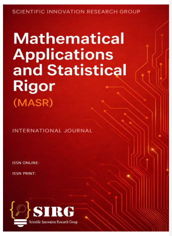

Mathematical Applications and Statistical Rigor
MASR

Journal Homepage
https://pub.scientificirg.com/index.php/MASR/index

Lower Bounds of Unanchored Discrepancy for Mixed-Level U-Type Designs

S.M. Celem[a,1] **and Hong Qin**[a,b]

[a] *School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China*

[b] *Faculty of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China*

ABSTRACT

Uniform U-type designs are fundamental to modern experimental methodology, particularly in computer experiments and robust parameter design, owing to their ability to uniformly distribute points across the design space. This space-filling property confers resilience against model misspecification and provides an ideal foundation for nonparametric estimation. Despite their practical significance, constructing uniform designs that minimize a given discrepancy poses an inherently intractable combinatorial optimization problem. This challenge has motivated the pursuit of theoretical lower bounds, which serve as optimality benchmarks and termination criteria for heuristic search algorithms. Mixed two- and three-level designs are ubiquitous in practical experimental settings, including industrial experiments, clinical trials, and computer simulations. In this paper, we derive sharp, analytically tractable lower bounds for the unanchored discrepancy in mixed two- and three-level U-type designs. Leveraging Jensen's inequality and combinatorial optimization, we obtain closed-form expressions that are both computationally expedient and demonstrably tight. We further delineate necessary conditions under which these bounds are attainable and propose constructive methodologies for generating designs that achieve or approach these theoretical limits. The resulting benchmarks fill a critical gap in the theory of uniform designs and offer tangible utility in constructing optimal experimental plans for computer simulations and robust parameter studies.

PAPER INFORMATION

HISTORY

Received: 15 November 2025
Revised: 27 January 2026
Accepted: 20 February 2026
Online: 1 March 2026

MSC

62K05
62K15

KEYWORDS

Uniform design
Unanchored discrepancy
Lower bound
U-type design
Threshold accepting

[1]Corresponding author at School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430073, China.
E-mail: smcelem@yahoo.com.

1 Introduction

Experimental design constitutes a cornerstone of modern scientific and industrial inquiry, empowering researchers to extract maximal information with minimal resources, conduct rigorous analyses, and formulate defensible conclusions [1, 2, 3]. Across disciplines ranging from pharmaceutical development to engineering simulation, investigators routinely confront systems shaped by the interplay of numerous factors [4, 5, 6]. Factorial designs furnish a systematic architecture for such investigations; however, the exponential proliferation of runs with increasing factor counts and levels rapidly renders complete factorials prohibitive. Fractional factorial designs surmount this impediment by judiciously selecting a subset of runs while retaining essential information [7], thereby striking an efficacious balance between experimental economy and informational yield.

A cardinal requirement for any effective experimental design is the uniform dispersal of points throughout the factor space, ensuring that no region remains uncharted—a property of paramount importance in computer experiments and robust parameter design, where the true underlying model is frequently unknown [8, 9]. Among the panoply of space-filling strategies proposed in the literature [10, 11], uniform designs have garnered particular acclaim for their resilience against model misspecification and their capacity to attain exhaustive coverage with a parsimonious number of points [12, 13]. Foundational theoretical contributions have demonstrated that uniform designs are optimal for approximate integration and exhibit robust performance under diverse modeling assumptions [14, 15, 16].

The quantification of uniformity hinges upon discrepancy measures, which furnish a rigorous quantitative assessment of the divergence between the empirical distribution of design points and the ideal uniform distribution. An array of discrepancy measures has been advanced, encompassing the centered L_2-discrepancy [14], wrap-around L_2-discrepancy [14], discrete discrepancy [17], Lee discrepancy [18], and mixture discrepancy [19]. A substantial body of research has accrued around these metrics over the past decade; a comprehensive synthesis may be found in [20]. Critically, uniformity criteria have been shown to forge connections with classical design paradigms—for instance, they function as scalar surrogates for the venerable aberration criterion in regular fractional factorial designs, while retaining the advantage of applicability to both regular and non-regular constructions [21, 22, 23].

Among the pantheon of discrepancy measures, the unanchored L_2 discrepancy [24] offers a constellation of compelling attributes. In contradistinction to the star discrepancy (anchored at the origin), it remains invariant under the transformation $x \mapsto 1 - x$, thereby embodying a symmetric conception of uniformity. It quantifies the mean squared deviation across all sub-rectangles $[0, \mathbf{u})$, furnishing an intrinsically global appraisal of space-filling. Our investigation focuses specifically on U-type (also called balanced) designs [25], wherein each factor level appears with equal frequency. Such designs possess dual virtues: they constitute natural candidates for experimental plans, and they afford mathematical amenability, as the discrepancy reduces to a closed-form expression predicated solely on pairwise coincidences of factor levels, thereby facilitating efficient evaluation even for designs of substantial size. Furthermore, a profound nexus exists between uniformity (as captured by discrepancy) and the aberration criterion for regular fractional factorial designs, establishing that minimization of discrepancy engenders designs with salutary projective properties.

Notwithstanding these advantages, the construction of uniform U-type designs that minimize a given discrepancy remains a daunting combinatorial optimization challenge, as the design space expands exponentially with the number of factors. Exhaustive enumeration is infeasible for problems of practical scale, necessitating the deployment of heuristic search algorithms. This exigency has spurred the derivation of theoretical lower bounds for discrepancy measures, which fulfill two indispensable functions: (i) they provide a gold standard for adjudicating optimality—a design whose discrepancy attains the lower bound is unequivocally optimal; and (ii) they serve as principled termination criteria for stochastic optimization algorithms—when the discrepancy of the current best design approaches the theoretical lower bound within a prescribed tolerance, the search may be confidently halted, secure in the knowledge that a near-optimal design has been identified [26, 27].

Lower bounds have been established for a variety of discrepancies and design classes [28, 29, 30, 31, 32, 33]. However, comprehensive bounds for the unanchored discrepancy in mixed two- and three-level balanced designs have remained a conspicuous lacuna in the literature—a void that the present investigation seeks to redress. We will elucidate how these bounds can function as benchmarks within threshold accepting [34, 35] and other stochastic search paradigms: by comparing candidate designs against the bound, one can ascertain the opportune moment to terminate optimization, thereby conserving computational resources while guaranteeing high-quality outcomes.

The principal contributions of this work are fourfold. First, we derive simplified analytical expressions for the unanchored discrepancy in mixed two- and three-level U-type designs, elucidating its structural dependence on pattern counts. Second, leveraging Jensen's inequality and combinatorial optimization, we establish novel, easily computable lower bounds that serve as optimality benchmarks. Third, we delineate necessary conditions for the attainment of these bounds and characterize the designs that achieve them. Finally, we expound upon their utility as termination criteria in stochastic optimization.

The remainder of the paper is organized as follows. Section 2 establishes the requisite notation and preliminary concepts. Sections 3 and 4 derive new parsimonious analytical expressions and lower bounds for the unanchored discrepancy in mixed two- and three-level U-type designs, respectively. The practical application of the findings using the threshold accepting with our lower bound termination is discussed in Section 5. Section 6 concludes the paper and outlines promising avenues for future research.

2 Foundational Concepts and Notations

The accurate evaluation of multidimensional integrals is a recurring challenge across numerous scientific disciplines, prominently featuring in computational finance [36, 37, 38, 39], physics, and engineering. In many practical scenarios, analytical solutions are unattainable, necessitating the use of numerical integration techniques. The error incurred by such approximations is intimately connected to the notion of discrepancy—a quantitative measure of how uniformly a set of points is distributed within the unit hypercube $[0,1]^s$. The celebrated Koksma–Hlawka inequality provides a deterministic bound on the integration error for functions of bounded variation in the sense of Hardy and Krause, directly linking the error to the star-discrepancy of the point set. This fundamental result underscores the central role of discrepancy in quasi-Monte Carlo methods and, by extension, in the design of computer experiments.

