Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

Design of Computer Experiments: A Review

Sushant S Garud [1], Iftekhar A Karimi [1], Markus Kraft [2,3]

released: 31 March 2017

[1] Department of Chemical and Biomolecular Engineering
National University of Singapore
4 Engineering Drive 4
Singapore, 117576
Singapore
E-mail: cheiak@nus.edu.sg

[2] Department of Chemical Engineering and Biotechnology
University of Cambridge
New Museums Site, Pembroke Street
Cambridge, CB2 3RA
United Kingdom
E-mail: mk306@cam.ac.uk

[3] School of Chemical and Biomedical Engineering
Nanyang Technological University
62 Nanyang Drive
Singapore, 637459
Singapore

Preprint No. 182

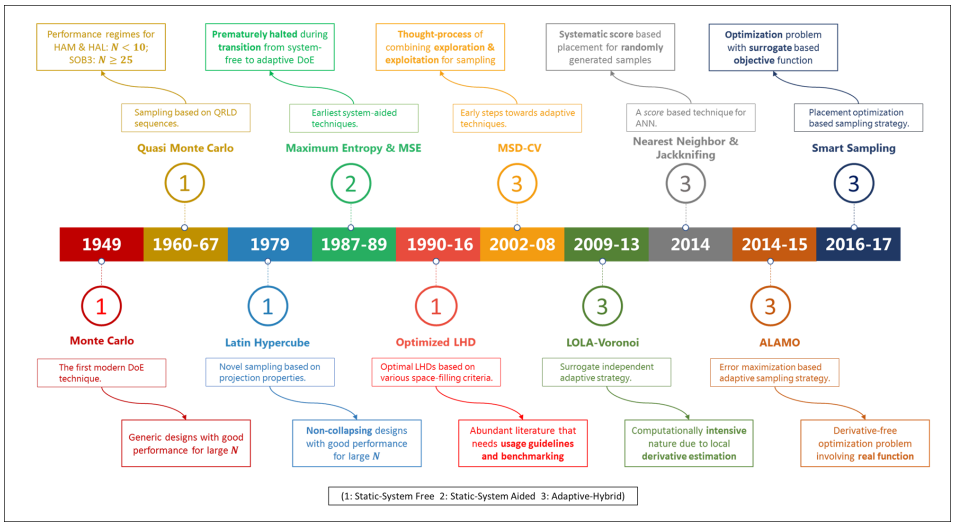
Keywords: Design of experiments, Computer experiments, Surrogate development, Adaptive sampling, Space-filling

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
New Museums Site
Pembroke Street
Cambridge CB2 3RA
United Kingdom

Fax: + 44 (0)1223 334796
E-Mail: c4e@cam.ac.uk
World Wide Web: http://como.ceb.cam.ac.uk/

Abstract

In this article, we present a detailed overview of the literature on design of computer experiments. We classify the existing literature broadly into two categories *viz.* static and adaptive design of experiments (DoE). We discuss the abundant literature available on static DoE, their chronological evolution, and their pros and cons. Our numerical and visual analyses reveal the excellent performance of Sobol sampling based on recent work of Joe and Kuo (SOB3) at higher dimensions while showing that Hammersley (HAM) and Halton (HAL) sampling are suited for lower dimensions. Our investigation of these techniques highlight the vital challenges that are dealt by adaptive DoE techniques, an upcoming class of modern DoE. They employ intelligent and iterative strategies that combine system knowledge and space-filling for sample placement. Adaptive DoE literature is critically analyzed based on the key features of their placement strategies. Finally, we provide several potential opportunities for future modern DoE research.



Highlights:

- Modern DoE techniques are comprehensively reviewed.
- A detailed classification and chronological evolution of modern DoE research is presented.
- Our numerical and visual analyses revealed the excellent high dimensional performance of SOB3.
- Rapidly growing class of adaptive DoE is critically discussed.
- Several potential opportunities for future research in modern DoE are discussed.

Contents

1 Introduction

The earliest footprints of design of experiments (DoE), a procedure to plan and define the conditions for carrying out controlled experimental trials, can be traced back to the era of the *Old Testament*. The first known example of DoE predates to 2^{nd} century B.C. in the first chapter of the *Book of Daniel* [156]. In this text, Daniel showed the superiority of his diet while using king's servants as a control group. In 150 A.D., Galen [170] discussed the importance and effect of sample size for medical studies. In this work, Galen presented a debate between an empiricist and a dogmatist discussing the experience and theory of medical research. Avicenna, the eminent Arabic researcher and philosopher, formally proposed many principles for designing the trials in the 2^{nd} volume of *Canon of Medicine* predating to the 11^{th} century [30]. These earlier efforts stemmed various theories and applications of DoE during the mid 17^{th} and late 18^{th} centuries [39, 68, 127]. In spite of these primal contributions, the work of Sir R. A. Fisher [57] on designing agricultural experiments is widely recognized as the first milestone in the field of DoE. Moreover, he proposed the "Fisher Information Matrix" [2, 56] which is used to reduce variance via the development of alphabetical optimal DoE [3]. Subsequently, various designs such as Full and Half Factorial [58], Central Composite, Plackett Burman [129], and Box Behken [15] were developed for real/physical experiments. These designs are commonly known as classical designs. While the classical DoE (a branch of DoE) essentially caters to physical experiments, the advent of computers has spawned a new branch of DoE, namely the modern DoE.

Researchers are increasingly replacing time-consuming and monetarily expensive physical experiments by faster and cheaper computer simulations. Often, computers enable experimentation that is not feasible in practice. In this paradigm, a computer code, commonly a high-fidelity simulator, generates data in lieu of real, physical systems. The primary aim of DoE in such cases is to decide the points at which the system behavior should be simulated. Although the classical DoE methods are well studied in the literature, their straightforward application to computer experiments is not appropriate due to the fundamental differences between physical and computer experiments. Most physical experiments are stochastic in nature due to a variety of unknown (hidden) and/or uncontrolled variables resulting in random errors. Thus, the classical DoE methods incur unavoidable randomness. On the other hand, computer experiments involving deterministic models are free of randomness. In addition, the most classical DoE methods typically assume a linear/quadratic approximation for the system response. To understand the impact of random errors in physical experiments, consider the following. The measured/observed response $y(x)$ in an experiment can be modeled as $y(x) = y_t(x) + \varepsilon$ where y_t is the true response, and ε represents random error. The primary aim of DoE is to derive the best possible approximation $\hat{y}(x)$ for $y_t(x)$ in spite of random error. Typically, when a linear/quadratic response model is assumed for a physical experiment, the vertices or points on the faces of the domain are the best sample points as explained by [69] and [117]. This can be directly inferred from the fundamental assumptions of the classical DoE. On the other hand, the modern DoE does not assume a linear/quadratic response. In this case, as explained by Giunta *et al.* [69] and Myers and Montgomery [117] the best sample points are those that are distributed within the domain. In other words, space-filling becomes a primary consideration for the modern DoE. This is also made necessary due to

the inherent mismatch between true and an approximate model. Hereafter, the term DoE implies the modern DoE unless explicitly mentioned. Besides classical and modern DoE, many researchers study complex systems based on semi-empirical models involving an amalgamation of computer and physical experiments. Such experimental designs are not considered in this article, however, interested readers can refer to [13, 112, 113, 115].

In 1996, Kohler and Owen published a chapter [96] that thoroughly reviewed DoE. Subsequently, Giunta *et al.* briefly introduced common DoE techniques in [69], however, it lacked a thorough discussion of these techniques and several of their variants. An article by Chen *et al.* [23] reviewed designs and modeling techniques for computer experiments in a statistical point of view. Levy and Steinberg presented a brief review on computer experiments, however, they devoted a very short section on DoE. Recently, Pronzato and Müller published a review article [130] aiming to detail advances beyond Kohler and Owen [96] and Chen *et al.* [23]. Although they discussed space filling techniques, they did not discuss the upcoming field of adaptive sampling (discussed later in Section 5). With this motivation, we capture the following key aspects in this article:

1. Elaborate classification and chronological evolution of DoE (see Section 3).
2. Comprehensive overview of research in DoE (see Sections 4 and 5).
3. Literature as well as numerical and visual analyses of the prominent DoE techniques (see Section 4.3).
4. Thorough discussion on adaptive sampling techniques (see Section 5).
5. Potential opportunities for further developments and the future of DoE (see Section 6).

This article is organized as follows. Section 2 presents the definitions and notations followed by various metrics to quantify space-filling. Section 3 presents a detailed classification and chronological evolution of DoE. Sections 4 and 5 discuss the static and adaptive DoE techniques respectively. Finally, in section 6, we conclude our discussion with future directions followed by a list of possible opportunities and unexplored fields of DoE.

2 Background

2.1 Definitions and Notations

Experiments, whether physical or computer, involve attributes and parameters that are varied to study the system response. These are called design/input variables or factors (commonly used by statisticians). Let $x = \{x_n \,|\, n = 1, 2, ..., N\} \in \mathbb{R}^N$ denote the N-dimensional vector of design variables. Typically, each design variable has user-specified bounds: $x_n^L \leq x_n \leq x_n^U$. The space defined by these bounds, namely $\mathcal{D} : x^L \leq x \leq x^U$, is called the domain. This is typically scaled as $[-1, 1]^N$ or $[0, 1]^N$ to avoid numerical ill-conditioning [141]. Throughout this article, we use $[0, 1]^N$ normalized domain space.

A sample or sample point is a specific instance of x in the domain. The collection of sample points, $\mathcal{X}_N^{(K)} = \{x^{(k)} \,|\, k = 1, 2, ..., K\}$ is a sample set of size K. Let the system response at $x^{(k)}$ be described by M output variables as $y = \{y_s \,|\, s = 1, 2, ..., S\}$. The collection of all such responses is a response set, $\mathcal{Y}_S^{(K)} = \{y^{(k)} \,|\, k = 1, 2, ..., K\}$. K samples and the respective responses are subsequently used to build an approximation, $\tilde{f}^{(K)}(x)$, $\tilde{f}^{(K)} : \mathcal{D} \to \mathbb{R}^S$ for the system response surface. This is called as surrogate/meta model. The literature has various surrogate modeling techniques like Polynomial, Kriging, Radial Basis Functions (RBF), Artificial Neural Networks (ANN), Support Vector Machine (SVM) *etc.* [145, 149].

2.2 Space-Filling Criteria

As discussed earlier in Section 1, the key aim of the DoE is to generate sample points for filling the domain. This requires metrics that can quantify the space-filling ability of any given sample set. Several space-filling criteria (SFC) have been proposed in the literature and there are two broad categories *viz.* (a) uniformity-based (b) distance-based.

2.2.1 Uniformity-based SFC

Discrepancy quantifies the departure of a given sample set/design from a uniform design. For this, a uniform design is defined as the one where the number of sample points in a subspace $\Delta\mathcal{D}$ of the domain $\mathcal{D}$ is proportional to the hyper-volume $V(\Delta\mathcal{D}) = \Delta x_1 \times \Delta x_2 \times ... \times \Delta x_N$ of the subspace [78].

A discrepancy that measures the maximum departure for this number for a given sample set is known as Star discrepancy given as follows.

$$D^*(\mathcal{X}_N^{(K)}) = \sup_{x \in \mathcal{D}} \left| \frac{1}{K} \#\{x^{(k)} \in \Delta\mathcal{D}\} - \prod_{n=1}^{N} \Delta x_j \right| \tag{1}$$

where $x^\intercal = (x_1, x_2, ..., x_N)$ and $\#\{x^{(k)} \in \Delta\mathcal{D}\}$ is the number of samples in $\Delta\mathcal{D}$. A modification of star discrepancy is L_2 discrepancy where L_2 norm of the departure is used instead of the absolute departure (L_∞ norm) [53, 103] as shown below.

$$D_2^*(\mathcal{X}_N^{(K)}) = \left\{ \int_{\mathcal{D}} \left[\frac{1}{K} \#\{x^{(k)} \in \mathcal{D}\} - \prod_{n=1}^{N} x_j \right]^2 \mathrm{d}x \right\}^{\frac{1}{2}} \tag{2}$$

The discrepancies based on L_2 norm are very popular due to ease of calculations and the availability of their closed form expressions. Two commonly used variations of L_2

discrepancy are Centered L_2 (Eq. (3)) and Wrap-around L_2 (Eq. (4)) [78, 79].

$$C_2(\mathcal{X}_N^{(K)}) = \left(\frac{13}{12}\right)^N - \frac{2}{K}\sum_{k=1}^{K}\prod_{n=1}^{N}\left(1+\frac{1}{2}|x_n^{(k)}-0.5| - \frac{1}{2}|x_n^{(k)}-0.5|^2\right) + \frac{1}{K^2}\sum_{j,k=1}^{K}\prod_{n=1}^{N}\left(1+\frac{1}{2}|x_n^{(k)}-0.5| + \frac{1}{2}|x_n^{(j)}-0.5| - \frac{1}{2}|x_n^{(k)}-x_n^{(j)}|\right) \quad (3)$$

$$W_2(\mathcal{X}_N^{(K)}) = \left(\frac{4}{3}\right)^N - \frac{1}{K^2}\sum_{j,k=1}^{K}\prod_{n=1}^{N}\left(\frac{3}{2} + |x_n^{(k)}-x_n^{(j)}| \times (1-|x_n^{(k)}-x_n^{(j)}|)\right) \quad (4)$$

These are a few commonly used types of discrepancies, however, the literature discusses a variety of discrepancies and their properties in terms of uniformity and projections of samples [52, 53, 83].

In information theory, Kullback-Leibler information (Eq. (5)) quantifies the difference between two density functions f and g [100, 101]. Consider $x^{(1)}, x^{(2)}, ..., x^{(K)}$ as K independent observations of the random x with density function f over the unit cube $[0,1]^N$. In this case, we call f as "*design*" density while g as "*target*" density.

$$I_{KL}(f,g) = \int_E f(x)\ln\left(\frac{f(x)}{g(x)}\right)\mathrm{d}x \quad (5)$$

When target density, g, is uniform, Eq. (5) reduces to Eq. (6).

$$I_{KL}(f) = \int_E f(x)\ln\left(f(x)\right)\mathrm{d}x \quad (6)$$

Minimizing $I_{KL}(f)$ in Eq. (6) gives a design that is close to uniform. While the discrepancies and information theory mainly focus on the distribution of sample points within a domain, the distance-based criteria consider the inter-point or inter-sample distances to quantify space-filling as discussed next.

2.2.2 Distance-based SFC

Audze and Eglajs proposed a distance based criterion for space filling [43] which is also known as potential energy criterion given in Eq. (7).

$$PE(\mathcal{X}_N^{(K)}) = \sum_{k=1}^{K}\sum_{j=k+1}^{K}\frac{1}{d(x^{(k)},x^{(j)})^2} \quad (7)$$

where $d(x^{(k)}, x^{(j)})$ is the Euclidean distance between points $x^{(k)}$ and $x^{(j)}$. By minimizing PE an evenly spread design can be obtained. This criterion has been widely used for constructing space-filling designs especially the optimal Latin Hypercube Designs (LHDs) discussed later.

Minimum Spanning Tree (MST) was proposed by Dussert *et al.* for studying orders and disorders [40]. MST uses samples as vertices to construct a tree by connecting them and minimizing the sum of edge lengths. Once the MST-based design is developed, the mean edge-length μ_e and standard deviation σ_e are computed. The designs with large μ_e and small σ_e, known as *quasi-periodic* designs, perform well in terms of space-filling since large μ_e implies large inter-point distance and small σ_e means low variations in inter-sample distances. Therefore, MST-based designs, say D_1 and D_2, can be partially ordered as follows: if $\mu_e(D_1) > \mu_e(D_2)$ and $\sigma_e(D_1) < \sigma_e(D_2)$, then D_1 may be better than D_2 in terms of space-filling.

Johnson *et al.* [88] proposed the two distance-based criteria, namely *maximin* (Mm) and *minimax* (mM) to spread sample points within the domain. The maximin criterion maximizes the minimum distance between two sample points. This can be given mathematically as shown in Eq. (8).

$$Mm(\mathcal{X}_N^{(K)}) = \max_{x \in \mathcal{D}} \left[\min_{j \neq k} \left[d(x^{(j)}, x^{(k)}) \right] \right] \tag{8}$$

On the other hand, the *minimax* criterion minimizes the *maximin* distance between two

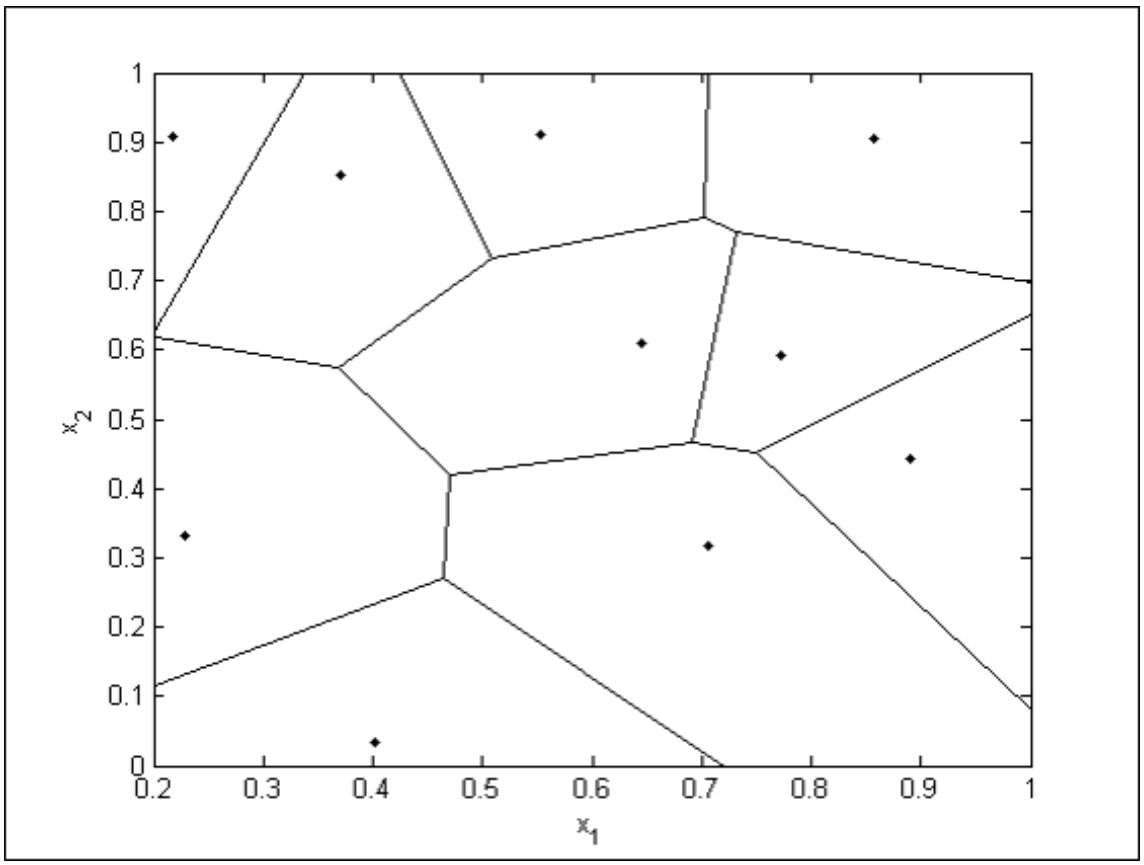


Figure 1: *Illustration of Voronoi diagram in* 2 *dimensional domain for* $K = 10$*. (Developed in Matlab using in built Voronoi function)*

points and can be given by Eq. (9).

$$mM(\mathcal{X}_N^{(K)}) = \min\left[\max_{x\in\mathcal{D}}\left[\min_{j\neq k}\left[d(x^{(j)}, x^{(k)})\right]\right]\right] \tag{9}$$

Strictly speaking, Eq. (9) denotes "*minimaximin*" design. However, for the sake of convenience, it is generally called as just *minimax*.

Morris and Mitchell proposed ϕ_p criterion [111] that has become a popular space-filling criterion.

$$\phi_p(\mathcal{X}_N^{(K)}) = \left[\sum_{k=1}^{K-1}\sum_{j=k+1}^{K}\left(d(x^{(j)}, x^{(k)})\right)^{-p}\right]^{\frac{1}{p}} \tag{10}$$

Minimizing the ϕ_p maximizes the point-to-point distance, hence, the better sample spreading.

