RESEARCH STATEMENT

DMITRIY BILYK

A significant part of my current research is devoted to exploring fundamental problems and conjectures of discrepancy theory through the prism of harmonic and functional analysis. Discrepancy theory is concerned with various forms of the following question: How well can uniform distribution be approximated by a discrete set of \( N \) points? And what are the errors and limitations that necessarily arise in such approximations? This area is closely related to numerous fields of mathematics (probability, number theory, approximation theory) and has direct applications to numerical methods in mathematics, science, engineering, and finance.

The strategic goals of this research are two-fold: to infuse discrepancy theory with new methods stemming from analysis, as well as to enrich the field of analysis with new problems and ideas. Historically, analysis played a remarkably important role in discrepancy theory. Every pivotal result in the subject hinged on the application of an analytic concept, such as Fourier series, Fourier transform, Haar wavelets, Walsh-Paley functions, Riesz products. At the same time, many other techniques, common for an analyst (Littlewood-Paley theory, BMO estimates), have been completely overlooked by experts in discrepancy.

The introduction of these techniques to the theory has recently led to several major breakthroughs, to which I have made some significant contributions. It is my intention to broaden and elevate this approach and to innovatively apply the methods of analysis to various branches and aspects of discrepancy.

In what follows, I discuss some of the main results, ideas, problems, and conjectures of the theory, describe my contributions to the field, and outline the plans for future research.

In §1 I focus on a set of problems in irregularity of distributions. It has been known for a long time that finite point sets cannot be distributed ‘too uniformly’. However, the most important precise quantitative versions of this thesis (i.e., sharp uniform lower bounds for the discrepancy) still remain elusive. Motivated by the recent achievements, I intend to further evolve this direction of research and to tackle numerous problems and conjectures arising in this area.

Finite point sets with low discrepancy – the theme of §2 – provide efficient cubature formulas for numerical integration and hence are tremendously important in applied mathematics and science. Despite decades of research, many basic questions are still open, especially in higher dimensions. I present a program addressing the use of analytic techniques in various methods of producing such sets.

In §3 I explore subtle relations between discrepancy bounds and geometry of the underlying sets. The discrepancy estimates, in a strong parallel with objects of harmonic analysis, differ dramatically depending on whether or not curvature or rotations are present. I plan to study the precise effect of measure-theoretic and geometric properties of sets on discrepancy.

Finally, in §4 I discuss the small ball inequality – a lower bound for the sup-norm of the hyperbolic sums of multidimensional Haar functions, which delivers most methods of attack for problems and conjectures of §1. Moreover, this inequality provides a strong link between several problems in analysis, discrepancy, probability, and approximation. I strive to thoroughly investigate these connections, formalize them, and to improve our understanding of the relations between discrepancy, analysis, and other areas of mathematics.

1. Irregularities of distribution: lower bounds for the discrepancy function.

One of the most standard ways to measure the equidistribution of a finite point set is by means of the discrepancy function. Let \( \mathcal{P}_N \) be a set of \( N \) points in \([0, 1]^d\). Its discrepancy function is defined as the difference between the actual and expected numbers of points of \( \mathcal{P}_N \) in the box \([0, x_1) \times \ldots \times [0, x_d)\), i.e.

\[
D_N(x_1, \ldots, x_d) = \#(\mathcal{P}_N \cap [0, x_1) \times \ldots \times [0, x_d) - N x_1 \cdot \ldots \cdot x_d.
\]
Various norms of this function quantify the uniformity of the distribution of \( P_N \), the most natural being its \( L^\infty \) norm (called the star-discrepancy), which represents the worst possible error.

The discrepancy function is intimately related to numerical integration. By the famous Koksma-Hlawka inequality [45], [32], under certain assumptions, the error of the cubature formula given by \( P_N \) can be estimated as

\[
\left| \int_{[0,1]^d} f(x) \, dx - \frac{1}{N} \sum_{p \in P_N} f(p) \right| \leq \frac{1}{N} \left\| \frac{\partial^d}{\partial x_1 \cdots \partial x_d} f \right\|_1 \cdot \|D_N\|_\infty.
\]

Various analogous duality-type estimates involving other \( L^p \) norms of \( D_N \) are also known. Such a connection makes discrepancy estimates invaluable for countless applications which call for approximate evaluation of multivariate integrals [52].

The common wisdom of the theory of irregularities of distribution states that no finite point set can be distributed too well, in other words, discrepancy has to satisfy certain uniform lower estimates. The first precise quantification of this principle was obtained by Roth [55, 1954]. He proved that for any \( N \)-point set

\[
\|D_N\|_2 \gtrsim (\log N)^{\frac{d-1}{2}}.
\]

This result, which according to Roth’s own words ‘started a new theory’ [24], also provided the first connection between irregularities of distribution and dyadic harmonic analysis. Essentially, the idea consisted of projecting the discrepancy function onto the span of the Haar functions supported by dyadic rectangles of volume \( 2^{-n} = \frac{1}{N} \). Inequality (1.3) has been extended to other \( L^p \) norms, \( 1 < p < \infty \), only over 20 years later [60, 1977] (although, a harmonic analyst may realize that such an extension can be derived in several lines using the Littlewood-Paley inequalities [54]).