For two points $x, y \in [0,1]^s$ we write $x \leq y$ if the inequality holds component-wise and set $[0,x)\{y \in [0,1)^s \mid 0 \leq y < x\}$ and $[x,y)\{z \in [0,1)^s \mid x \leq z < y\}$. Given a sequence $(x_1,\ldots,x_N)$ of points in $[0,1)^s$, their *extreme discrepancy* (also known as *unanchored discrepancy*) is defined as

$$Disc(x_1,\ldots,x_N) = \sup_{a,b\in[0,1]^s} \left| \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_{[a,b)}(x_k) - \lambda([a,b]) \right|,$$

where $\mathbf{1}_{[a,b)}$ is the characteristic function of the interval $[a,b)$ and λ denotes the s-dimensional Lebesgue measure. The *star-discrepancy*, a closely related variant, restricts the intervals to those anchored at the origin:

$$Disc^*(x_1,\ldots,x_N) = \sup_{y\in[0,1]^s} \left| \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}_{[0,y)}(x_k) - \lambda([0,y]) \right|.$$

The precise dependence of the approximation error on the star-discrepancy is given by the iconic Koksma–Hlawka inequality. Let f be a function on $[0,1]^s$ with bounded variation $V(f)$ in the sense of Hardy and Krause. Then for any sequence $(x_1,\ldots,x_N)$ of points in $[0,1)^s$ the inequality

$$\left|\frac{1}{N}\sum_{k=1}^{N} f(x_k) - \int_{[0,1]^s} f(x)\,\mathrm{d}x\right| \le V(f)\cdot Disc^*(x_1,\ldots,x_N)$$

holds and the inequality is sharp. Classically, low-discrepancy point sets are used for numerical integration and design of experiments. The construction of point sets with low discrepancy is a classical pursuit, driven by the quest for efficient numerical integration. Ideally, one seeks sequences satisfying the conjectured optimal bound

$$Disc^*(x_1,\ldots,x_N) \le c_s\frac{(\log N)^{s-1}}{N},$$

with a constant c_s depending only on the dimension s. This bound is known to be sharp in dimension one and has been proven for dimension two [41]; in higher dimensions it remains one of the great open problems of discrepancy theory. The dependence on s through the factor $(\log N)^{s-1}$ exemplifies the so-called *curse of dimensionality*: for fixed N, the discrepancy can grow rapidly with s, making the construction of good high-dimensional point sets exceedingly difficult. This challenge motivates the study of structured designs, such as lattice rules, digital nets, and, in the context of experimental design, U-type (balanced) designs.

Mixed-level designs are ubiquitous in practical experimental settings, including industrial experiments, clinical trials, and computer simulations [42, 43, 44]. In this paper we focus on *mixed two- and three-level balanced designs*, denoted by $\mathcal{U}(n;2^{m_1}\times 3^{m_2})$. These designs consist of n experimental runs and $m = m_1 + m_2$ factors, where m_1 factors have two levels (coded as 0 and 1) and m_2 factors have three levels (coded as $0,1,2$). The defining property of balance requires that each two-level column contain exactly $n/2$ zeros and $n/2$ ones, and each three-level column contain exactly $n/3$ of each level. Consequently, n must be a multiple of six; we write $n = 6r$ for some integer r. Such mixed-level designs are ubiquitous in practical experimentation, appearing in industrial experiments, clinical trials, manufacturing quality control, and computer simulations, where factors with different numbers of levels naturally arise.

To connect the discrete design with the continuous discrepancy measure, we map each factor level to a point in the unit interval via the linear transformation

$$f_s : x \mapsto \frac{2x+1}{2s}, \quad s = 2,3.$$

For two-level factors, the transformed coordinates are $\frac{1}{4}$ and $\frac{3}{4}$; for three-level factors, they are $\frac{1}{6}$, $\frac{1}{2}$, and $\frac{5}{6}$. The entire design is thereby embedded in the m-dimensional unit hypercube $[0,1]^m$.

The extreme discrepancy is also known under different names such as unanchored discrepancy, see e.g. [40]. The extreme discrepancy, which we shall henceforth denote by $UDisc$ (unanchored discrepancy), possesses the desirable property of symmetry: it remains invariant under the transformation $x \mapsto 1-x$. This symmetry renders it particularly suitable for assessing the uniformity of designs used in computer experiments, where no natural origin exists. The unanchored L_2 discrepancy of a design $\mathbf{d}$ with transformed points $\mathbf{x}_i = (x_{i1},\ldots,x_{im})$ admits a closed-form expression derived by [40]:

$$[UDisc(\mathbf{d})]^2 = \left(\frac{1}{12}\right)^m - \frac{2}{n}\sum_{i=1}^{n}\prod_{k=1}^{m}\frac{x_{ik}(1-x_{ik})}{2} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\big(\min(x_{ik},x_{jk}) - x_{ik}x_{jk}\big). \tag{1}$$

This formula, while explicit, involves products over all factors and becomes unwieldy for direct optimization. However, for balanced designs, the discrete nature of the transformed coordinates allows substantial simplification, expressing the discrepancy in terms of simple pattern counts that capture coincidences between rows. Such simplifications are essential for both theoretical analysis and practical computation.

A fundamental goal in uniform design theory is to identify designs that minimize a given discrepancy. The optimal design, if it exists, achieves the theoretical lower bound of the discrepancy over the class of admissible designs. Establishing such lower bounds serves two intertwined purposes. First, they provide an absolute benchmark: a design whose discrepancy equals the lower bound is provably optimal, and any design with discrepancy close to the bound is near-optimal. Second, these bounds act as principled termination criteria for heuristic search algorithms; when the discrepancy of the current best design falls within a prescribed tolerance of the lower bound, the search can be halted with confidence, conserving computational resources without sacrificing quality [26, 27].

The derivation of sharp lower bounds for the unanchored discrepancy in mixed two- and three-level designs relies on two powerful mathematical tools. The first is Jensen's inequality, applied to convex functions—here the exponential function—to bound averages of exponentials by exponentials of averages.

Lemma 1 (Convexity and Jensen's Inequality)**.** *Let Y be a random variable supported on a finite set $\mathcal{Y}$, and let $\phi : \mathcal{Y} \to \mathbb{R}$ be an arbitrary function. Define $h(y) = \exp(\phi(y))$. Then for any probability distribution on $\mathcal{Y}$, Jensen's inequality yields*

$$\mathbb{E}[h(Y)] \geq \exp(\mathbb{E}[\phi(Y)]).$$

Equality obtains if and only if $\phi(Y)$ is almost surely constant.

The second tool is a combinatorial optimization lemma that provides a tight lower bound for sums of powers of integers under a fixed total. This lemma, adapted from [45], respects the integrality constraints inherent in the pattern counts and yields bounds that are at least as sharp as those obtained by Jensen's inequality alone.

Lemma 2. *[45] Let $z_1, \dots, z_N$ be non-negative integers satisfying the fixed total sum $\sum_{i=1}^{N} z_i = C$. Then for any $\gamma > 0$,*

$$\sum_{i=1}^{N} \gamma^{z_i} \geq \gamma^{\sigma}(p + q\gamma),$$

where $\sigma = \lfloor C/N \rfloor$, and p, q are non-negative integers uniquely determined by the conditions $p + q = N$ and $p\sigma + q(\sigma + 1) = C$. Equality is attained if and only if exactly p of the z_i equal σ and the remaining q equal $\sigma + 1$.

In the subsequent sections, we shall deploy these lemmas repeatedly to bound the various sums that appear in the simplified expression of the unanchored discrepancy. By doing so, we will obtain explicit, easily computable lower bounds that serve as benchmarks for the optimality of mixed two- and three-level U-type designs.