Apart from the SFCs discussed above, there are space-filling criteria like Delaunay Triangulation [37] and Voronoi Tessellations/Diagram [168] which are geometrical in nature, however, they implicitly incorporate Euclidean distance in their definition as follows. Consider a set of points $\mathcal{X}_N^{(K)} = \left\{x^{(1)}, x^{(2)}, ..., x^{(K)}\right\}$. For every sample $x^{(k)}, k =$

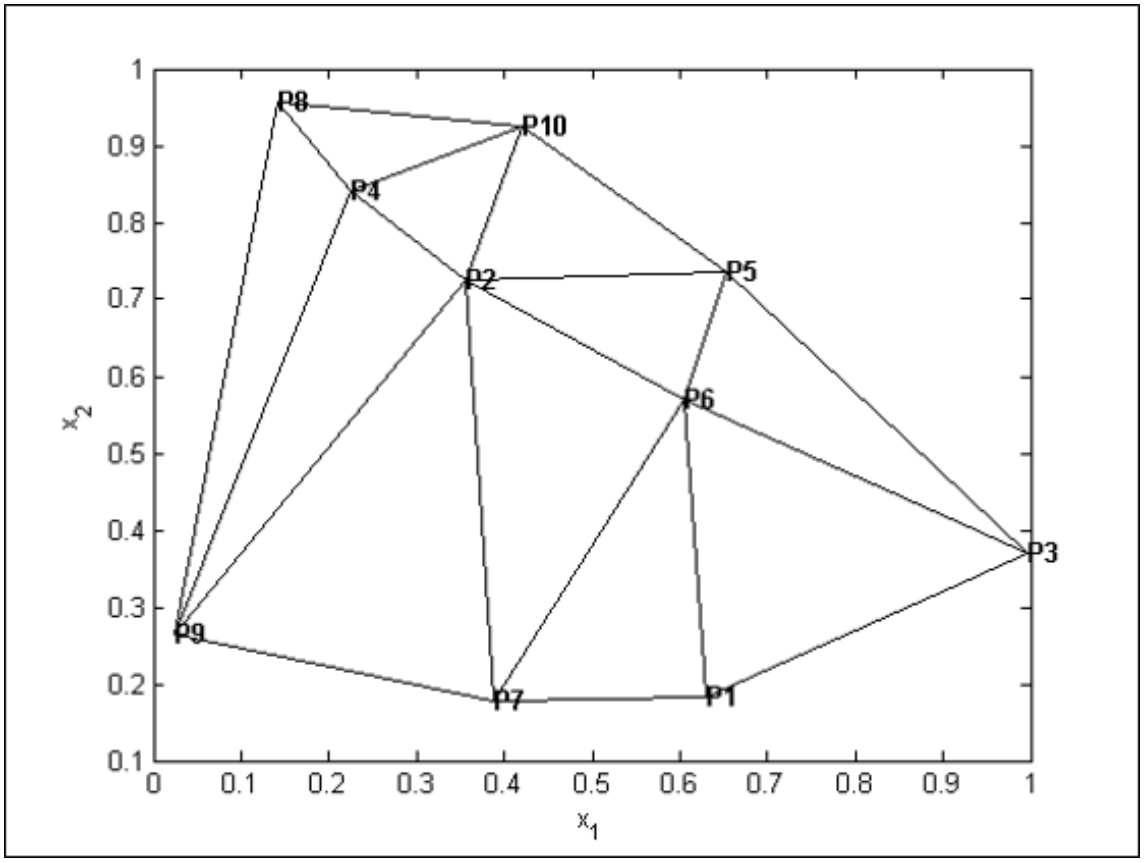


Figure 2: *Illustration of Delaunay triangulation in 2 dimensional domain for $K = 10$. (Developed in Matlab using in built Delaunay function)*

$1, 2, ..., K$, it has a corresponding Voronoi cell $V^{(k)}$ given as shown in Eq. (11).

$$V^{(k)} = \{x^{(k)} \in \mathcal{X}_N^{(K)} | d(x, x^{(k)}) \leq d(x, x^{(j)}) \, \forall \, j \neq k\} \tag{11}$$

Figure 1 illustrates a simple Voronoi diagram construction based on 10 randomly generated sample points in 2 dimensional domain shown with black dots. Each sample point has a surrounding cell (V_k) bounded by the boundary shown with the solid black line in Figure 1. Qualitatively, Voronoi diagram aims to fill the space uniformly by placing samples based on cell ($V^{(k)}$) size [41, 167]. The larger the Voronoi cell size, the more unexplored the region. Hence, samples can be placed in large Voronoi cells to enhance homogeneity of the sample spreading.

Delaunay triangulation is the dual of Voronoi diagram. Delaunay triangulation in 2 dimensional domain for a set of K points $\mathcal{X}_N^{(K)}$ is a triangulation $DT(\mathcal{X}_N^{(K)})$ such that no point in $\mathcal{X}_N^{(K)}$ is inside the circumcirle of any triangle formed by points in $\mathcal{X}_N^{(K)}$. In other words, this aims to minimize the maximum angle for all triangles in the triangulation. Figure 2 shows an illustration of Delaunay triangulation in 2 dimensional domain for 10 points $\mathrm{P}_1, \mathrm{P}_2, ..., \mathrm{P}_{10}$. This can be generalized to N dimensional case by generalizing the concept of triangle to simplex and circumcircle of a triangle to circum-hypersphere of a simplex [37]. For further details on Voronoi Diagram and Delaunay triangulation, readers may refer to Aurenhammer *et al.* [7]. Most of these criteria are often used to analyze as well as enhance the performance of various DoE techniques.

3 Evolution and Classification of DoE

Revolutionary developments in the fields of computers and associated technologies have motivated both academic scholars and industrial personnel to opt for computer experi-

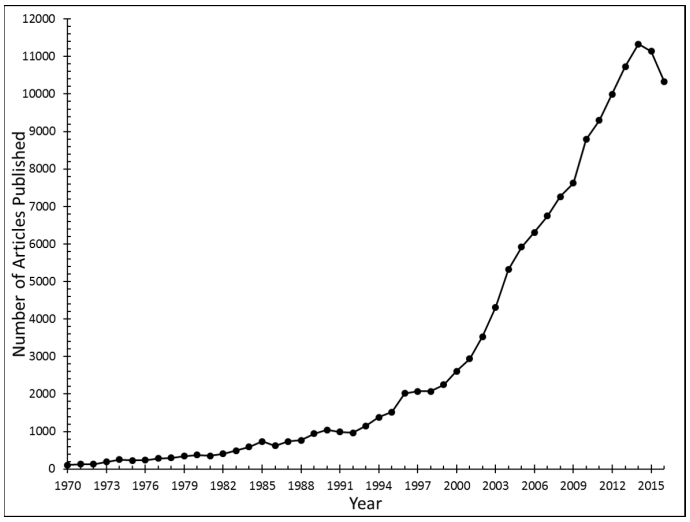


Figure 3: *Number of articles published on DoE over the years. (Retrieved from Scopus on* 31 *January* 2017*; Search: "design of experiments" OR "experimental design" OR sampling AND computer)*

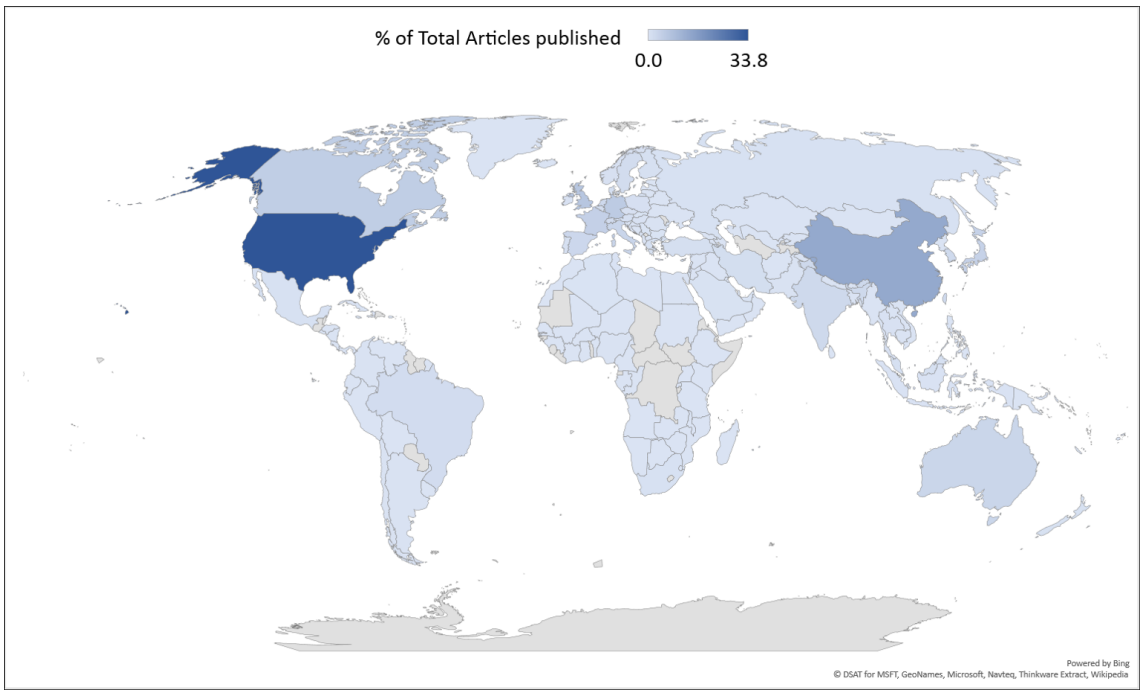


Figure 4: *Geographical distribution of articles published on DoE. (Retrieved from Scopus on* 31 *January* 2017*; Search: "design of experiments" OR "experimental design" OR sampling AND computer. Countries colored in black do not have any contribution.)*

ments. Computer experiments have proven to be essential for a variety of applications in the fields of molecular physics, product design, electronics and communications, process design and operation, automobiles, aeronautics, structures, and so on. Moreover, the scale of applications ranges widely from nano (molecular simulations) to mega scale (structural analysis of buildings and bridges). This inspired various researchers to work on the design of computer experiments and its applications which is evident from Figure 3. It shows the trend of approximate number of articles published in the field of design of experiment over the years. Though the earliest research in the field of modern DoE can be traced back to late 1940s, it attracted significant attention in late 1970s. Hence, Figure 3 provides a trend of number of articles published (> 100) from 1970 onwards. Thenceforth, the field is booming which is evident from a strong growth trend in Figure 3.

Figure 4 describes the overview of DoE research in the different parts of the world. The highest contribution is by the United States of America (33.8%) followed by China (13.9%), the United Kingdom (7.0%), Germany (5.7%) and Canada (5.1%). All the other countries contribute less than 5%. The literature consists of approximately $149,000$ articles that can be further categorized based on the fields of publication shown in Figure 5. Out of the total literature on DoE, engineering contributes 34%, computer science contributes 31%, medicine contributes 17%, mathematics contributes 14%, and rest of the fields contribute 10% or less.

Regardless of the field of application, any DoE technique is employed for one or more of the following four major objectives [96]:

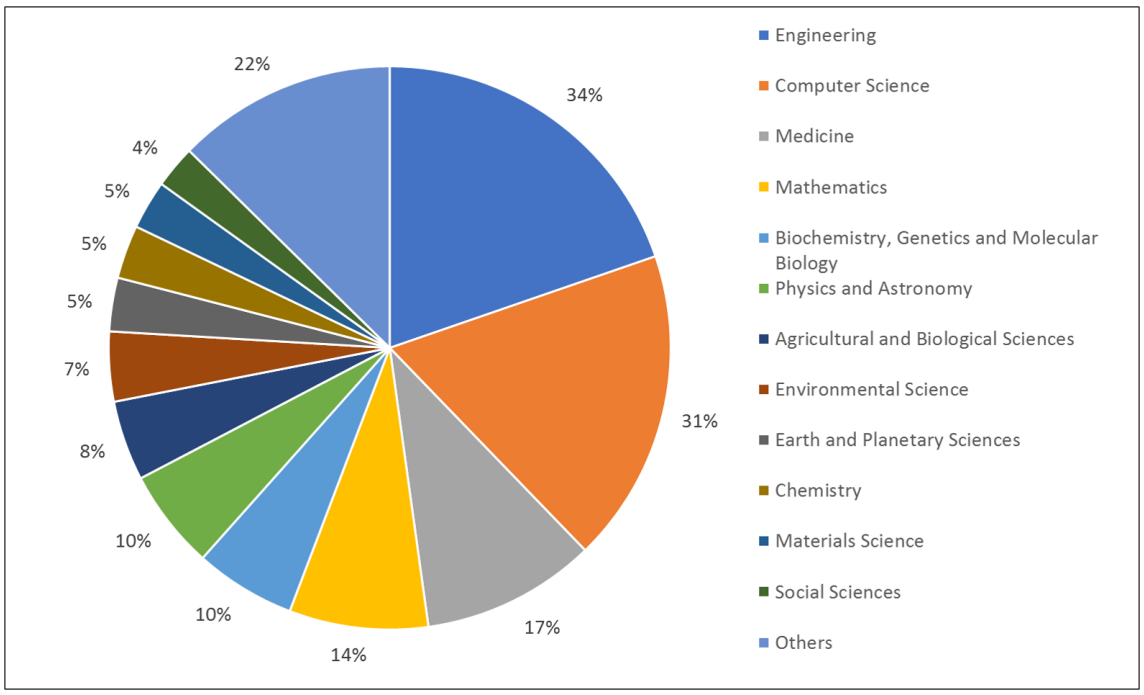


Figure 5: *Distribution of articles published on DoE in various fields. (Retrieved from Scopus on* 31 *January* 2017*; Search: "design of experiments" OR "experimental design" OR sampling AND computer. Others include Chemical Engineering, Decision Sciences, Energy, Neuroscience, Pharmacology, Toxicology and Pharmaceutics, Health Professions, Psychology, Business, Management and Accounting, Immunology and Microbiology, Arts and Humanities, Nursing, Multidisciplinary, Economics, Econometrics and Finance, Veterinary, Dentistry, and Undefined)*

1. System Approximation/Prediction
2. System Optimization
3. System Visualization
4. Numerical Integration

The historical development of the DoE points to a natural evolution with two major categories, namely Static and Adaptive. The earliest DoE techniques ignored the system under study, and focused on the spatial distribution of sample points in a domain ($\mathcal{D}$). In that sense, these designs were generic and fixed. Once generated, they could be repeatedly used for any system. Such techniques come under a subcategory of system-free techniques. Over time, researchers realized the obvious shortcomings of this generic approach, and began incorporating system knowledge to tailor the DoE for a given system. These techniques were also static and come under a subcategory of system-aided techniques. Recent research focus has moved towards a more flexible and dynamic approach in which system knowledge is integrated and used in an incremental or progressive manner during sample generation. This is called adaptive or sequential approach and offers

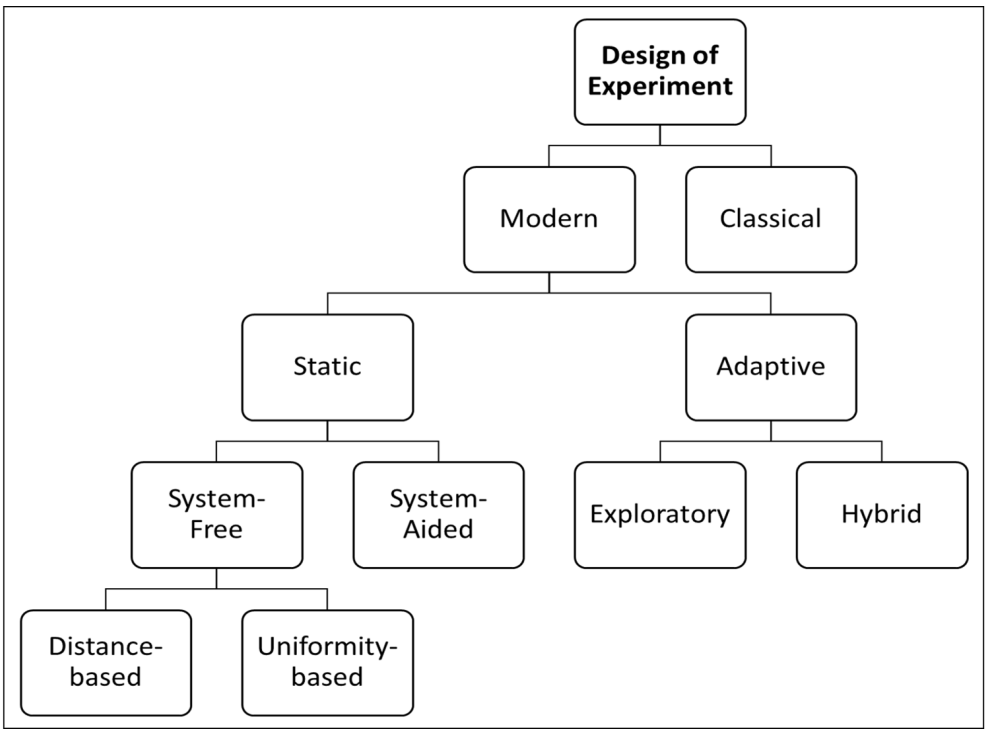


Figure 6: *Flowchart describing the classification of DoE.*

the hope of getting the best design with the smallest sample set. Figure 6 shows the classification chart of DoE and we now discuss each class in detail.

4 Static DoE

4.1 System-Free DoE

The early DoE methods were largely inspired by the classical DoE literature. Since the latter dealt with experimental systems whose characteristics were unknown and complex, they neglected the system knowledge altogether. Therefore, early DoE merely aimed to fill the domain as uniformly as possible. This is most easily achieved by randomly placing sample points in the domain. Thus, the Monte Carlo sampling was the first formal method for the DoE.

4.1.1 Monte Carlo and Stratified Monte Carlo Sampling

Classical Monte Carlo Sampling simply known as Monte Carlo Sampling (MCS) or pseudo-random sampling was proposed by Metropolis and Ulam in 1949 [110]. MCS uses pseudo-random numbers to generate K sample points hoping that its random actions will result in space-filling. This has motivated efforts to develop a variety of pseudo-random number generators [66]. The finite sample size and the shortcomings of pseudo-

random number generators may result in clustering and unrepresented regions within the domain. To overcome these difficulties, researchers proposed the idea of Stratified Monte Carlo Sampling (SMCS) where the space-filling is achieved by dividing the domain space into non-random strata and employing MCS for each stratum. Note that, although MCS and SMCS were initially developed for multidimensional numerical integration, they have a wide range of applications in designing computer experiments [108]. MCS and SMCS have been studied thoroughly by various researchers from the fields of mathematics, statistics, sciences, and engineering. This has resulted in an abundant literature [44, 66, 67, 94, 108, 136, 137, 154]. It is clear that SMCS, by introducing strata, attempts to infuse an element of deterministic design in the purely chaotic MCS framework. Subsequent DoE techniques exploited this idea further which will be evident in the following sections.

4.1.2 Quasi-Monte Carlo Sampling

Quasi-Monte Carlo Sampling (QMCS) uses a *quasi-random low discrepancy* (QRLD) sequence for sample generation. The term quasi-random implies a deterministic nature of the sequence, while the term low discrepancy implies its nearness to a uniform distribution of sample points in the domain. According to the Koksma-Hlawka inequality [80, 118], the approximation error is upper-bounded by the star discrepancy discussed earlier (Eq. 1). Moreover, this error bound is valid independent of the dimensions (N) of the underlying domain. Hence, deterministic sequences with low star discrepancy are attractive for sampling. Several QRLD sequences like Halton, Hammersley, Sobol, and their variations exist in the literature. We now discuss these sequences, their implementations and their key features.

QRLD sequences use the concepts of *inverse radix number* and prime numbers to generate a deterministic sequence. For this, let us first understand *inverse radix number*. Any given integer I can be written in radix (base) R (where R is an positive integer) as follows:

$$I = I_0 + I_1 R + I_2 R^2 + ... + I_Q R^Q \tag{12}$$

where $I \equiv I_Q I_{Q-1} ... I_1 I_0$, $\mathcal{Q} = \{1, 2, ..., Q\}$ and $\mathcal{Q} \subseteq \mathbb{N}$. Then a unique fraction between 0 and 1 is called the *inverse radix number* and can be generated by reversing the order of the digits of I as follows:

$$\theta_R(I) = .I_0 I_1 I_2 ... I_q \tag{13}$$

$$\theta_R(I) = I_0 R^{-1} + I_1 R^{-2} + ... + I_q R^{-q-1} \tag{14}$$

In 1935, van Der Corput proposed the following 1-D sequence [163]:

$$\mathcal{S} = \{\theta_R(k) \,|\, k = 1, 2, ..., K\} \tag{15}$$

Thus, for $R = 10$ (decimal base), Eq. (15) gives

$$\mathcal{S} = \left\{ \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10}, \frac{5}{10}, \frac{6}{10}, \frac{7}{10}, \frac{8}{10}, \frac{9}{10}, \frac{1}{100}, \frac{11}{100}, \frac{21}{100}, ... \right\} \tag{16}$$

This sequence generation philosophy inspired scholars like Hammersley, Halton and Sobol to propose more elegant QRLDs.

In 1960, J. M. Hammersley proposed a set of sample points called Hammersley points, also known as Hammersley sequence, to solve multidimensional Monte Carlo integration [74]. This is one of the first high dimensional QRLD sequences proposed to tackle the issues associated with MCS and SMCS. The Hammersley points/samples in an N dimensional space are given by following sequence:

$$z_N^{(k)} = \left(\frac{k}{K}, \theta_2(k), \theta_3(k), ..., \theta_{R_{N-1}}(k) \right), \; k = 1, 2, ..., K \tag{17}$$

where $2, 3, ..., R_{N-1}$ are the first $N-1$ prime numbers. The Hammersley points are $x_N^{(k)} = 1 - z_N^{(k)}$ [91].

Hammersley sequence has been employed by many scholars [32, 38, 61, 76, 91, 93, 102, 119, 171, 172] for a variety of applications. In spite of the popularity, it has some limitations when it comes to implementation flexibility. A user need to specify the sample size (K) a priori in order to generate a Hammersley sequence (Eq. (17)). Moreover, as the dimensions of the domain increase, the largest base (R_{N-1}) used for sequence generation increases, resulting in slower space filling in some dimensions. Thus, for high dimensional domains Hammersley sequence does not fill the space homogeneously in all the dimensions.

