alone is sufficient for trading-off randomness to maximize coverage, and performs better than trading-off randomness by increasing oscillations in the PCF.

$$G(r; r_0, r_1, P_0) = f(r - r_1) + P_0(f(r - r_0) - f(r - r_1)), \quad (19)$$

$$\text{with } f(r - r_0) = \begin{cases} 0 & \text{if } r \leq r_0 \\ 1 & \text{if } r > r_0 \end{cases},$$

where $r_0 \leq r_1$ and $P_0 \geq 1$.

This family of space-filling spectral designs has three interesting properties:

- except for a single peak in the region $r_0 \leq r \leq r_1$, the PCF is flat, thus, does not compromise randomness entirely,
- both the height and width of the peak can be optimized to maximize coverage,
- the Step PCF based spectral design can be derived as a special case of this construction.

Now, the problem boils down to finding the combinations of the three parameters (r_0, r_1, P_0) that are realizable and yield a good sample design (discussed in Section 7.2). A representative example of Stair PCF is shown in Figure (7(a)).

Stair PCF in the Spectral Domain: Following the analysis in the earlier sections, we derive the power spectral density of Stair PCF based space-filling spectral designs.

Proposition 12 *The power spectral density of a Stair PCF based space-filling spectral designs, $G(r; r_0, r_1, P_0)$, with N samples in the sampling region with volume V is given by*

$$P(k) = 1 - \frac{N}{V} P_0 \left(\frac{2\pi r_0}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_0) - \frac{N}{V} (1 - P_0) \left(\frac{2\pi r_1}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_1).$$

Proof Using results from Section 4.2, we have

$$P(k) = 1 + \frac{N}{V} (2\pi)^{\frac{d}{2}} k^{1-\frac{d}{2}} H_{\frac{d}{2}-1} \left(r^{\frac{d}{2}-1} (G(r) - 1) \right). \quad (20)$$

To derive the PSD of a Stair function, we first evaluate the Hankel transform of $f(r) = (G(r) - 1)$ where $G(r)$ is a Stair function.

$$\begin{aligned} H_{\frac{d}{2}-1} \left(r^{\frac{d}{2}-1} (G(r) - 1) \right) &= \int_0^\infty r^{\frac{d}{2}} J_{\frac{d}{2}-1}(kr) (G(r) - 1) dr \\ &= -P_0 \int_0^{r_0} r^{\frac{d}{2}} J_{\frac{d}{2}-1}(kr) dr - (1 - P_0) \int_0^{r_1} r^{\frac{d}{2}} J_{\frac{d}{2}-1}(kr) dr \\ &= -P_0 \frac{r_0^{\frac{d}{2}}}{k} J_{\frac{d}{2}}(kr_0) - (1 - P_0) \frac{r_1^{\frac{d}{2}}}{k} J_{\frac{d}{2}}(kr_1) \end{aligned}$$

Using this expression in (20),

$$P(k) = 1 - \frac{N}{V} P_0 \left(\frac{2\pi r_0}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_0) - \frac{N}{V} (1 - P_0) \left(\frac{2\pi r_1}{k} \right)^{\frac{d}{2}} J_{\frac{d}{2}}(kr_1). \quad (21)$$

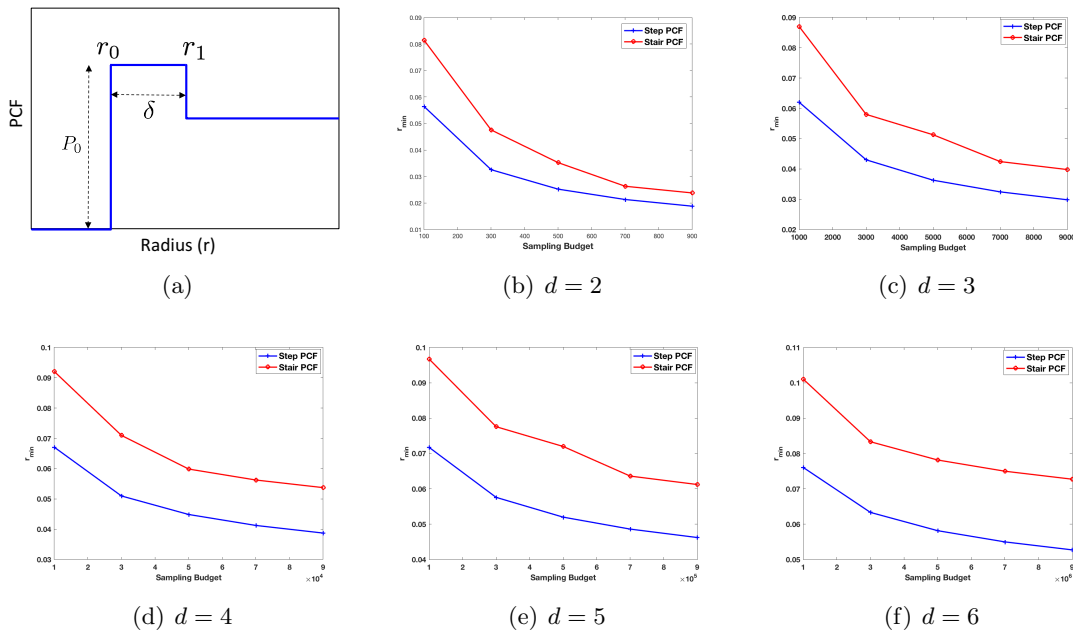


Figure 7: (a) Pair correlation function of Stair PCF based designs, (b)-(f) Maximum Disk Radius For Step and Stair PCF for dimensions 2 to 6.

■

Next, we empirically evaluate the gain in coverage achieved by Stair PCF based designs compared to the Step PCF based designs.

7.2 Coverage Gain with Stair PCF

Ideally, the optimal Stair PCF should be obtained by simultaneously maximizing r_0 ($:= r_{min}$) and minimizing P_0 . Furthermore, not all PCFs in the Stair PCF family are realizable. Due to the realizability conditions, the parameters cannot be adjusted independently. The main challenge, therefore, is to find the combinations of the three parameters (r_0, r_1, P_0) that is realizable and yield a good sample design. Unlike Step PCF, the closed form expression for the optimal parameters (r_0, r_1, P_0) are difficult to obtain, and, therefore, we explore this family of PCF patterns empirically by searching configurations for which:

- the disk radius r_0 is as high as possible, and
- the PCF is flat with minimal increase in the peak height P_0 .

7.2.1 DISK RADIUS r_{min} VS. SAMPLE BUDGET N

In this section, we show the increase in coverage (or r_{min}) obtained by compromising randomness by increasing peak height in the PCF. We constrain the peak height to be below $P_0 \leq 1.5$ and analyze the gain in r_{min} due to this small compromise in randomness. Furthermore, we assume that $r_{min}^{step} \leq r_0 \leq 2 \times r_{min}^{step}$ and $r_0 \leq r_1 \leq 1.5 \times r_0$. In Figures 7(b)

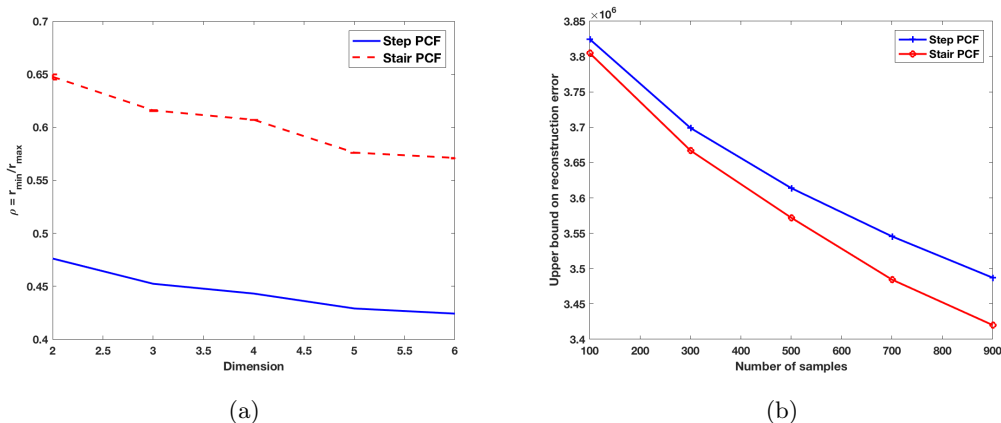


Figure 8: (a) Gain in the relative radius ρ achieved with the Stair PCF constructions, in comparison to the Step PCF constructions; (b) Upper bound on the reconstruction error of Step and Stair PCF based constructions.

through 7(f), we compare the maximum r_{min} achieved by the Step and Stair PCF designs, for varying sample sizes in dimensions 2 to 6. It can be seen that introducing a small peak in the PCF results in a significant increase in the coverage. This gain can be observed for all sampling budgets in all dimensions. Furthermore, as expected, for low sampling budgets maximal gain is observed, and should decrease with increasing N as the r_{min} for both the families will asymptotically (in N) converge to zero.

7.2.2 RELATIVE RADIUS ρ VS. DIMENSION d

In this section, we study the increase in relative radius ρ due to the introduction of a peak in the PCF. Again, we assume that $P_0 \leq 1.5$, $r_{min}^{step} \leq r_0 \leq 2 \times r_{min}^{step}$ and $r_0 \leq r_1 \leq 1.5 \times r_0$. In Figure 8(a), we show the maximum $\rho = r_{min}/r_{max}$ achieved by the Step and Stair PCFs for different dimensions d . For Stair PCF, we do not have a closed form expression of ρ , thus, we obtain the maximum achievable r_{min} empirically for various sampling budgets and plot the mean (with standard deviation) behavior of the ρ . It can be seen that introducing a small peak in the PCF results in a significant increase in the relative radius. This gain can be observed at all sampling budgets in all dimensions. This also corroborates the recommendation of using $0.65 \leq \rho \leq 0.85$ in practice for coverage based designs and suggests that in higher dimensions ρ should be higher.

7.2.3 ANALYSIS OF RECONSTRUCTION ERROR UPPER BOUND

We also assess the reconstruction quality of the Step and Stair PCF based spectral designs, on the class of periodic functions considered in Section 5.1, for varying sampling budgets. Here, we consider the setup where $0 \leq k \leq 1000$ and $0 \leq f \leq 1000$. In Figure 8(b), we plot the average reconstruction error upper bounds as given in (11) for Step and Stair PCF. As expected, for both sample designs, the reconstruction error decreases with an

increase in the sampling budget. More interestingly, the reconstruction error of Stair PCF is lower compared to the reconstruction error of Step PCF, thus showing the effectiveness of increased coverage in sample designs.