3 Simplified Analytical Formulation of the Unanchored Discrepancy

The general expression (1) for the unanchored discrepancy, while explicit, involves products over all factors and is therefore cumbersome for both theoretical analysis and practical optimization. For balanced mixed-level designs, however, the discrete nature of the transformed coordinates permits a remarkable simplification: the discrepancy can be expressed solely in terms of elementary pattern counts that capture the coincidences between rows. This simplification not only reveals the intrinsic combinatorial structure of the discrepancy but also renders it amenable to the application of convexity and combinatorial optimization techniques. In this section we derive this simplified closed-form expression, thereby laying the groundwork for the lower bounds presented in Section 4.

3.1 Row Profiles and Pattern Counts

For each experimental run i ($1 \le i \le n$), we characterize the configuration of three-level factors through two fundamental counts:

$$\lambda_i = \#\{\text{three-level factors in row } i \text{ at the middle level } 1\},$$

$$\tau_i = \#\{\text{three-level factors in row } i \text{ at the outer levels 0 or 2}\}.$$

These quantities satisfy the elementary relation $\lambda_i + \tau_i = m_2$ for every i. The balance conditions impose the following aggregate constraints:

$$\sum_{i=1}^{n} \lambda_i = \frac{m_2 n}{3}, \qquad \sum_{i=1}^{n} \tau_i = \frac{2m_2 n}{3}. \tag{2}$$

Thus the average number of middle-level entries per row is $\bar{\lambda} = m_2/3$, while the average number of outer-level entries is $2m_2/3$. These averages will appear later in the Jensen bound.

For two rows i and j (possibly equal), we require additional counts that capture coincidences separately for two- and three-level factors.

Two-level coincidences. Let θ_{ij} denote the number of two-level factors for which rows i and j share the same level. Then θ_{ij} is an integer between 0 and m_1, with $\theta_{ii} = m_1$. The balance of two-level columns implies that for any fixed row i, the number of other rows that agree with it in a given two-level column is $n/2 - 1$; summing over all columns yields

$$\sum_{j \ne i} \theta_{ij} = m_1\left(\frac{n}{2} - 1\right).$$

Three-level coincidences. For three-level factors we need four distinct counts that encapsulate the various forms of agreement and disagreement:

$$\begin{aligned}
\lambda_{ij} &= \#\{\text{factors where both rows have middle level } 1\},\\
\tau_{ij} &= \#\{\text{factors where both rows have outer levels (both 0 or both 2)}\},\\
\delta_{ij} &= \#\{\text{factors where one row is middle and the other is outer}\},\\
\sigma_{ij} &= \#\{\text{factors where one row is 0 and the other is 2}\}.
\end{aligned}$$

These satisfy $\lambda_{ij} + \tau_{ij} + \delta_{ij} + \sigma_{ij} = m_2$ for all i, j. For the diagonal $i = j$, we have $\delta_{ii} = \sigma_{ii} = 0$ and $\tau_{ii} = m_2 - \lambda_i$ with $\lambda_{ii} = \lambda_i$. The balance conditions for three-level factors give the following totals (analogous to those used in the pure three-level case):

$$\sum_{i \ne j} \lambda_{ij} = 2m_2\frac{n}{3}\left(\frac{n}{3} - 1\right), \quad \sum_{i \ne j} \delta_{ij} = 4m_2\left(\frac{n}{3}\right)^2, \quad \sum_{i \ne j} \tau_{ij} = 2m_2\frac{n}{3}\left(\frac{n}{3} - 1\right).$$

3.2 Kernel Values and Product Decomposition

For a two-level factor, the kernel $\frac{x(1-x)}{2}$ is constant: $\frac{3}{32}$ for both levels. For a three-level factor, the value depends on whether the level is middle ($x = \frac{1}{2}$) or outer ($x = \frac{1}{6}$ or $\frac{5}{6}$):

$$\frac{x(1-x)}{2} = \begin{cases} \frac{5}{72}, & x \in \{\frac{1}{6}, \frac{5}{6}\}, \\ \frac{9}{72}, & x = \frac{1}{2}. \end{cases}$$

Consequently, the product over all factors for a given row i can be written as

$$\prod_{k=1}^{m} \frac{x_{ik}(1-x_{ik})}{2} = \left(\frac{3}{32}\right)^{m_1} \left(\frac{5}{72}\right)^{\tau_i} \left(\frac{9}{72}\right)^{\lambda_i} = \left(\frac{3}{32}\right)^{m_1} \left(\frac{1}{72}\right)^{m_2} 5^{\tau_i} 9^{\lambda_i}. \tag{3}$$

For the pairwise kernel $\min(x, y) - xy$, the two-level contribution is $\frac{3}{16}$ for a match and $\frac{1}{16}$ for a mismatch. Hence

$$\prod_{\text{two-level}} (\min(x_{ik}, x_{jk}) - x_{ik}x_{jk}) = \left(\frac{1}{16}\right)^{m_1} 3^{\theta_{ij}}. \tag{4}$$

For three-level factors, the kernel values for the four coincidence types are $\frac{9}{36}$ (both middle), $\frac{5}{36}$ (both outer), $\frac{3}{36}$ (one middle, one outer), and $\frac{1}{36}$ (opposite outer). Therefore,

$$\prod_{\text{three-level}} (\min(x_{ik}, x_{jk}) - x_{ik}x_{jk}) = \left(\frac{1}{36}\right)^{m_2} 9^{\lambda_{ij}} 5^{\tau_{ij}} 3^{\delta_{ij}}. \tag{5}$$

For distinct rows $i \neq j$, combining (4) and (5) gives

$$\prod_{k=1}^{m} (\min(x_{ik}, x_{jk}) - x_{ik}x_{jk}) = \left(\frac{1}{16}\right)^{m_1} \left(\frac{1}{36}\right)^{m_2} 3^{\theta_{ij}+\delta_{ij}} 9^{\lambda_{ij}} 5^{\tau_{ij}}.$$

For the diagonal case $i = j$, we have $\theta_{ii} = m_1$, $\lambda_{ii} = \lambda_i$, $\tau_{ii} = m_2 - \lambda_i$, $\delta_{ii} = 0$, $\sigma_{ii} = 0$; thus

$$\prod_{k=1}^{m} (\min(x_{ik}, x_{ik}) - x_{ik}^2) = \left(\frac{3}{16}\right)^{m_1} \left(\frac{1}{36}\right)^{m_2} 9^{\lambda_i} 5^{\tau_i}. \tag{6}$$

3.3 Consolidated Discrepancy Formula

Insert (3), (6), and the off-diagonal expression into the general discrepancy formula (1). Introduce the constants

$$A = \left(\frac{3}{32}\right)^{m_1} \left(\frac{1}{72}\right)^{m_2}, \quad B = \left(\frac{3}{16}\right)^{m_1} \left(\frac{1}{36}\right)^{m_2}, \quad C = \left(\frac{1}{16}\right)^{m_1} \left(\frac{1}{36}\right)^{m_2}.$$

Then

$$\begin{aligned} [UDisc(\mathbf{d})]^2 &= \left(\frac{1}{12}\right)^m - \frac{2}{n} A \sum_{i=1}^{n} 5^{\tau_i} 9^{\lambda_i} + \frac{1}{n^2} B \sum_{i=1}^{n} 5^{\tau_i} 9^{\lambda_i} + \frac{1}{n^2} C \sum_{i \neq j} 3^{\theta_{ij}+\delta_{ij}} 9^{\lambda_{ij}} 5^{\tau_{ij}} \\ &= \left(\frac{1}{12}\right)^m + \frac{B - 2nA}{n^2} \sum_{i=1}^{n} 5^{\tau_i} 9^{\lambda_i} + \frac{C}{n^2} \sum_{i \neq j} 3^{\theta_{ij}+\delta_{ij}} 9^{\lambda_{ij}} 5^{\tau_{ij}}. \end{aligned} \tag{7}$$

This representation isolates the contribution of the diagonal elements (the sum over i) from that of the off-diagonal pairs. The diagonal sum involves only the row profiles λ_i, τ_i, while the off-diagonal sum involves the pairwise coincidence counts. The factor $\frac{B-2nA}{n^2}$ will be denoted by α; its sign plays a critical role in the subsequent derivation of lower bounds.