Inspired by Hammersley sequence, Halton proposed a QRLD sequence that can also be viewed as an N dimensional generalization of van Der Coput sequence. He suggested various modifications to the Hammersley sequence to ameliorate its demerits. First, he removed the dependence on sample size K (Eq. (17)) by rewriting the sequence as follows [19].

$$\mathcal{S} = \{(\theta_{R_1}(k), \theta_{R_2}(k), ..., \theta_{R_N}(k)) \,|\, k = 1, 2, ..., K\} \tag{18}$$

Second, he showed that radices should only be mutually prime or co-prime in contrast to Hammersley sequence where radices are strictly prime numbers. Thus, Halton sequence tries avoid large bases and hence, slow space-filling. For further details on Halton sequence readers may refer to [73, 120].

Although Halton sequence is easy to construct and practically implementable, it faces serious limitations for $N \geq 14$. Thus, in practice Halton sequence is avoided for $N \geq 8$. Moreover, in some cases there might be a linear correlation between the generated samples of the sequence [99]. Thus, it is a common practice to drop the first K^* samples from the sequence to assure uncorrelated samples [62]. To address these issues with Halton sequence, researchers like Braaten and Weller [17], Hellekalek [77], Faure [55], Krommer and Ueberhuber [98], and Kocis and Whiten [95] proposed several variations and modifications of them. Many researchers have employed Halton sequence in a variety of studies [33, 59, 60, 72, 121, 161].

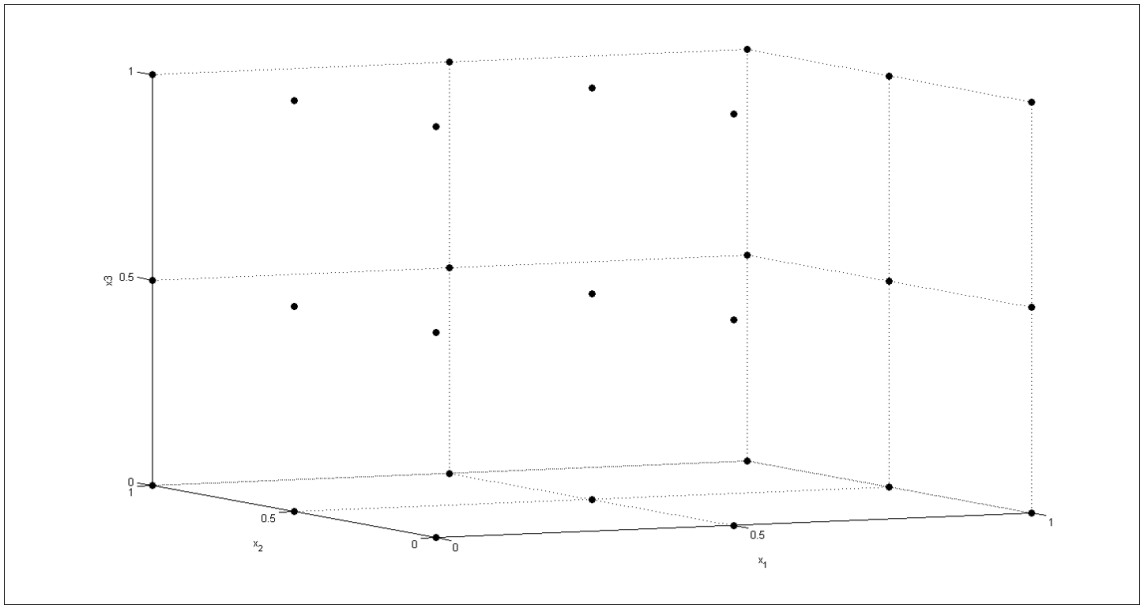
In 1967, Sobol proposed a QRLD sequence to approximate high dimensional integrals [152]. It employs radix 2 to form finer partitions and achieves faster space-filling over other sequences. Moreover, it uses a concept of *direction numbers* to generate a sequence. Sobol sequence is a vast and very well studied topic [18, 86, 120] and we do

not discuss the sequence generation here, however, reader can refer to [18, 86] for understanding algorithm implementation. Apart from fast and better space-filling, Sobol introduced additional conditions for the uniformity of sample distribution (properties **A** and **A'**) [153]. Thus, Sobol sequence can be constructed to satisfy either one or both of these properties along with the discrepancy criterion giving them an edge over the other QRLD sequences. The literature is replete with the articles that employ Sobol sequence for various applications [9, 20, 63, 114, 148].

Note that although these methods inherently aim to space-fill, they do not incorporate formal quantification of space-filling during the placement. Thus, various deterministic methods based on a variety of space-filling criteria (discussed in Sections 2.2.1 and 2.2.2) have been proposed to complement QMCS. They typically use a space-filling criterion as an objective function in their placement optimization problem.

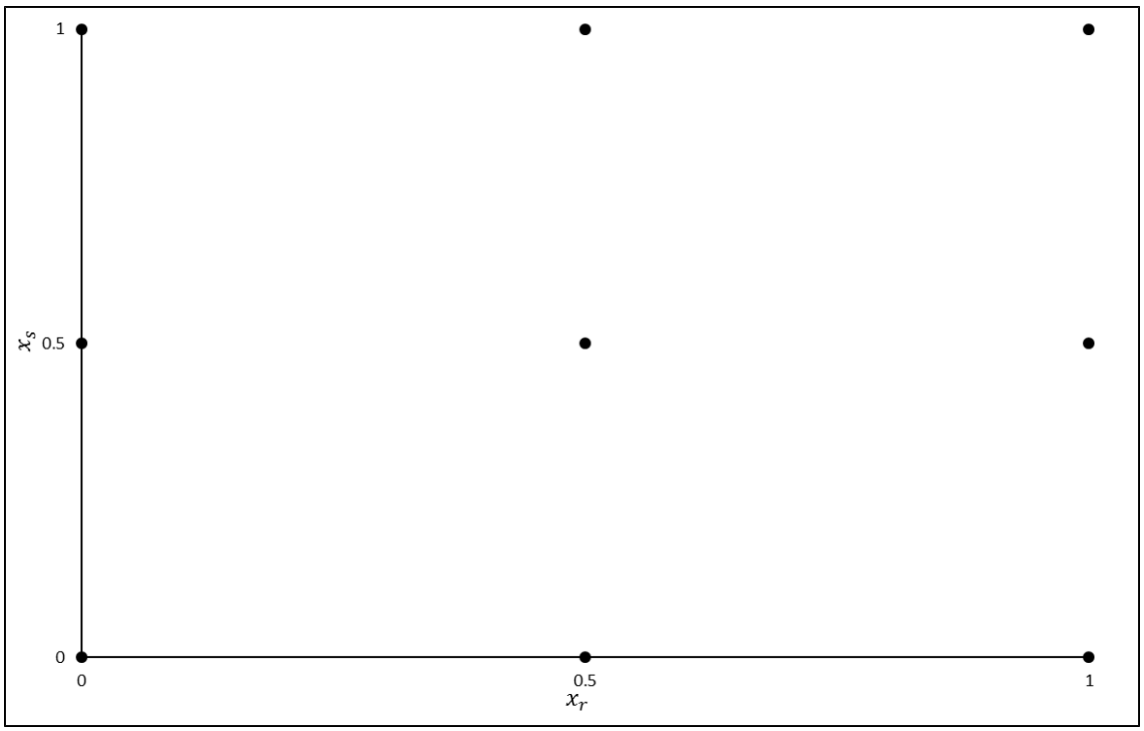
4.1.3 SFC-based Sampling

The simplest method in this category constructs a fully deterministic, symmetric and systematic grid of sample points using the Cartesian co-ordinate system. The sample size is manageable for small N values (typically ≤ 3) but it explodes exponentially as N increases. Consider an N dimensional grid design with K sample points in each dimension. This amounts to K^N samples. When this grid is projected on to any one of the dimensions, it results in a grid of $(N-1)$ dimensions with K sample points in each dimension.



(a) 3 *dimensional grid with* 3 *sample points per dimension.*

Figure 7: *Illustration of grid design in* 3 *dimensional domain and its projections in* 2 *dimensional plane. (Figure 7a shows* 27 *sample points in* 3 *dimensional space and Figure 7b shows a* 2 *dimensional projection of Figure 7a on* x_t *plane for* $r = \{1, 2, 3\}$*,* $s = \{2, 3, 1\}$*, and* $t = \{3, 1, 2\}$*)*



(b) 2 *dimensional projection on* x_t.

Figure 7: *Illustration of grid design in* 3 *dimensional domain and its projections in* 2 *dimensional plane. (Figure 7a shows* 27 *sample points in* 3 *dimensional space and Figure 7b shows a* 2 *dimensional projection of Figure 7a on* x_t *plane for* $r = \{1, 2, 3\}$*,* $s = \{2, 3, 1\}$*, and* $t = \{3, 1, 2\}$*)*

This can be viewed as $K^N - K^{N-1}$ wasteful computations due to the assumption of equi-spaced samples. Let us understand this with a simple illustration shown in Figure 7. Figure 7a shows a $3 \times 3 \times 3$ grid design. This amounts to 3^3 *i.e.* 27 samples points.

Figures 7b shows the projection of this 3 dimensional grid on x_t plane for $t = \{3, 1, 2\}$. In each case, it results in 9 sample points. This clearly shows 18 wasteful sample points in each case. The projection property is a key aspect of DoE and its importance will be clearer in the discussion of LHDs in Section 4.1.4. This realization led to a series of methods that aimed to avoid such duplication namely, Latin Hypercube Designs, Good Lattice Sampling, *etc*. The simplest modification of grid that overcomes the above discussed issue is good lattice point sampling (GLS). For the N dimensional space, the good lattice points are given by

$$x_n^{(k)} = \left\{ \frac{2k(n + h_n) - 1}{2K} \right\} \quad n = 1, 2, ..., N \text{ and } k = 1, 2, ..., K \tag{19}$$

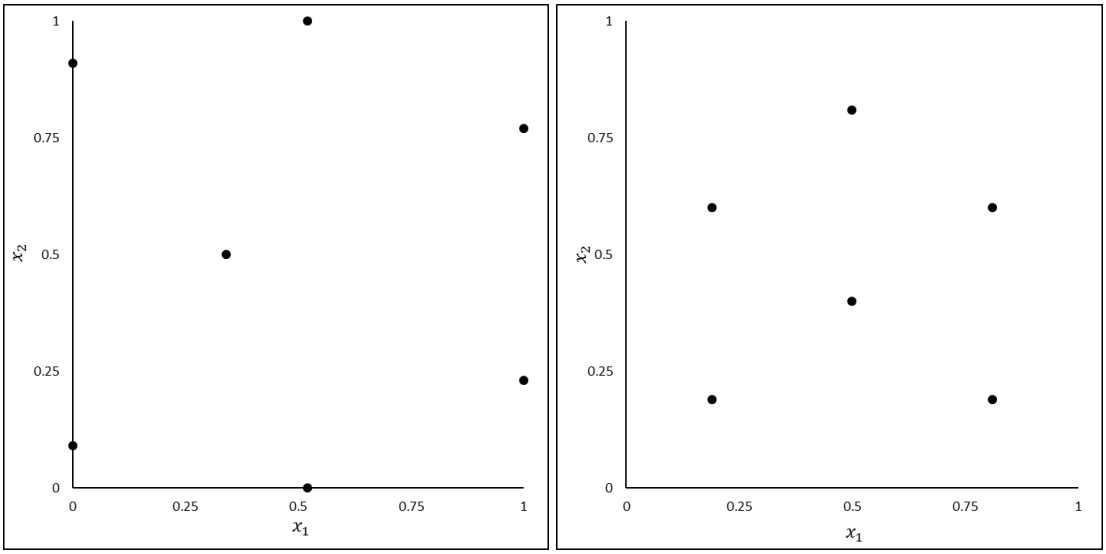
where h_n is any integer between 0 and $K - N$ and $\{x\}$ is a fractional part of x. Thus, $(K - N + 1)^N$ distinct GLS designs are possible with $K - N + 1$ values of h_n for every n. For $h_n = N - n$, GLS results in a uniform grid. The GLS corresponding to $N = 2$, $K = 34$, $h_1 = 0$, and $h_2 = 19$ is known as Fibonacci lattice [151]. Several complex GLS designs can be found in [50]. Although the GLS works better than a Cartesian

grid, it still suffers from the *curse of dimensionality*. These two sampling methods do not explicitly incorporate space-filling criterion, however, they implicitly consider space-filling by placing geometrically equi-distant samples.

In 1990, Johnson *et al.* proposed two distance-based designs namely maximin and minimax [88]. In case of the maximin design, the smallest distance between any two samples is maximized as shown in Eq.(8). In 3 dimensional space, this can be viewed as filling the unit cube ($\mathcal{D}$) with K non-intersecting rigid spheres with centers as sample sites located within $\mathcal{D}$ [65]. Similarly, minimax designs minimize the maximin criterion for sample placement as given in Eq.(9). Figure 8a shows the maximin design (obtained from `http://www.packomania.com/`) and figure 8b shows the minimax design [88] for the 2 dimensional case.

These designs can be qualitatively explained by using an interesting practical example given by Pronzato and Müller [130]. A maximin design can be viewed as a placement of tables in a restaurant such that chance of hearing the talks from the neighboring table by a party at a given table is minimal. On the other hand, a minimax design can be viewed as a placement of tables in a restaurant such that a waiter is as close as possible to a table when he is in the restaurant. A large compilation of maximin and minimax designs for various settings can be found at `http://www.packomania.com/`. Moreover, interested reader may refer to [4, 36, 70, 126, 142, 143, 157, 160] for further insights and applications.

Since 1980s uniform design (UD) *i.e.* design based on sample placements over the domain minimizing the uniformity criterion, has become very popular [45]. Typically, UD is constructed using star discrepancy (Eq.(1)) discussed earlier in section 2.2.1. Generation of UD based on D^* for 1-D case is straight forward and is given as $\mathcal{X}_1^{(K)} = \left\{\frac{1}{2K}, \frac{3}{2K}, ..., \frac{2K-1}{2K}\right\}$ [50]. However, this paradigm becomes arduous as the dimensions increase [51]. Although D^* is a very commonly used discrepancy (uniformity measure),



(a) *Maximin Design* **(b)** *Minimax Design*

Figure 8: *Illustration of Maximin and Minimax Designs in 2 dimensional domain.*

there are various designs that uses variation of star discrepancy such as D_2^* (Eq.(2)), C_2 (Eq.(3)) and W_2 (Eq.(4)). For further details, reader can refer to [46–49, 52].

4.1.4 Latin Hypercube Design

Latin Hypercube Design (LHD) also known as Latin Hypercube Sampling (LHS) developed by McKay *et al.* [109] is one of the most popular DoE techniques to overcome the issues associated with the MCS and its variations. Although the article by McKay *et al.* is considered as a milestone in the field of DoE, its underlying concepts were existing for a long time. It is an extension of the quota sampling proposed by Steinberg in 1963 [155] and a work inspired by Latin Square Sampling [133]. Let us look at the LHD construction. Consider an N-dimensional design space $[0,1]^N$. Divide each dimension into K equal bins of edge length $\frac{1}{K}$. This results in K^N hypercubes. Arrange K sample points as a $K \times N$ matrix $\boldsymbol{L} = [x^{(1)}, x^{(2)}, ..., x^{(K)}]^\top$, where each column represents a variable and each row represents a sample point. Then, $\boldsymbol{L}$ is an N dimensional LHD of size K, if for each column of $\boldsymbol{L}$, no two elements in that column fall in the same bin. For instance, consider an LHD for $N = 2$ and $K = 10$ as shown in Figure 9. Both dimensions have 10 equal bins. It is clear that the samples are placed such that no row or column has more than one sample point. The above condition on the placement of elements in each column of $\boldsymbol{L}$ is also known as non-collapsing design condition. This condition gives LHD the following special characteristics. First, any LHD in N dimensional domain reduces to an LHD in $N - N'$ dimensions when N' dimensions are removed from the former. Second, this also allows LHD to perform uniformly well over a range of dimensions. However, the LHD configuration and the sample placement within bins are still random. This is precisely the reason why LHD may not guarantee adequate space-filling. For instance, Figure 10 is a well-known example of the worst possible LHD. This necessitates enhancing the ability of LHD for space-filling. Therefore, researchers in early 1990s began optimiz-

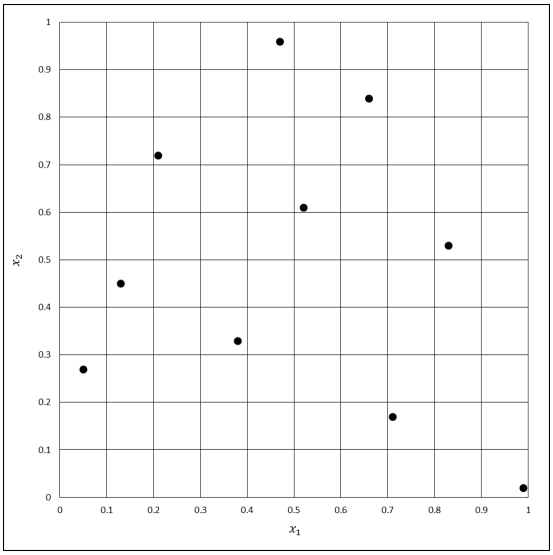


Figure 9: *Illustration of LHD for $K = 10$ and $N = 2$.*

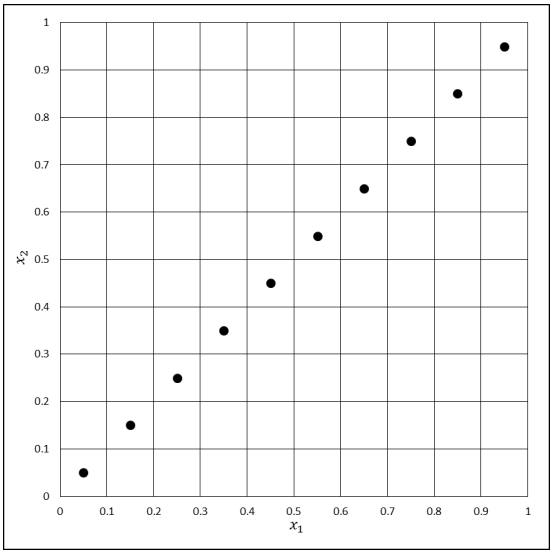


Figure 10: *Worst case of LHD for $K = 10$ and $N = 2$.*

ing LHDs for space-filling. This is a computationally challenging task since $(K!)^{N-1}$ configurations are possible for K sample points and N dimensional domain. Moreover, a variety of space-filling metrics are available in the literature. Hence, the literature on LHD optimization can be classified primarily based on space-filling criteria.

Table 1: *Overview LHD optimization literature based on space-filling criteria and optimization algorithms. (IMSE: Integrated Mean Squared Error, ME: Maximum Entropy, OA: Optimization Algorithms, SA: Simulated Annealing, RCE: Row Column Exchange, QN: Quasi Newton search, CP: Columnwise-Pairwise, TA: Threshold Accepting based global search, PerGA: Permuted Genetic Algorithm, ESE: Enhanced Stochastic Evolutionary, BB: Branch and Bound, ILS: Iterative Local Search, TP: Translational Propagation, QLHD: Quasi-LHD, PSO: Particle Swarm Optimization, AMPSO: Adaptive Memetic Particle Swarm Optimization, SOBSA: Sequencing Optimization based on Simulated Annealing)*

Author/s	OA	PE	MST	Mm	ϕ_p	D_2^*	C_2	IMSE	ME
Morris and Mitchell [111]	SA				✓				
Park [125]	RCE, QNS							✓	✓
Ye *et al.* [92]	CP				✓				✓
Fang *et al.* [52]	TA-GS						✓		
Bates *et al.* [11]	PerGA	✓							
Jin *et al.* [83]	ESE				✓	✓			✓
Liefvendahl and Stocki [105]	CP	✓		✓					
van Dam *et al.* [162]	BB			✓					
Grosso *et al.* [71]	ILS			✓	✓				
Viana *et al.* [166]	TP				✓				
Xiong *et al.* [174]	QLHD			✓					
Hussalage *et al.* [81]	ESE	✓		✓					
Zhu *et al.* [178]	SLE				✓				
Chen *et al.* [22]	PSO				✓				
Damblin *et al.* [34]	ESE, SA		✓						
Pan *et al.* [124]	TP-SLE	✓		✓	✓		✓		
Aziz *et al.* [10]	AMPSO	✓							
Pholdee *et al.* [128]	SOBSA				✓				
Long *et al.* [107]	S-SLE			✓					

Table 1 presents a chronological list of the efforts on LHD optimization along with the respective space-filling criteria and optimization algorithms. Clearly, there is no basis/recommendation for selecting one criterion over the other, since there seems no way to compare one criterion with another. Of the 14 space-filling criteria known in the literature (discussed in Sections 2.2.1 and 2.2.2), only 8 have been used as objectives for LHD optimization. To the best of our knowledge, no work has used KL information, star discrepancy, wrap around discrepancy, minimax, Delaunay triangulation, and Voronoi tessellations for optimizing LHDs. ϕ_p, Mm and PE have been widely studied. The most works have focused on developing/comparing algorithms for LHD optimization. Evolutionary

algorithms such as SA, GA, ESE and their variants have received the most attention due to their simplicity and ease of application. Van Dam *et al.* [162] are the only one who have employed branch and bound algorithm for LHD optimization. Their optimal designs are archived online and freely available at http://www.spacefillingdesigns.nl. Most of the works in the literature have done selective comparison of their proposed algorithm with existing algorithm/s. Though there is no thorough comparison of designs in the literature, articles by Pan *et al.* [124] and Jin *et al.* [83] attempt to be comprehensive.