1.1. The \( L^\infty \) norm: the star-discrepancy. It is natural to expect that the \( L^\infty \) (supremum) norm of the discrepancy function is much greater in order of magnitude than its \( L^2 \) (averaging) norm. Indeed, in [59, 1972] Schmidt proved that in dimension \( d = 2 \),

\[
\|D_N\|_\infty \gtrsim \log N.
\]

A proof, related to the orthogonal function method, was given by Halász [38]. However, high dimensional analogues of this inequality turned out to be extremely proof-resistant. In fact, until recently, the only gain over the \( L^2 \) estimate was by a factor of \((\log \log N)^{\frac{c}{2}}\), just in dimension \( d = 3 \) [5, 1989].

In 2008, jointly with M. Lacey and A. Vagharshakyan ([12, Duke Math. J.] for \( d = 3; [13, J. Funct. Anal.] \) for \( d \geq 4 \)) we were able to obtain the first significant improvement of the \( L^\infty \) bound in all dimensions \( d \geq 3 \):

**Theorem 1.5** (Bilyk, Lacey, Vagharshakyan). For all \( d \geq 3 \), there exists some \( \eta = \eta(d) > 0 \), such that for all \( P_N \subset [0,1]^d \) with \( \#P_N = N \) we have the estimate:

\[
\|D_N\|_\infty \gtrsim (\log N)^{\frac{d-1}{2} + \eta}.
\]

The exact rate of growth of the star-discrepancy in higher dimensions remains an intriguing open problem; in their famous book on the subject [6] Beck and Chen named it “the great open problem” and called it “excruciatingly difficult”. The precise form of the conjecture is a subject of debate among the experts in the field. Motivated by the connections of this theory to other areas, we state the conjecture in the form, which agrees with the sharp conjectures discussed in §4. I strongly believe this form of the conjecture, which seems to be the most accepted nowadays.

**Conjecture 1.7.** In dimensions \( d \geq 3 \), for any distribution of \( N \) points we have the estimate

\[
\|D_N\|_\infty \gtrsim (\log N)^{\frac{d}{2}}.
\]
I am actively involved in the ongoing work on this conjecture and plan to continue pursuing this line of investigation. The approach and the methodology are drawn from the related problems and are described below in Section 4. I would like to note that any partial progress on this problem will constitute a huge leap forward in the theory. A natural suggestion for the next step toward Conjecture 1.7 is to attempt to obtain (1.6) with a gain \( \eta \) independent of \( d \). Certain evidence supporting the validity of Conjecture 1.7 is presented in §4.1 and §2.2.

1.2. The discrepancy function in \( L^1 \) and below. The situation at the opposite endpoint of the \( L^p \) scale, \( L^1 \), is even less understood than the \( L^\infty \) estimates. Halász [38] has exhibited that in dimension \( d = 2 \),

\[
\|D_N\|_1 \geq (\log N)^\frac{1}{2},
\]

and his argument can be easily extended to obtain the same bound in \( d \geq 3 \). It is conjectured, however, that the \( L^1 \) norm should behave similarly to the \( L^2 \) norm, i.e.

Conjecture 1.10. In dimensions \( d \geq 3 \), for any \( P_N \subset [0, 1]^d \),

\[
\|D_N\|_1 \geq (\log N)^{\frac{d-1}{2}}
\]

Unfortunately, at this point, it is not even known whether the exponent of \( \log N \) in the \( L^1 \) bound increases with the dimension. As a reasonable first step, one should prove that there exists \( \alpha(d) \), an increasing function of the dimension, such that

\[
\|D_N\|_1 \geq (\log N)^{\alpha(d)}.
\]

The only available information pertaining to this problem in higher dimensions is due to Lacey [46], who showed that Conjecture 1.10 holds if one replaces the \( L^1 \) norm with the close \( L^1(\log N)^{\frac{d-2}{2}} \). Conjecture 1.10 and inequality (1.12) have a natural counterpart in the case when the integrability index \( p \) drops below one:

Problem 1.13. Prove an analogue of (1.12) for the \( L^p \) norms of \( D_N \), when \( p < 1 \). Eventually, prove that \( \|D_N\|_p \geq (\log N)^{(d-1)/2} \) for \( p < 1 \), in dimensions \( d \geq 2 \).

Relevant inequalities are only known for the \( H^p \) norms with \( p < 1 \), in which case the lower bound of \( (\log N)^{(d-1)/2} \) can be established by an application of the Littlewood-Paley square function and an adaptation of the \( L^2 \) argument. However, neither duality, nor orthogonality arguments are applicable in the setting of Problem 1.13, thus the problem requires an invention of completely new methods. A reasonable direction of the first attack would be to try to produce a non-duality proof of Halász’s two-dimensional \( L^1 \) estimate or to establish lower bounds for special classes of point sets, e.g. lattices. These questions appear connected to the problem about the distribution of \( D_N \) proposed by Schmidt [60], which I also plan to tackle:

Problem 1.14. Is it true that in all dimensions \( d \geq 2 \), for all \( c > 0 \), \( |\{x \in [0, 1]^d : |D_N(x)| < c\}| \) approaches 0 as \( N \to \infty \) uniformly over \( N \)-point sets \( P_N \subset [0, 1]^d \)? Quantify this convergence.

Sublevel set estimates are a familiar concept in harmonic analysis. In particular, suitable extensions of the methods of Carbery, Christ, and Wright [22], which relate multidimensional estimates of this type to the lower bounds of mixed derivatives, may be applicable in this situation.