8. Synthesis of Space-filling Spectral Designs

In this section, we describe the proposed approach for synthesizing sample designs that match the optimal (Stair or Step) PCF characteristics. Existing approaches for PCF matching such as (Oztireli and Gross, 2012; Kailkhura et al., 2016b) rely on kernel density estimators to evaluate the PCF of a point set. A practical limitation of these approaches is the lack of an efficient PCF estimator in high dimensions. More specifically, these estimators are biased due to lack of an appropriate edge correction strategy. This bias in PCF estimation arises due to the fact that sample hyper-spheres used in calculating point-pattern statistics may fall partially outside the study region and will produce a biased estimate of the PCF unless a correction is applied. The effect of this bias is barely noticeable in 2 dimensions and hence existing PCF matching algorithms have ignored this. However, this problem becomes severe in higher dimensions, thus, making the matching algorithm highly inaccurate. To address this crucial limitation, we introduce an edge corrected estimator for computing the PCF of sample designs in arbitrary dimensions. Following this, we describe a gradient descent based optimization technique to synthesize samples that match the desired PCF.

8.1 PCF Estimation in High Dimensions with Edge Correction

In order to create an unbiased PCF estimator, we propose to employ an edge corrected kernel density estimator, defined as follows:

$$\hat{G}(r) = \frac{V_W}{\gamma_W} \frac{V_W}{N} \frac{1}{S_E(N-1)} \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N k(r - |x_i - x_j|) \quad (22)$$

where $k(\cdot)$ denotes the kernel function; here we use the classical Gaussian kernel

$$k(z) = \frac{1}{\sqrt{\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right). \quad (23)$$

In the above expression, V_W indicates the volume of the sampling region. When the sampling region is a hyper-cube with length 1, we have $V_W = 1$. Let S_E denote the area of hyper-sphere with radius r which is given by

$$S_E = \frac{dr^{d-1} \pi^{\frac{d}{2}}}{\Gamma(1 + \frac{d}{2})}.$$

Also, we denote the surface area of the sampling region by S_W , which is expressed as

$$S_W = r^{d-1} \sin^{d-2} \phi_1 \sin^{d-3} \phi_2 \cdots \sin \phi_{d-2}.$$

The term $\frac{V_W}{\gamma_W}$ performs edge correction to handle the unboundedness of the estimator, where γ_W is an isotropic set covariance function given by

$$\gamma_W = \frac{1}{S_E} \int_{\substack{0 \leq \phi_{d-1} \leq 2\pi \\ 0 \leq \phi_i \leq \pi, i=1 \text{ to } d-2}} S_W \gamma d\phi_1 \cdots d\phi_{d-1} \quad (24)$$

where $\gamma = \prod_{p=1}^d (1 - |x^p|)$ with

$$\begin{aligned} x^1 &= r \cos \phi_1 \\ x^2 &= r \sin \phi_1 \cos \phi_2 \\ x^3 &= r \sin \phi_1 \sin \phi_2 \cos \phi_3 \\ &\vdots \\ x^{d-1} &= r \sin \phi_1 \cdots \sin \phi_{d-2} \cos \phi_{d-1} \\ x^d &= r \sin \phi_1 \cdots \sin \phi_{d-2} \sin \phi_{d-1}. \end{aligned}$$

In Figure 9(a), we show that by using an approximate edge correction factor (using the same factor as $d = 2$), the PCF is wrongly estimated. Moreover, as the dimension increases, the estimated PCF moves farther away from the true PCF very quickly.

Note that the calculation of the correct edge correction factor requires the evaluation of a multi-dimensional integral which is computationally expensive in high dimensions. In this paper, we provide a closed form approximation of γ_W (using polynomial regression of order 2) in different dimensions $d = 2$ to 6 when $r \leq 1.0$. More specifically, we have the following approximation $\gamma_{\hat{W}} = 1 - a_1 r + a_2 r^2$ where a_1 and a_2 are as given below.

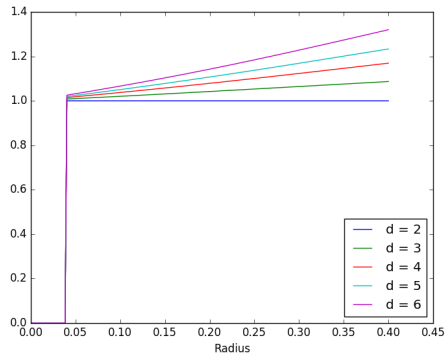
Dimension	$d = 2$	$d = 3$	$d = 4$	$d = 5$	$d = 6$
a_1	$4/\pi$	1.47	1.63	1.75	1.89
a_2	$1/\pi$	0.54	0.72	0.87	1.04

It can be observed from Figures 9(b) through 9(f) that the proposed approximations are quite tight.

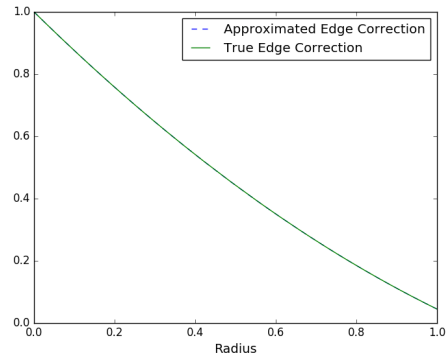
8.2 Synthesis Algorithm

The underlying idea of the proposed algorithm is to iteratively transform an initial random input sample design such that its PCF matches the target PCF. In particular, we propose a non-linear least squares formulation to optimize for the desired space-filling properties. Let us denote the target PCF by $G^*(r)$. We discretize the radius r into m points $\{r_j\}_{j=1}^m$ and minimize the sum of the weighted squares of errors between the target PCF $G^*(r_j)$ and the curve-fit function (kernel density estimator of PCF) $G(r_j)$ over m points. This scalar-valued goodness-of-fit measure is referred to as the *chi-squared error* criterion and can be posed as a non-linear weighted least squares problem as follows.

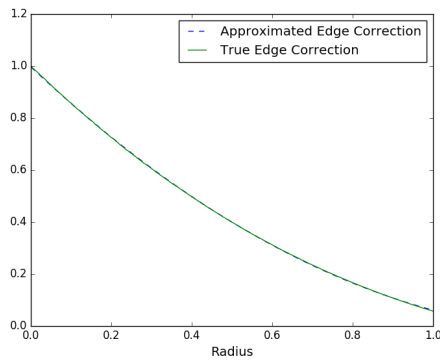
$$\arg \min \sum_{j=1}^M \left(\frac{G(r_j) - G^*(r_j)}{w_j} \right)^2,$$



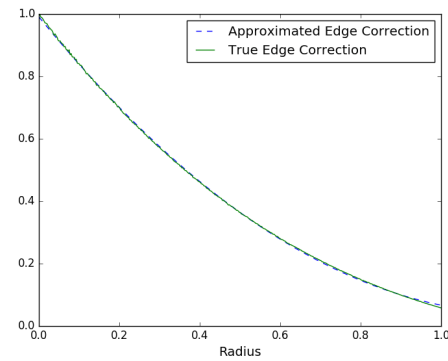
(a)



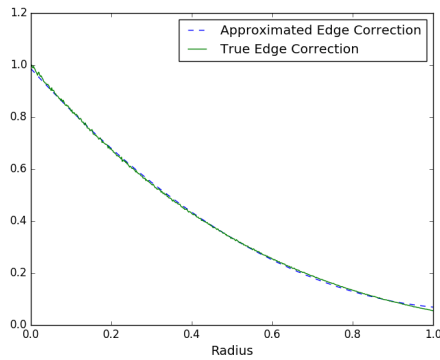
(b) $d = 2$



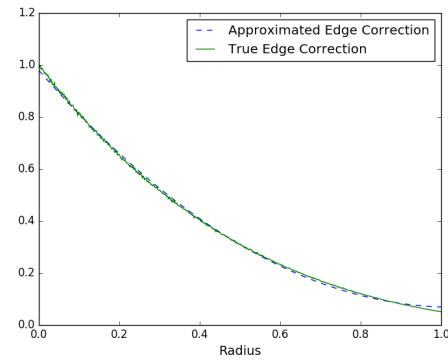
(c) $d = 3$



(d) $d = 4$



(e) $d = 5$



(f) $d = 6$

Figure 9: (a) Incorrect PCF estimation due to the use of an approximate edge correction factor, (b)-(f) Effectiveness of the approximation edge correction, obtained using polynomial regression, in comparison to the true edge correction from the evaluation of a multi-dimensional integral, for dimensions 2 to 6.

where w_j indicates the weight (importance) assigned to the fitting error at radius r_j . This optimization problem can be efficiently solved using a variant of gradient descent algorithm (discussed next), that in our experience converges quickly. In the simplest cases of uniform weights the solution tends to produce a higher fitting error at lower radii r_j . To address this challenge we use a non-uniform distribution for the weights $\{w_j\}$. These weights are initialized to be uniform and are updated in an adaptive fashion in the gradient descent iterations. The weight w_j at gradient descent iteration $t + 1$ is given by (Kailkhura et al., 2016b):

$$w_j = \frac{1}{|G^t(r_j) - G^*(r_j)|}$$

where $G^t(r_j)$ is the value of the PCF at radius r_j during the gradient descent iteration t . Note that PCF matching is a highly non-convex problem. We found that the following trick further helps solve PCF matching problem more efficiently.

8.2.1 ONE SIDED PCF SMOOTHING

We propose to perform one sided smoothing of the target PCF which is given as follows:

$$\hat{G}^*(r) = \begin{cases} (cr)^b & \text{if } r < r_{\min} \\ 1 & \text{if } r \geq r_{\min}. \end{cases}$$

where c is some pre-specified constant and $b > 1$ is the smoothing constant obtained via cross-validation. More specifically, we add polynomial noise in the low radius region of the PCF. This can also be interpreted as polynomial approximation of the PCF in the low radii regime. We have noticed that sometimes adding a controlled amount of Gaussian noise instead of polynomial noise also improves the performance.