Note that an optimized LHD for a given dimension remains an LHD for reduced dimension but it loses optimality.

4.1.5 Orthogonal Array Sampling

Fundamentally, orthogonal array sampling (OAS) shares many similarities with LHD. While it retains the basic idea of random placement within bins as in LHD, it results in uniform sampling in any T dimensional projection ($T < N$) of an N dimensional domain. In case of LHD, $T = 1$ and hence, OAS can be viewed as a generalization of LHD. OAS employs four parametric integers: the number of sample points (K), the dimensions of the domain (N), the number of bins per dimension (B), and the strength of the OAS (T). Then the array index (λ) is given as $\lambda = K \times B^{-T}$. The index λ denotes the sample points per bin after projection. Note that the term orthogonal array is not related to the notion of orthogonality in linear algebra [69].

Let us understand OAS construction with a simple illustrative example of $OAS(4, 3, 2, 2)$ *i.e.* 4 sample points in 3 dimensional space with 2 bins per dimension resulting in an OAS of strength 2. Clearly, $\lambda = 1$. Thus, the domain is a cube where each dimension is divided into two equal sized bins resulting in 8 bins. Four samples are placed in diagonally opposite bins so that after projecting them into any 2 dimensions, each bin will have only one sample point. Readers can refer to [75, 122, 123, 158, 159] for further details on the construction and properties of OAS.

Although OAS is a generalization of LHD in some sense, it still has a few practical disadvantages [69]. One primary disadvantage is that its construction depends on four carefully chosen parametric integers. This makes OAS less flexible than the other system-free designs. Moreover, OAS construction is a nontrivial task as there are many possible permutations in choosing bins and then for placements of samples within each bin. These limitations have severely impacted the OAS literature over the years as discussed by Viana [165].

4.2 System-Aided DoE

From late 1940s till mid 1980s researchers strived to develop generic system-free designs. The lack of fast computing made those designs attractive. However, several researchers (like Shewry, Wynn, Currin *etc.*) realized the vital importance of incorporating system information while generating experimental designs. As system specific designs typically perform better than the generic space-filling designs, in late 1980s, many scholars and

researchers proposed model based designs that incorporated system information in various ways.

4.2.1 Maximum Entropy Sampling

Lindley [106] proposed a measure to quantify information provided by an experiment based on the concept of Shannon's entropy [146]. This entropy criterion was first employed by Shewry and Wynn [147] and subsequently by Currin *et al.* [31] to construct system-based experimental designs. Consider a sample set $\mathcal{X}^{(K)}$ of size K over a domain $\mathcal{D}$. Let $\mathcal{Y}^{(K)}$ be the response set for $\mathcal{X}^{(K)}$ and $\mathcal{Y}^{(\zeta)}$ for $\mathcal{X}^{(\zeta)}$. Then, for any random $\mathcal{Y}$ and its p.d.f. (probability distribution function) $\psi(.)$, $H(\mathcal{Y})$ denotes the entropy of ψ. Thus,

$$H(\mathcal{Y}^{(K)}) = H(\mathcal{Y}^{(\zeta)}) + H(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)}) \tag{20}$$

and $H(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)})$ denotes the conditional entropy given as follows:

$$H(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)}) = -\int \psi\left(\mathcal{Y}^{(\zeta)}\right)\left(\int \psi(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)})\log[\psi(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)})]\mathrm{d}\mathcal{Y}^{(\zeta)}\right)\mathrm{d}\mathcal{Y}^{(\zeta)} \tag{21}$$

Readers may refer to [6] for detailed a discussion and proofs. To this end, Shewry and Wynn argued as follows: since $H(\mathcal{Y}^{(K)})$ in Eq. (20) is fixed, the objective is to minimize $H(\mathcal{Y}^{(K\setminus\zeta)}|\mathcal{Y}^{(\zeta)})$ or maximize $H(\mathcal{Y}^{(\zeta)})$, hence the name maximum entropy sampling. In case of a Gaussian process model or Kriging, this results in a D-optimal design [144, 173]. Thus, it is commonly employed with Kriging [96].

4.2.2 MSE (Mean Squared Error)-based Designs

The prediction accuracy of a surrogate (predictor) can be improved by minimizing its integrated mean squared error (IMSE Eq.(23)) over $\mathcal{D}$ [16]. This is employed by Sacks and Schiller [138] and Sacks *et al.* [139] in case of Kriging. They formulated a problem of finding $\mathcal{X}^{(K)}_{N,\text{IMSE}}$, IMSE based optimal design, that is given as follows:

$$J(\mathcal{X}^{(K)}_{N,\text{IMSE}}) = \min_{\mathcal{X}^{(K)}_N} J(\mathcal{X}^{(K)}_N) \tag{22}$$

where

$$J(\mathcal{X}^{(K)}_N) = \frac{1}{\sigma^2}\int_{\mathbb{R}^N} E[y(x) - \hat{y}(x)]^2\mathrm{d}x \tag{23}$$

The minimization problem in Eq. (22) is solved by Sacks and Schiller [138] with the help of simulated annealing while Sacks *et al.* [139] used a quasi-Newton optimizer. Although IMSE based designs are optimized using system information, their performance still depends on the parameters associated with surrogate modeling techniques. For example, in case of Kriging it depends on θ, a spatial correlation parameter. Koehler and Owen [96] discussed this with various designs as illustrations.

Sacks *et al.* [140] discusses the use of maximum mean squared error (MMSE) instead of IMSE for generating designs. Such designs can be generated by minimizing MMSE (Eq.

(24)).

$$\mathcal{X}^{(K)}_{N,\text{MMSE}} = \min \left[\max_{\mathcal{D}} MSE \right] \tag{24}$$

where MSE is a mean squared error of the predictor. Sacks and Schiller [138] compared the performance of IMSE and MMSE for the discrete designs spaces. However, MMSE designs become computationally intensive for the continuous spaces, especially for higher dimensional cases as the resulting objective function turns out to be multimodal with numerous local minima.

4.3 Comparative Analysis

Thus far, we discussed various static DoE techniques developed over the last six decades and their key features. We now use literature statistics to understand the evolution of each of these techniques and identify the prominent techniques. Furthermore, we perform numerical analyses to formulate recommendations on their usage.

4.3.1 Literature Statistics

We now consider five static DoE techniques namely, MCS, QMCS, LHD, OAS and Maximum Entropy and trace their evolution over the years and across various fields (Figures

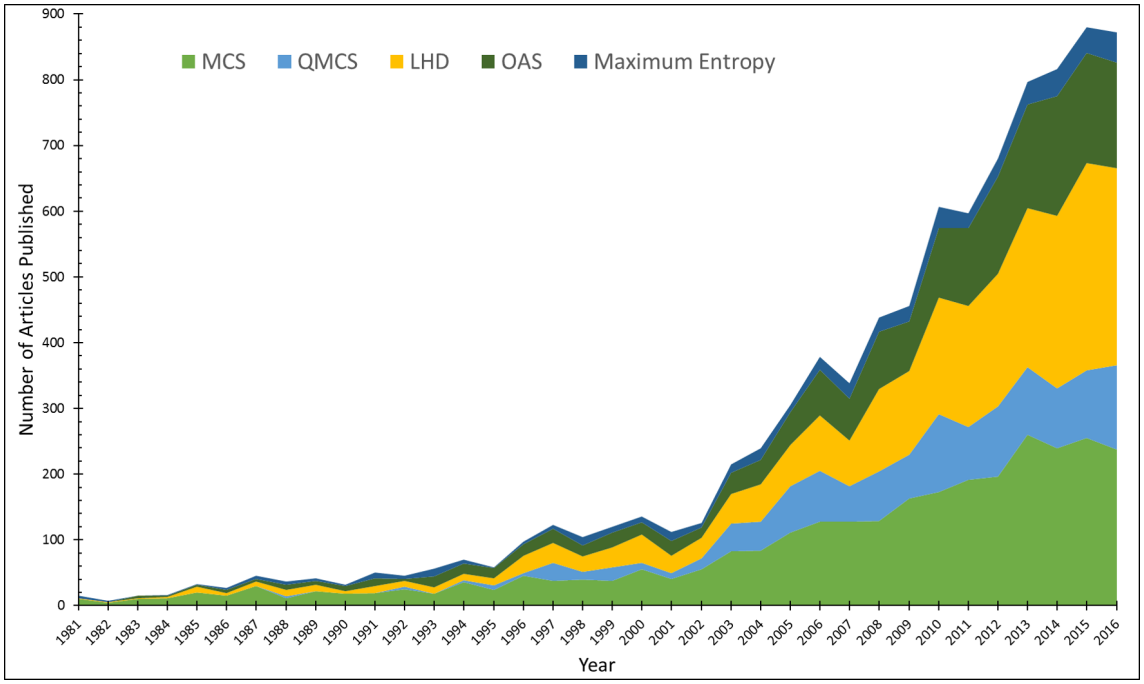


Figure 11: *Literature statistics on number of articles published for various DoE techniques over the years. (QMCS consists of Hammersley, Halton, and Sobol sampling) (Retrieved from Scopus on 1st Feb. 2017 with following search across title, abstract and keywords: "design of experiment" OR "experimental design" OR sampling AND "ST", ST: Name of Sampling Technique)*

11 and 12). We use this literature analysis as a qualitative indicator of popularity of these techniques. Figure 11 shows the trend of number of articles published on each of these techniques over the years. Though the earliest examples of DoE appeared first in late 1940s, they did not receive significant attention from research community until late 1970s (discussed earlier in Section 3). Thus, Figure 11 considers the literature from 1981 to 2016. It is clear that the literature has seen large number of articles for MCS and LHD followed by QMCS and OAS. Maximum Entropy sampling has seen the least number of articles over the years. This clearly indicates that system-free techniques are more popular than system-aided ones. This is mainly due to the generic nature and wide applicability of the system-free techniques. On the other hand, system-aided techniques are not flexible as they are mainly available for Kriging models. Furthermore, system-free techniques have become easily accessible to users through various DoE software tools. This highlights the necessity of DoE techniques that have wide range applicability while integrating system information in sample placement paradigm. This thought-process led to the birth of a new class of sampling *viz.* adaptive sampling (Section 5).

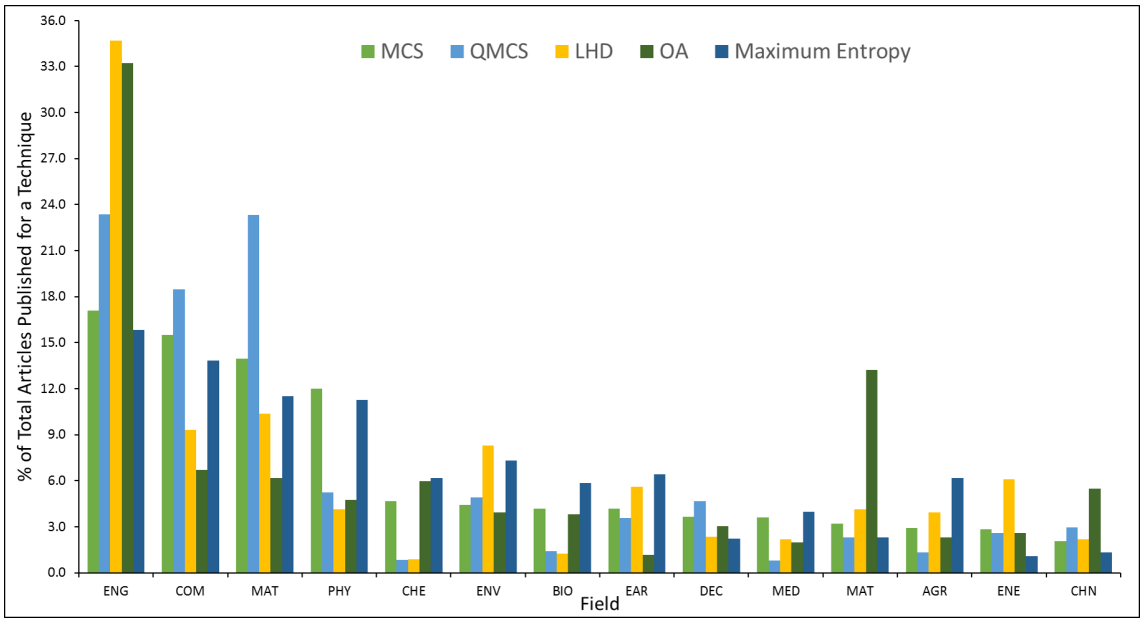
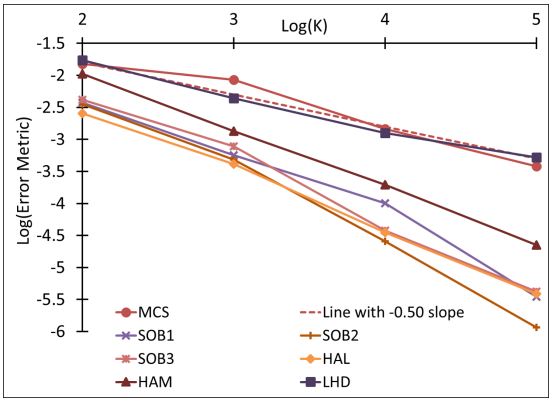


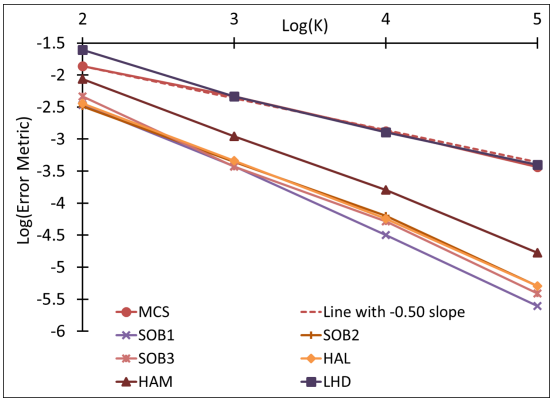
Figure 12: *Distribution of the DoE literature over the various fields. (QMCS consists of Hammersley, Halton, and Sobol sampling; ENG: Engineering, COM: Computer Science, MAT: Mathematics, PHY: Physics and Astronomy, CHE: Chemistry, ENV: Environmental Science, BIO: Biochemistry, Genetics and Molecular Biology, EAR: Earth and Planetary Sciences, DEC: Decision Sciences, MED: Medicine, MAT: Material Science, AGR: Agricultural and Biological Sciences, ENE: Energy, CHN: Chemical Engineering) (Retrieved from Scopus on 1st Feb. 2017 with following search across title, abstract and keywords: "design of experiment" OR "experimental design" OR sampling AND "ST", ST: Name of Sampling Technique)*

4.3.2 Numerical Analysis

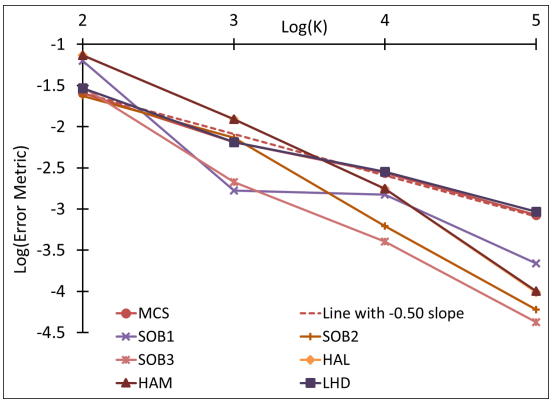
Though the literature statistics qualitatively ranked the popularity of various static techniques, we now present a quantitative comparison of sampling techniques like MCS, HAM (Hammersley sampling), HAL (Halton sampling), SOB (Sobol sampling), and LHD. We use three variations of Sobol sampling *viz.* SOB1, SOB2, and SOB3. SOB1 and SOB2 are based on [86] while SOB3 is based on a recently proposed Sobol implementation [87]. We use Matlab for generating MCS, HAM, HAL, and SOB1 while Model Development Suite (MoDS) [1] for generating SOB2 and SOB3. *Curse of dimensionality*



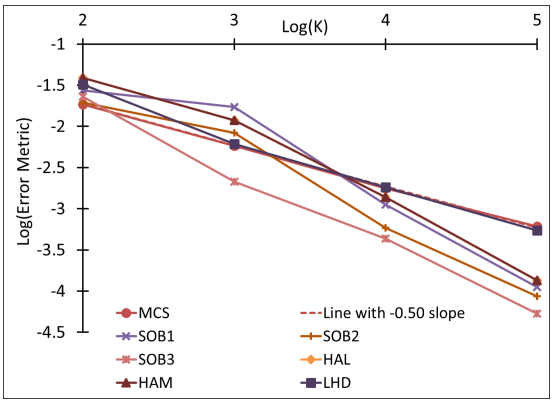
(a) $N = 5$, $m = 1$



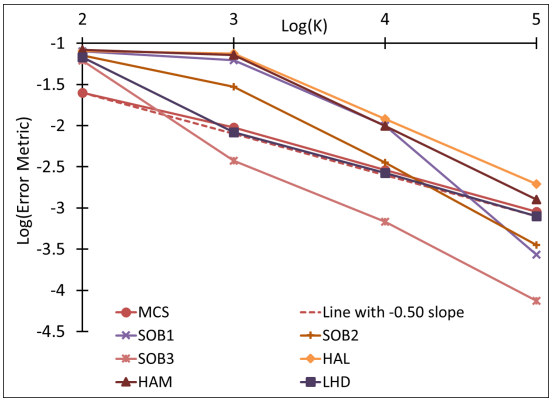
(b) $N = 5$, $m = 5$



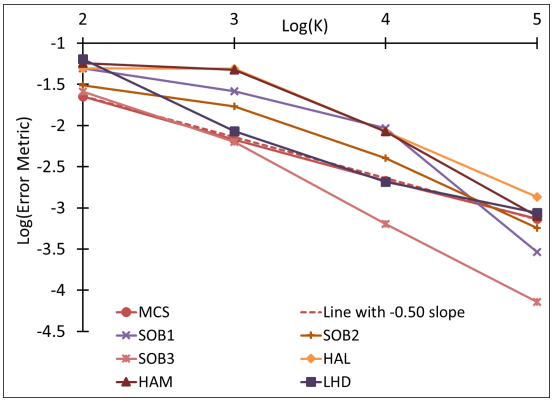
(c) $N = 25$, $m = 1$



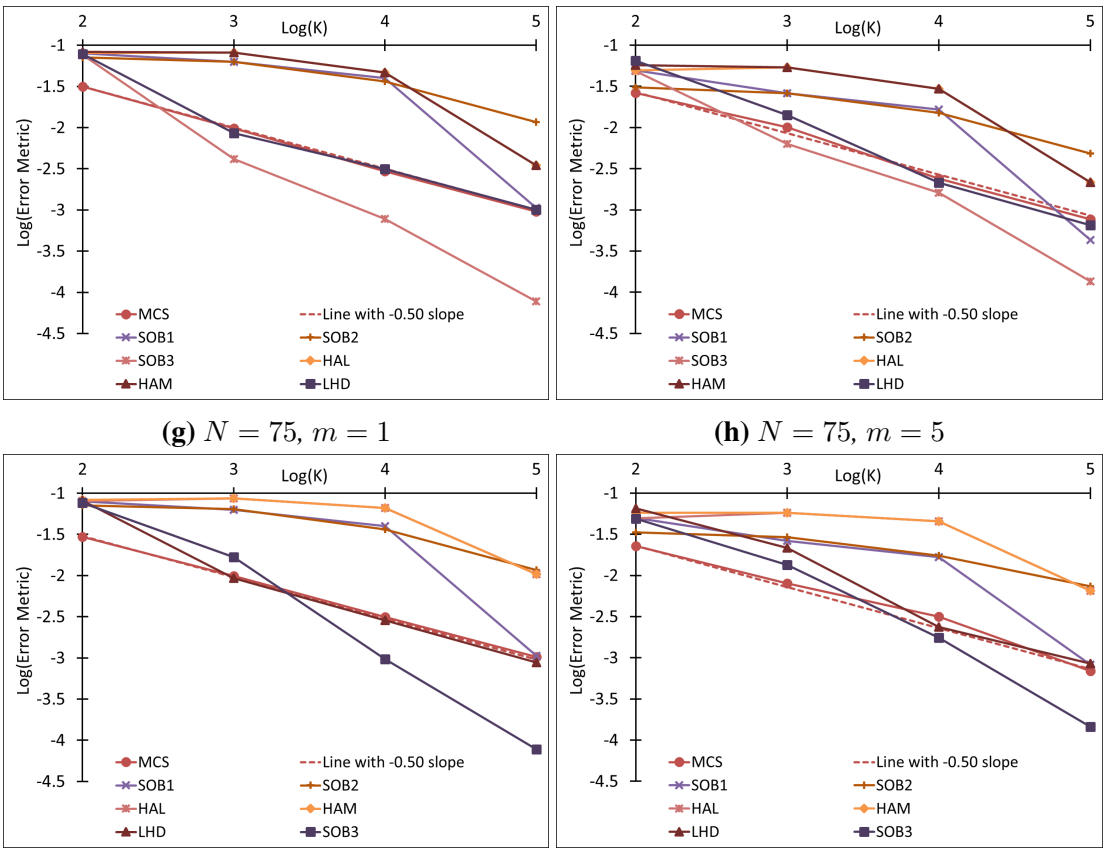
(d) $N = 25$, $m = 5$



(e) $N = 50$, $m = 1$



(f) $N = 50$, $m = 5$



(g) $N = 75$, $m = 1$ **(h)** $N = 75$, $m = 5$

(i) $N = 100$, $m = 1$ **(j)** $N = 100$, $m = 5$

Figure 13: *Error analysis accessing effects of dimensions (N), sample sizes (K), and moments (m) on various sampling techniques. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

is a fundamental challenge faced by all the techniques, especially the static system-free techniques. Thus, robustness to the effects of dimensionality is a desirable feature in a DoE technique. Robustness of the static system-free DoE techniques can be tested based on their ability to fill space in higher dimensions. For this, we perform three analyses *viz.* distribution based error analysis, visualization analysis, and distance based space-filling (Section 2.2.2) analysis. Let us discuss them one by one.