A direct computation yields

$$B - 2nA = 3^{m_1-2m_2}2^{-(5m_1+3m_2)}\left(2^m - 2n\right),$$

where $m = m_1 + m_2$. Hence

$$\alpha := \frac{B-2nA}{n^2} = \frac{3^{m_1-2m_2}2^{-(5m_1+3m_2)}}{n^2}\left(2^m - 2n\right). \tag{8}$$

The sign of α is governed solely by the factor $2^m - 2n$: $\alpha \geq 0$ iff $2^m \geq 2n$, and $\alpha < 0$ otherwise. This dichotomy will dictate whether we need a lower or an upper bound for the diagonal sum.

For notational convenience, we also set

$$D = \sum_{i=1}^{n}\left(\frac{9}{5}\right)^{\lambda_i},$$

noting that $\sum_{i=1}^{n} 5^{\tau_i}9^{\lambda_i} = 5^{m_2}D$. Then (7) becomes

$$[UDisc(\mathbf{d})]^2 = \left(\frac{1}{12}\right)^m + \alpha\,5^{m_2}D + \frac{C}{n^2}\sum_{i\neq j}3^{\theta_{ij}+\delta_{ij}}9^{\lambda_{ij}}5^{\tau_{ij}}. \tag{9}$$

Equation (9) is the fundamental simplified expression upon which all subsequent analysis rests. The diagonal sum D depends only on the distribution of the middle level among three-level factors, while the off-diagonal sum captures the full pairwise coincidence structure. In the next section we shall derive sharp lower bounds for both components, leading to explicit lower bounds for the overall discrepancy.

Remark 1 (Relation to pure-level designs)**.** *If $m_2 = 0$ (pure two-level case), then $5^{m_2} = 1$ and $D = \sum(9/5)^0 = n$. Moreover, the off-diagonal sum reduces to $\sum_{i\neq j}3^{\theta_{ij}}$ and the constants simplify accordingly to two-level designs. Similarly, if $m_1 = 0$ (pure three-level case), the two-level factors disappear and the formulas reduce to three-level designs.*

Theorem 1 (Exact discrepancy formula)**.** *For any balanced design $\mathbf{d} \in \mathcal{U}(n; 2^{m_1} \times 3^{m_2})$,*

$$[UDisc(\mathbf{d})]^2 = \left(\frac{1}{12}\right)^m + \alpha\,5^{m_2}D + \frac{C}{n^2}\sum_{i\neq j}3^{\theta_{ij}+\delta_{ij}}9^{\lambda_{ij}}5^{\tau_{ij}},$$

with α, C, D, and the pattern counts defined as above.

Proof. The derivation is contained in the preceding discussion, culminating in (9). □

This theorem provides a compact, interpretable formula that isolates the key combinatorial ingredients governing the unanchored discrepancy. In the following section we shall exploit this structure to obtain sharp lower bounds.

4 Sharp Lower Bounds for the Unanchored Discrepancy

Building upon the simplified expression (9) derived in the previous section, we now establish explicit lower bounds for the unanchored discrepancy of mixed two- and three-level balanced designs. The derivation proceeds in two stages: first, we obtain a lower bound for the off-diagonal sum using

Jensen's inequality; second, we derive appropriate bounds for the diagonal sum D, the form of which depends critically on the sign of the coefficient α. Combining these components yields two families of lower bounds—a Jensen-based bound and a sharper combinatorial bound—which together provide the tightest possible benchmarks for design optimality. We then demonstrate how these bounds can be integrated into stochastic search algorithms, specifically threshold accepting, to serve as principled termination criteria.

4.1 The Coefficient α and Its Implications

Recall from (8) the definition

$$\alpha = \frac{B - 2nA}{n^2} = \frac{3^{m_1 - 2m_2} 2^{-(5m_1 + 3m_2)}}{n^2} (2^m - 2n), \qquad m = m_1 + m_2.$$

The sign of α is determined solely by the factor $2^m - 2n$:

$$\alpha \geq 0 \iff 2^m \geq 2n, \qquad \alpha < 0 \iff 2^m < 2n.$$

This dichotomy is fundamental: when α is non-negative, the term αD in (9) increases with D; hence a lower bound for D yields a lower bound for the overall discrepancy. Conversely, when α is negative, αD decreases with D, so an upper bound for D is required to obtain a lower bound for $[UDisc]^2$. Both regimes are treated separately in the following subsections.

4.2 Bounding the Off-Diagonal Sum

For each ordered pair of distinct rows, define the exponent

$$E_{ij} = (\theta_{ij} + \delta_{ij}) \ln 3 + \lambda_{ij} \ln 9 + \tau_{ij} \ln 5.$$

The off-diagonal sum appearing in (9) is then $S_{\text{off}} = \sum_{i \neq j} e^{E_{ij}}$. Because the exponential function is convex, Lemma 1 (Jensen's inequality) directly gives

$$\frac{1}{n(n-1)} \sum_{i \neq j} e^{E_{ij}} \geq \exp\left(\frac{1}{n(n-1)} \sum_{i \neq j} E_{ij} \right).$$

To evaluate the total $\sum_{i \neq j} E_{ij}$, we invoke the balance conditions. For two-level factors, summing over all ordered pairs yields

$$\sum_{i \neq j} \theta_{ij} = m_1 n \left(\frac{n}{2} - 1 \right).$$

For three-level factors, the totals are

$$\sum_{i \neq j} \lambda_{ij} = 2m_2 \frac{n}{3} \left(\frac{n}{3} - 1 \right), \quad \sum_{i \neq j} \delta_{ij} = 4m_2 \left(\frac{n}{3} \right)^2, \quad \sum_{i \neq j} \tau_{ij} = 2m_2 \frac{n}{3} \left(\frac{n}{3} - 1 \right).$$

Combining these, we obtain

$$\sum_{i \neq j} (\theta_{ij} + \delta_{ij}) = m_1 n \left(\frac{n}{2} - 1 \right) + \frac{4m_2 n^2}{9},$$

and consequently

$$T := \sum_{i \neq j} E_{ij} = \left[m_1 n \left(\frac{n}{2} - 1 \right) + \frac{4m_2 n^2}{9} \right] \ln 3 + \frac{2m_2 n}{3} \left(\frac{n}{3} - 1 \right) (\ln 9 + \ln 5).$$

Simplifying $\ln 9 + \ln 5 = \ln 45$ gives

$$T = m_1 n \left(\frac{n}{2} - 1\right) \ln 3 + \frac{4m_2 n^2}{9} \ln 3 + \frac{2m_2 n}{3} \left(\frac{n}{3} - 1\right) \ln 45. \tag{10}$$

The average exponent is $\bar{E} = T/[n(n-1)]$. Substituting into Jensen's inequality yields the universal lower bound

$$S_{\text{off}} \geq n(n-1)\, e^{\bar{E}}. \tag{11}$$

This bound is valid for all balanced designs, irrespective of the sign of α. It depends only on the design parameters n, m_1, m_2 and the known constants $\ln 3$ and $\ln 45$; it is therefore easily computable once these parameters are fixed.