8.2.2 EDGE CORRECTED GRADIENT DESCENT

The non-linear least squares problem is solved iteratively using gradient descent. Starting with a random point set $X = \{x_i\}_{i=1}^N$, we iteratively update x_i in the negative gradient direction of the objective function. At each iteration k , this can be formally stated as

$$x_i^{k+1} = x_i^k - \lambda \frac{\Delta_i}{|\Delta_i|},$$

where λ is the step size and $\Delta_i = \{\Delta_i^k\}_{k=1}^d$ in the normalized edge corrected gradient is given by

$$\Delta_i^p = \sum_{l \neq i} \frac{(x_l^p - x_i^p)}{|x_l - x_i|} \sum_{j=1}^m \frac{G(r_j)^k - G^*(r_j)}{w_j(1 - a_1 r_j + a_2 r_j^2) r_j^{d-1}} (|x_l - x_i| - r_j) k (r_j - |x_i - x_l|). \quad (25)$$

We re-evaluate the PCF $G(r_j)^k$ of the updated point set after each iteration using the unbiased estimator from the previous section.

The pseudocode of the algorithm is provided in Algorithm 1.

Algorithm 1 Space-filling Spectral Sample Design using PCF Matching Algorithm

- 1: **Input:** Number of samples N , dimension d , Smoothed target PCF $\hat{G}^*(r_j)$, weights w_j , step size λ , edge correction factors (a_1, a_2)
 - 2: $\mathbf{X} \leftarrow \text{Random}(N, d)$ ▷ Initial random sample design
 - 3: $G \leftarrow \text{PCF}(\mathbf{X})$ ▷ Calculate initial PCF using Eq. (22)
 - 4: **for** $t = 1$ to T **do** ▷ Total T gradient descent iterations
 - 5: **for** $i = 1$ to N **do** ▷ Update each sample at a time
 - 6: $\Delta_i^p \leftarrow \frac{\partial}{\partial x_i^p} \sum_{j=1}^M \left(\frac{G^t(r_j) - G^*(r_j)}{w_j} \right)^2$ for $p \in \{1, \dots, d\}$ ▷ Calculate gradients using (25)
 - 7: $x_i^{t+1} \leftarrow x_i^t - \lambda \frac{\Delta_i}{|\Delta_i|}$ ▷ Update the sample
 - 8: $G^t \leftarrow \text{PCF}(\mathbf{X})$ ▷ Update the PCF
 - 9: $w_j \leftarrow \frac{1}{|G^t(r_j) - \hat{G}^*(r_j)|}$ ▷ Update weights
 - 10: **return** \mathbf{X} ▷ Space-filling Spectral Samples
-

In Figure 10, we compare the behavior of the proposed PCF matching algorithm with and without the one sided PCF smoothing. The target PCF is designed using a Step PCF design with r_{min} as given in Proposition 10. PCF matching is carried out with varying sampling budget, $N = 100, 200, 400, 600, 800$ for $d = 2, 3, 4, 5, 6$, respectively. The variances of the Gaussian kernel were set at $\sigma^2 = 0.0065, 0.007, 0.01, 0.01, 0.01$ for $d = 2, 3, 4, 5, 6$, respectively and the step size for the gradient descent algorithm was fixed at 0.001. The value of b was obtained using cross-validation. The initial point set was generated randomly (uniform) in the unit hyper-cube and matching was carried for 100 gradient descent iterations. It can be observed that the proposed algorithm produces an accurate fit to the target, and that the smoothing actually leads to improved performance.

In Figure 11, we demonstrate the synthesis of a Stair PCF based spectral design, using parameters $P_0 = 1.2, \delta = 0.025$. Similar to the previous case, PCF matching is carried out with varying sampling budget, $N = 100, 200, 400, 600, 800$, for $d = 2, 3, 4, 5, 6$ respectively. The variances of the Gaussian kernel were set at $\sigma^2 = 0.0065, 0.007, 0.01, 0.01, 0.01$ for $d = 2, 3, 4, 5, 6$, respectively and the step size for the gradient descent algorithm was fixed at 0.001. We found that matching the Stair PCF is more challenging for a gradient descent optimization compared to the Step PCF. When a random point set is used for initialization, reaching convergence takes much longer. However, choosing the initial point set intelligently improves the quality of matching significantly. In all our experiments, we used the maximal PDS (Ebeida et al., 2012) to initialize the optimization and matching was carried for 100 gradient descent iterations. We observed that another reasonable choice for the initialization is a regular grid sample, and interestingly in most cases it matches the performance of the MPDS initialization. Furthermore, one sided PCF smoothing does not provide significant improvements in this case, particularly in higher dimensions.

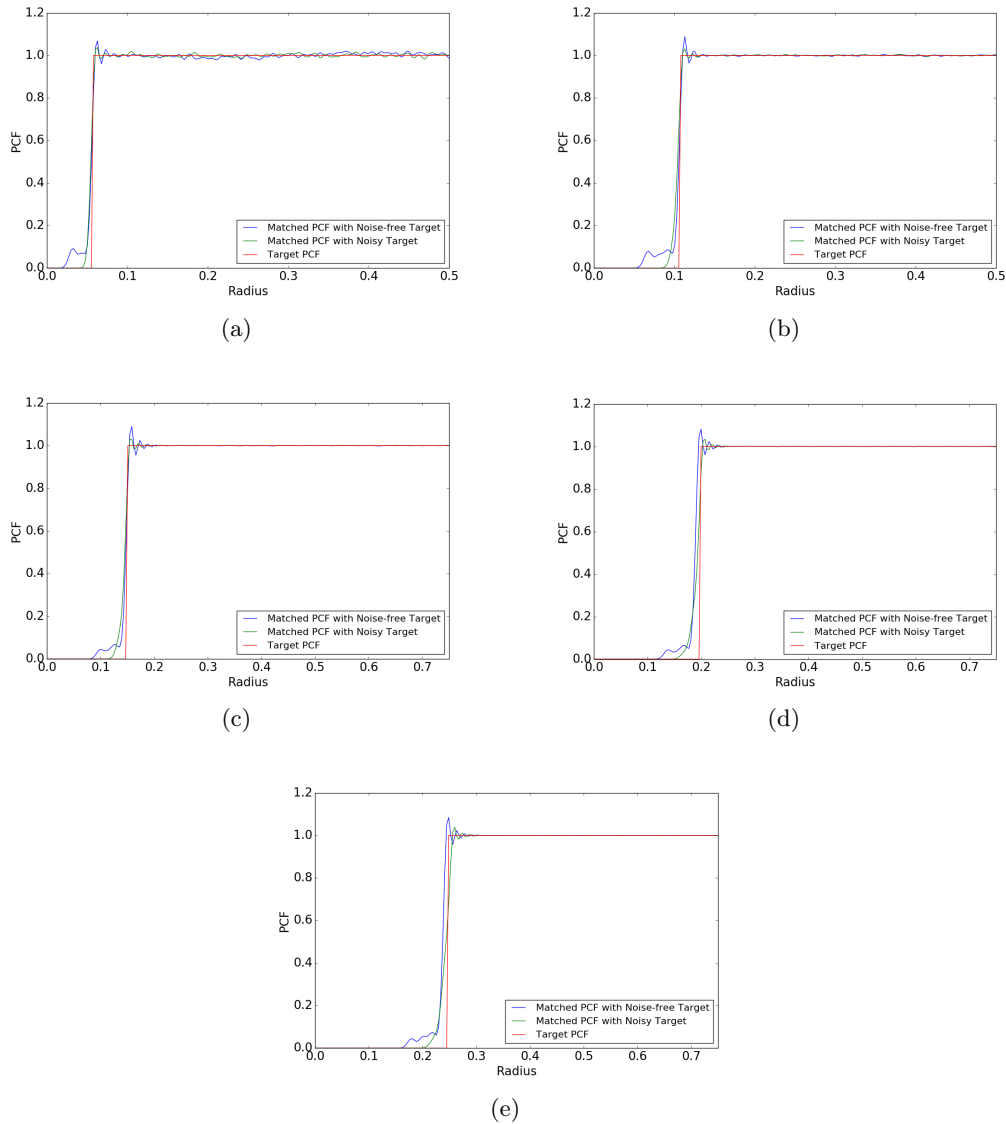


Figure 10: Step PCF synthesis using one sided PCF smoothing technique. (a) $d = 2$ (b) $d = 3$ (c) $d = 4$ (d) $d = 5$ (e) $d = 6$.

9. Experiments

In this section, we evaluate the qualitative performance of proposed space-filling spectral designs and present comparisons to popularly adopted space-filling designs, such as LHS, QMC and MPDS. Note that currently there does not exist any PDS synthesis approach which can generate sample sets with a desired size N while achieving user-specified spatial characteristics (e.g. relative radius). In all PDS synthesis approaches, there is no control over the number of samples generated by the algorithm which makes the use of these al-

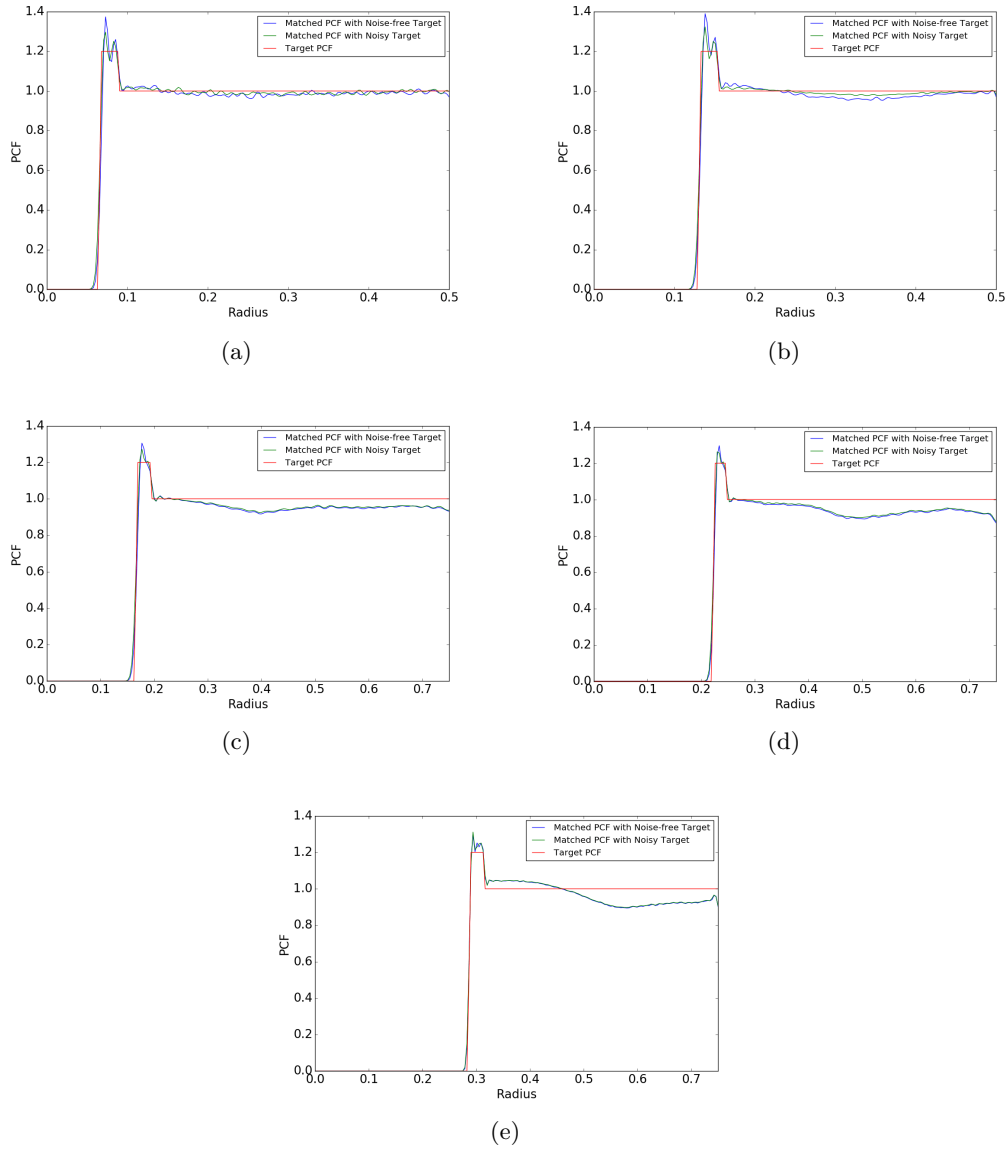
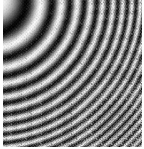
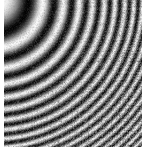