1.3. Discrepancy estimates in other function spaces. In [15], together with M. Lacey, I. Parissis, and A. Vagharshakyan, we have initiated the study of the discrepancy function in norms other than \( L^p \). The endeavor was aimed at understanding the precise nature of the kink that occurs at the passage from the case \( p < \infty \) to \( p = \infty \). In particular, the following Orlicz space estimate has been obtained in dimension 2:

Theorem 1.15 (Bilyk, Lacey, Parissis, Vagharshakyan). For any \( N \) element point set and \( \alpha \in [2, \infty) \),

\[
\|D_N\|_{\exp(L^\alpha)} \geq (\log N)^{1-\frac{1}{\alpha}},
\]
This estimate yields a smooth interpolation between the $L^p$ and $L^\infty$. Indeed, in the subgaussian case ($\alpha = 2$) one obtains the $L^2$ growth (1.3) $- \sqrt{\log N}$, whereas, as $\alpha \to \infty$, the bound approaches $\log N$ - Schmidt’s $L^\infty$ estimate (1.4). In addition, with the same coauthors, we have proved a sharp BMO estimate (where BMO stands for the dyadic Chang-Fefferman product BMO [17]):

**Theorem 1.17** (Bilyk, Lacey, Parissis, Vagharshakyan). *In dimension $d = 2$, we have*

\begin{equation}
(1.18) \quad \|D_N\|_{BMO} \gtrsim (\log N)^{\frac{1}{2}}.
\end{equation}

This shows that the BMO norm of the discrepancy function tends to behave like the $L^p$ norms rather than $L^\infty$, which stresses the delicacy of the problem. The sharpness of the estimates (1.16), (1.18) above is also obtained in [15] (see §2.1 for a more detailed discussion).

The techniques employed to obtain these inequalities are based on the aforementioned Haar function methods. While the BMO estimate can be extended to higher dimensions easily, the exponential estimates present difficulties comparable to the ones that arise in the $L^\infty$ case. It is conceivable that the approach, developed by myself and coauthors [12], [13] for the proof of (1.6), may yield non-trivial results for the $\exp(L^2)$ norms of the higher-dimensional discrepancy function, although this transition is far from being obvious. *I am currently working on acquiring such results.*

Our recent work on the $L^\infty$ and other norms of $D_N$ [12], [13], [46], [15] has provoked a surge of activity in the investigation of the discrepancy in the framework of the function space theory: for example, the $L^p(\omega)$ estimates for $A_p$ weights $\omega$ [53], Sobolev and Besov space estimates [72], [42] (see also the upcoming book of H. Triebel [73]). *I intend to undertake a systematic study of the discrepancy function in this context* and expect that it will open the door to various extrapolation techniques and enhance our understanding of the main conjectures.

### 1.4. Combinatorial discrepancy.

Discrepancy has a natural combinatorial companion. Let $\mathcal{P}_N \subset [0, 1]^d$ be an $N$ point set and let the function $\lambda : \mathcal{P}_N \rightarrow \{\pm 1\}$ represent a “red-blue” coloring of $\mathcal{P}_N$. The *combinatorial discrepancy* of $\mathcal{P}_N$ with respect to a family of sets $\mathcal{B}$ is defined as $T(\mathcal{P}_N) = \inf_{\lambda} \sup_{B \in \mathcal{B}} \left| \sum_{p \in \mathcal{P}_N \cap B} \lambda(p) \right|$, i.e. the minimization of the largest disbalance of colors in sets from $\mathcal{B}$ over all possible colorings.

In [2], Beck discovered that, when $\mathcal{B}$ is the family of axis-parallel boxes, the quantity $T(N) := \sup_{\mathcal{P}_N} T(\mathcal{P}_N)$ is tightly related to the discrepancy function estimates. He proved, in particular, that in $d = 2$, one has $T(N) \gtrsim \log N$. In [58] Roth has extended this result to the case when the function $\lambda$ takes arbitrary real values (continuous coloring).

Building on these results and the development of the higher-dimensional techniques for the small ball inequality in [12], [13], I have obtained new combinatorial discrepancy estimates in dimensions three and above. More precisely, I prove [11]:

**Theorem 1.19** (Bilyk). *a) In all dimensions $d \geq 3$, for some $\eta(d) > 0$, we have*

\begin{equation}
(1.20) \quad T(N) \gtrsim (\log N)^{\frac{d-1}{2} + \eta}
\end{equation}

in the “red-blue” case $\lambda(p) = \pm 1$. This inequality improves Beck’s prior estimates by combining the results of [13] and [2].

*b) In dimension $d = 3$, for arbitrary real-valued colorings $\lambda$, we have*

\begin{equation}
(1.21) \quad T(N) \gtrsim \frac{(\log N)^{1+\eta}}{N} \sum_{p \in \mathcal{P}_N} |\lambda(p)|.
\end{equation}

This is the first estimate in this general setting in dimension greater than two.
2. Low discrepancy point distributions

As mentioned earlier, point sets with low discrepancy play an important role in numerical integration, giving rise to powerful “Quasi-Monte Carlo” methods [52], [65]. Hence, constructions of points whose discrepancy is small in various senses are extremely valuable in applied mathematics, science, finance, etc. In this topic, despite its long history and a great number of beautiful results, many fundamental questions still remain open, especially in higher dimensions. Their successful resolution would require a mixture of harmonic analytic techniques with ideas from number theory, combinatorics, and geometry.