In distribution based error analysis, we access the performance of various techniques based on a error metric that quantifies the departure of a given sample distribution from the uniform sample distribution. This metric is based on covariance matrix and Appendix A provides a detailed development of this metric. The larger the error metric, the larger the departure, hence, the poorer the space-filling ability. Figure 13 shows the error metric $\xi^{(t)}$ (Eq. (A.8)) computed for a sample set $\mathcal{X}_N^{(K)}$ for $K = \{100, 1000, 10000, 100000\}$, and $N = \{5, 25, 50, 75, 100\}$ using $t = \{\text{MCS}, \text{SOB1}, \text{SOB2}, \text{SOB3}, \text{HAM}, \text{HAL}, \text{LHD}\}$. Order of magnitude of $\xi^{(\text{MCS})}$ decreases linearly with increasing sample size ($\log K$) and is independent of dimensions (N). This follows the expected theoretical trend where order

Table 2: *Correlation analysis of various sampling techniques using $\rho_{max}^{(t)}$ and $m = 1$. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

$N = 5$							
K	MCS	SOB1	SOB2	SOB3	HAL	HAM	LHD
100	0.1264	0.0444	0.0421	0.0504	**0.0299**	0.1260	0.1353
1000	0.0862	0.0068	0.0058	0.0093	**0.0050**	0.0162	0.0511
10000	0.0240	0.0012	**0.0003**	0.0005	0.0004	0.0023	0.0195
100000	0.0060	**0.0000**	**0.0000**	0.0001	**0.0000**	0.0003	0.0055
$N = 25$							
100	0.4156	0.7525	**0.2797**	0.3657	0.8824	0.8824	0.3363
1000	0.1218	**0.0201**	0.0878	0.0255	0.1465	0.1465	0.0962
10000	0.0249	0.0180	0.0074	**0.0048**	0.0211	0.0211	0.0261
100000	0.0095	0.0026	0.0007	**0.0005**	0.0012	0.0012	0.0113
$N = 50$							
100	**0.3336**	0.9497	0.8524	0.7499	1.0000	1.0000	0.8098
1000	0.1254	0.7494	0.3545	**0.0449**	0.8852	0.8335	0.1199
10000	0.0328	0.1196	0.0429	**0.0082**	0.1441	0.1195	0.0355
100000	0.0127	0.0033	0.0043	**0.0009**	0.0234	0.0152	0.0109
$N = 75$							
100	**0.3486**	0.9497	0.8524	0.9336	1.0000	1.0000	0.9305
1000	0.1147	0.7532	0.7515	**0.0500**	0.9618	0.9618	0.1107
10000	0.0358	0.4788	0.4356	**0.0094**	0.5621	0.5621	0.0364
100000	0.0124	0.0127	0.1399	**0.0009**	0.0417	0.0417	0.0105
$N = 100$							
100	**0.3649**	0.9497	0.8524	0.9336	1.0000	1.0000	0.9659
1000	**0.1161**	0.7532	0.7704	0.2012	0.9871	0.9871	0.1254
10000	0.0406	0.4788	0.4356	**0.0116**	0.7859	0.7859	0.0336
100000	0.0113	0.0127	0.1399	**0.0009**	0.1250	0.1250	0.0119

of magnitude of error is proportional to $\frac{1}{\sqrt{K}}$. LHD shows the performance trend similar to MCS since LHD places sample points randomly within each bin. For $N \geq 50$, MCS and LHD performs better than all the techniques except for SOB3 in some cases. For $N \geq 25$ and $K \geq 1000$, SOB3 outperforms all the other techniques making it the most robust technique to the effects of dimensionality. However, SOB1 and SOB2 are moderate performers and their performance worsens compared to MCS, LHD, and SOB3 with increasing N and m. HAL is the best performer for $N = 5$ and $K \leq 1000$, however, its performance steadily drops with increasing dimensions. HAM follows the similar performance trend that of HAL. HAM and HAL are the worst performers for $N \geq 50$ and for $N = 25$, $K \leq 1000$.

Apart from $\xi^{(t)}$, we use Pearson coefficient based metric $\rho_{max}^{(t)}$ (Eq. (A.10)) to quantify the maximum correlation in the samples generated using various sampling techniques.

Table 3: *Correlation analysis of various sampling techniques using $\rho_{max}^{(t)}$ and $m = 5$. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

$N = 5$							
K	MCS	SOB1	SOB2	SOB3	HAL	HAM	LHD
100	0.1661	0.0588	**0.0522**	0.0783	0.0636	0.1467	0.1295
1000	0.0520	0.0060	0.0070	**0.0059**	0.0074	0.0175	0.0639
10000	0.0174	**0.0005**	0.0010	0.0008	0.0009	0.0025	0.0199
100000	0.0077	**0.0000**	0.0001	0.0001	0.0001	0.0003	0.0031
$N = 25$							
100	**0.2954**	0.4492	0.3090	0.3738	0.8172	0.8172	0.3802
1000	0.0816	0.2741	0.1321	**0.0339**	0.1991	0.1991	0.0983
10000	0.0274	0.0178	0.0093	**0.0069**	0.0221	0.0221	0.0311
100000	0.0096	0.0018	0.0014	**0.0008**	0.0021	0.0021	0.0091
$N = 50$							
100	**0.3872**	0.8361	0.4910	0.4230	1.0000	1.0000	0.9998
1000	**0.0993**	0.4151	0.2727	0.1002	0.8427	0.7806	0.1352
10000	0.0365	0.1477	0.0637	**0.0101**	0.1364	0.1364	0.0370
100000	0.0125	0.0046	0.0091	**0.0011**	0.0215	0.0126	0.0103
$N = 75$							
100	**0.3896**	0.8361	0.4910	0.7782	1.0000	1.0000	1.0000
1000	0.1154	0.4162	0.4128	**0.1011**	0.9411	0.9411	0.2256
10000	0.0338	0.2634	0.2401	**0.0257**	0.4835	0.4835	0.0385
100000	0.0098	0.0068	0.0771	**0.0021**	0.0343	0.0343	0.0100
$N = 100$							
100	**0.4041**	0.8361	0.5473	0.7782	1.0000	1.0000	1.0000
1000	**0.1181**	0.4172	0.4614	0.2146	0.9784	0.9784	0.3428
10000	0.0461	0.2634	0.2764	**0.0278**	0.7294	0.7294	0.0376
100000	0.0122	0.0131	0.1169	**0.0023**	0.1038	0.1038	0.0127

Appendix A discusses a detailed development of $\rho_{max}^{(t)}$. The lower the metric value, the lesser the correlation, the better the sampling technique. Tables 2 and 3 enlist the metric values computed for $m = 1$ and $m = 5$ respectively, and the least correlated sample sets are highlighted with boldface. SOB3 is the best performer for $N \geq 25$ and $K \geq 10000$. On the other hand, MCS performs the best for $N \geq 50$ and $K \leq 1000$. LHD follows similar performance trend as that of MCS. SOB1, SOB2, HAM, and HAL results in highly correlated samples with increasing N. Overall, SOB3 is the most prominent choice of sampling technique for high dimensional cases followed MCS and LHD, while HAL and HAM are limited to lower dimensions.

Apart from distribution based metrics like $\xi^{(t)}$ and $\rho_{max}^{(t)}$, we use distance based SFC like MST and Mm to analyze the performance of various sampling techniques. MST is employed to quantify the overall space-filling ability of a technique while Mm (Eq. (8)) is

used to identify the maximum of minmum inter-sample distance. We normalize MST and Mm so that $1 \leq \text{MST} < \infty$ and $1 \leq \text{Mm} < \infty$. We follow the normalization procedure described in [12, 64].

MST criterion quantifies the spreading of the sample points by generating a tree and taking average value of the edge lengths (Section 2.2.2). The larger the MST value, the better the space-filling. Here, we compute MST for the seven equal sized ($K = 1000$) sample sets generated using seven different techniques. We choose $2 \leq N \leq 40$. Figure 14 shows that for $N \leq 8$, HAM, HAL, and variants of SOB perform better than MCS and LHD. On the other hand, for $N \geq 10$, MST values for HAM and HAL drop consistently, worsening their performance. For $10 \leq N \leq 23$, SOB3 shows the best performance, however, for $N > 23$, SOB2 steadily outperforms rest of the techniques. MCS and LHD follow a steadily increasing trend similar to SOB3 (for $N \geq 10$). For $N \geq 33$, MCS and LHD consistently outperform all the techniques except SOB2.

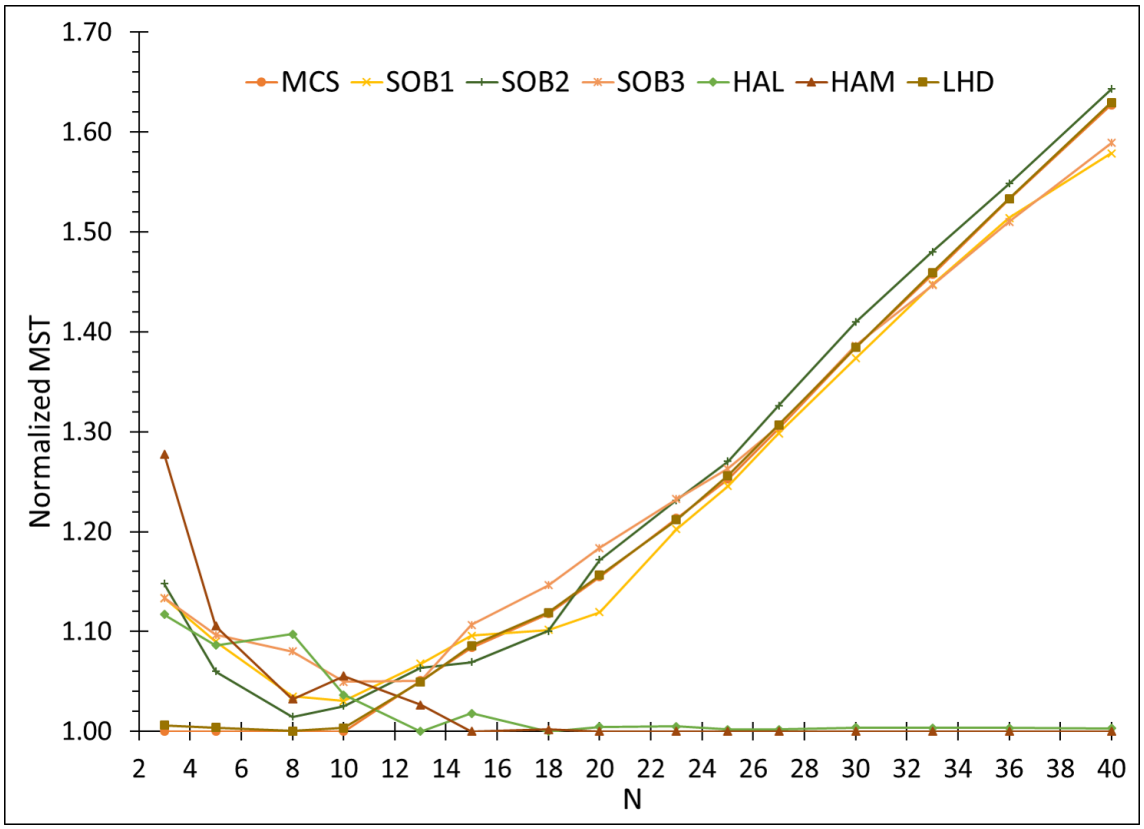


Figure 14: *Effects of dimensionality on the space-filling ability of various techniques based on normalized MST. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

Table 4: *Numerical Analysis of various sampling techniques based on normalized MST criteria. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

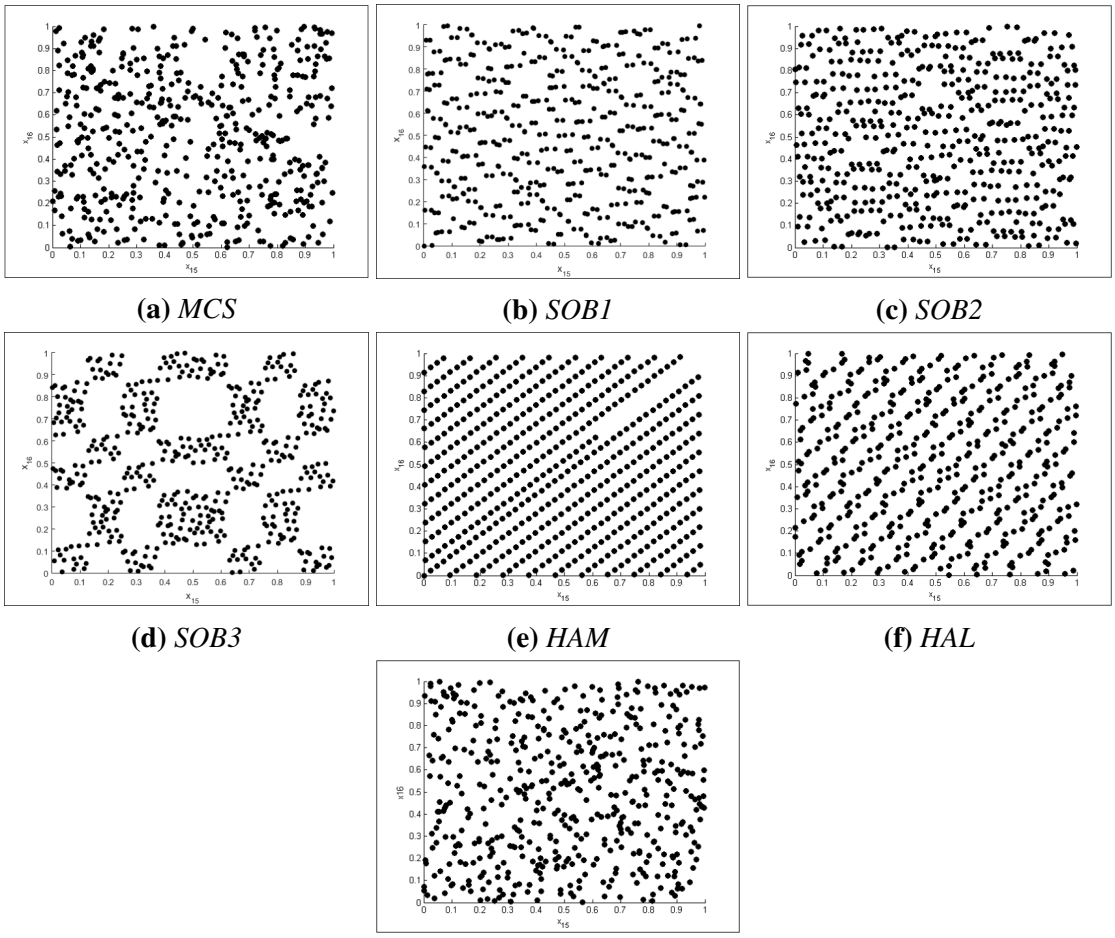
N	MST						
	MCS	SOB1	SOB2	SOB3	HAM	HAL	LHD
3	1.00	1.13	1.15	1.13	**1.28**	1.12	1.01
5	1.00	1.09	1.06	1.10	**1.11**	1.09	1.00
8	1.00	1.03	1.01	1.08	1.03	**1.10**	1.00
10	1.00	1.03	1.03	1.05	**1.06**	1.04	1.00
13	1.05	**1.07**	1.06	1.05	1.03	1.00	1.05
15	1.08	1.10	1.07	**1.11**	1.00	1.02	1.09
18	**1.12**	1.10	1.10	1.15	1.00	1.00	1.12
20	1.15	1.12	1.17	**1.18**	1.00	1.00	1.16
23	1.21	1.20	**1.23**	**1.23**	1.00	1.00	1.21
25	1.25	1.25	**1.27**	1.26	1.00	1.00	1.26
27	1.30	1.30	**1.33**	1.31	1.00	1.00	1.31
30	1.38	1.37	**1.41**	1.39	1.00	1.00	1.38
33	1.46	1.45	**1.48**	1.45	1.00	1.00	1.46
36	1.53	1.51	**1.55**	1.51	1.00	1.00	1.53
40	1.63	1.58	**1.64**	1.59	1.00	1.00	1.63

Unlike MST, Mm criterion only quantifies the maximin inter-sample distance. Thus, the larger the Mm value, the better the space-filling potential of a technique. For $3 \leq N < 13$, HAM and HAL show the best performance, however, their performance drop severely and become the worst with increasing dimensions. For $N \geq 15$, variants of SOB, especially SOB2, exhibit the best performance. MCS and LHD are moderate performers and they show steadily increasing performance trend with increasing dimensions. Tables 4 and 5 summarize the normalized MST and Mm values for selected sampling techniques over the range of dimensions. The best performer for a given dimension is denoted with boldface.

Table 5: *Numerical Analysis of various sampling techniques based on normalized Mm criteria. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

N	Mm						
	MCS	SOB1	SOB2	SOB3	HAM	HAL	LHD
3	1.00	4.14	4.77	4.14	**4.96**	3.11	1.45
5	1.00	1.88	1.79	2.28	**2.75**	2.24	1.05
8	1.00	1.36	1.32	1.39	1.72	**1.84**	1.03
10	1.14	1.29	1.00	1.52	**1.92**	1.84	1.05
13	1.00	1.07	1.00	1.22	**1.38**	1.31	1.04
15	1.01	**1.24**	1.13	1.15	1.07	1.07	1.00
18	1.16	1.29	1.38	**1.39**	1.00	1.00	1.23
20	1.35	1.37	**1.54**	1.41	1.00	1.00	1.38
23	1.62	1.62	**1.74**	1.70	1.00	1.00	1.57
25	1.78	1.67	1.80	**1.97**	1.00	1.00	1.73
27	1.89	1.87	1.88	**2.06**	1.00	1.00	1.89
30	2.17	2.12	**2.31**	2.22	1.00	1.00	2.10
33	2.33	2.36	**2.64**	2.39	1.00	1.00	2.35
36	2.55	2.51	**2.73**	2.49	1.00	1.00	2.48
40	2.79	2.78	**2.84**	2.67	1.00	1.00	2.79

Apart from the SFC, we use 2-D projection plots (Figures 15-17) to visualize space-filling abilities of various techniques. For this, we use $N = 40$ dimensional domain and generate $K = 500$ sample points. Although $\binom{40}{2}$ different 2 dimensional projection plots are possible for every DoE technique, we have chosen only three of them for illustration purposes. We can qualitatively judge robustness of a DoE technique based on the *extent of clustering*. The lesser the clustering, the more robust the DoE technique to the effects of dimensionality. Figures 15e, 16e, 17e, 15f, 16f, and 17f clearly show the linearly correlated samples generated by HAM and HAL respectively. This subsequently leads to severe clustering at higher dimensions for HAM and correlated clusters for HAL. Variants of SOB exhibit clustering in some cases (Figures 15d, 16f, 17c, 17b, 17d). Though MCS and LHD face certain amount clustering, they neither present massive unexplored regions like in HAM nor have correlated samples like in HAL.