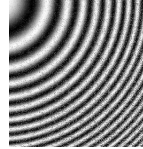
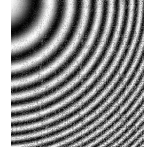
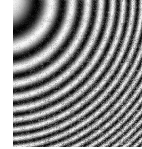
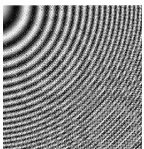
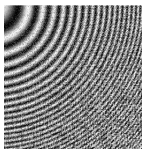
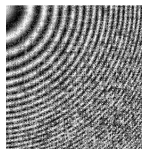
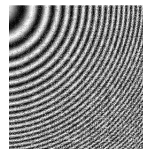
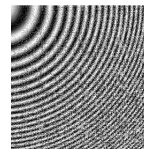
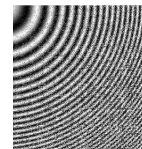
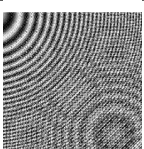
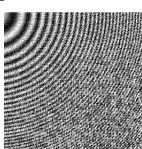

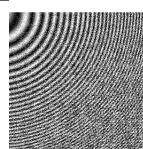
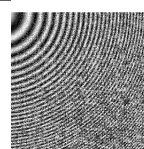
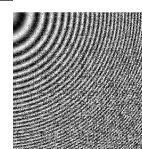
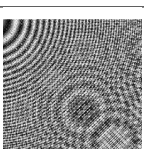
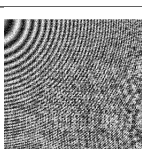
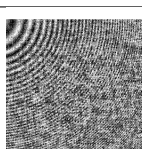
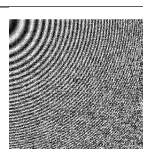
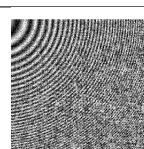

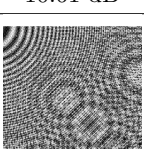
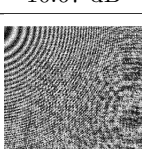
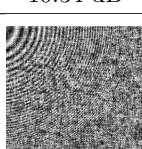
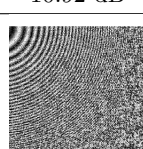
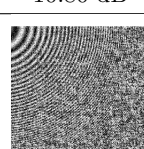
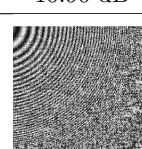
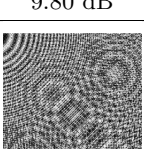
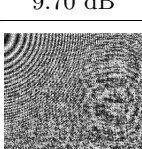
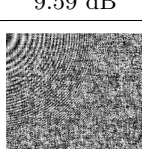
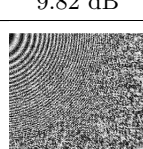
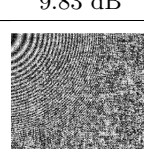
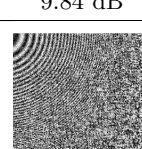
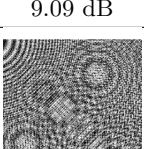
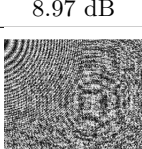
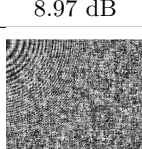
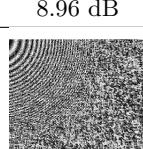
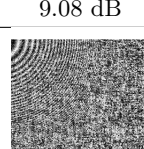
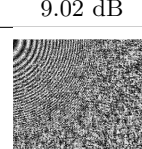


Figure 11: Stair PCF synthesis using one sided PCF smoothing technique. (a) $d = 2$ (b) $d = 3$ (c) $d = 4$ (d) $d = 5$ (e) $d = 6$.

gorithms difficult in practice. However, the proposed approach can control both N and r_{min} simultaneously. For our qualitative comparison, we perform three empirical studies, in dimensions 2 to 6 : (a) image reconstruction, (b) regression on several benchmark optimization functions, and (c) surrogate modeling for an inertial confinement fusion (ICF) simulation code.

Table 1: Impact of different space-filling designs on image reconstruction performance. In all cases, we show the reconstructed images and their PSNR values.

α	Sobol	Halton	LHS	MPDS	Step	Stair
0.001	 14.82 dB	 15.75 dB	 14.29 dB	 16.90 dB	 16.00 dB	 16.46 dB
0.002	 12.10 dB	 12.34 dB	 11.39 dB	 13.16 dB	 12.58 dB	 12.96 dB
0.003	 11.22 dB	 11.38 dB	 10.75 dB	 11.88 dB	 11.55 dB	 11.78 dB
0.004	 10.61 dB	 10.67 dB	 10.34 dB	 10.92 dB	 10.80 dB	 10.90 dB
0.005	 9.80 dB	 9.70 dB	 9.59 dB	 9.82 dB	 9.83 dB	 9.84 dB
0.006	 9.09 dB	 8.97 dB	 8.97 dB	 8.96 dB	 9.08 dB	 9.02 dB
0.007	 8.49 dB	 8.39 dB	 8.51 dB	 8.32 dB	 8.49 dB	 8.42 dB

9.1 Image Reconstruction

In this experiment, we consider the problem of designing sample distributions for image reconstruction. More specifically, we consider the commonly used zone plate test function:

$$z(r) = (1 + \cos(\alpha r^2))/2,$$

with varying levels of complexity (or frequency content) α . Note that we choose the zone plate for our study over natural images, since it shows the response for a wide range of frequencies and aliasing effects that are not masked by image features. For all zone plate renderings in this paper, we have tiled toroidal sets of 1000 2-dimensional points over the image-plane and utilized a Lanczos filter with a support of width 4 for resampling. Further, we also report the peak signal-to-noise ratio (PSNR) as a quantitative error measure:

$$\text{PSNR} = 20 \log_{10} \frac{1}{\text{MSE}},$$

where MSE is the mean squared error. However, it is well known in the image processing community that PSNR can be a weak surrogate for visual quality (as we will see later) and, therefore, we also show the reconstructed images.

Table 1 illustrates the reconstructions obtained using different space-filling designs, for varying values of α . It can be observed from the results that the QMC sequences produce a large amount of aliasing artifacts in the high frequency regions, which can be directly linked to the oscillations in their corresponding PCFs. On the other hand, LHS design recovers a small amount of low-frequencies, and maps most of the frequencies to white noise due to its small r_{min} and near-constant PCF. In contrast, sample designs which attempt to trade-off between coverage and randomness properties, i.e., MPDS and the proposed spectral space-filling designs (as seen in Figure 3), have superior reconstruction quality. These designs reduce the aliasing artifacts, have cleaner low frequency content (upper left corner) and map all high frequencies (bottom right corner) to white noise. More interestingly, we see that for low complexity cases, i.e., lower α , the MPDS performs the best followed by the proposed Stair and Step PCF respectively. For moderately complex images, the Stair PCF performs the best followed by the Step and the MPDS. Finally, for highly complex images, the Step PCF performs the best followed by the Stair and the MPDS. These observations corroborate our discussion in Section 5.2 that an increase in r_{min} (coverage) in the PCF results in an increase in the range of low frequencies that can be recovered without aliasing, and equivalently reduction in the amount of oscillations (or an increase in randomness) in the PCF leads to reduced oscillations in the PSD, which in turn indicates a systematic mapping of high frequency content to white noise. Note that when $\alpha = 0.007$, both LHS and Sobol designs have PSNR greater than (or equal to) the PSNR of Step PCF design. However, the quality of the reconstructed image by Step PCF is far superior compared to the one by LHS and Sobol designs. This further corroborates our claim on PSNR being a weak surrogate and justifies the use of reconstructed images itself as a performance metric.