In shocking contrast to Conjecture 1.7, the best known higher dimensional point distributions satisfy

\[
\|D_N\|_{\infty} \leq (\log N)^{d-1},
\]

which matches the lower bound in \( d = 2 \). A classical example in dimension 2 is the ‘digit-reversing’ van der Corput set [28, 1935], i.e., the set of \( N = 2^n \) points, represented in its binary expansion as

\[
\mathcal{V}_n = \{(0.x_1x_2...x_n, 0.x_nx_{n-1}...x_1) : x_k = 0, 1\}.
\]

An even older example [47, 1904] is the irrational lattice \( \{(k/N, (k\alpha)) : k = 1, ..., N\} \), where \( \{\} \) denotes the fractional part and \( \alpha \) is a badly approximable irrational number. Examples satisfying (2.1) in dimensions 3 and above were first constructed by Halton [39, 1960]. I intend to work on closing the gap between the upper (2.1) and lower (1.6) bounds:

**Problem 2.2.** Prove that in dimensions \( d \geq 3 \) there exist sets with \( \|D_N\|_{\infty} = o(\log^{d-1} N) \).

The resolution of this problem would produce tremendous impact on the methods of numerical integration and a variety of applications. A plausible plan of attack is outlined in §2.2.

2.1. Points with low \( L^2 \) discrepancy. It is well known that many standard sets with low \( L^\infty \) discrepancy fail to meet the optimal \( L^2 \) discrepancy bounds and require certain adjustments to overcome this deficiency. Most available remedies are probabilistic in nature, especially in higher dimensions. In fact, the first deterministic constructions of sets with \( L^2 \) discrepancy of the order \( \log^{(d-1)/2} N \) in dimensions \( d \geq 3 \) have been obtained only recently by Chen and Skriganov [25, 2002]. Generalizations of these results to \( L^p \) norms for \( p \neq 2 \) are even more scarce and usually require highly non-trivial Littlewood-Paley arguments [62].

**Symmetrization.** The first example demonstrating the sharpness of the \( L^2 \) lower bound (1.3) was constructed by Davenport [30] in 1956 as a symmetrized irrational lattice in dimension 2. Similar results were obtained for the van der Corput set by Chen and Skriganov [26].

In [20], with V. Temlyakov and R. Yu, we have proved the same fact for the Fibonacci lattice \( \mathcal{F}_n = \left\{ \left( \frac{i}{b_n}, \frac{j}{b_n} \right) : i = 1, ..., b_n \right\} \), where \( b_n \) are Fibonacci numbers. This set closely resembles the irrational lattice, but is more practical and implementable as its coordinates are rational. In addition to demonstrating the optimal order of magnitude, we have discovered a precise formula for the \( L^2 \) discrepancy of this set, which allows for numerical computations and the evaluation of the arising constants. Our numerical experiments indicate that the Fibonacci lattice has the lowest \( L^2 \) discrepancy among all known two-dimensional sets. Furthermore, we prove that a more sophisticated symmetrization of the Fibonacci set yields the sharp order of magnitude of \( L^p \) discrepancy for all \( p \in (1, \infty) \).

I plan to continue working in this direction, in particular, generalizing these results to other lattices with the goal of finding the optimal constant in the two-dimensional \( L^2 \) discrepancy estimates. The approach to these problems is based on harmonic analysis techniques, in particular, a rigorous examination of the Fourier coefficients of the discrepancy function. It would also be of interest to understand higher dimensional generalizations of the symmetrization method, which are currently completely lacking.

**Derandomization.** In [56, 1979], Roth has proved that random shifts of the van der Corput set \( \mathcal{V}_n = \{(x + \alpha) \mod 1, y) : (x, y) \in \mathcal{V}_n \} \) satisfy the discrepancy estimate

\[
E_0 \left\| D_{\mathcal{V}_n} \right\|_2 \leq (\log N)^{1/2},
\]

\( E_0 \) represents the expected value.
thus guaranteeing an existence of a shift with asymptotically optimal $L^2$ discrepancy $\sqrt{\log N}$; in [57], he extended the scope of this method to certain higher dimensional point sets. The method was purely probabilistic and no “derandomization” of his idea was known until recently.

In [8], using the methods of Fourier analysis, I have constructed the first concrete example of such a shift, enhancing the method later [9] to obtain a more practical algorithm.

I expect that it may be possible to extend this method to point distributions in dimensions three and higher and propose to start with the case $d = 3$. Direct siblings of the van der Corput set – dyadic nets (i.e. sets of $2^n$ points with exactly 1 point per any dyadic box of size $2^{-n}$) exist in $d = 3$, but not when $d \geq 4$ [50], which makes the method more adaptable to this situation. However, it is likely that these techniques can also be extended to non-binary analogs of the van der Corput set, which may eventually lead to novel constructions of sets with optimal $L^2$ discrepancy in all dimensions.

2.2. Other norms and low star-discrepancy. In addition to establishing the lower bounds for the $\exp(L^\alpha)$ and BMO norms of the discrepancy function (1.16), (1.18), in [15] my coauthors and myself demonstrated their sharpness by computing these quantities for the modified van der Corput set. The analysis leading to these results included evaluating the Haar coefficients of $D_N$ for this set and exploiting the sharp form of the Littlewood-Paley inequalities [54]. The next important step in the theory would be:

**Problem 2.4.** Establish similar upper bounds in dimension $d \geq 3$, in other words, show that there exist point distributions for which $\|D_N\|_{\text{BMO}}$ or $\|D_N\|_{\exp(L^2)}$ are of the order $\log^{d+1/2} N$.