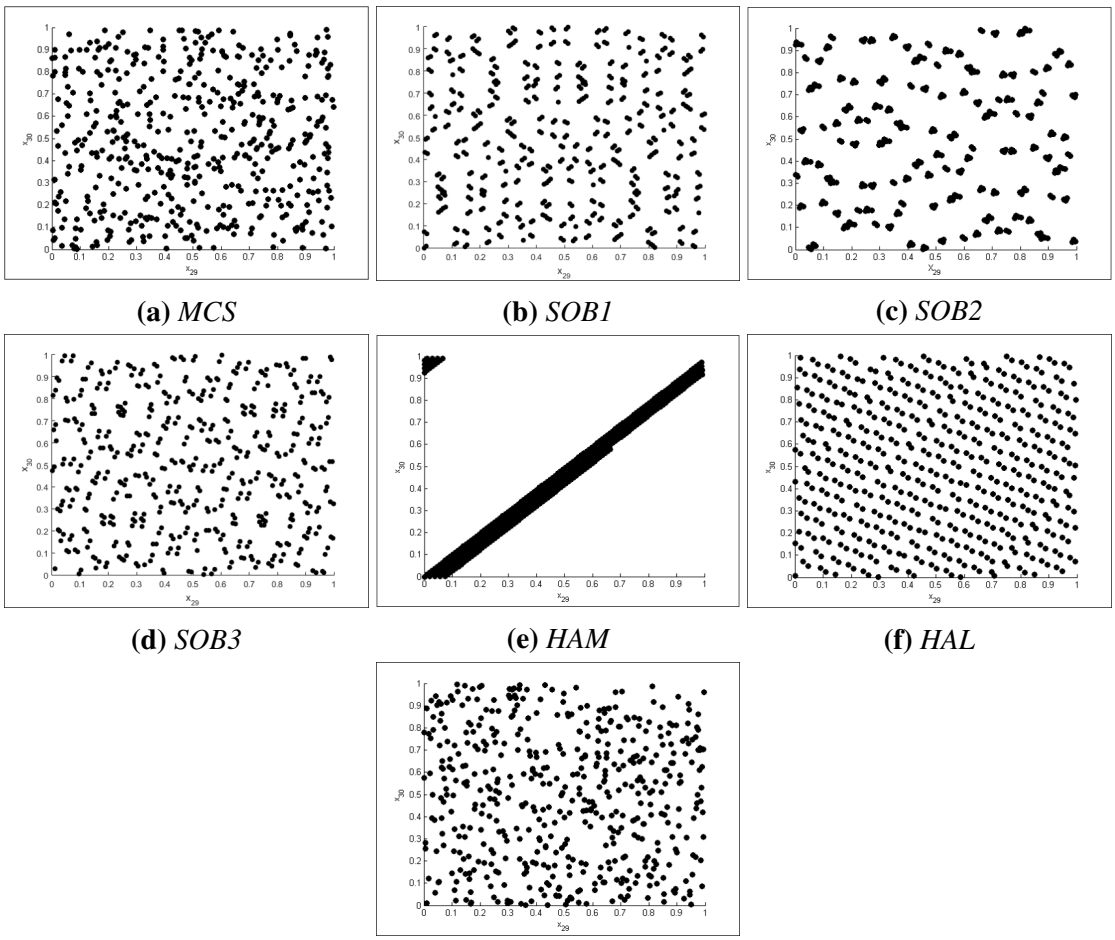


(a) *MCS* **(b)** *SOB1* **(c)** *SOB2*

(d) *SOB3* **(e)** *HAM* **(f)** *HAL*

(g) *LHD*

Figure 15: 2 *dimensional projection plots (x_{15} vs. x_{16}) illustrating clustering in various sampling techniques for $K = 500$ in $[0, 1]^{40}$. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

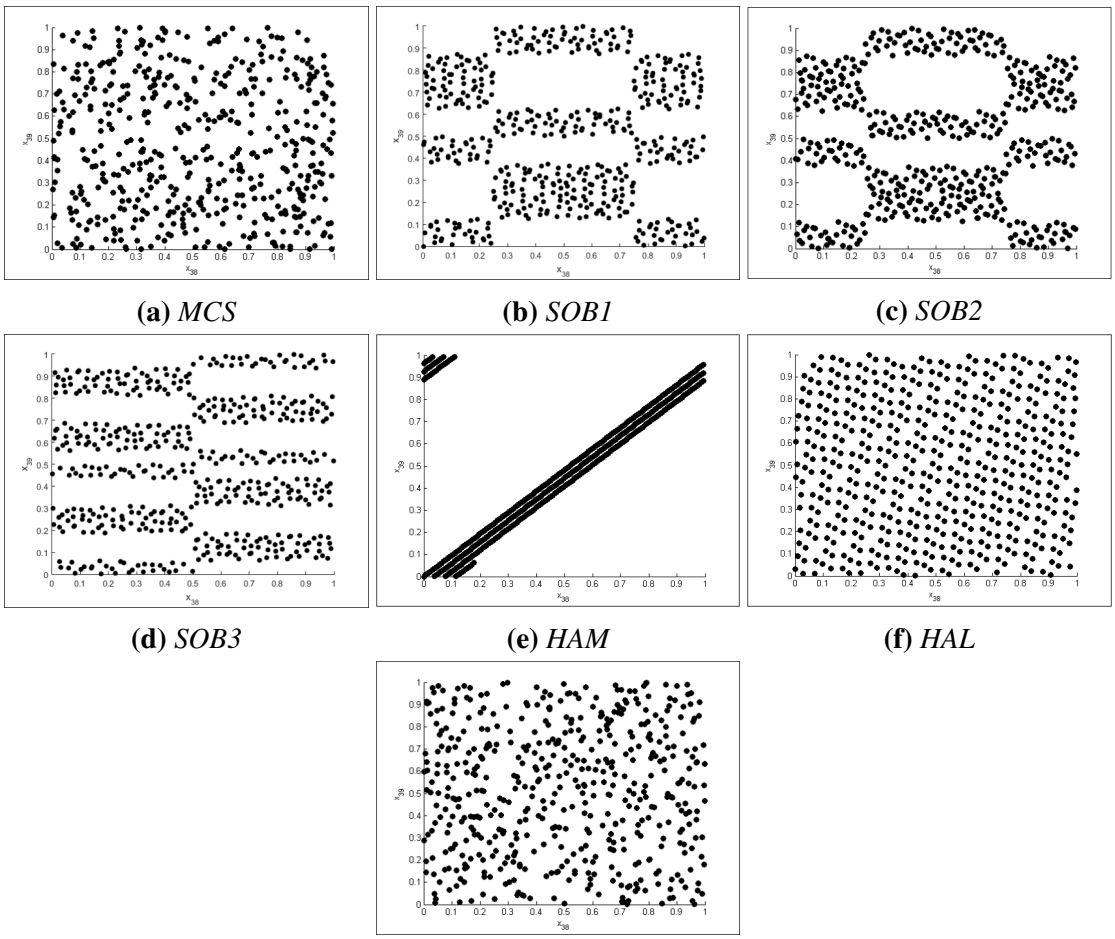


(a) *MCS* **(b)** *SOB1* **(c)** *SOB2*

(d) *SOB3* **(e)** *HAM* **(f)** *HAL*

(g) *LHD*

Figure 16: 2 *dimensional projection plots (x_{29} vs. x_{30}) illustrating clustering in various sampling techniques for $K = 500$ in $[0, 1]^{40}$. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*



(a) *MCS* **(b)** *SOB1* **(c)** *SOB2*

(d) *SOB3* **(e)** *HAM* **(f)** *HAL*

(g) *LHD*

Figure 17: 2 *dimensional projection plots (x_{38} vs. x_{39}) illustrating clustering in various sampling techniques for $K = 500$ in $[0, 1]^{40}$. (MCS: Monte Carlo Sampling, SOB1: Sobol Sampling type 1, SOB2: Sobol Sampling type 2, SOB3: Sobol Sampling type 3, HAM:Hammersley Sampling, HAL: Halton Sampling, LHD: Latin Hypercube Design)*

5 Adaptive DoE/Sampling

Sampling techniques discussed so far are one-shot or static in nature where all sample points are generated at once. Although these techniques are very popular, they can result in under/oversampling and thus, poor system approximation [28, 64]. In order to tackle these issues, a dynamic class of DoE called adaptive sampling (also known as sequential sampling) has attracted attention from research community. Adaptive sampling has two major advantages over one shot techniques *viz.* low computational expenses and better approximations [28, 29]. The adaptive sampling techniques discussed here aim to enhance surrogate quality globally. Thus, they are different from adaptive sampling techniques for surrogate assisted optimization [14, 35, 84, 89, 90, 131, 134, 135, 169].

Exploration and exploitation are the two basic concepts underlying adaptive sampling. Exploration of the design space during sample placement aims to cover poorly-represented or *under-sampled* sub-spaces. In other words, exploration aims for homogeneous sample placement via space-filling. On the other hand, exploitation aims to place sample points in the complex/highly nonlinear regions. While pure exploitation may result in local oversampling, pure exploration may cause under-sampling in complex/nonlinear regions and/or oversampling in simple/linear regions. Adaptive sampling may solely target exploration only or a balance between exploration and exploitation. The number of sample points that an adaptive sampling algorithm generates at each iteration is called its granularity. Fine grained algorithms are preferred over coarse grained as they completely avoid any oversampling [29].

5.1 Adaptive Exploratory Sampling

The simplest case of adaptive exploratory sampling (AES) is adaptive grid. Consider a 2 dimensional grid of size 3×3 . If these 9 points are insufficient for constructing an acceptable approximation, then we can add additional points (*e.g.* at midpoints) between the existing grid points. Although this is the easiest AES, it may typically be coarse grained and it rapidly faces *curse of dimensionality*. The term adaptive in AES suggests incorporation of system information, however, they are just sequential/iterative space-filling techniques. Thus, any space-filling technique can be made adaptive. Crombecq *et al.* [29] proposed various AES techniques based on nested LHD and MCS by optimizing the inter-sample distances and projected inter-site distances. Their numerical comparison showed that the adaptive techniques outperformed the conventional one shot techniques like LHD, Sobol, Halton [26].

5.2 Adaptive Hybrid Sampling

Though adaptive exploratory designs do perform better than the existing one-shot space-filling designs, they still do not incorporate system information in sample placement. Thus, over the last decade, researchers have shifted their focus to *adaptive hybrid sampling* (AHS) techniques that exploit the system information for sample placement in an intelligent and adaptive manner.

In 2002, Jin *et al.* [82] proposed two approaches namely the Maximin Scaled Distance (MSD) and the Cross Validation (CV). The former is a modification of maximin distance based sampling that utilizes system information by assigning weights to important variables. On the other hand, the latter uses CV error [97] to direct exploitation, and maximin distance to direct exploration. Their analyses showed that an MSD approach with Kriging performed better than a CV approach with Radial Basis Function (RBF) when the underlying systems are not excessively nonlinear. In contrast, their CV approach with RBF outperformed their MSD approach with Kriging when the underlying system was highly nonlinear or wavy. Farhang-Mehr and Azarm [54] proposed an adaptive variation of maximum entropy sampling, which aims to place samples in *irregular* regions of the response. Although their work lacked a thorough numerical comparison, they showed that the adaptive version performed better than the one-shot versions. Busby *et al.* [21] proposed a Kriging based hierarchical nonlinear approximation that employs adaptive grids and considers entropy minimization. Their hierarchical approach outperforms the conventional one-shot approach. Li *et al.* [104] proposed an adaptive design for Kriging approximation based on prediction error. Their algorithm places new samples in *CAMM i.e.* continuous and multi-modal region with the hope to improve poorly approximated parts of the response. While they impose a clustering constraint to avoid local oversampling, they do not use any explicit exploration criterion. Their approach performed better than the MSD approach. Aute *et al.* [8] extended this further and employed it for heat exchanger design optimization. Though all of the techniques discussed are adaptive in nature, they are mainly suited for Kriging due to easy access to prediction errors. Thus, all these techniques are surrogate specific. Moreover, Kriging performance drops, and it becomes computationally costly with increase in dimensions and non-linearity of the problem [116]. Hence, such techniques may not be viable for a wide range of problems.

In 2009, Crombecq *et al.* [27, 28] proposed a generic, robust and reliable adaptive sampling strategy to overcome some shortcomings of the literature discussed above. Their strategy, namely LOLA-Voronoi involves a combination of Voronoi tessellations and local linear approximation (LOLA). Voronoi tessellations target domain exploration while LOLA guides local exploitation. Their algorithm outperformed the static techniques for all surrogate model types [26]. Singh *et al.* [150] proposed three trade-off schemes to balance exploration versus exploitation for LOLA-Voronoi strategy. Although LOLA-Voronoi strategy has shown promising results, it has some drawbacks. Its major drawback is its compute-intensive nature due to Voronoi tessellations and LOLA computations. Voronoi tessellations become computationally costly as dimensions increase. Furthermore, LOLA requires estimation of local derivatives which is computationally expensive and becomes further complex with increasing N. Van der Herten *et al.* [164] attempted to tackle these issues with a fuzzy variation of LOLA called as FLOLA. This provides the benefits of the original algorithm and reduces computational burden significantly.

Xu *et al.* [175] employed Voronoi tessellations for dividing the domain into smaller regions and then placed samples based on cross validation errors (exploitation). Zhou *et al.* employed a similar concept using support vector machines [176, 177]. Eason and Cremaschi [42] employed an adaptive strategy based on some scores for ANN surrogates. This *score* consists of the normalized nearest neighbor distance of a new potential sample and its normalized expected variance based on jackknifing [132]. Though their sample placement is systematic, the sample points are still generated randomly similar to LOLA-

Voronoi. Jin *et al.* [85] essentially extended the work of Eason and Cremaschi [42] with two modifications. First, they enhanced ANN modelling with auto-node selection, and second, they used maximum predicted error instead of expected variance. Ajdari and Mahlooji proposed Delaunay-Hybrid Adaptive Sequential Design (DHASD) [5] which can be viewed as a variation of [175]. It employs Delaunay triangulation which is a dual of Voronoi tessellations (discussed in Section 2.2.2) for exploration and cross validation error for exploitation. Their algorithm outperformed LOLA-Voronoi. Though most of the adaptive sampling techniques discussed so far place sample points systematically, the placement is still random.

On the other hand, works by Cozad *et al.* [24, 25] and Garud *et al.* [63, 64] formulated optimization problems to place new samples. Cozad *et al.* proposed adaptive sampling for their surrogate modeling tool called ALAMO [24, 25]. They add sample points one at a time to an initial sample set. For each sample point, they solve a derivative-free optimization problem to maximize the deviation of the surrogate from the real function. This can become compute-intensive, as it requires the evaluation of the real function during optimization. Garud *et al.* presented an adaptive sampling strategy based on point placement optimization [64]. They formulated a nonlinear programming problem (NLP) based on crowding distance metric for exploration and departure function for exploitation. Table 6 summarizes distinguished adaptive sampling strategies in the literature, key characteristics of their strategies, and their dependence on type of surrogate models.

Table 6: *Adaptive hybrid sampling techniques in the literature. (Mm: Maximin Distance, CVE: Cross Validation Error, MD: Mahalanobis Distance, ME: Maximum Entropy, VT: Voronoi Tessellation, LOLA: Local Linear Approximation, CC: Clustering Constraint, EE: Expected Error, NN: Nearest Neighbor, JK: Jackknifing, DT: Delaunay Triangulation, MSE: Maximum Sampling Error, CDM: Crowding Distance Metric, DF: Departure Function)*

Author	Exploration	Exploitation	Surrogate Dependence	Approach
Jin *et al.* [82]	Mm	CVE	×	Optimization
Busby *et al.* [21]	MD	ME	✓(Kriging)	Score
Crombecq *et al.* [27, 28]	VT	LOLA	×	Score
Li *et al.* [104]	CC	EE	✓(Kriging)	Optimization
Xu *et al.* [175]	VT	CVE	×	Optimization
Eason and Cremaschi [42]	NN	JK	×	Score
Ajdari and Mahlooji [5]	DT	CVE	×	Score
Cozad *et al.* [24]	-	MSE	×	Optimization
Garud *et al.* [64]	CDM	DF	×	Optimization

6 Conclusions and Future Prospects

Historically, DoE literature began with system-free approaches, and gradually evolved into system-aided techniques. Recently, adaptive techniques are receiving much attention. In this comprehensive review, we critically classified the DoE literature following its natural evolution and described key developments within respective sub-classes.

System-free techniques are the most popular to date due to their simplicity, easy commercial availability, and flexibility across applications. Our numerical analysis using distance based SFC and distribution error metrics highlighted the excellent space-filling abilities of SOB3 for $\mathbb{R}^N \forall\, N > 16$. In contrast, QMCS techniques such as HAL and HAM are well suited for low dimensional systems ($N \leq 8$). Moreover, our visualization analysis revealed that they face a massive clustering and/or correlated samples resulting in a very poor space-filling for large N. Overall, SOB3 performs the best for $\mathbb{R}^N \forall\, N > 16$, and large K while MCS and LHD tend to perform better for $\mathbb{R}^N \forall\, N > 16$, and smaller K. HAM and HAL perform well for $\mathbb{R}^N \forall\, N \leq 8$.

System-aided techniques have received limited attention due to three major reasons. First, although they employ system information using surrogates to enhance performance, their availability in commonly used software tools has been limited. Second, their reliance on one-shot surrogates limits their scalability across a range of dimensions. Finally and most importantly, their growth was prematurely halted by the more attractive and rapidly advancing science of adaptive techniques. The idea of enhancing system-free techniques via optimization understandably and quickly evolved into sequential DoE aided by system information. By promoting the idea of using system knowledge to improve DoE, these techniques laid the foundation for the modern adaptive approaches.

The current DoE research focuses largely on adaptive techniques that generate samples intelligently by incorporating exploration and exploitation in a sequential manner. The earlier techniques were surrogate dependent. However, researchers soon realized the importance of generic and surrogate-independent strategies. This resulted in many generic sequential techniques mostly based on some scores, while maintaining the conventional idea of random placement. Recent works [24, 25, 64] have strived to use optimization strategies to achieve sample placement. The class of adaptive techniques is still in infancy and holds much promise for the future.

Our work has identified several key gaps in DoE research and suggests various opportunities for further developments. The literature has a variety of space-filling criteria, but it lacks clear guidelines on their usage or merits. This has resulted in a plethora of works exploring various combinations of these criteria. The consistency in the performance of SFC across a range of dimensions remains unclear. Thus, analyzing existing SFC and making recommendations on their usage can be a useful contribution. Novel comprehensive SFC can play a vital role in making better sense of system-free DoE techniques.

While several advanced variants of system-free techniques exist in the literature, very few of them have found their way into commercial tools or user-friendly avenues. An exhaustive benchmarking of these techniques over ranges of dimensions, sample sizes, and test beds (function sets) will help streamline the DoE literature. The same holds for the rapidly growing adaptive DoE.

Though optimization-based adaptive sampling strategies hold much promise, several vital research topics need attention. First, efficient methods for placement optimization and obtaining globally optimal solutions for multi-modal objective functions (due to nonlinear surrogates) is a clear challenge. Second, minimizing the computational burden of executing adaptive strategies, especially for solving point placement optimization problems and surrogate refitting, is another challenge. Third, there is a need for strategies that are flexible with respect to surrogates and are scalable to higher dimensions. Fourth, optimization-based sample placement strategies for multiple/hybrid surrogates would be useful. Finally, the inclusion of these methods into commercial tools remains a concern that will hopefully be resolved with time and technological advances. As the current computational research moves towards modeling, simulation, analysis, and visualization of complex interdisciplinary systems, the application of adaptive DoE techniques to high dimensional systems needs much attention. Here, machine or active learning can play a significant role.

Apart from system approximation and prediction, optimization becomes paramount in system design and operation. Hence, surrogate-based optimization of complex systems is a growing area where adaptive sampling can play a principal role.

Acknowledgements

This publication is made possible by the Singapore National Research Foundation (NRF) under its Campus for Research Excellence And Technological Enterprise (CREATE) programme.