9.2 Regression Modeling for Benchmark Optimization Functions

In this study, we consider the problem of fitting regression models to analytical functions and perform a comparative study of different sample designs, in terms of their generalization

performance. More specifically, we consider a set of benchmark analytic functions between dimensions 2 and 6, that are commonly used in global optimization tests (Jamil and Yang, 2013). They are chosen due to their diversity in terms of their complexity and shapes. We compare the performance of proposed space-filling spectral designs (Step, Stair) with coverage based designs (MPDS), low-discrepancy designs (Halton and Sobol), latin hypercube sampling and random sampling. Appendix A lists the set of functions used in our experiments. In each case, we fit a random forest regressor with 30 trees and repeated for 20 independent realizations of sample designs. We evaluate the generalization performance on 10^6 regular grid based test samples. Finally, we report mean (horizontal lines) and standard deviation (vertical lines) of 3 popular quality metrics (over 20 realizations) to quantify the performance of the resulting regression models: mean squared error (MSE), relative average absolute error (AAE), and the R^2 -statistic. The metrics are defined as follows:

$$MSE(\mathbf{y}, \hat{\mathbf{y}}) = \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}, \quad (26)$$

$$AAE(\mathbf{y}, \hat{\mathbf{y}}) = \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{N * STD(\mathbf{y})}, \quad (27)$$

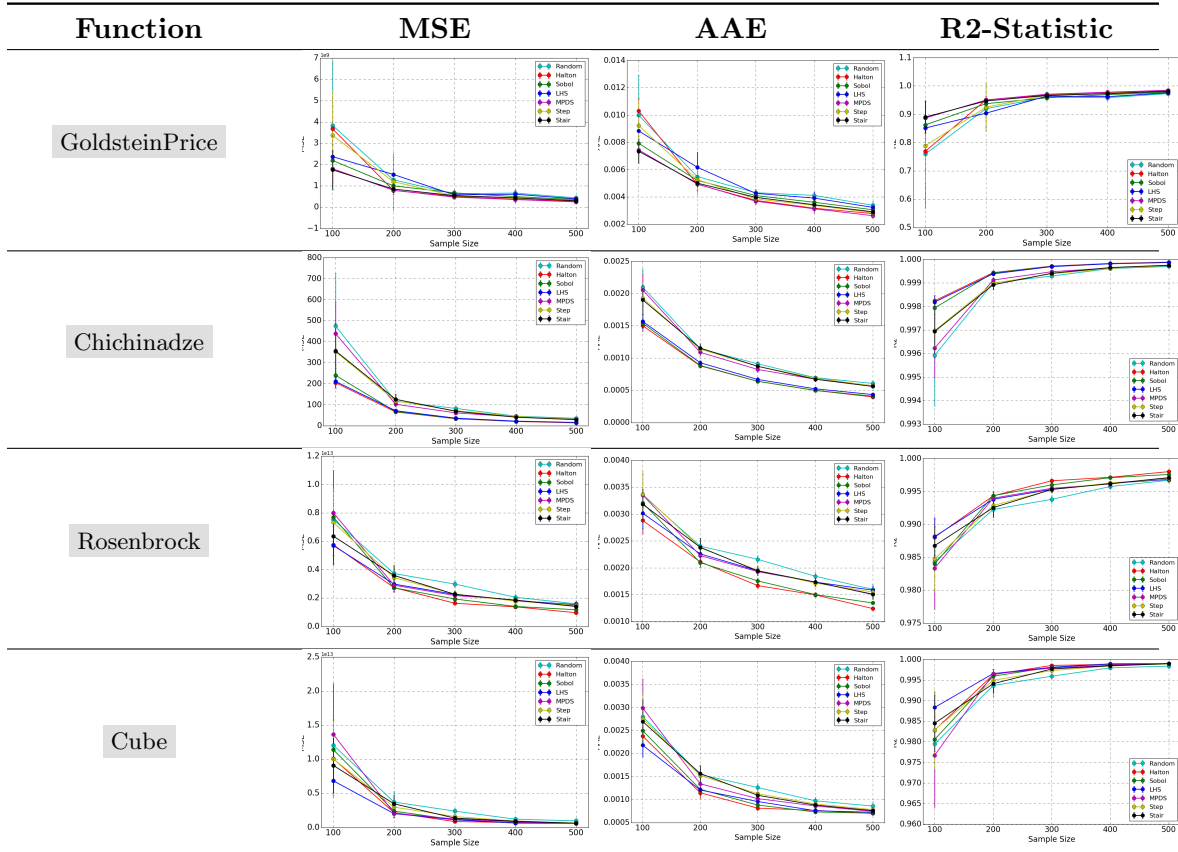
$$R^2(\mathbf{y}, \hat{\mathbf{y}}) = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - MEAN(\mathbf{y}))^2} \quad (28)$$

where $\mathbf{y} = f(\mathbf{x})$ are the true function values and $\hat{\mathbf{y}}$ are the predicted values.

Tables 2 through 6 show the performance of different space-filling designs for various analytic functions in dimensions 2 to 6, respectively¹⁰. We see that, for $d = 2$ (Table 2), LHS and Halton sequences perform better compared to the rest of the sample designs on most of the test functions. However, on some functions, e.g., GoldsteinPrice, Stair PCF and MPDS perform better. Therefore, none of the sample designs consistently guarantee superior performance. For $d = 3$ (Table 3), we see that Stair PCF design and MPDS (followed by Sobol sequences) perform consistently better compared to the rest of the approaches. As we go higher in dimensions, i.e., $d > 3$, we notice a significant gain in the performance of Stair PCF based space-filling spectral designs. Interesting, the amount of performance gain of Stair PCF based design increases as we go higher in dimensions. The reason for the poor regression performance of QMC sequences and LHS for $d > 3$ is due to their poor space-filling properties in high dimensions (Wang and Sloan, 2008). In comparison, both space-filling spectral designs and MPDS have good space-filling properties. We found that Stair PCF design and MPDS have similar coverage characteristics (r_{min}). However, the difference in their performance can be attributed to the fact that MPDS designs have significantly more oscillations in their PCF compared to an equivalent Stair PCF based space-filling spectral design, i.e. violation of the randomness objective.

10. Non-monotonicity of the error curves represents over-fitting and is more prominent with conventional sample designs.

Table 2: Impact of sample design on generalization performance of regression models fit to benchmark analytical functions in 2 dimensions. LHS and Halton sequences perform slightly better compared to rest of the sample designs.

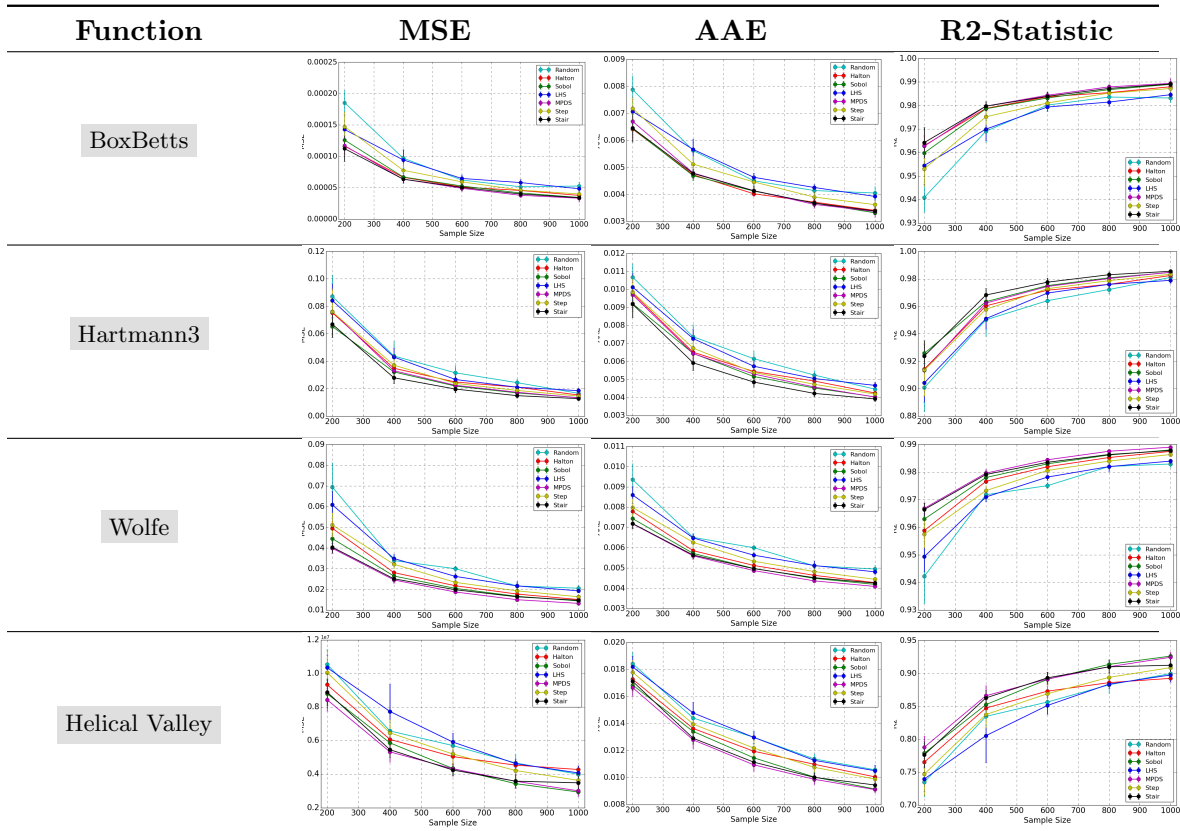


9.3 Surrogate Model Design for an Inertial Confinement Fusion (ICF) Simulator

In this subsection, we consider the problem of designing surrogate models for an inertial confinement fusion (ICF) simulator developed at the National Ignition Facility (NIF). The NIF is aimed at demonstrating inertial confinement fusion (ICF), that is, thermonuclear ignition and energy gain in a laboratory setting. The goal is to focus 192 beams of the most energetic laser built so far onto a tiny capsule containing frozen deuterium. Under the right conditions, the resulting pressure will collapse the target to the point of ignition where hydrogen starts to fuse and produce massive amounts of energy, effectively creating a small star which can be harnessed for energy production. Though significant progress has been made, the ultimate goal of “ignition” has not yet been reached.