I would suggest (modifications of) the sets with low $L^2$ discrepancy constructed by Chen and Skriganov [25] as natural candidates to be the minimizers of this problem. The resulting BMO bounds may have a direct impact on Problem 2.2 – the grand problem about low-discrepancy sets. While BMO estimates by itself do not imply the $L^\infty$ bounds, additional structural information about $D_N$ can be extracted, making it amenable to the local oscillation inequalities, e.g. a recent inequality of Lerner [48], which has already been successfully employed to obtain an array of weighted estimates in harmonic analysis [29]. Under mild structural assumption on the point distribution (e.g. coordinates are dyadic rationals), the oscillation of the discrepancy function on small dyadic cubes arises only from the ‘volume’ part of the function and is not very large. The smallness of local oscillations can be then combined with the global BMO estimate to acquire an upper bound for the $L^\infty$ norm. When Problem 2.4 is resolved, this approach would likely yield

(2.5) $\|D_N\|_{\infty} \lesssim (\log N)^{d+1/2}.$

This is better than (2.1) whenever $d \geq 4$ and would constitute a huge step forward in the theory.

An alternative approach to Problem 2.2, which circumvents the BMO estimates, but is similarly connected to Problem 2.4, lies in obtaining pointwise estimates for the Littlewood-Paley square function of the discrepancy function. Inequalities of Chang-Wilson-Wolff type [54] can then be exploited to produce estimates for the exponential norms of $D_N$. As above, additional information about the point sets can assist in turning them into $L^\infty$ estimates. At this point, both approaches look extremely promising and I am determined to pursue this exciting direction of research.

2.3. Star-discrepancy in large dimensions. In numerous applications which require numerical integration the dimension $d$ is forced to be very large. For instance, in financial mathematics one often has to deal with $d = 360$ (the number of months in a typical 30 year mortgage) [33]. This fact produces several important complications. In particular, the implied constants in the ‘logarithmic’ discrepancy estimates (2.1) are exponential in $d$. Secondly, the expression $\log^{d-1} N$ arising in the cubature error estimates (see (1.2)), is increasing in $N$ for $N \leq e^{d-1}$, rendering logarithmic bounds useless for these applications. In this connection, it becomes extremely important to study the star-discrepancy as a function of both the number of points $N$ and the dimension $d$. A significant breakthrough in this direction was achieved in [40]: for some
absolute constants $e_0$, $c_1$, $c_2$:

\[
\min \left\{ e_0, \ e^{-\frac{c_2}{N}} \right\} \leq \inf_{\mathcal{P}_N \subset [0,1]^d, \ |\mathcal{P}_N| = N} \frac{1}{N} \cdot \|D_N\|_\infty \leq c_2 \sqrt{\frac{d}{N}}.
\]

The upper bound in estimate (2.6) was obtained by purely probabilistic methods and is completely non-constructive. Its probable sharpness is suggested by the Monte Carlo rate. Besides, the value of $c_2$, which may be important for applications, is not known. The constants may be made explicit at the cost of losing a factor of $\sqrt{\log(d) \log(N)}$ [40]. The best known constructive upper bound of the above type is $d^{3/2}N^{-1/2} \log^{1/2}(dN)$, [71]. This raises several questions [41] which I plan to explore:

\textbf{Problem 2.7.} (a) Provide a constructive proof of the upper inequality in (2.6) and evaluate the arising constants. In particular, find out if any of the known low discrepancy point sequences satisfy this estimate.

(b) Show that this bound is best possible, i.e. prove a matching lower bound (at least up to a small error).


Thus far, we have only discussed discrepancy with respect to the family of axis-parallel parallelepipeds $[0,x_1] \times ... \times [0,x_d]$. However, other families of sets (balls, convex sets, polytopes, etc) arise naturally in numerous problems that require a quantification of the uniformity of discrete point distributions, leading to different notions of geometric discrepancy. Given a family $\mathcal{A}$ of subsets of $[0,1]^d$ and an $N$-point set $\mathcal{P}_N \subset [0,1]^2$, the discrepancy of $\mathcal{P}_N$ with respect to $\mathcal{A}$ is defined as:

\[
D(\mathcal{P}_N, \mathcal{A}) = \sup_{A \in \mathcal{A}} \left| \#(A \cup \mathcal{P}_N) - N \cdot |A| \right|.
\]

(Suitable averages may be taken in place of the supremum in the expression above.) Such quantities, in particular, give rise to various estimates for the error of numerical integration in the spirit of the Koksma-Hlawka inequality (1.2), see [52]. It is thus important to determine the asymptotics of the quantity $\inf_{\mathcal{P}_N} D(\mathcal{P}_N, \mathcal{A})$ for large $N$. In a broad perspective, the ultimate goal of our project is to understand the precise correlation between the behavior of the discrepancy and the geometric features of the underlying family of sets.