Nomenclature

Abbreviations

AES: Adaptive Exploratory Sampling
AHS: Adaptive Hybrid Sampling
AMPSO: Adaptive Memetic Particle Swarm Optimization
ANN: Artificial Neural Networks
BB: Branch and Bound
CAMM: Continuous and Multi-Modal
CC: Clustering Constraint
CDM: Crowding Distance Metric
CP: Columnwise-Pairwise
CV: Cross Validation
CVE: Cross Validation Error
DF: Departure Function
DHASD: Delaunay Hybrid Adaptive Sequential Design
DoE: Design of Experiments
DT: Delaunay Triangulation
EE: Expected Error
ESE: Enhanced Stochastic Evolutionary
FLOLA: Fuzzy Local Linear Approximation
GLS: Good Lattice Sampling
HAL: Halton sampling
HAM: Hammersley sampling
ILS: Iterative Local Search
IMSE: Integrated Mean Squared Error
JK: Jackknifing
KL: Kullback-Leibler
LHD: Latin Hypercube Designs
LHS: Latin Hypercube Sampling
LOLA: Local Linear Approximation
MCS: Monte Carlo Sampling
MD: Mahalanobis Distance
ME: Maximum Entropy
Mm: Maximin distance
mM: miniMax distance
MMSE: Maximum Mean Squared Error

MSD: Maximin Scaled Distance
MSE: Mean Squared Error
MSE: Maximum Sampling Error
MST: Minimum Spanning Tree
NLP: Nonlinear Programming
NN: Nearest Neighbor
OA: Optimization Algorithms
OAS: Orthogonal Array Sampling
PerGA: Permuted Genetic Algorithm
PE: Potential Energy
PSO: Particle Swarm Optimization
QLHD: Quasi Latin Hypercube Design
QMCS: Quasi-Monte Carlo Sampling
QNS: Quasi Newton Search
QRLD: Quasi-Random Low Discrepancy
RBF: Radial Basis Functions
RCE: Row Column Exchange
SA: Simulated Annealing
SFC: Space-Filling Criteria
SMCS: Stratified Monte Carlo Sampling
SOB1: Sobol sampling in Matlab (based on [86])
SOB2: Sobol Sampling in MoDS (based on [86])
SOB3: Sobol Sampling in MoDS (based on [87])
SOBSA: Sequencing Optimization Based on Simulated Annealing
SVM: Support Vector Machines
TA: Threshold Accepting based global search
TP: Translational Propagation
UD: Uniform Designs
VT: Voronoi Tessellation

Notations

Subscripts

n: Index for elements of design/input variables' vector
s: Index for elements of response/output variables' vector
R: Radix or base

Superscripts

j: Index for elements of set
k: Index for elements of set
t: Index for elements in set of sampling techniques
L: Lower bound
U: Upper bound

Parameters

m: Moment of x
p: ordering parameter in ϕ_p
B: Number of bins in orthogonal array
K: Total number of sample points in a sample set
N: Total number of input domain dimensions
S: Total number of output domain dimensions
T: Strength of orthogonal array
λ: Orthogonal array index

Continuous Variables

x: Vector of input/design variables
y: Vector of output/response variables

Symbols

d: Euclidean distance
DT: Delaunay Triangulation
$\tilde{f}$: Surrogate model form
H: Entropy
$\boldsymbol{L}$: $K \times N$ matrix
$\mathbb{1}_{[0,1]^N}$: Indicator function
$\mathcal{D}$: Real bounded domain
$\mathbb{E}$: Expectation
$\mathbb{N}$: Set of natural numbers
$\mathcal{Q}$: Set of integers
$\mathbb{R}$: Real space
$\mathcal{S}$: Sequence of numbers
$\mathcal{U}_{[0,1]^N}$: Uniform distribution over $[0,1]^N$ domain
V: Voronoi cell
$\mathcal{X}$: Sample set
$\mathcal{Y}$: Response set
L_2: Order 2 norm
D^*: Star discrepancy
D_2^*: L_2 discrepancy
C_2: Centered L_2 discrepancy
W_2: Wrap-around L_2 discrepancy
I_{KL}: Kullback-Leibler information
ε: Random error
ξ: Covariance based error metric
ρ: Pearson correlation coefficient
Σ: Variance-Covariance matrix
σ_e: Standard deviation in edge length for MST

σ_n^2: Variance of n^{th} component of x
σ_{np}: Covariance between n^{th} and p^{th} component of x
θ: Inverse radix number
μ_e: Mean edge length for MST
ψ: Probability distribution function

A Distribution based Error Metrics

Consider $x \in \mathbb{R}^N$, $\Psi(x)$; $\Psi : \mathbb{R}^N \to \mathbb{R}^N$, and $x \sim \mathcal{U}_{[0,1]^N}$. Without loss of generality, let us assume the domain of x as follows, $x \in \mathcal{D}$ and $\mathcal{D} : [0,1]^N$. Then, expectation of $\Psi(x)$ is given as follows.

$$\begin{aligned}\mathbb{E}(\Psi(x)) &= \int_{\mathcal{D}} \frac{1}{K} \sum_{k=1}^{K} \Psi(x) \delta_{x^{(k)}}(\mathrm{d}x) \\ &= \int_{\mathcal{D}} \Psi(x) \mathbb{1}_{[0,1]^N}(x) \mathrm{d}x\end{aligned} \tag{A.1}$$

For $\Psi(x) = x^m$, Eq. (A.1) results in expectation of m^{th} moment of x,

$$\mathbb{E}(x^m) = \int_{\mathcal{D}} x^m \mathbb{1}_{[0,1]^N}(x) \mathrm{d}x \tag{A.2}$$

Note that x is a N dimensional vector and hence, the integral in Eq. (A.2) can be expanded to N integrals in each dimension $n = \{1, 2, ..., N\}$.

$$\begin{aligned}\mathbb{E}(x^m) &= \int_0^1 \int_0^1 ... \int_0^1 x^m \mathrm{d}x_1 \mathrm{d}x_2 ... \mathrm{d}x_N \\ &= \left[\left[\frac{1}{m+1} x_1^{m+1}, \frac{1}{m+1} x_2^{m+1}, ..., \frac{1}{m+1} x_N^{m+1} \right]_0^1 \right]^\intercal \\ &= \left[\frac{1}{m+1}, \frac{1}{m+1}, ..., \frac{1}{m+1} \right]^\intercal\end{aligned} \tag{A.3}$$

Variance σ_n^2 of any $x_n^m \, \forall \, n = \{1, 2, ..., N\}$ is as,

$$\begin{aligned}\sigma_n^2 &= \mathbb{E}((x_n^m)^2) - (\mathbb{E}(x_n^m))^2 \\ &= \mathbb{E}((x_n^{2m}) - (\mathbb{E}(x_n^m))^2 \\ &= \frac{1}{1+2m} - \frac{1}{(1+m)^2} \\ &= \frac{m^2}{(1+2m)(1+m)^2}\end{aligned} \tag{A.4}$$

Covariance σ_{np} between x_n^m and $x_p^m \, \forall \, n, p = \{1, 2, ..., N\}$ and $n \neq p$ can be written as,

$$\begin{aligned}\sigma_{np} &= \mathbb{E}((x_n^m - \mathbb{E}(x_n^m))((x_p^m - \mathbb{E}(x_p^m))) \\ &= \mathbb{E}(x_n^m x_p^m) - \mathbb{E}(x_n^m)\mathbb{E}(x_p^m) \\ &= 0\end{aligned} \tag{A.5}$$

A covariance matrix consisting of variances and covariances of components of x^m is given in Eq. A.6. For uniform distributions x_n and x_p are uncorrelated and hence, $\sigma_{np} = 0$ (Eq. (A.5)) reducing Σ to a diagonal matrix given in Eq. (A.7).

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_N^2 \end{bmatrix} \tag{A.6}$$

$$= \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_N^2 \end{bmatrix} \tag{A.7}$$

To this end, we define error metric $\xi^{(t)}$ for a given sample set $\mathcal{X}_N^{(K)}$ generated using technique $t = \{\text{MCS}, \text{SOB1}, \text{SOB2}, \text{SOB2}, \text{HAM}, \text{HAL}, \text{LHD}\}$ as the maximum absolute departure of its covariances from that of the uniform distribution (given in Eq. (A.7)). In other words, we characterize the potential correlation based on covariance of the samples that can be given as follows.

$$\begin{aligned} \xi^{(t)} &= \max_{n,\, p \in \mathcal{N}, n \neq p} |\sigma_{np}^{(t)} - \sigma_{np}| \\ &= \max_{n,\, p \in \mathcal{N}, n \neq p} |\sigma_{np}^{(t)} - 0| \\ &= \max_{n,\, p \in \mathcal{N}, n \neq p} |\sigma_{np}^{(t)}| \end{aligned} \tag{A.8}$$

where $\mathcal{N} = \{1, 2, ..., N\}$ and $t = \{\text{MCS}, \text{SOB1}, \text{SOB2}, \text{HAM}, \text{HAL}, \text{LHD}\}$.

Moreover, we define coefficient of correlation also known as Pearson coefficient as shown in Eq. (A.9) and $-1 \leq \rho_{np}^{(t)} \leq 1$. $\rho_{np}^{(t)} = 0$ denotes uncorrelated samples while positive and negative values of $\rho_{np}^{(t)}$ imply positively and negatively correlated samples respectively.

$$\rho_{np}^{(t)} = \frac{\sigma_{np}^{(t)}}{\sigma_n^{(t)} \sigma_p^{(t)}} \tag{A.9}$$

We use this coefficient to compare the correlations between sample sets generated by various techniques. For this, we compute $\rho_{max}^{(t)}$ (Eq. (A.9)) for a sample set $\mathcal{X}_N^{(K)}$ generated using $t = \{\text{MCS}, \text{SOB1}, \text{SOB2}, \text{SOB2}, \text{HAM}, \text{HAL}, \text{LHD}\}$.

$$\rho_{max}^{(t)} = \max_{n,\, p \in \mathcal{N}, n \neq p} |\rho_{np}^{(t)}| \tag{A.10}$$

The larger the value of $\rho_{max}^{(t)}$, the greater the correlation ($0 \leq \rho_{max}^{(t)} \leq 1$).

B Computational Tools

Tables 7 and 8 provide a list of tools available for design and analysis of computer experiments. Note that this list is just a indicator of available resources for users and is far from complete.

Table 7: *Open source computational tools for design and analysis of computer experiments and their fields of applicability.*

Software/Tool	Fields of Application
Advanced Simulation Library (ASL)	Transport Phenomenon, Chemical Reactions, Elasticity, Interface tracking
FreeMat	Rapid Engineering, Scietific prototyping, Data processing
GNU Octave	Numerical Computations
SageMath	Algebra, Combinatorics, Numerical mathematics, Number theory, Calculus
Scilab	Cross Platform Numerical Computations
SU2	Computational Fluid Dynamics, Aerodynamics optimization

Table 8: *Proprietary computational tools for design and analysis of computer experiments and their fields of applicability.*

Software/Tool	Fields of Application
GoldSim	Decision and Risk analysis
HyperWorks	Product design and development, Engineering
LMS Imagine.Lab Amesim	Multidomain Mechatronic Systems
Maple	Symbolic and Numeric Computing
Mathematica	Mathematical Symbolic Computing
MATLAB	Multi-paradigm Numerical Computing
Model Development Suite (MoDS)	Computer Aided Engineering
ModelCenter	System Design and Optimization
Plant Simulation	Production Systems and Process Operations

References

[1] CMCL Innovations. http://www.cmclinnovations.com/. Accessed: 2017-01-31.

[2] *3. Information Matrices*, pages 61–97. . doi:10.1137/1.9780898719109.ch3.

[3] *9. D, A, E, T-Optimality*, pages 210–246. . doi:10.1137/1.9780898719109.ch9.

[4] B. Addis, M. Locatelli, and F. Schoen. Packing circles in a square: New putative optima obtained via global optimization. *Optimization Online*, 155, 2005.

[5] A. Ajdari and H. Mahlooji. An adaptive exploration-exploitation algorithm for constructing metamodels in random simulation using a novel sequential experimental design. *Communications in Statistics-Simulation and Computation*, 43(5): 947–968, 2014. doi:10.1080/03610918.2012.720743.

[6] R. Ash. Information theory. 1965. *Interscience, New York*.

[7] F. Aurenhammer, R. Klein, and D. T. Lee. *Voronoi diagrams and Delaunay triangulations*. World Scientific Publishing Co Inc, 2013.

[8] V. Aute, K. Saleh, O. Abdelaziz, S. Azarm, and R. Radermacher. Cross-validation based single response adaptive design of experiments for kriging metamodeling of deterministic computer simulations. *Structural and Multidisciplinary Optimization*, 48(3):581–605, 2013. doi:10.1007/s00158-013-0918-5.

[9] P. Azadi, G. Brownbridge, I. Kemp, S. Mosbach, J. S. Dennis, and M. Kraft. Microkinetic modeling of the Fischer–Tropsch synthesis over cobalt catalysts. *ChemCatChem*, 7(1):137–143, 2015. doi:10.1002/cctc.201402662.

[10] M. Aziz and M.-H. Tayarani-N. An adaptive memetic particle swarm optimization algorithm for finding large-scale Latin hypercube designs. *Engineering Applications of Artificial Intelligence*, 36:222–237, 2014. doi:http://dx.doi.org/10.1016/j.engappai.2014.07.021.

[11] S. J. Bates, J. Sienz, and V. V. Toropov. Formulation of the optimal Latin hypercube design of experiments using a permutation genetic algorithm. *AIAA Paper*, (2004-2011):1–7, 2004. doi:10.2514/6.2004-2011.

[12] S. Bhushan and I. Karimi. Heuristic algorithms for scheduling an automated wet-etch station. *Computers & chemical engineering*, 28(3):363–379, 2004.

[13] G. Blau, M. Lasinski, S. Orcun, S.-H. Hsu, J. Caruthers, N. Delgass, and V. Venkatasubramanian. High fidelity mathematical model building with experimental data: A bayesian approach. *Computers & Chemical Engineering*, 32(4): 971–989, 2008. doi:10.1016/j.compchemeng.2007.04.008.

[14] F. Boukouvala and M. G. Ierapetritou. Surrogate-based optimization of expensive flowsheet modeling for continuous pharmaceutical manufacturing. *Journal of Pharmaceutical Innovation*, 8(2):131–145, 2013. doi:10.1007/s12247-013-9154-1.

[15] G. E. Box and D. W. Behnken. Some new three level designs for the study of quantitative variables. *Technometrics*, 2(4):455–475, 1960. doi:10.1080/00401706.1960.10489912.

[16] G. E. Box and N. R. Draper. A basis for the selection of a response surface design. *Journal of the American Statistical Association*, 54(287):622–654, 1959.

[17] E. Braaten and G. Weller. An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration. *Journal of Computational Physics*, 33(2): 249–258, 1979. doi:10.1016/0021-9991(79)90019-6.

[18] P. Bratley and B. L. Fox. Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software (TOMS)*, 14(1): 88–100, 1988. doi:10.1145/42288.214372.

[19] M. Broadie, P. Glasserman, and S. Kou. A continuity correction for discrete barrier options. *Mathematical Finance*, 7(4):325–349, 1997. doi:10.1111/1467-9965.00035.

[20] G. P. E. Brownbridge, A. Smallbone, W. Phadungsukanan, S. Mosbach, M. Kraft, and B. Johansson. Automated IC engine model development with uncertainty propagation. Technical report, SAE Technical Paper, 2011.

[21] D. Busby, C. L. Farmer, and A. Iske. Hierarchical nonlinear approximation for experimental design and statistical data fitting. *SIAM Journal on Scientific Computing*, 29(1):49–69, 2007. doi:10.1137/050639983.

[22] R. B. Chen, D. N. Hsieh, Y. Hung, and W. Wang. Optimizing Latin hypercube designs by particle swarm. *Statistics and computing*, 23(5):663–676, 2013. doi:10.1007/s11222-012-9363-3.

[23] V. C. Chen, K. L. Tsui, R. R. Barton, and M. Meckesheimer. A review on design, modeling and applications of computer experiments. *IIE transactions*, 38(4):273–291, 2006. doi:10.1080/07408170500232495.

[24] A. Cozad, N. V. Sahinidis, and D. C. Miller. Learning surrogate models for simulation-based optimization. *AIChE Journal*, 60(6):2211–2227, 2014. doi:10.1002/aic.14418.

[25] A. Cozad, N. V. Sahinidis, and D. C. Miller. A combined first-principles and data-driven approach to model building. *Computers & Chemical Engineering*, 73:116–127, 2015. doi:10.1016/j.compchemeng.2014.11.010.

[26] K. Crombecq. *Surrogate modeling of computer experiments with sequential experimental design*. PhD thesis, Ghent University, 2011.

[27] K. Crombecq, L. De Tommasi, D. Gorissen, and T. Dhaene. A novel sequential design strategy for global surrogate modeling. In *Simulation Conference (WSC), Proceedings of the 2009 Winter*, pages 731–742. IEEE, 2009. doi:10.1109/WSC.2009.5429687.

[28] K. Crombecq, D. Gorissen, D. Deschrijver, and T. Dhaene. A novel hybrid sequential design strategy for global surrogate modeling of computer experiments. *SIAM Journal on Scientific Computing*, 33(4):1948–1974, 2011. doi:10.1137/090761811.

[29] K. Crombecq, E. Laermans, and T. Dhaene. Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling. *European Journal of Operational Research*, 214(3):683–696, 2011. doi:10.1016/j.ejor.2011.05.032.

[30] A. C. Crombie. Avicenna on medievel scientific tradition. In G. M. Wickens, editor, *Avicenna: Scientist and Philosopher. A Millenary Symposium.* Luzac and Co., London, 1952.

[31] C. Currin, T. Mitchell, M. Morris, and D. Ylvisaker. A bayesian approach to the design and analysis of computer experiment. *ORNL-6498. Springfield, VA: National Technical Information Service*, 1988.

[32] H. Dai and W. Wang. Application of low-discrepancy sampling method in structural reliability analysis. *Structural Safety*, 31(1):55–64, 2009. doi:10.1016/j.strusafe.2008.03.001.

[33] R. Daling and W. van Haeringen. Hartree-fock and lowest-order vertex-correction contribution to the direct gap of the semiconductor silicon. *Physical Review B*, 40 (17):11659, 1989. doi:10.1103/PhysRevB.40.11659.

[34] G. Damblin, M. Couplet, and B. Iooss. Numerical studies of space-filling designs: optimization of Latin hypercube samples and subprojection properties. *Journal of Simulation*, 7(4):276–289, 2013. doi:10.1057/jos.2013.16.

[35] E. Davis and M. Ierapetritou. A centroid-based sampling strategy for kriging global modeling and optimization. *AIChE journal*, 56(1):220–240, 2010. doi:10.1002/aic.11881.

[36] C. de Groot, R. Peikert, and D. Würtz. The optimal packing of ten equal circles in a square, 1990.

[37] B. Delaunay. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7(793-800):1–2, 1934.

[38] U. M. Diwekar and J. R. Kalagnanam. Efficient sampling technique for optimization under uncertainty. *AIChE Journal*, 43(2):440–447, 1997. doi:10.1002/aic.690430217.

[39] P. M. Dunn. James Lind (1716-94) of Edinburgh and the treatment of scurvy. *Archives of Disease in Childhood-Fetal and Neonatal Edition*, 76(1):F64–F65, 1997. doi:10.1136/fn.76.1.F64.

[40] C. Dussert, G. Rasigni, M. Rasigni, J. Palmari, and A. Llebaria. Minimal spanning tree: A new approach for studying order and disorder. *Physical Review B*, 34(5): 3528, 1986. doi:10.1103/PhysRevB.34.3528.

[41] R. Dyer, H. Zhang, and T. Möller. Surface sampling and the intrinsic Voronoi diagram. In *Computer Graphics Forum*, volume 27, pages 1393–1402. Wiley Online Library, 2008. doi:10.1111/j.1467-8659.2008.01279.x.

[42] J. Eason and S. Cremaschi. Adaptive sequential sampling for surrogate model generation with artificial neural networks. *Computers & Chemical Engineering*, 68:220–232, 2014. doi:10.1016/j.compchemeng.2014.05.021.

[43] V. Eglajs and P. Audze. New approach to the design of multifactor experiments. *Problems of Dynamics and Strengths*, 35:104–107, 1977.

[44] M. Evans and T. Swartz. *Approximating integrals via Monte Carlo and deterministic methods*, volume 20. OUP Oxford, 2000.

[45] K.-T. Fang. The uniform design: application of number-theoretic methods in experimental design. *Acta Math. Appl. Sinica*, 3(4):363–372, 1980.

[46] K. T. Fang. Theory, method and applications of the uniform design. *International journal of reliability, quality and safety engineering*, 9(04):305–315, 2002. doi:10.1142/S0218539302000858.

[47] K. T. Fang and D. K. Lin. Ch. 4. uniform experimental designs and their applications in industry. *Handbook of statistics*, 22:131–170, 2003. doi:10.1016/S0169-7161(03)22006-X.

[48] K. T. Fang and C. X. Ma. Wrap-around L_2-discrepancy of random sampling, Latin hypercube and uniform designs. *journal of complexity*, 17(4):608–624, 2001. doi:10.1006/jcom.2001.0589.

[49] K. T. Fang and H. Qin. A note on construction of nearly uniform designs with large number of runs. *Statistics & probability letters*, 61(2):215–224, 2003. doi:10.1016/S0169-7161(03)22006-X.

[50] K. T. Fang, Y. Wang, and P. M. Bentler. Some applications of number-theoretic methods in statistics. *Statistical Science*, pages 416–428, 1994. doi:10.2307/2246355.

[51] K. T. Fang, D. K. Lin, P. Winker, and Y. Zhang. Uniform design: Theory and Application. *Technometrics*, 42(3):237–248, 2000. doi:10.1080/00401706.2000.10486045.

[52] K. T. Fang, C.-X. Ma, and P. Winker. Centered L_2-discrepancy of random sampling and Latin hypercube design, and construction of uniform designs. *Mathematics of Computation*, 71(237):275–296, 2002. doi:10.1090/S0025-5718-00-01281-3.

[53] K. T. Fang, R. Li, and A. Sudjianto. *Design and modeling for computer experiments*. CRC Press, 2005.

[54] A. Farhang-Mehr and S. Azarm. Bayesian meta-modelling of engineering design simulations: a sequential approach with adaptation to irregularities in the response behaviour. *International Journal for Numerical Methods in Engineering*, 62(15): 2104–2126, 2005. doi:10.1002/nme.1261.

[55] H. Faure. On the star-discrepancy of generalized Hammersley sequences in two dimensions. *Monatshefte für Mathematik*, 101(4):291–300, 1986. doi:10.1007/BF01559392.

[56] V. V. Fedorov. *Theory of optimal experiments*. Elsevier, 1972.

[57] R. Fisher. *The design of experiments*. Oliver and Boyd, Edinburgh, 1935.

[58] R. A. Fisher. The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain*, 33(4):503–513, 1926. doi:10.1002/9781118763667.

[59] G. S. Fishman. Estimating network characteristics in stochastic activity networks. *Management Science*, 31(5):579–593, 1985. doi:10.1287/mnsc.31.5.579.

[60] K. Forsberg, N. He, and A. Massih. Probabilistic analysis of nuclear fuel rod behavior using a quasi-Monte Carlo method. *Nuclear science and engineering*, 122 (1):142–150, 1996.

[61] Y. Fu and U. M. Diwekar. An efficient sampling approach to multiobjective optimization. *Annals of Operations Research*, 132(1-4):109–134, 2004. doi:10.1023/B:ANOR.0000045279.46948.dd.

[62] S. Galanti and A. Jung. Low-discrepancy sequences: Monte Carlo simulation of option prices. *The Journal of Derivatives*, 5(1):63–83, 1997. doi:10.3905/jod.1997.407985.

[63] S. Garud, I. Karimi, and M. Kraft. Smart adaptive sampling for surrogate modelling. *Computer Aided Chemical Engineering*, 38:631–636, 2016. doi:10.1016/B978-0-444-63428-3.50110-7.

[64] S. S. Garud, I. Karimi, and M. Kraft. Smart sampling algorithm for surrogate model development. *Computers & Chemical Engineering*, 96:103–114, 2017. doi:10.1016/j.compchemeng.2016.10.006.