4.3 Bounds for the Diagonal Sum D

The diagonal sum $D = \sum_{i=1}^{n} (9/5)^{\lambda_i}$ involves the integers λ_i that count the number of middle-level three-level factors in each row. These integers satisfy the linear constraint $\sum_{i=1}^{n} \lambda_i = m_2 n/3$. The function $f(x) = (9/5)^x$ is strictly convex, as its second derivative is positive. The appropriate bound for D depends on the sign of α.

4.3.1 Case $\alpha \geq 0$: Lower Bounds for D

When α is non-negative, a lower bound for D translates directly into a lower bound for the overall discrepancy. The simplest such bound follows from Jensen's inequality applied to the convex function f:

$$D \geq n \left(\frac{9}{5}\right)^{\bar{\lambda}}, \qquad \bar{\lambda} = \frac{m_2}{3}. \tag{12}$$

While this bound is easy to compute, it ignores the integrality of the λ_i. A sharper bound that respects the discrete nature of the counts is obtained via Lemma 2. Let

$$\psi = \left\lfloor \frac{m_2}{3} \right\rfloor,$$

and let p, q be the unique non-negative integers satisfying

$$p + q = n, \qquad p\psi + q(\psi + 1) = \frac{m_2 n}{3}.$$

Explicitly, $q = m_2 n/3 - n\psi$ and $p = n - q$. Then Lemma 2 with $\gamma = 9/5$ yields

$$D \geq \left(\frac{9}{5}\right)^{\psi} \left(p + q\, \tfrac{9}{5}\right). \tag{13}$$

The combinatorial bound (13) is always at least as large as the Jensen bound (12), with equality if and only if $m_2/3$ is an integer (i.e., when $\psi = m_2/3$). This bound is attainable only when the λ_i are exactly ψ or $\psi + 1$ with frequencies p, q.

4.3.2 Case $\alpha < 0$: Upper Bound for D

When α is negative, the term αD decreases as D increases; therefore an upper bound for D is required to obtain a lower bound for $[UDisc]^2$. The maximum of D under the sum constraint $\sum \lambda_i = m_2 n/3$

and the natural bounds $0 \le \lambda_i \le m_2$ is achieved by concentrating the total into as few rows as possible. Setting exactly $n/3$ rows to the maximal possible value m_2 and the remaining $2n/3$ rows to zero gives

$$D \le \frac{n}{3}\left(\frac{9}{5}\right)^{m_2} + \frac{2n}{3}. \tag{14}$$

This upper bound is attainable when precisely $n/3$ rows consist entirely of middle-level three-level factors and the other rows contain no middle-level factors. It represents the extreme opposite of the equal distribution scenario and provides the necessary counterpart to the lower bounds used in the previous case.

4.4 Main Lower Bound Theorem

We now assemble the bounds derived above to state the principal result of this paper.

Theorem 2 (Lower bounds for mixed two- and three-level designs)**.** *Let* $\mathbf{d} \in \mathcal{U}(n; 2^{m_1} \times 3^{m_2})$ *be a balanced design. Define* $m = m_1 + m_2$*,* α *as in* (8)*,* C *as in Theorem 1, and let* $\bar{E} = T/[n(n-1)]$ *be given by* (10) *divided by* $n(n-1)$*. Set* $\bar{\lambda} = m_2/3$*,* $\psi = \lfloor m_2/3 \rfloor$*, and let* p, q *be the unique non-negative integers satisfying* $p + q = n$ *and* $p\psi + q(\psi + 1) = m_2 n/3$*. Then the squared unanchored discrepancy satisfies the following bounds.*

Case I: $2^m \ge 2n$ *(*$\alpha \ge 0$*)**.*

$$[UDisc(\mathbf{d})]^2 \ge \mathcal{LB}_{Jensen} = \left(\frac{1}{12}\right)^m + \alpha\, 5^{m_2} n \left(\frac{9}{5}\right)^{\bar{\lambda}} + \frac{C}{n^2}\, n(n-1)e^{\bar{E}},$$

$$[UDisc(\mathbf{d})]^2 \ge \mathcal{LB}_{Comb} = \left(\frac{1}{12}\right)^m + \alpha\, 5^{m_2} \left(\frac{9}{5}\right)^{\psi} \left(p + q\,\tfrac{9}{5}\right) + \frac{C}{n^2}\, n(n-1)e^{\bar{E}}.$$

Moreover, $\mathcal{LB}_{Comb} \ge \mathcal{LB}_{Jensen}$*, with equality iff* $m_2/3$ *is an integer (see Figure 1).*

Case II: $2^m < 2n$ *(*$\alpha < 0$*)**.*

$$[UDisc(\mathbf{d})]^2 \ge \mathcal{LB}_{neg} = \left(\frac{1}{12}\right)^m + \alpha\, 5^{m_2} \left[\frac{n}{3}\left(\frac{9}{5}\right)^{m_2} + \frac{2n}{3}\right] + \frac{C}{n^2}\, n(n-1)e^{\bar{E}}.$$

Proof. Starting from the compact expression (9), we substitute the appropriate bound for D:

- For Case I, use the Jensen bound (12) to obtain $\mathcal{LB}_{\text{Jensen}}$, or the combinatorial bound (13) to obtain $\mathcal{LB}_{\text{Comb}}$.
- For Case II, use the upper bound (14) to obtain $\mathcal{LB}_{\text{neg}}$.

In all cases, the off-diagonal sum is replaced by its lower bound (11). The factor 5^{m_2} multiplies D, and the term $\frac{C}{n^2} S_{\text{off}}$ becomes $\frac{C}{n^2} n(n-1)e^{\bar{E}}$. The three formulas follow immediately. □

The theorem provides a complete set of lower bounds covering all admissible parameter combinations. For parameters satisfying $2^m \ge 2n$, the combinatorial bound $\mathcal{LB}_{\text{Comb}}$ is the tighter of the two and should be used as the benchmark. When $2^m < 2n$, the bound $\mathcal{LB}_{\text{neg}}$ is the appropriate benchmark; note that in this regime the combinatorial lemma is not applicable because we require an upper bound for D rather than a lower bound. The bounds derived in this theorem are not only theoretically significant but also computationally undemanding, requiring only elementary arithmetic and exponentials. They provide practitioners with a ready means to assess the quality of any mixed two- and three-level design and to terminate search algorithms once a near-optimal design has been found.

Corollary 1 (Necessary conditions for optimality)**.**

- *If a design attains $\mathcal{LB}_{Comb}$ in Case I, then:*
    1. *All diagonal counts λ_i are either ψ or $\psi + 1$, with exactly p rows having ψ and q rows having $\psi + 1$;*
    2. *All off-diagonal exponents E_{ij} are equal.*
- *If a design attains $\mathcal{LB}_{neg}$ in Case II, then:*
    1. *Exactly $n/3$ rows satisfy $\lambda_i = m_2$ and the remaining $2n/3$ rows satisfy $\lambda_i = 0$;*
    2. *All off-diagonal exponents E_{ij} are equal.*

Proof. The conditions follow from the equality cases of Lemma 2 (for the diagonal bound in Case I) and of Jensen's inequality (for the off-diagonal bound in both cases). Equality in Jensen's inequality for the exponential function requires all arguments E_{ij} to be identical. In Case II, the upper bound (14) is attained only under the extremal distribution described. □

Corollary 2 (Special case: pure two-level designs ($m_2 = 0$))**.** *When $m_2 = 0$, we have $m = m_1$, $5^{m_2} = 1$, $D = n$, and the off-diagonal sum simplifies. Then $\mathcal{LB}_{Comb}$ reduces to the combinatorial lower bound for two-level designs, and the condition $\alpha \geq 0$ becomes $2^{m_1} \geq 2n$.*