NIF employs an adaptive pipeline: perform experiments, use post-shot simulations to understand the experimental results, and design new experiments with parameter settings that are expected to improve performance. From an analysis viewpoint, the goal is to search the parameter space to find the region that leads to near-optimal performance. The dataset

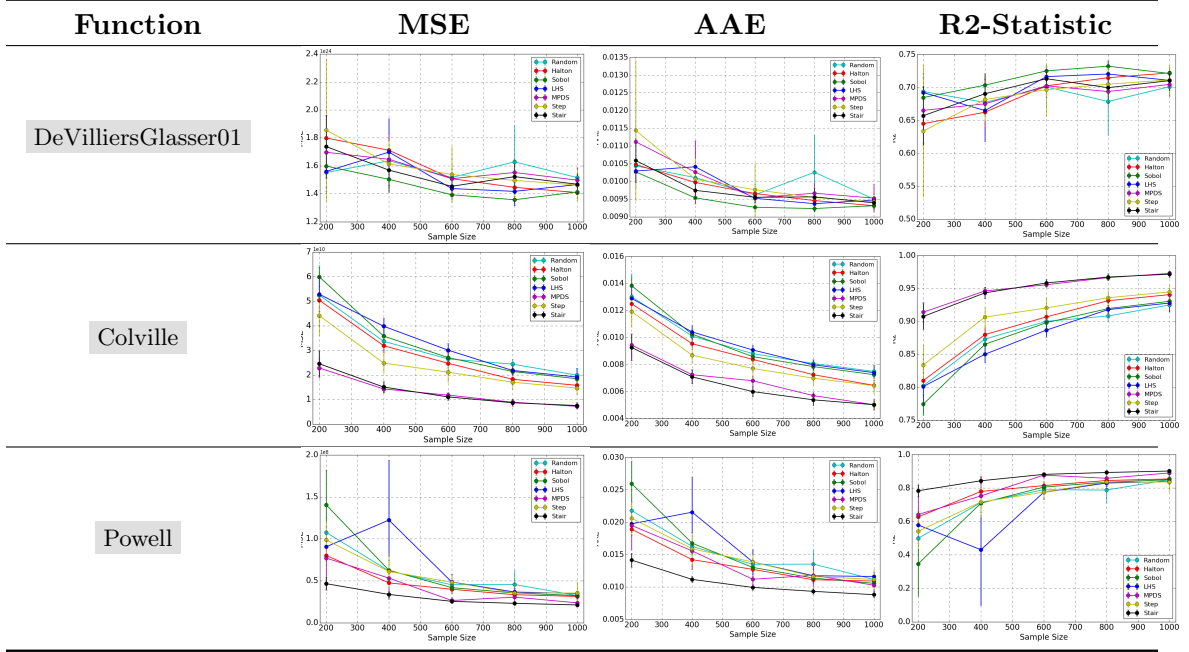
Table 3: Impact of sample design on generalization performance of regression models fit to benchmark analytical functions in 3 dimensions. While the Stair PCF and MPDS designs are consistently better than the other methods, the amount of performance gain is minimal.



considered here is a so called engineering or macro-physics simulation ensemble in which an implosion is simulated using different input parameters, such as, laser power, pulse shape etc. From these simulations, scientists extract a set of drivers, physical quantities believed to determine the behavior of the resulting implosion. These drivers are then analyzed with respect to the energy yield to better understand how to optimize future experiments. As one can expect, the success of this pipeline heavily depends on the quality of samples used for post-shot simulations.

We use the NIF 1-d HYDRA simulator (Marinak et al., 2001) and compare the performance of proposed space-filling spectral designs with existing approaches (random, LHS, Halton and Sobol and MPDS). For each simulation run, a large number of output quantities, such as peak velocity, yield, etc., are computed, and subsequently used to describe the resulting implosion. We vary the number of input parameters between 2 and 6, and fix the remaining variables to their default values. In each case, we fit a random forest regressor with 30 trees and repeated for 20 independent realizations of sample designs. We evaluated

Table 4: Impact of sample design on generalization performance of regression models fit to benchmark analytical functions in 4 dimensions. Stair PCF and MPDS designs demonstrate appreciable gains over popular sample design choices.



the reconstruction performance on 10^5 regular grid based test samples using the metrics in the previous experiment.

Table 7 shows the regression performance of the different sample designs for various output quantities in dimensions 2 to 6. We observe that regression error patterns are consistent with our observations in Section 9.2. The proposed Stair PCF based design consistently performs the best (followed by MPDS) for $d \geq 3$. Furthermore, the performance gain with the Stair PCF based design improves as we go higher in dimensions. This performance gain can be credited to their ability to achieve better space-filling properties in high dimensions by intelligently balancing the trade-off between coverage and randomness, and the effectiveness of the proposed metric (PCF) adopted for design and optimization.

10. Applications of Design of Experiments in Machine Learning Problems

In addition to the experiments presented in the paper, the proposed experiment design methodologies have a broader impact on several classical machine learning (ML) formulations. In its simplest form, the proposed sample designs can be used to create training data in supervised learning problems. In particular, optimized sample designs can provide significant performance gains in application areas where efficient data acquisition is required, e.g. machine learning for scientific data analysis (Karpatne et al., 2017). One such use case was considered in Section 9.3 of this paper for inertial confinement fusion (ICF) studies. We demonstrated that supervised models learned using samples from the proposed exper-

Table 5: Impact of sample design on generalization performance of regression models fit to benchmark analytical functions in 5 dimensions. In higher dimensions, conventional methods such as the LHS and QMC perform very poorly, while Stair PCF design significantly outperforms all competing methods, because of improved trade-off between coverage and randomness properties.

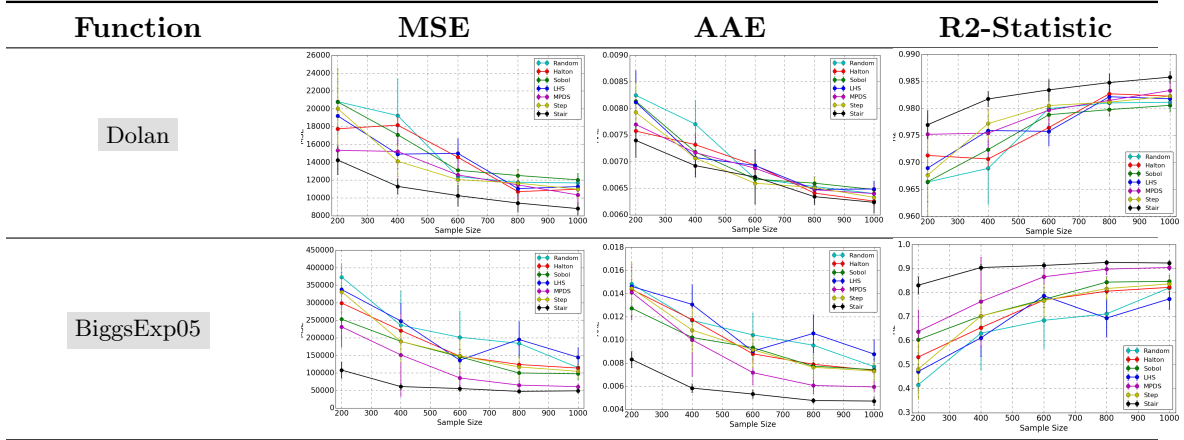
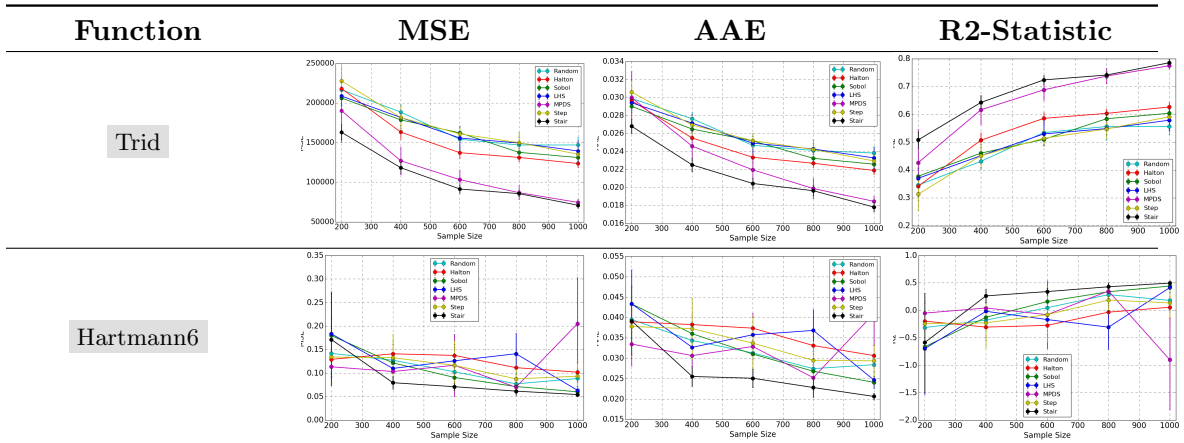
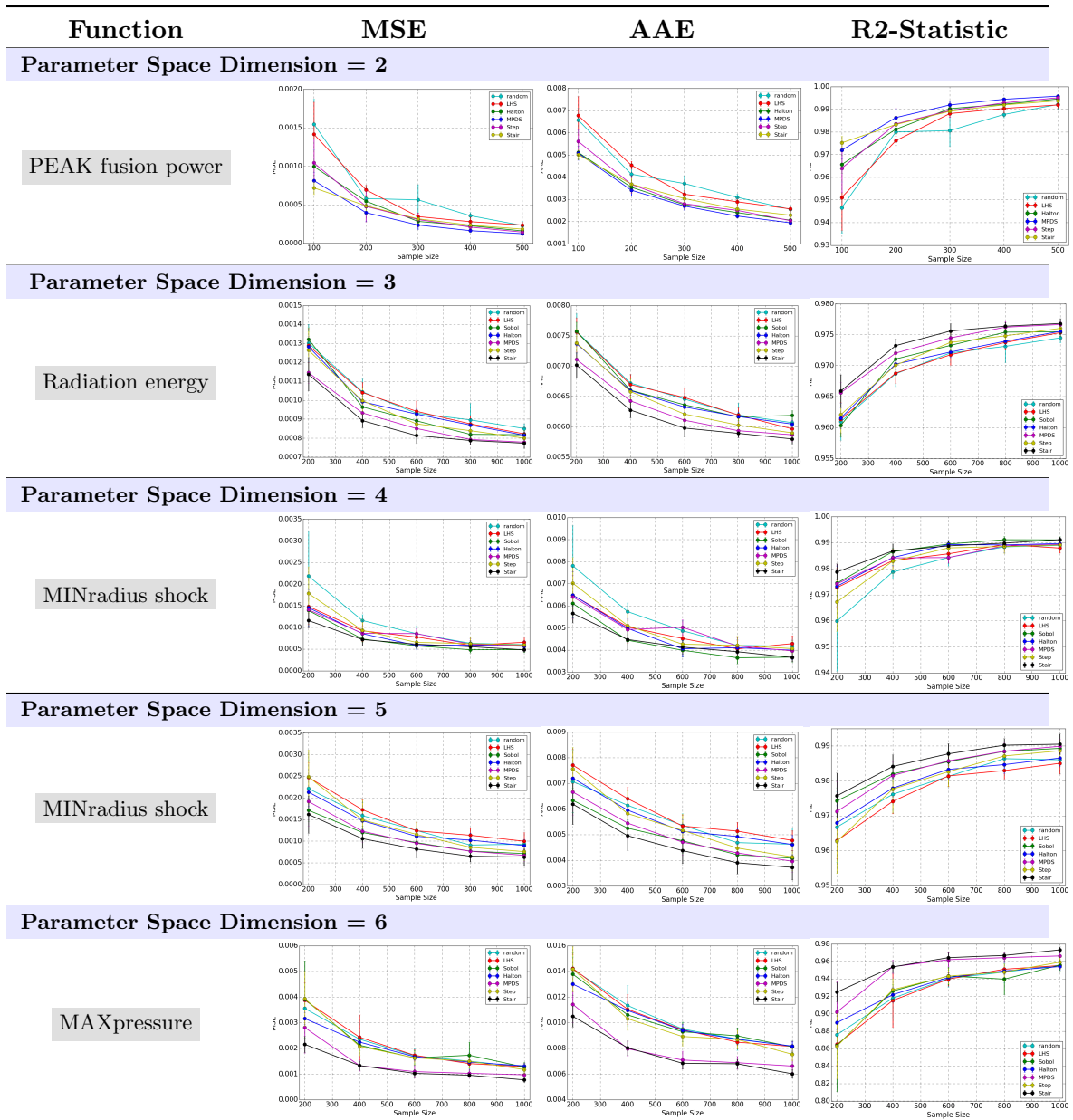