3.1. Directional discrepancy. The following classical results in dimension $d = 2$ serve as our starting point. As pointed out before, in the case of axis-parallel rectangles, the discrepancy is bounded below by $\log N$ [59] (matching upper bounds are well known [28], [47]). However, if one allows rectangles rotated in all possible directions, the situation changes drastically – the discrepancy estimates are now of the order $N^{1/4}$, [3], [4]. This leads one to some very natural questions. What happens ‘in between’, that is, if the rectangles are rotated in a partial set of directions? How do these rotations influence the size of discrepancy and where is the threshold between logarithmic and polynomial estimates?

Until recently, the only known results in this vein dealt with a finite set of directions, [7], [27], in which case the estimates are essentially the same as in the axis-parallel case (single direction). In 2009, we (together with X. Ma, J. Pipher, and C. Spencer) have started a collaboration aimed at understanding more complicated rotation sets. In [18] we studied several specific intermediate situations (lacunary sequences of directions, lacunary sets of finite order, sets with small Minkowski dimension) and obtained upper bounds for the discrepancy in these settings (powers of the logarithm in the former two case and powers of $N$ in the latter). This work involved non-trivial extensions of a number-theoretic lemma of Davenport [31] and Cassels [23] which lead to constructions of suitable rotations of the integer lattice.

Rotated lattices are an old example in the discrepancy theory [47]. If a scaled integer lattice is rotated by an angle $\theta$, where $\tan \theta$ is a badly approximable number, then its discrepancy with respect to aligned rectangles matches the lower bound (1.4): $\|D_N\|_\infty \gtrsim \log N$. In the case of a direction set $\Omega$, one wants to find a rotation $\theta$ so that $\tan(\theta - \omega)$ is uniformly badly approximable for all $\omega \in \Omega$, i.e. the diophantine inequality

\[
|\tan(\theta - \omega) - \frac{p}{q}| \gtrsim \frac{1}{q^2}
\]
holds for all integers $p$, $q$, and all $\omega \in \Omega$. For a finite $\Omega$, the existence of such a $\theta$ is guaranteed by the Cassels-Davenport lemma. We have been able to obtain generalizations of this lemma to infinite examples of $\Omega$ mentioned above by exploiting subtle stopping time arguments. Naturally, for infinite sets, $q^2$ in an analog of (3.2) is replaced by a larger function of $q$, depending on the geometry of $\Omega$. In [19], we generalized these results to arbitrary sets $\Omega$ with mild additional assumptions, giving the answer in terms of the covering function $N(x)$ of $\Omega$, i.e. the smallest number of intervals of length $x$ needed to cover $\Omega$. Using the Erdős-Turan inequality [45] and the machinery developed in [7], [27], [18], such results can be translated into discrepancy estimates for a given set of directions. We also demonstrated that these estimates are sharp in the class of rotated lattices. Many important related questions are still open.

Problem 3.3. a) Extend these results to higher dimensions.

b) Establish complementing lower bounds for the directional discrepancy for various sets of rotations.

c) It is easy to see that any lattice has discrepancy $N^{1/2}$ with respect to all rotations, while the optimal bound ([3], [4]) is close to $N^{1/4}$. At which point do rotated lattices cease to have the optimal discrepancy?

Numerous other questions arise in this project. For instance, it would be interesting to explore more general families of sets (e.g., polygons with infinitely many sides). Another intriguing path is to draw connections to problems and objects in harmonic analysis, e.g., the directional maximal function (see [1], [61]): geometric considerations, such as dependence of estimates on rotations and curvature, are a recurring theme in harmonic analysis, [63]. I am currently enthusiastically involved in ongoing work on these questions.

4. The small ball conjecture: connections between analysis, probability, approximation, and discrepancy

4.1. The small ball inequality. The small ball inequality, which arises naturally in probability and approximation, besides being important and significant in its own right, also serves as a model for the lower bounds for the star-discrepancy (1.8). This inequality is concerned with the lower estimates of the supremum norm of linear combinations of multivariate Haar functions supported by dyadic boxes of fixed volume (we call such sums ‘hyperbolic’) and can be viewed as a reverse triangle inequality. It is linked to the discrepancy function through the aforementioned Roth’s orthogonal function method. Although no formal connections are known, most arguments designed for this inequality can be transferred to the discrepancy setting.

For a dyadic interval $I$, the ($L^\infty$ normalized) Haar function is defined as $h_I = -1_{I_L} + 1_{I_R}$, where $I_L$ and $I_R$ are the left and right halves of $I$. In higher dimensions, for a dyadic box $R = R_1 \times \ldots \times R_d$, we set $h_R(x_1, \ldots, x_d) = h_{R_1}(x_1) \cdot \ldots \cdot h_{R_d}(x_d)$ and restrict our attention to $R \subset [0, 1]^d$.

Conjecture 4.1 (The Small Ball Conjecture). In dimensions $d \geq 2$, for any choice of the coefficients $\alpha_R$ one has the following inequality:

$$n^{d-2} \left\| \sum_{R: |R|=2^{-n}} \alpha_R h_R \right\|_\infty \geq 2^{-n} \sum_{R: |R|=2^{-n}} |\alpha_R|.$$

The point of interest here is the precise exponent of $n$ on the left-hand side. If one replaces $n^{(d-2)/2}$ by $n^{(d-1)/2}$, this inequality becomes almost trivial, and, in fact, holds for the $L^2$ norm – this should be compared to Roth’s $L^2$ discrepancy estimate (1.3). The presence of the quantity $d - 1$ in this context is absolutely natural, as it is, in fact, the number of ‘free’ parameters (dictated by the condition $|R| = 2^{-n}$). The passage to $d - 2$ for the $L^\infty$ norm requires a much deeper analysis and brings out a number of complications.