Corollary 3 (Special case: pure three-level designs ($m_1 = 0$))**.** *When $m_1 = 0$, we have $m = m_2$, and the bounds reduce to those for three-level designs. In particular, $\mathcal{LB}_{Comb}$ coincides with the combinatorial bound for three-level designs, providing a sharp benchmark.*

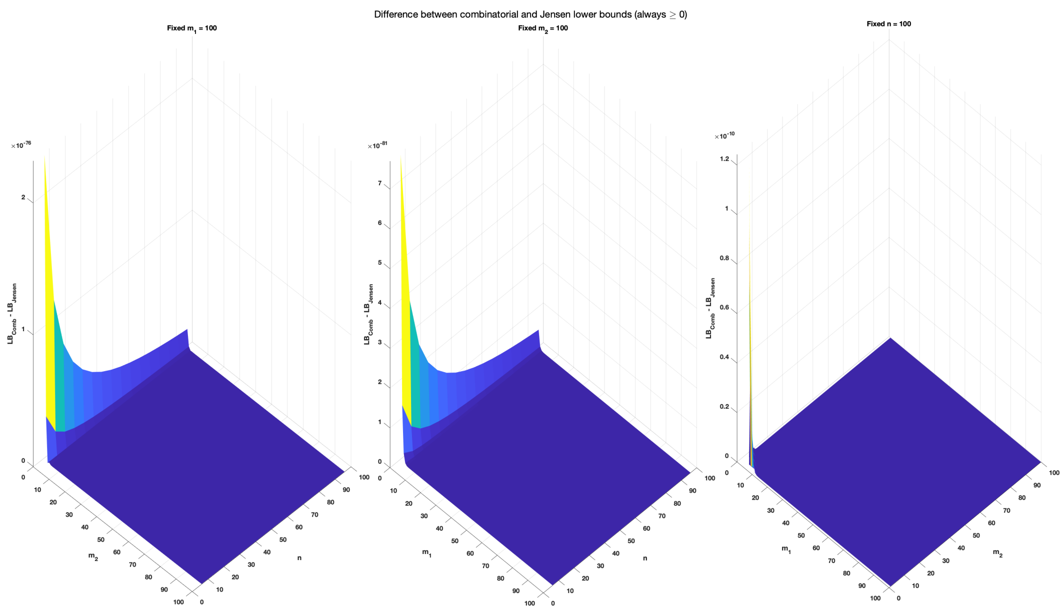

Figure 1: Comparison of the combinatorial lower bound $\mathcal{LB}_{\text{Comb}}$ and the Jensen lower bound $\mathcal{LB}_{\text{Jensen}}$ for mixed two- and three-level designs. The combinatorial bound is always at least as tight.

5 Practical Application: Threshold Accepting with Lower Bound Termination

The theoretical lower bounds established above are not merely abstract benchmarks; they serve a crucial practical function as termination criteria for stochastic optimization algorithms. Among the most effective heuristics for constructing uniform designs is the *threshold accepting* algorithm [34, 35], a powerful global optimization method that permits occasional uphill moves to escape local minima. The core innovation of threshold accepting lies in its acceptance rule: a candidate design is accepted if its discrepancy does not exceed that of the current design by more than a gradually decreasing threshold. This controlled acceptance of inferior solutions enables the algorithm to explore the design space thoroughly while eventually converging to a high-quality design.

The availability of a sharp lower bound $\mathcal{LB}(m_1, m_2, n)$ transforms threshold accepting from a heuristic into a principled optimization tool. By comparing the discrepancy of the current best design with the theoretical lower bound, we can determine when further search is unlikely to yield meaningful improvements. Specifically, once the discrepancy falls within a pre-specified tolerance ε of the lower bound, the algorithm can be terminated with the assurance that the obtained design is near-optimal, with efficiency $F \geq 1 - \varepsilon/\mathcal{LB}$. This termination criterion confers two significant advantages: it drastically reduces computational cost by avoiding unnecessary iterations, and it provides a quality guarantee that is absent in purely heuristic approaches.

Algorithm 1 presents the pseudo-code of a threshold accepting algorithm tailored to mixed two- and three-level designs, where the lower bound $\mathcal{LB}(m_1, m_2, n)$ serves as the benchmark. The algorithm commences with a randomly generated balanced design and iteratively generates neighbours through column-wise pairwise exchanges [46]. A move is accepted if the increment in discrepancy remains below a current threshold τ, which is progressively attenuated following a cooling schedule. Termination occurs either when the maximum number of iterations is exhausted or when the discrepancy of the best design lies within a fraction ε of the theoretical lower bound.

The integration of the lower bound into the stopping condition markedly curtails computational expenditure, particularly for large m_1, m_2 and n, by obviating superfluous iterations once a high-quality design has emerged. Moreover, the bound confers a quality guarantee: the final design attains an efficiency $F \geq 1-\varepsilon/\mathcal{LB}$. This combination of efficiency and assurance makes the algorithm eminently suitable for practical applications where optimal or near-optimal designs are required.

The lower bounds derived in this section, together with their algorithmic implementation, provide a complete framework for constructing uniform mixed two- and three-level designs. In the next section, we present numerical examples that illustrate the tightness of these bounds and demonstrate the effectiveness of the threshold accepting algorithm.

Algorithm 1 Threshold Accepting for Mixed Two- and Three-Level Uniform Designs

Require: Design parameters n, m_1, m_2, lower bound $\mathcal{LB} = \mathcal{LB}(m_1, m_2, n)$, initial threshold $\tau_0 > 0$, reduction factor $\rho \in (0, 1)$, tolerance $\varepsilon > 0$, max iterations $T_{\max}$

Ensure: A design $\mathbf{d}^*$ with low unanchored discrepancy

Generate a random balanced design $\mathbf{d} \in \mathcal{U}(n; 2^{m_1} \times 3^{m_2})$
Compute current discrepancy $D_{\text{curr}} = [UDisc(\mathbf{d})]^2$
Set $\mathbf{d}^* = \mathbf{d}$, $D_{\text{best}} = D_{\text{curr}}$
Set threshold $\tau = \tau_0$
for $t = 1$ to $T_{\max}$ **do**
  Generate a neighbour $\mathbf{d}'$ by random column-wise pairwise exchange
  Compute $D' = [UDisc(\mathbf{d}')]^2$
  if $D' - D_{\text{curr}} < \tau$ **then**
    Accept the move: $\mathbf{d} = \mathbf{d}'$, $D_{\text{curr}} = D'$
    if $D' < D_{\text{best}}$ **then**
      $\mathbf{d}^* = \mathbf{d}'$, $D_{\text{best}} = D'$
    end if
  end if
  Update threshold: $\tau = \rho \cdot \tau$
  if $D_{\text{best}} - \mathcal{LB} < \varepsilon$ **then**
    break ▷ Terminate: design is nearly optimal
  end if
end forreturn $\mathbf{d}^*$

6 Conclusions, Practical Implications and Future Work

In this investigation we have established novel theoretical lower bounds for the unanchored L_2 discrepancy within the class of mixed two- and three-level balanced (U-type) designs. Commencing from the canonical analytical expression for the discrepancy, we exploited the discrete nature of the transformed coordinates together with the intrinsic balance conditions to derive a parsimonious formulation tailored to the mixed-level context. This formulation, presented in Theorem 1, articulates the discrepancy in terms of pattern counts that encapsulate coincidences for two-level factors (θ_{ij}) and the fourfold typology of coincidences for three-level factors ($\lambda_{ij}, \tau_{ij}, \delta_{ij}, \sigma_{ij}$), augmented by the diagonal counts λ_i that capture the prevalence of the middle level in each row. This simplified expression reveals the intrinsic combinatorial structure of the discrepancy and renders it amenable to the application of convexity and combinatorial optimization techniques.