Table 6: Impact of sample design on generalization performance of regression models fit to benchmark analytical functions in 6 dimensions. Even with highly complex functions such as the *Hartmann6*, the proposed Stair PCF based spectral space-filling design produces more accurate regression models, thus evidencing the importance of improved space-filling characteristics.



iment design produce highly generalizable models, i.e., have significantly lower prediction error for unseen input conditions. Furthermore, utilizing the notion of space-filling spectral designs to incorporate the prior information on signal structure (e.g., low dimensionality, sparsity) will be particularly useful for image analysis and computer vision tasks. Though not discussed, the analytical framework developed in this paper, can be extended beyond

Table 7: Performance of surrogate models for the NIF 1-d HYDRA simulator using different sample design techniques, with varying number of input parameters. While the conventional sample designs achieve reasonable performance in low dimensions, the proposed Stair PCF based design is consistently superior as the dimension of the input space grows.



Euclidean spaces to non-linear, embedded manifolds. For example, Anirudh et.al. (Anirudh et al., 2017) proposed a dart throwing technique to generate Poisson disk samples on the

Grassmannian manifold of low-dimensional linear subspaces. Despite the effectiveness of such heuristic techniques over other randomized sampling strategies, generating space-filling spectral designs on embedded manifolds is challenging since it is non-trivial to create equivalent definitions of the PSD and the PCF metrics for non-Euclidean domains.

Furthermore, there is a connection between the sample design problem considered in this paper and the classic ML task of active learning. In many practical scenarios, it is possible to use information gleaned from previous observations to improve the sampling process. As more samples are obtained, one can learn how to improve the sampling process by deciding where to sample next. These sampling feedback techniques are more generally known as adaptive sampling in the statistics literature. Note that, several popular design of experiment techniques have been extended to adaptive sampling scenarios (Yan and Wonka, 2012b). A natural extension of our work is towards builds importance sampling techniques, guided by spectral properties. In addition to these conventional applications, more recently, optimized sample designs have been used to improve the convergence characteristics of neural network training process. Several efforts are currently being undertaken for effective mini-batch sampling and studying their effect on the convergence rate of training algorithms. Similarly, one could develop improved mini-batch sampling strategies through the analytical framework of space-filling spectral designs. Similarly, hyper-parameter optimization in deep learning is another application area where optimized sample design can be very useful (Bergstra and Bengio, 2012). Finally, space-filling spectral designs are also applicable in reinforcement learning (Sutton and Barto, 1998) and Bayesian optimization (Snoek et al., 2012) where a key requirement is to effectively balance between exploration and exploitation.

11. Conclusion and Future Directions

In this work, we considered the problem of constructing high quality space-filling designs. We proposed the use of pair correlation function (PCF) to quantify the space-filling property and systematically traded-off coverage and randomness in sample designs in arbitrary dimensions. Next, we linked PCF to the power spectral density (PSD) to analyze the qualitative measure of the design performance. Using the insights provided by this spatial-spectral analysis, we proposed novel space-filling spectral designs. We also provided an efficient PCF estimator to evaluate the space-filling properties of sample designs in arbitrary dimensions. Next, we devised a gradient descent based optimization algorithm to generate high quality space-filling designs. Superiority of proposed space-filling spectral designs were shown on two different applications in 2 to 6 dimensions: a) image reconstruction and b) surrogate modeling on several benchmark optimization functions and an inertial confinement fusion (ICF) simulation code. There are still many interesting questions that remain to be explored in the future work such as an analysis of the problem for non-linear manifolds. Note that some analytical methodologies used in this paper are certainly exploitable for studying and designing space-filling designs in different manifolds. Other questions such as PCF parameterizations for other variants of space-filling designs, adaptive and importance sampling, and optimization approaches to synthesize them can also be investigated.

Acknowledgments

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. LLNL-JRNL-743060

Appendix A. Benchmark Optimization Functions

A.1 2 Dimensional Functions

Rosenbrock: $\sum_{p=1}^2 (100((x^p)^2 - x^{p+1}) + (x^p - 1)^2)$

Cube: $100(x^2 - (x^1)^3)^2 + (1 - x^1)^2$

Chichinadze: $(x^1)^2 - 12x^1 + 8 \sin(2.5\pi x^1) + 10 \cos(0.5\pi x^1) + 11 - 0.2 \frac{\sqrt{5}}{\exp(0.5(x^2 - 0.5)^2)}$

GoldsteinPrice: $(1 + (x^1 + x^2 + 1)^2(19 - 14x^1 + 3(x^1)^2 - 14x^2 + 6x^1x^2 + 3(x^2)^2))$
 $(30 + (2x^1 - 3x^2)^2(18 - 32x^1 + 12(x^1)^2 + 48x^2 - 36x^1x^2 + 27(x^2)^2))$

A.2 3 Dimensional Functions

BoxBetts: $\sum_{p=1}^3 g(x^i)^2; g(x) = \exp(-0.1(p+1)x^1) - \exp(-0.1(p+1)x^2) - (\exp(-0.1(p+1)) - \exp(-(p+1))x^3)$

HelicalValley: $100(x^3 - 10\psi(x^1, x^2))^2 + (\sqrt{(x^1)^2 + (x^2)^2} - 1)^2 + (x^3)^2; 2\pi\psi(x^1, x^2) = \tan^{-1}(x^2/x^1)$ if $x^1 \geq 0$, **and** $\pi + \tan^{-1}(x^2/x^1)$ **otherwise.**

Wolfe: $\frac{4}{3}((x^1)^2 + (x^2)^2 - x^1x^2)^{0.75} + x^3$

Hartmann3: $-\sum_{i=1}^4 c_i \exp(-\sum_{j=1}^3 a_{ij}(x^j - p_{ij})^2)$

A.3 4 Dimensional Functions

DeVilliersGlasser01: $\sum_{i=1}^{24} (x^1(x^2)^{0.1(i-1)} \sin(x^3(0.1(i-1)) + x^4) - y_i)^2$

Powell: $(x^3 + 10x^1)^2 + 5(x^2 - x^4)^2 + (x^1 - 2x^2)^4 + 10(x^3 - x^4)^4$

Colville: $(x^1 - 1)^2 + 100((x^1)^2 - x^2)^2 + 10.1(x^2 - 1)^2 + (x^3 - 1)^2 + 90((x^3)^2 - x^4)^2 + 10.1(x^4 - 1)^2 + 19.8 \frac{x^4 - 1}{x^2}$

A.4 5 Dimensional Functions

BiggsExp05: $\sum_{i=1}^{11} (x^3 e^{-t_i x^1} - x^4 e^{-t_i x^2} + 3e^{-t_i x^5} - y_i)^2; t_i = 0.1i; y_i = e^{-t_i} - 5e^{-10t_i} + 3e^{-4t_i}$

Dolan: $|(x^6 + 1.7x^2) \sin(x^1) - 1.5x^3 - 0.1x^4 \cos(x^5 - x^1) + 0.2(x^5)^2 - x^2 - 1|$

A.5 6 Dimensional Functions

Trid: $\sum_{p=1}^6 (x^p - 1)^2 - \sum_{p=2}^6 x^p x^{p-1}$

Hartmann6: $-\sum_{i=1}^4 c_i \exp(-\sum_{j=1}^6 a_{ij}(x^j - p_{ij})^2)$

References

- R. Anirudh, B. Kailkhura, J. J. Thiagarajan, and P. T. Bremer. Poisson disk sampling on the grassmannian: Applications in subspace optimization. *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1(1):690–698, July 2017. ISSN 2160-7516.
- Michael Balzer, Thomas Schlomer, and Oliver Deussen. Capacity-constrained point distributions: A variant of lloyd’s method. *ACM Trans. Graph.*, 28(3):86:1–86:8, July 2009. ISSN 0730-0301.
- James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(Feb):281–305, 2012.
- Robert Bridson. Fast poisson disk sampling in arbitrary dimensions. *ACM SIGGRAPH 2007 Sketches*, 2007.
- Russel E. Caffisch. Monte carlo and quasi-monte carlo methods. *Acta Numerica*, 7:149, 1998.
- Henry Cohn, Abhinav Kumar, Stephen D Miller, Danylo Radchenko, and Maryna Viazovska. The sphere packing problem in dimension 24. *Annals of Mathematics*, 185(3):1017–1033, 2017.
- Robert L. Cook. Stochastic sampling in computer graphics. *ACM Trans. Graph.*, 5(1):51–72, January 1986. ISSN 0730-0301.
- Fernando de Goes, Katherine Breeden, Victor Ostromoukhov, and Mathieu Desbrun. Blue noise through optimal transport. *ACM Trans. Graph.*, 31(6):171:1–171:11, November 2012. ISSN 0730-0301.
- Mark A. Z. Dippe and Erling Henry Wold. Antialiasing through stochastic sampling. *SIGGRAPH Comput. Graph.*, 19(3):69–78, July 1985. ISSN 0097-8930.
- Daniel Dunbar and Greg Humphreys. A spatial data structure for fast poisson-disk sample generation. *ACM Trans. Graph.*, 25(3):503–508, July 2006. ISSN 0730-0301.
- Mohamed S. Ebeida, Andrew A. Davidson, Anjul Patney, Patrick M. Knupp, Scott A. Mitchell, and John D. Owens. Efficient maximal poisson-disk sampling. *ACM Trans. Graph.*, 30(4):49:1–49:12, July 2011. ISSN 0730-0301.
- Mohamed S. Ebeida, Scott A. Mitchell, Anjul Patney, Andrew A. Davidson, and John D. Owens. A simple algorithm for maximal poisson-disk sampling in high dimensions. *Computer Graphics Forum*, 31(2pt4):785–794, 2012. ISSN 1467-8659.
- Mohamed S. Ebeida, Anjul Patney, Scott A. Mitchell, Keith R. Dalbey, Andrew A. Davidson, and John D. Owens. K-d darts: Sampling by k-dimensional flat searches. *ACM Trans. Graph.*, 33(1):3:1–3:16, February 2014. ISSN 0730-0301.
- Ronald A Fisher. The design of experiments. 1935. *Oliver and Boyd, Edinburgh*, 1935.