Choosing $\alpha_R$’s to be independent gaussian random variables verifies that this conjecture is sharp, which provides one of the reasons to believe in the sharpness of Conjecture 1.7 about the star-discrepancy.

Conjecture 4.1 has been proved in $d = 2$ by M. Talagrand [64] in 1994. In 1995, V. Temlyakov [67–69] has given another, very elegant proof of this inequality in two dimensions, which closely resembled Halász’s discrepancy argument [38] for (1.4). In dimensions three and above however, there have been virtually no improvements over the $L^2$ bound until in 2008, with M. Lacey and A. Vagharshakyan, we have obtained a remarkable improvement over the ‘trivial’ estimate in all dimensions greater than two [12], [13]:
Theorem 4.3. In all dimensions $d \geq 3$ there exists $\eta(d) > 0$ such that for all choices of coefficients we have the inequality:

$$n^{d-1} \eta(d) \left\| \sum_{R: |R| = 2^{-\alpha}} \alpha_R h_R \right\|_\infty \geq 2^{-n} \sum_{R: |R| = 2^{-\alpha}} |\alpha_R|.$$  

It is the proof of this result that was modified to the discrepancy framework to obtain (1.6). In fact, the $\eta(d)$ in (1.6) is the same $\eta$ as in Theorem 4.3.

Most approaches to Conjecture 4.1 follow the two-dimensional arguments of Halász and Temlyakov. In [67], an $L^1$ test function is constructed as a Riesz product $\Psi = \prod (1 + f_j)$, where the product is extended over all vectors $\vec{r} = \vec{r} \in \mathbb{Z}_d$ and $|\vec{r}|_1 = r_1 + \ldots + r_d = n$, and $f_j = \sum_{R: |R| = 2^{-\alpha} \gamma} \text{sgn} (\alpha_R) h_R$ are the generalized Rademacher functions corresponding to dyadic boxes of volume $2^{-n}$, whose configuration is defined by $\vec{r}$: $|R_j| = 2^{-\alpha(j)}$. The success of this method in dimension $d = 2$ is rooted in the fact that the condition $|R| = 2^{-n}$ effectively leaves only one free parameter, creating ‘lacunarity’ which validates the use of the Riesz product. The absence of a natural ordering in higher dimensions presents substantial complications. In $d \geq 3$ (but not in $d = 2$) two different dyadic boxes of the same volume can still have the same sidelength in one of the coordinates, which leads to loss of orthogonality ($h_1 \cdot h_1 = 1$). High combinatorial complexity of such ‘coincidences’ in large dimensions aggravates the difficulty of the problem. A major part of our work in [12], [13] was devoted to the study of analytic and combinatorial aspects of these coincidences. Further success of the Riesz product method would require inequalities of the type:

$$\left\| \sum_{\vec{r}_1, \ldots, \vec{r}_k} f_{\vec{r}_1} \cdots f_{\vec{r}_k} \right\|_p \lesssim (pn)^{\frac{d}{2}},$$

where the sum is extended over all $k$-tuples $\vec{r}_1, \ldots, \vec{r}_k$ with a specified configuration of coincidences and $M$ is the number of ‘free’ parameters imposed by this configuration. Such estimates, which essentially say that free parameters behave orthogonally, first appeared in [5]. In [12], [13], we obtained sharp results for $k = 2$ and partial for $k > 2$. I intend to prove (4.5) in its sharp form, while at the same time searching for other approaches which would avoid Riesz products (e.g., based on greedy algorithms or $\ell^1$-minimization/compressed sensing), thus being more adaptable to the case $d \geq 3$.

Other versions of the small ball conjecture. Various other forms of Conjecture 4.1 are important both for a deeper understanding of the problem, as well as for applications. Signed small ball inequality. The signed small ball inequality, i.e. a version with $\alpha_R = \pm 1$ for each $R$, may be viewed as a toy model of the conjecture. It avoids numerous technicalities, while preserving most of the complications arising from the combinatorial complexity of the higher dimensional dyadic boxes. In this setting the conjecture bears an even stronger resemblance to the star-discrepancy conjecture (1.8):

$$\left\| \sum_{R: |R| = 2^{-\alpha}} \alpha_R h_R \right\|_\infty \geq n^{d/3}, \quad \alpha_R = \pm 1,$$

In [14], we came up with a significant simplification of the arguments in [12], [13] for the signed case, which yielded the estimate $\left\| \sum_{R: |R| = 2^{-\alpha}} \alpha_R h_R \right\|_\infty \geq n^{d/3} \eta$ for $\alpha_R = \pm 1$ in all dimensions with $\eta(d) = \frac{1}{8d} - \varepsilon$. In [15], Lacey, Parissis, Vagharshakayan, and myself initiated a different approach to the signed small ball inequality based on delicate conditional expectation arguments. At the moment, it is restricted to the three-dimensional case, but already yields the best currently known gain: $\eta(3) = \frac{1}{8}$. I am currently working on extensions of this approach to higher dimensions.