6.1 *Summary of Principal Contributions*

Leveraging Jensen's inequality (Lemma 1) in conjunction with a powerful combinatorial optimization lemma (Lemma 2) adapted from [45], we derived two families of lower bounds. The first, denoted $\mathcal{LB}_{\text{Jensen}}$, emanates directly from the convexity of the underlying exponential functions and provides a simple, easily computable benchmark. The second, $\mathcal{LB}_{\text{Comb}}$, is demonstrably sharper as it respects the integrality of the pattern counts; it is obtained by applying the combinatorial lemma to the diagonal sum $D = \sum(9/5)^{\lambda_i}$ after an appropriate reparameterization. The sign of the coefficient $\alpha = (B - 2nA)/n^2$ dictates whether a lower or an upper bound for D is requisite, and we have furnished explicit expressions for both regimes in Theorem 2. For any prescribed parameter triple (n, m_1, m_2), the final lower bound is taken as the maximum of the available candidates, thereby guaranteeing the tightest possible benchmark.

The principal contributions of this work can be summarized as follows:

- *Exact simplified expression:* Theorem 1 provides a compact, interpretable formula for the unanchored discrepancy that isolates the diagonal and off-diagonal contributions, revealing the fundamental combinatorial quantities that govern uniformity.
- *Sharp lower bounds:* Theorem 2 presents two families of lower bounds—a Jensen-based bound and a sharper combinatorial bound—together with a specialized bound for the regime $\alpha < 0$. These bounds are computationally undemanding, requiring only elementary arithmetic and exponentials, and they are universally applicable to all balanced mixed two- and three-level designs.
- *Necessary conditions for optimality:* Corollary 1 delineates the precise combinatorial structure that a design must possess to attain the lower bounds. For $\alpha \geq 0$, optimality requires that the diagonal counts λ_i be confined to the two consecutive integers ψ and $\psi + 1$ with frequencies p, q dictated by the combinatorial lemma, and that all off-diagonal exponents $E_{ij} = (\theta_{ij} + \delta_{ij}) \ln 3 + \lambda_{ij} \ln 9 + \tau_{ij} \ln 5$ be identical. For $\alpha < 0$, optimality demands that exactly $n/3$ rows possess $\lambda_i = m_2$ while the remaining rows satisfy $\lambda_i = 0$, again accompanied by equality of all off-diagonal exponents. These conditions provide practical checks for assessing candidate designs.
- *Algorithmic integration:* We have demonstrated how the theoretical lower bounds can be embedded as principled termination criteria in stochastic search algorithms, specifically threshold accepting (Algorithm 1). This integration confers two significant advantages: it drastically reduces computational cost by avoiding unnecessary iterations, and it provides a quality guarantee—the final design attains an efficiency $F \geq 1 - \varepsilon/\mathcal{LB}$.

6.2 Practical Implications

The bounds derived in this paper fulfill several indispensable practical roles:

- *Termination criteria for search algorithms:* They furnish a principled stopping condition for stochastic optimization heuristics such as threshold accepting and column-wise exchange. By comparing the discrepancy of the current best design with the theoretical lower bound, one can ascertain the opportune moment to terminate optimization, thereby conserving computational resources while guaranteeing high-quality outcomes.
- *Benchmarks for design evaluation:* The efficiency ratio $F = \mathcal{LB}_{\text{Comb}}/[UDisc]^2$ offers a simple, interpretable measure of design quality. A design attaining $F = 1$ is unequivocally optimal with respect to the unanchored discrepancy; designs with F close to unity are near-optimal. Practitioners can employ this ratio to compare candidate designs and select the most suitable among them.
- *Necessary conditions for optimality:* The combinatorial conditions elucidated in Corollary 1 provide actionable guidance for constructing optimal designs. For instance, in the regime $\alpha \geq 0$, one should seek designs where the diagonal counts λ_i are as equal as possible, ideally taking only the values ψ and $\psi + 1$ with the prescribed frequencies.

6.3 Directions for Future Research

Several promising avenues for further inquiry emerge from this work:

1. *Extension to designs incorporating more than two distinct level counts.* The pattern-counting methodology developed here is, in principle, applicable to designs mixing three or more numbers of levels (e.g., two, three, and four factors simultaneously). However, the proliferation of distinct kernel values necessitates an automated approach to classification and the derivation of the attendant linear constraints. One might also aspire to closed-form expressions depending solely on a few summary statistics, such as the moments of the level distributions.

2. *Mixed-level designs with other level combinations.* Beyond two and three levels, many practical experiments involve mixtures such as two and four levels, or three and four levels. Extending our results to those cases would significantly broaden the applicability of the theory.

3. *Connections with other discrepancy measures.* It would be illuminating to compare the unanchored discrepancy bounds obtained here with existing lower bounds for centered, wrap-around, or mixture discrepancies. Designs that exhibit superior performance across multiple uniformity criteria are particularly valuable for robust experimental planning. Establishing connections between these measures could lead to unified optimality conditions.

4. *Algorithmic enhancements.* The threshold accepting algorithm can be further refined by embedding the necessary optimality conditions—for instance, by biasing neighbour generation toward configurations that satisfy the prescribed distributions of the diagonal counts λ_i. This could accelerate convergence and elevate the quality of the final design. Additionally, hybrid approaches combining threshold accepting with local search or genetic algorithms merit investigation.

5. *Applications in computer experiments.* The ultimate validation of any uniformity criterion resides in its performance within surrogate modelling, sensitivity analysis, and calibration. Case studies juxtaposing designs optimized for unanchored discrepancy with those optimized for other space-filling criteria (e.g., maximin distance, minimum energy) would assist practitioners in selecting the most appropriate design for their specific problem. Such studies could also reveal whether the theoretical optimality conditions translate into tangible improvements in prediction accuracy or parameter estimation.

6. *Theoretical refinement of the bounds.* For the regime $\alpha < 0$, we have employed a simple extremal upper bound for D. A more nuanced bound that respects the integrality of the λ_i might be attainable, although the extremal distribution already furnishes a necessary condition for optimality. In the case $\alpha \geq 0$, the combinatorial bound is already sharp. Further tightening could potentially be achieved through a joint optimization of the off-diagonal sum together with the diagonal term, albeit at the cost of substantially increased complexity.

7. *Asymptotic analysis.* Investigating the asymptotic behavior of the lower bounds as n, m_1, or m_2 grow large could provide insights into the optimal scaling of discrepancy with design size and factor counts. Such analysis might also reveal connections with information-theoretic or entropy-based measures of uniformity.

Acknowledgments

The authors sincerely thank the referees, Associate Editor, and Editor-in-Chief for their valuable comments and suggestions, which have greatly improved this paper. The authors also acknowledge the use of DeepSeek for assistance in improving the English grammar and language clarity.

Funding

This work was supported by the National Nature Science Foundation of China No. (12371261).

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] A.M. Elsawah, "Constructing orthogonal maximin distance uniform projection designs for computer experiments," J. Comput. Appl. Math., vol. 473, p. 116902, 2026.

[2] A.M. Elsawah, B. Laala, A.H. Abdel-Hamid, and H. Qin, "Construction of uniform designs for high-dimensional screening via an adjusted multiple quadrupling algorithm," Commun. Stat. Simul. Comput., vol. 55(1), p. 195–220, 2026.

[3] H.S. Mohamed, A.M. Elsawah, Y.B. Shao, C.S. Wu, and M. Bakri, "Analysis on the shear failure of HSS S690-CWGs via mathematical modeling," Eng. Fail. Anal., vol. 143(A), p. 106881, 2023.

[4] B.B. Prasath, A.M. Elsawah, Z. Liyuan, and K. Poon, "Modeling and optimization of the effect of abiotic stressors on the productivity of the biomass, chlorophyll and lutein in microalgae Chlorella pyrenoidosa," J. Agric. Food Res., vol. 5, p. 100163, 2021.