- Manuel N. Gamito and Steve C. Maddock. Accurate multidimensional poisson-disk sampling. *ACM Trans. Graph.*, 29(1):8:1–8:19, December 2009. ISSN 0730-0301.
- Sushant S. Garud, Iftekhar A. Karimi, and Markus Kraft. Design of computer experiments: A review. *Computers and Chemical Engineering*, 106(Supplement C):71 – 95, 2017. ISSN 0098-1354. ESCAPE-26.
- Bo Geng, HuiJuan Zhang, Heng Wang, and GuoPing Wang. Approximate poisson disk sampling on mesh. *Science China Information Sciences*, 56(9):1–12, 2013. ISSN 1674-733X.
- Jianwei Guo, Dong-Ming Yan, Guanbo Bao, Weiming Dong, Xiaopeng Zhang, and Peter Wonka. Efficient triangulation of poisson-disk sampled point sets. *The Visual Computer*, 30(6-8):773–785, 2014. ISSN 0178-2789.
- J. H. Halton. Algorithm 247: Radical-inverse quasi-random point sequence. *Commun. ACM*, 7(12):701–702, December 1964. ISSN 0001-0782.
- Daniel Heck, Thomas Schlomer, and Oliver Deussen. Blue noise sampling with controlled aliasing. *ACM Trans. Graph.*, 32(3):25:1–25:12, July 2013. ISSN 0730-0301.
- Wenguang Hou, Xuming Zhang, Xin Li, Xudong Lai, and Mingyue Ding. Poisson disk sampling in geodesic metric for dem simplification. *International Journal of Applied Earth Observation and Geoinformation*, 23(1):264 – 272, 2013. ISSN 0303-2434.
- Janine Illian, Antti Penttinen, Helga Stoyan, and Dietrich Stoyan. *Statistical analysis and modelling of spatial point patterns*, volume 70. John Wiley and Sons, 2008.
- Cheuk Yiu Ip, M. Adil Yalin, David Luebke, and Amitabh Varshney. Pixelpie: Maximal poisson-disk sampling with rasterization. *Proceedings of the 5th High-Performance Graphics Conference*, pages 17–26, 2013.
- Momin Jamil and Xin-She Yang. A literature survey of benchmark functions for global optimisation problems. *International Journal of Mathematical Modelling and Numerical Optimisation*, 4(2):150–194, 2013.
- Ruichen Jin, Wei Chen, and Agus Sudjianto. An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134(1): 268 – 287, 2005. ISSN 0378-3758.
- V Roshan Joseph. Space-filling designs for computer experiments: A review. *Quality Engineering*, 28(1):28–35, 2016.
- B. Kailkhura, J. J. Thiagarajan, P. T. Bremer, and P. K. Varshney. Theoretical guarantees for poisson disk sampling using pair correlation function. *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2589–2593, March 2016a.
- Bhavya Kailkhura, Jayaraman J. Thiagarajan, Peer-Timo Bremer, and Pramod K. Varshney. Stair blue noise sampling. *ACM Trans. Graph.*, 35(6):248:1–248:10, November 2016b. ISSN 0730-0301.

- Anuj Karpatne, Gowtham Atluri, James H Faghmous, Michael Steinbach, Arindam Banerjee, Auroop Ganguly, Shashi Shekhar, Nagiza Samatova, and Vipin Kumar. Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*, 29(10):2318–2331, 2017.
- JR Koehler and AB Owen. Computer experiments. *Handbook of statistics*, 13:261–308, 1996.
- Ares Lagae and Philip Dutr. A comparison of methods for generating poisson disk distributions. *Computer Graphics Forum*, 27(1):114–129, 2008. ISSN 1467-8659.
- Stephen Leary, Atul Bhaskar, and Andy Keane. Optimal orthogonal-array-based latin hypercubes. *Journal of Applied Statistics*, 30(5):585–598, 2003.
- Pierre LEcuyer and Christiane Lemieux. Recent advances in randomized quasi-monte carlo methods. *Modeling uncertainty*, pages 419–474, 2005.
- MM Marinak, GD Kerbel, NA Gentile, O Jones, D Munro, S Pollaine, TR Dittrich, and SW Haan. Three-dimensional hydra simulations of national ignition facility targets. *Physics of Plasmas*, 8(5):2275–2280, 2001.
- Michael McCool and Eugene Fiume. Hierarchical poisson disk sampling distributions. *Proceedings of the Conference on Graphics Interface '92*, pages 94–105, 1992.
- Michael D. McKay. Latin hypercube sampling as a tool in uncertainty analysis of computer models. *Proceedings of the 24th Conference on Winter Simulation*, pages 557–564, 1992.
- William J. Morokoff and Russel E. Caflisch. Quasi-random sequences and their discrepancies. *SIAM J. Sci. Comput*, 15:1251–1279, 1994.
- Max D. Morris and Toby J. Mitchell. Exploratory designs for computational experiments. *Journal of Statistical Planning and Inference*, 43(3):381 – 402, 1995. ISSN 0378-3758.
- Victor Ostromoukhov, Charles Donohue, and Pierre-Marc Jodoin. Fast hierarchical importance sampling with blue noise properties. *ACM Trans. Graph.*, 23(3):488–495, August 2004. ISSN 0730-0301.
- Art B. Owen. *Randomly Permuted (t,m,s) -Nets and (t, s) -Sequences*. Springer New York, New York, NY, 1995. ISBN 978-1-4612-2552-2.
- Art B Owen. Monte carlo and quasi-monte carlo for statistics. *Monte Carlo and Quasi-Monte Carlo Methods 2008*, pages 3–18, 2009.
- Art B Owen and Seth D Tribble. A quasi-monte carlo metropolis algorithm. *Proceedings of the National Academy of Sciences of the United States of America*, 102(25):8844–8849, 2005.
- A. Cengiz Oztireli and Markus Gross. Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph.*, 31(6):170:1–170:10, November 2012. ISSN 0730-0301.

- N. Packham. Combining latin hypercube sampling with other variance reduction techniques. *Wilmott*, 2015(76):60–69, 2015. ISSN 1541-8286.
- Thomas Schlomer, Daniel Heck, and Oliver Deussen. Farthest-point optimized point sets with maximized minimum distance. *Proceedings of the ACM SIGGRAPH Symposium on High Performance Graphics*, pages 135–142, 2011.
- Edwin W. Schreiber. Mathematical snapshots*. *School Science and Mathematics*, 43(9):795–799, 1943. ISSN 1949-8594.
- Jasper Snoek, Hugo Larochelle, and Ryan P Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, pages 2951–2959, 2012.
- I.M Sobol. On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, 7(4):86 – 112, 1967. ISSN 0041-5553.
- Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.
- Boxin Tang. Orthogonal array-based latin hypercubes. *Journal of the American statistical association*, 88(424):1392–1397, 1993.
- maryna s viazovska. the sphere packing problem in dimension 8. *annals of mathematics*, 185(3):991–1015, 2017.
- Florent Wachtel, Adrien Pilleboue, David Coeurjolly, Katherine Breeden, Gurprit Singh, Gael Cathelin, Fernando de Goes, Mathieu Desbrun, and Victor Ostromoukhov. Fast tile-based adaptive sampling with user-specified fourier spectra. *ACM Trans. Graph.*, 33(4):56:1–56:11, July 2014. ISSN 0730-0301.
- Xiaoqun Wang and Ian H. Sloan. Low discrepancy sequences in high dimensions: How well are their projections distributed? *Journal of Computational and Applied Mathematics*, 213(2):366 – 386, 2008. ISSN 0377-0427.
- Li-Yi Wei. Parallel poisson disk sampling. *ACM Trans. Graph.*, 27(3):20:1–20:9, August 2008. ISSN 0730-0301.
- Li-Yi Wei. Multi-class blue noise sampling. *ACM Trans. Graph.*, 29(4):79:1–79:8, July 2010. ISSN 0730-0301.
- Qingtong Xu, Jing Wang, and Xuandong An. A pipeline for surface reconstruction of 3-dimensional point cloud. *Audio, Language and Image Processing (ICALIP), 2014 International Conference on*, pages 822–826, July 2014.
- Dong-Ming Yan and Peter Wonka. Adaptive maximal poisson-disk sampling on surfaces. *SIGGRAPH Asia 2012 Technical Briefs*, pages 21:1–21:4, 2012a.
- Dong-Ming Yan and Peter Wonka. Adaptive maximal poisson-disk sampling on surfaces. *SIGGRAPH Asia 2012 Technical Briefs*, page 21, 2012b.

- Dong-Ming Yan and Peter Wonka. Gap processing for adaptive maximal poisson-disk sampling. *ACM Trans. Graph.*, 32(5):148:1–148:15, October 2013. ISSN 0730-0301.
- Ji Yellott. Spectral consequences of photoreceptor sampling in the rhesus retina. *Science*, 221(4608):382–385, 1983. ISSN 0036-8075.
- Xiang Ying, Zhenhua Li, and Ying He. A parallel algorithm for improving the maximal property of poisson disk sampling in r2 and r3. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pages 179–179, 2013a.
- Xiang Ying, Shi-Qing Xin, Qian Sun, and Ying He. An intrinsic algorithm for parallel poisson disk sampling on arbitrary surfaces. *Visualization and Computer Graphics, IEEE Transactions on*, 19(9):1425–1437, Sept 2013b. ISSN 1077-2626.
- Xiang Ying, Zhenhua Li, and Ying He. A parallel algorithm for improving the maximal property of poisson disk sampling. *Computer-Aided Design*, 46(1):37 – 44, 2014. ISSN 0010-4485.
- Yahan Zhou, Haibin Huang, Li-Yi Wei, and Rui Wang. Point sampling with general noise spectrum. *ACM Trans. Graph.*, 31(4):76:1–76:11, July 2012. ISSN 0730-0301.