Walsh-Paley small ball inequality. It would be illustrative to understand the signed inequality, when the choice of signs is not arbitrary, but has some underlying structure, e.g., if $f_{\vec{r}} = \sum_{R: |R| = 2^{-\alpha} \gamma} \alpha_R h_R$ are Walsh-Paley functions. The analysis of the Riesz product should be simpler in this case, since products of Walsh-Paley functions are again Walsh-Paley. This special case may be applicable to discrepancy estimates, since Walsh-Paley functions naturally appear in the analysis of certain low-discrepancy sets [26], [10].
Trigonometric small ball inequality. This is an analog of (4.2) for trigonometric polynomial with frequencies in the “hyperbolic cross” \( \Gamma(N) = \{ k \in \mathbb{Z}^d : \prod_{j=1}^d \max(|k_j|, 1) \leq N \} \). It arises in approximation theory and is known in \( d = 2 \), [69], [68]. Extensions to \( d \geq 3 \) would likely depend upon a trigonometric analog of (4.5).

4.2. Connections between fields. While the connection of the Small Ball Conjecture to discrepancy function is indirect, it does have important formal implications in probability and approximation theory.

Approximation theory: Entropy of mixed smoothness classes. Let \( MW^p([0, 1]^d) \) be the space of functions on \([0, 1]^d\) with mixed derivative \( \frac{\partial^d f}{\partial x_1^{s_1} \partial x_2^{s_2} \cdots \partial x_d^{s_d}} \) in \( L^p \) and consider its unit ball \( B(MW^p) \). It is compact in the \( L^\infty \) metric and its compactness may be quantified through the device of covering numbers. Define \( N(\varepsilon, p, d) \) to be the least number \( N \) of \( L^\infty \) balls of radius \( \varepsilon \) needed to cover \( B(MW^p) \). The task at hand is to determine the correct order of growth of these numbers as \( \varepsilon \downarrow 0 \).

**Conjecture 4.7.** For \( d \geq 2 \), we have \( \log N(\varepsilon, 2, d) \approx \varepsilon^{-1} (\log 1/\varepsilon)^{d-1/2} \), as \( \varepsilon \downarrow 0 \).

The case \( d = 2 \) follows from the work of Talagrand [64], and the upper bound is known in full generality [35,68]. It is well known [66] that inequalities akin to the Small Ball Conjecture (4.2) imply lower bounds on the covering numbers.

Probability: The small ball problem for the Brownian sheet is concerned with finding the exact behavior of the small deviation probability \( \mathbb{P}(|B|_{C([0,1]^d)} < \varepsilon) \), where \( B \) is the Brownian sheet, i.e. a centered multiparameter Gaussian process characterized by the covariance relation \( \mathbb{E} X_s \cdot X_t = \prod_{j=1}^d \min(s_j, t_j) \).

Kuelbs and Li [43] have discovered a tight connection between the Small Ball probabilities and the properties of the reproducing kernel Hilbert space, which in the case of the Brownian Sheet is \( WM^2([0, 1]^d) \). Their result, applied to the setting of the Brownian sheet in [35], yields an equivalent conjecture:

**Conjecture 4.8.** In dimensions \( d \geq 2 \), for the Brownian Sheet \( B \) we have

\[
- \log \mathbb{P}(|B|_{C([0,1]^d)} < \varepsilon) \approx \varepsilon^{-2} (\log 1/\varepsilon)^{2d-1}, \quad \varepsilon \downarrow 0.
\]

The case \( d = 2 \) has been resolved by Talagrand [64]; in \( d \geq 3 \), the upper bounds are established, see [35].

Our inequality (4.4) directly translates [12,13] into the improvement of lower bounds in both problems.

Relations between the aforementioned objects and discrepancy are constantly lurking in the background. Mixed smoothness classes appear naturally in error estimates for cubature formulas involving the discrepancy function [70], [66], [52]. A striking connection of discrepancy to the Brownian sheet was obtained by Woźniakowski [74]: if \( \omega \) is the Wiener sheet measure on \( C([0,1]^d) \) arising from this process, then

\[
\left( \int_{\mathcal{C}([0,1]^d)} \left( \int_{[0,1]^d} f \, dx - \frac{1}{N} \sum_{p \in \mathcal{P}_{\mathbb{N}}} f(p) \right)^2 \, d\omega(f) \right)^{1/2} = \frac{1}{N^2} \| D_{\mathcal{P}_{\mathbb{N}}} \|_{1,2}^2,
\]

where \( \mathcal{P}_{\mathbb{N}}^* \) is a reflection of \( \mathcal{P}_{\mathbb{N}} \). Generalizations of this equality may shed more light on the interplay between the Brownian sheet and discrepancy function. Furthermore, it would be interesting to explore other random processes in this context, e.g. the fractional Brownian sheet [34], Slepian fields [37], and study the arising analytical estimates. More general ties of discrepancy to probability (empirical distributions, Kolmogorov-Smirnov statistic) and approximation theory (cubature formulas, interpolation) are self-evident.

While the link of the Small Ball Conjecture 4.1 to probability and approximation theory is established, the connection to discrepancy is not direct. In some sense, inequality (4.2) can be viewed as the first order term in the discrepancy estimate (1.8), although this analogy is yet to be fully understood. **Does the Small Ball Conjecture 4.1 imply the discrepancy estimate (1.8)? Are discrepancy bounds directly related to the described concepts in probability and approximation theory?** These and numerous other questions form our last problem, which would help to perceive the true place and role of discrepancy theory among other disciplines of mathematics:

**Problem 4.10.** Establish connections between discrepancy, harmonic analysis, probability, and approximation theory: formalize and quantify them.
References