[65] T. Gensane. Dense packings of equal spheres in a cube. *the electronic journal of combinatorics*, 11(1):R33, 2004.

[66] J. E. Gentle. *Random number generation and Monte Carlo methods*. Springer Science & Business Media, 2006. doi:10.1007/b97336.

[67] J. E. Gentle, W. K. Härdle, and Y. Mori. *Handbook of computational statistics: concepts and methods*. Springer Science & Business Media, 2012. doi:10.1007/978-3-642-21551-3.

[68] J. Gergonne. Application de la méthode des moindres quarrésa linterpolation des suites. *Annales de Math. Pures et Appl*, 6:242–252, 1815.

[69] A. A. Giunta, S. F. Wojtkiewicz, and M. S. Eldred. Overview of modern design of experiments methods for computational simulations. In *Proceedings of the 41st AIAA aerospace sciences meeting and exhibit, AIAA-2003-0649*, 2003. doi:10.2514/6.2003-649.

[70] M. Goldberg. The packing of equal circles in a square. *Mathematics Magazine*, 43 (1):24–30, 1970.

[71] A. Grosso, A. Jamali, and M. Locatelli. Finding maximin Latin hypercube designs by iterated local search heuristics. *European Journal of Operational Research*, 197 (2):541–547, 2009. doi:10.1016/j.ejor.2008.07.028.

[72] D. Guo and X. Wang. Quasi-Monte Carlo filtering in nonlinear dynamic systems. *IEEE transactions on signal processing*, 54(6):2087–2098, 2006. doi:10.1109/TSP.2006.873585.

[73] J. H. Halton. Algorithm 247: Radical-inverse quasi-random point sequence. *Communications of the ACM*, 7(12):701–702, 1964. doi:10.1145/355588.365104.

[74] J. M. Hammersley. Monte Carlo methods for solving multivariable problems. *Annals of the New York Academy of Sciences*, 86(3):844–874, 1960. doi:10.1111/j.1749-6632.1960.tb42846.x.

[75] A. S. Hedayat, N. J. A. Sloane, and J. Stufken. *Orthogonal arrays: theory and applications*. Springer Science & Business Media, 2012. doi:10.1007/978-1-4612-1478-6.

[76] S. Heinrich and A. Keller. Quasi-Monte Carlo methods in computer graphics, part ii: The radiance equation. Technical report, Citeseer, 1994.

[77] P. Hellekalek. Regularities in the distribution of special sequences. *Journal of Number Theory*, 18(1):41–55, 1984. doi:10.1016/0022-314X(84)90041-6.

[78] F. Hickernell. A generalized discrepancy and quadrature error bound. *Mathematics of Computation of the American Mathematical Society*, 67(221):299–322, 1998. doi:10.1090/S0025-5718-98-00894-1.

[79] F. J. Hickernell. Lattice rules: how well do they measure up? In *Random and quasi-random point sets*, pages 109–166. Springer, 1998.

[80] E. Hlawka. Funktionen von beschränkter Variation in der Theorie der Gleichverteilung. *Annali di Matematica Pura ed Applicata*, 54(1):325–333, 1961. doi:10.1007/BF02415361.

[81] B. G. Husslage, G. Rennen, E. R. van Dam, and D. den Hertog. Space-filling Latin hypercube designs for computer experiments. *Optimization and Engineering*, 12 (4):611–630, 2011. doi:10.1007/s11081-010-9129-8.

[82] R. Jin, W. Chen, and A. Sudjianto. On sequential sampling for global metamodeling in engineering design. In *ASME 2002 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, pages 539–548. American Society of Mechanical Engineers, 2002. doi:10.1115/DETC2002/DAC-34092.

[83] R. Jin, W. Chen, and A. Sudjianto. An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134(1):268–287, 2005. doi:10.1016/j.jspi.2004.02.014.

[84] Y. Jin. Surrogate-assisted evolutionary computation: Recent advances and future challenges. *Swarm and Evolutionary Computation*, 1(2):61–70, 2011. doi:10.1016/j.swevo.2011.05.001.

[85] Y. Jin, J. Li, W. Du, and F. Qian. Adaptive sampling for surrogate modelling with artificial neural network and its application in an industrial cracking furnace. *The Canadian Journal of Chemical Engineering*, 94(2):262–272, 2016. doi:10.1002/cjce.22384.

[86] S. Joe and F. Y. Kuo. Remark on algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software (TOMS)*, 29(1):49–57, 2003. doi:10.1145/641876.641879.

[87] S. Joe and F. Y. Kuo. Constructing sobol sequences with better two-dimensional projections. *SIAM Journal on Scientific Computing*, 30(5):2635–2654, 2008. doi:10.1137/070709359.

[88] M. E. Johnson, L. M. Moore, and D. Ylvisaker. Minimax and maximin distance designs. *Journal of statistical planning and inference*, 26(2):131–148, 1990. doi:10.1016/0378-3758(90)90122-B.

[89] D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization*, 21(4):345–383, 2001. doi:10.1023/A:1012771025575.

[90] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global optimization*, 13(4):455–492, 1998. doi:10.1023/A:1008306431147.

[91] J. R. Kalagnanam and U. M. Diwekar. An efficient sampling technique for off-line quality control. *Technometrics*, 39(3):308–319, 1997. doi:10.1080/00401706.1997.10485122.

[92] Q. Y. Kenny, W. Li, and A. Sudjianto. Algorithmic construction of optimal symmetric Latin hypercube designs. *Journal of statistical planning and inference*, 90 (1):145–159, 2000. doi:10.1016/S0378-3758(00)00105-1.

[93] K. J. Kim and U. M. Diwekar. Hammersley stochastic annealing: efficiency improvement for combinatorial optimization under uncertainty. *IIE Transactions*, 34 (9):761–777, 2002. doi:10.1080/07408170208928910.

[94] J. P. Kleijnen. Design and analysis of Monte Carlo experiments. In *Handbook of Computational Statistics*, pages 529–547. Springer, 2012.

[95] L. Kocis and W. J. Whiten. Computational investigations of low-discrepancy sequences. *ACM Transactions on Mathematical Software (TOMS)*, 23(2):266–294, 1997. doi:10.1145/264029.264064.

[96] J. Koehler and A. Owen. 9 Computer experiments. *Handbook of statistics*, 13: 261–308, 1996. doi:10.1016/S0169-7161(96)13011-X.

[97] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2*, IJCAI'95, pages 1137–1143, San Francisco, CA, USA, 1995. Morgan Kaufmann Publishers Inc. ISBN 1-55860-363-8.

[98] A. R. Krommer and C. W. Ueberhuber. *Numerical Integration: On Advanced Computer Systems*, volume 848. Springer Science & Business Media, 1994. doi:10.1007/BFb0025796.

[99] I. Krykova. *Evaluating of path-dependent securities with low discrepancy methods*. PhD thesis, Worcester Polytechnic Institute, 2003.

[100] S. Kullback. *Information theory and statistics*. Courier Corporation, 1997.

[101] S. Kullback and R. A. Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86, 1951. doi:10.2307/2236703.

[102] G. Lee, J. Mou, and Y. Shen. Sampling strategy design for dimensional measurement of geometric features using coordinate measuring machine. *International Journal of Machine Tools and Manufacture*, 37(7):917–934, 1997. doi:10.1016/S0890-6955(96)00096-X.

[103] C. Lemieux. *Monte Carlo and quasi-Monte Carlo sampling*. Springer Science & Business Media, 2009. doi:10.1007/978-0-387-78165-5.

[104] G. Li, V. Aute, and S. Azarm. An accumulative error based adaptive design of experiments for offline metamodeling. *Structural and Multidisciplinary Optimization*, 40(1-6):137–155, 2010. doi:10.1007/s00158-009-0395-z.

[105] M. Liefvendahl and R. Stocki. A study on algorithms for optimization of Latin hypercubes. *Journal of Statistical Planning and Inference*, 136(9):3231–3247, 2006. doi:10.1016/j.jspi.2005.01.007.

[106] D. V. Lindley. On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, pages 986–1005, 1956.

[107] T. Long, D. Wu, X. Chen, X. Guo, and L. Liu. A deterministic sequential maximin Latin hypercube design method using successive local enumeration for metamodel-based optimization. *Engineering Optimization*, 48(6):1019–1036, 2016. doi:10.1080/0305215X.2015.1081518.

[108] S. Mark and S. Mordechai. Application of Monte Carlo methods in science and engineering, 2011.

[109] M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 42(1):55–61, 1979. doi:10.1080/00401706.1979.10489755.

[110] N. Metropolis and S. Ulam. The Monte Carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949. doi:10.1080/01621459.1949.10483310.

[111] M. D. Morris and T. J. Mitchell. Exploratory designs for computational experiments. *Journal of statistical planning and inference*, 43(3):381–402, 1995. doi:10.1016/0378-3758(94)00035-T.

[112] S. Mosbach and M. Kraft. Influence of experimental observations on n-propylbenzene kinetic parameter estimates. *Proceedings of the Combustion Institute*, 35(1):357–365, 2015. doi:10.1016/j.proci.2014.05.061.

[113] S. Mosbach, A. Braumann, P. L. Man, C. A. Kastner, G. P. Brownbridge, and M. Kraft. Iterative improvement of bayesian parameter estimates for an engine model by means of experimental design. *Combustion and Flame*, 159(3):1303–1313, 2012. doi:10.1016/j.combustflame.2011.10.019.

[114] S. Mosbach, J. H. Hong, G. P. Brownbridge, M. Kraft, S. Gudiyella, and K. Brezinsky. Bayesian error propagation for a kinetic model of n-propylbenzene oxidation in a shock tube. *International Journal of Chemical Kinetics*, 46(7):389–404, 2014. doi:10.1002/kin.20855.

[115] S. Mosbach, W. J. Menz, and M. Kraft. Outlier analysis for a silicon nanoparticle population balance model. *Combustion and Flame*, 177:89 – 97, 2017. ISSN 0010-2180. doi:10.1016/j.combustflame.2016.12.006.

[116] J. Müller and C. A. Shoemaker. Influence of ensemble surrogate models and sampling strategy on the solution quality of algorithms for computationally expensive black-box global optimization problems. *Journal of Global Optimization*, 60(2):123–144, 2014. doi:10.1007/s10898-014-0184-0.

[117] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. John Wiley & Sons, 2016.

[118] H. Niederreiter. Quasi-Monte Carlo methods and pseudo-random numbers. *Bulletin of the American Mathematical Society*, 84(6):957–1041, 1978. doi:10.1090/S0002-9904-1978-14532-7.

[119] H. Niederreiter. Low-discrepancy and low-dispersion sequences. *Journal of number theory*, 30(1):51–70, 1988. doi:10.1016/0022-314X(88)90025-X.

[120] H. Niederreiter. *Random Number Generation and quasi-Monte Carlo Methods*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992. ISBN 0-89871-295-5. doi:10.1137/1.9781611970081.bm.

[121] S. Ninomiya and S. Tezuka. Toward real-time pricing of complex financial derivatives. *Applied Mathematical Finance*, 3(1):1–20, 1996. doi:10.1080/13504869600000001.

[122] A. B. Owen. Orthogonal arrays for computer experiments, integration and visualization. *Statistica Sinica*, pages 439–452, 1992. doi:10.2307/24304869.

[123] A. B. Owen. Controlling correlations in Latin hypercube samples. *Journal of the American Statistical Association*, 89(428):1517–1522, 1994. doi:10.1080/01621459.1994.10476891.

[124] G. Pan, P. Ye, and P. Wang. A novel Latin hypercube algorithm via translational propagation. *The Scientific World Journal*, 2014, 2014. doi:10.1155/2014/163949.

[125] J. S. Park. Optimal Latin-hypercube designs for computer experiments. *Journal of statistical planning and inference*, 39(1):95–111, 1994. doi:10.1016/0378-3758(94)90115-5.

[126] R. Peikert, D. Würtz, M. Monagan, and C. de Groot. *Packing circles in a square: a review and new results*. Springer, 1992.

[127] C. S. Peirce. A theory of probable inference. *Studies in Logic by Members of the Johns Hopkins University*, pages 126–181, 1883. doi:10.1037/12811-007.

[128] N. Pholdee and S. Bureerat. An efficient optimum Latin hypercube sampling technique based on sequencing optimisation using simulated annealing. *International Journal of Systems Science*, 46(10):1780–1789, 2015. doi:10.1080/00207721.2013.835003.

[129] R. L. Plackett and J. P. Burman. The design of optimum multifactorial experiments. *Biometrika*, 33(4):305–325, 1946. doi:10.2307/2332195.

[130] L. Pronzato and W. G. Müller. Design of computer experiments: space filling and beyond. *Statistics and Computing*, 22(3):681–701, 2012. doi:10.1007/s11222-011-9242-3.

[131] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker. Surrogate-based analysis and optimization. *Progress in aerospace sciences*, 41(1): 1–28, 2005. doi:10.1016/j.paerosci.2005.02.001.

[132] M. H. Quenouille. Notes on bias in estimation. *Biometrika*, 43(3/4):353–360, 1956. doi:10.2307/2332914.

[133] D. Raj. Sampling theory. 1968.

[134] R. G. Regis. Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points. *Engineering Optimization*, 46(2):218–243, 2014. doi:10.1080/0305215X.2013.765000.

[135] R. G. Regis and C. A. Shoemaker. Constrained global optimization of expensive black box functions using radial basis functions. *Journal of Global Optimization*, 31(1):153–171, 2005. ISSN 1573-2916. doi:10.1007/s10898-004-0570-0.

[136] C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, 2013. doi:10.1007/978-1-4757-4145-2.

[137] R. Y. Rubinstein and D. P. Kroese. *Simulation and the Monte Carlo method*, volume 707. John Wiley & Sons, 2011.

[138] J. Sacks and S. Schiller. Spatial designs. *Statistical decision theory and related topics IV*, 2(S S):385–399, 1988.

[139] J. Sacks, S. B. Schiller, and W. J. Welch. Designs for computer experiments. *Technometrics*, 31(1):41–47, 1989.

[140] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical science*, pages 409–423, 1989.

[141] T. J. Santner, B. J. Williams, and W. I. Notz. *The design and analysis of computer experiments*. Springer Science & Business Media, 2013. doi:10.1007/978-1-4757-3799-8.

[142] J. Schaer. The densest packing of nine circles in a square. *Canad. Math. Bull*, 8: 273–277, 1965.

[143] J. Schaer. On the densest packing of spheres in a cube. *Canad. Math. Bull*, 9(3), 1966.

[144] P. Sebastiani and H. P. Wynn. Maximum entropy sampling and optimal bayesian experimental design. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(1):145–157, 2000. doi:10.1111/1467-9868.00225.

[145] S. Shan and G. G. Wang. Survey of modeling and optimization strategies to solve high-dimensional design problems with computationally-expensive black-box functions. *Structural and Multidisciplinary Optimization*, 41(2):219–241, 2010. doi:10.1007/s00158-009-0420-2.

[146] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 1948.

[147] M. C. Shewry and H. P. Wynn. Maximum entropy sampling. *Journal of applied statistics*, 14(2):165–170, 1987. doi:10.1080/02664768700000020.

[148] J. J. Sikorski, G. Brownbridge, S. S. Garud, S. Mosbach, I. A. Karimi, and M. Kraft. Parameterisation of a biodiesel plant process flow sheet model. *Computers & Chemical Engineering*, 95:108–122, 2016. doi:10.1016/j.compchemeng.2016.06.019.

[149] T. W. Simpson, T. M. Mauery, J. J. Korte, and F. Mistree. Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA journal*, 39(12):2233–2241, 2001. doi:10.2514/2.1234.

[150] P. Singh, D. Deschrijver, and T. Dhaene. A balanced sequential design strategy for global surrogate modeling. In *Simulation Conference (WSC), 2013 Winter*, pages 2172–2179. IEEE, 2013. doi:10.1109/WSC.2013.6721594.

[151] I. H. Sloan and S. Joe. *Lattice methods for multiple integration*. Oxford University Press, 1994.

[152] I. M. Sobol'. On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 7(4): 784–802, 1967. doi:10.1016/0041-5553(67)90144-9.

[153] I. M. Sobol. Uniformly distributed sequences with an additional uniform property. *USSR Computational Mathematics and Mathematical Physics*, 16(5):236–242, 1976. doi:10.1016/0041-5553(76)90154-3.

[154] I. M. Sobol. *A primer for the Monte Carlo method*. CRC press, 1994.

[155] H. A. Steinberg. Generalized quota sampling. *Nuclear Science and Engineering*, 15(2):142–145, 1963.

[156] S. M. Stigler. Gergonne's 1815 paper on the design and analysis of polynomial regression experiments. *Historia Mathematica*, 1(4):431–439, 1974. doi:10.1016/0315-0860(74)90033-0.

[157] P. G. Szabó, M. C. Markót, T. Csendes, E. Specht, L. G. Casado, and I. García. *New Approaches to Circle Packing in a Square: With Program Codes*, volume 6. Springer Science & Business Media, 2007. doi:10.1007/978-0-387-45676-8.

[158] B. Tang. *Latin hypercubes and supersaturated designs*. 1992.

[159] B. Tang. Orthogonal array-based Latin hypercubes. *Journal of the American statistical association*, 88(424):1392–1397, 1993. doi:10.1080/01621459.1993.10476423.

[160] M. Tatarevic. On limits of dense packing of equal spheres in a cube. *arXiv preprint arXiv:1503.07933*, 2015.

[161] N. Q. Uy, N. X. Hoai, R. I. McKay, and P. M. Tuan. Initialising PSO with randomised low-discrepancy sequences: The comparative results. In *2007 IEEE Congress on Evolutionary Computation*, pages 1985–1992. IEEE, 2007. doi:10.1109/CEC.2007.4424717.

[162] E. R. Van Dam, B. Husslage, D. Den Hertog, and H. Melissen. Maximin Latin hypercube designs in two dimensions. *Operations Research*, 55(1):158–169, 2007. doi:10.1287/opre.1060.0317.

[163] J. van der Corput. Verteilungsfunktionen. i. mitt. *Proc. Akad. Wet. Amsterdam*, 38: 813–821, 1935.

[164] J. van der Herten, I. Couckuyt, D. Deschrijver, and T. Dhaene. A fuzzy hybrid sequential design strategy for global surrogate modeling of high-dimensional computer experiments. *SIAM Journal on Scientific Computing*, 37(2):A1020–A1039, 2015. doi:10.1137/140962437.

[165] F. A. Viana. A tutorial on Latin hypercube design of experiments. *Quality and Reliability Engineering International*, 32(5):1975–1985, 2016. doi:10.1002/qre.1924.

[166] F. A. Viana, G. Venter, and V. Balabanov. An algorithm for fast optimal Latin hypercube design of experiments. *International journal for numerical methods in engineering*, 82(2):135–156, 2010. doi:10.1002/nme.2750.

[167] A. Villagran, G. Huerta, M. Vannucci, C. S. Jackson, and A. Nosedal. Non-parametric sampling approximation via Voronoi tessellations. *Communications in Statistics-Simulation and Computation*, 45(2):717–736, 2016. doi:10.1080/03610918.2013.870798.

[168] G. Voronoï. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième mémoire. recherches sur les parallélloèdres primitifs. *Journal für die reine und angewandte Mathematik*, 134:198–287, 1908.

[169] K. K. Vu, C. D'Ambrosio, Y. Hamadi, and L. Liberti. Surrogate-based methods for black-box optimization. *International Transactions in Operational Research*, 24 (3):393–424, 2017. doi:10.1111/itor.12292.

[170] R. Walzer. Galen, on medical experience. *Philosophy*, 21(78):93–94, 1946.

[171] R. Wang, U. Diwekar, and C. E. Grégoire Padró. Efficient sampling techniques for uncertainties in risk analysis. *Environmental Progress*, 23(2):141–157, 2004. doi:10.1002/ep.10020.

[172] T. T. Warnock. Computational investigations of low-discrepancy point sets. 1971.

[173] H. Wynn. Maximum entropy sampling and general equivalence theory. In *mODa 7Advances in Model-Oriented Design and Analysis*, pages 211–218. Springer, 2004.

[174] F. Xiong, Y. Xiong, W. Chen, and S. Yang. Optimizing Latin hypercube design for sequential sampling of computer experiments. *Engineering Optimization*, 41(8): 793–810, 2009. doi:10.1080/03052150902852999.

[175] S. Xu, H. Liu, X. Wang, and X. Jiang. A robust error-pursuing sequential sampling approach for global metamodeling based on Voronoi diagram and cross validation. *Journal of Mechanical Design*, 136(7):071009, 2014. doi:10.1115/1.4027161.

[176] Q. Zhou, X. Shao, P. Jiang, H. Zhou, and L. Shu. An adaptive global variable fidelity metamodeling strategy using a support vector regression based scaling function. *Simulation Modelling Practice and Theory*, 59:18–35, 2015. doi:10.1016/j.simpat.2015.08.002.

[177] Q. Zhou, X. Shao, P. Jiang, Z. Gao, H. Zhou, and L. Shu. An active learning variable-fidelity metamodelling approach based on ensemble of metamodels and objective-oriented sequential sampling. *Journal of Engineering Design*, 27(4-6): 205–231, 2016. doi:10.1080/09544828.2015.1135236.

[178] H. Zhu, L. Liu, T. Long, and L. Peng. A novel algorithm of maximin Latin hypercube design using successive local enumeration. *Engineering Optimization*, 44(5): 551–564, 2012. doi:10.1080/0305215X.2011.591790.