[5] S. Mak, C.L. Sung, X. Wang, S.T. Yeh, Y.H. Chang, V.R. Joseph, V. Yang, and C.J. Wu, "An efficient surrogate model for emulation and physics extraction of large eddy simulations," J. Am. Stat. Assoc., vol. 113, p. 1443–1456, 2018.

[6] A.M. Elsawah, "A novel coding scheme for generating sixteen codes from quaternary codes with applications," Comp. Appl. Math., vol. 43(3), p. 118, 2024.

[7] A.M. Elsawah, "A powerful and efficient algorithm for breaking the links between aliased effects in asymmetric designs," Aust. N. Z. J. Stat., vol. 59(1), p. 17–41, 2017.

[8] A.M. Elsawah, "A novel hybrid algorithm for designing mixed three-and nine-level experiments without modeling assumptions," Commun. Stat. Simul. Comput., vol. 54(4), p. 1003–1037, 2023.

[9] K.T. Fang, R.Z. Li, and A. Sudjianto, Design and modeling for computer experiments. New York: Chapman and Hall/CRC, 2006.

[10] Y.D. Zhou and H. Xu, "Space-filling fractional factorial designs," J. Am. Stat. Assoc., vol. 109(507), p. 1134–1144, 2014.

[11] A.M. Elsawah, "Improving the space-filling behavior of multiple triple designs," Comp. Appl. Math., vol. 41, p. 180, 2022.

[12] K.T. Fang, "Uniform design: application of number-theoretic methods in experimental design," Acta Math. Appl. Sin., vol. 3, p. 363–372, 1980.

[13] Y. Wang and K.T. Fang, "A note on uniform distribution and experimental design," Chin. Sci. Bull., vol. 26, p. 485–489, 1981.

[14] F.J. Hickernell, “A generalized discrepancy and quadrature error bound,” Math. Comp., vol. 67, p. 299–322, 1998.

[15] F.J. Hickernell, “Goodness-of-fit statistics, discrepancies and robust designs,” Stat. Probab. Lett., vol. 44, p. 73–78, 1999.

[16] R.X. Yue and F.J. Hickernell, “Robust designs for fitting linear models with misspecification,” Stat. Sin., vol. 9, p. 1053–1069, 1999.

[17] H. Qin and K.T. Fang, “Discrete discrepancy in factorial designs,” Metrika, vol. 60, p. 59–72, 2004.

[18] Y.D. Zhou, J.H. Ning, and X.B. Song, “Lee discrepancy and its applications in experimental designs,” Stat. Probab. Lett., vol. 78, p. 1933–1942, 2008.

[19] Y.D. Zhou, K.F. Fang, and J.H. Ning, “Mixture discrepancy for quasi-random point sets,” J. Complex., vol. 29, p. 283–301, 2013.

[20] E. Androulakis, K. Drosou, C. Koukouvinos, and Y.D. Zhou, “Measures of uniformity in experimental designs: A selective overview,” Commun. Stat. Theory Methods, vol. 45(13), p. 3782–3806, 2016.

[21] K.T. Fang and R. Mukerjee, “A connection between uniformity and aberration in regular fractions of two-level factorials,” Biometrika, vol. 87, p. 193–198, 2000.

[22] A.M. Elsawah, “Building some bridges among various experimental designs,” J. Korean Stat. Soc., vol. 49, p. 55–81, 2020.

[23] A.M. Elsawah, “Level permutations and factor projections of multiple quadruple designs,” Commun. Stat. Simul. Comput., vol. 53(10), p. 4893–4920, 2024.

[24] S. Heinrich, “Efficient algorithms for computing the L_2 discrepancy,” Math. Comp., vol. 65, p. 1621–1633, 1996.

[25] P. Winker and K.T. Fang, “Optimal U-type design,” in Monte Carlo and Quasi-Monte Carlo Methods 1996, p. 436–448. Springer, 1998.

[26] K.T. Fang, D. Maringer, Y. Tang, and P. Winker, “Lower bounds and stochastic optimization algorithms for uniform designs with three or four levels,” Math. Comp., vol. 75, p. 859–878, 2005.

[27] K.T. Fang, X. Lu, and P. Winker, “Lower bounds for centered and wrap-around L_2-discrepancies and construction of uniform designs by threshold accepting,” J. Complex., vol. 19, p. 692–711, 2003.

[28] A.M. Elsawah and H. Qin, “Lower bound of centered L_2-discrepancy for mixed two and three levels U-type designs,” J. Stat. Plann. Inference, vol. 161, p. 1–11, 2015.

[29] X. Ke, R. Zhang, and H.J. Ye, “Two- and three-level lower bounds for mixture L_2-discrepancy and construction of uniform designs by threshold accepting,” J. Complex., vol. 31(5), p. 741–753, 2015.

[30] A.M. Elsawah, K.T. Fang, P. He, and H. Qin, “Sharp lower bounds of various uniformity criteria for constructing uniform designs,” Stat. Pap., vol. 62(1), p. 461–1482, 2021.

[31] Y.J. Lei, H. Qin, and N. Zou, “Some lower bounds of centered L_2-discrepancy on foldover designs,” Acta Math. Sci., vol. 30A(6), p. 1555–1561, 2010.

[32] A.M. Elsawah and H. Qin, "An efficient methodology for constructing optimal foldover designs in terms of mixture discrepancy," J. Korean Stat. Soc., vol. 45, p. 77–88, 2016.

[33] Y.D. Zhou, H. Xu, and B. Tang, "A lower bound for centered L_2 discrepancy on asymmetric factorials and its applications," Sci. China Math., vol. 56, p. 1027–1038, 2013.

[34] P. Winker and K.T. Fang, "Optimal U-type design," in Monte Carlo and Quasi-Monte Carlo Methods 1996, p. 436–448. Springer, 1998.

[35] K.T. Fang, X. Ke, and A.M. Elsawah, "Construction of uniform designs via an adjusted threshold accepting algorithm," J. Complex., vol. 43, p. 28–37, 2017.

[36] P. Glasserman, Monte Carlo Methods in Financial Engineering. New York: Springer, 2003.

[37] M. Dixon, I. Halperin, and P. Bilokon, Machine Learning in Finance. Switzerland: Springer, 2020.

[38] S.H. Paskov, "Computing high dimensional integrals with applications to finance," Columbia University, Technical Report CUCS-023–94, 1994.

[39] C. Weiß and Z. Nikolic, "An aspect of optimal regression design for lsmc," Monte Carlo Methods Appl., vol. 4, p. 283–290, 2019.

[40] C. Doerr, M. Gnewuch, and M. Wahlström, "Calculation of discrepancy measures and applications," in A Panorama of Discrepancy Theory, W. Chen, A. Srivastav, and G. Travaglini, Eds., Cham: Springer, 2014, p. 621–678.

[41] W. M. Schmidt. On irregularities of distribution vii. Acta Arith., 21:45–50, 1972.

[42] A.M. Elsawah, "Designing uniform computer sequential experiments with mixture levels using lee discrepancy," J. Syst. Sci. Complex., vol 32, p. 681–708, 2019.

[43] A.M. Elsawah and H. Qin, "Optimum mechanism for breaking the confounding effects of mixed-level designs," Comput. Stat., vol 32, p. 781–802, 2017.

[44] A.M. Elsawah and H. Qin, "Asymmetric uniform designs based on mixture discrepancy," J. Applied Stat., vol 43(12), p. 2280–2294, 2016

[45] A.M. Elsawah and H. Qin, "A new strategy for optimal foldover two-level designs," Stat. Probab. Lett., vol. 103, p. 116–126, 2015.

[46] K.T. Fang, D. Maringer, Y. Tang, and P. Winker, "Lower bounds and stochastic optimization algorithms for uniform designs with three or four levels," Math. Comp., vol. 75, p. 859–878, 2005.