BOUNDS FOR DISCREPANCIES IN THE HAMMING SPACE

ALEXANDER BARG\textsuperscript{1} AND MAXIM SKRIGANOV\textsuperscript{2}

ABSTRACT. We derive bounds for the ball $L_p$-discrepancies in the Hamming space for $0 < p < \infty$ and $p = \infty$. Sharp estimates of discrepancies have been obtained for many spaces such as the Euclidean spheres and more general compact Riemannian manifolds. In the present paper, we show that the behavior of discrepancies in the Hamming space differs fundamentally because the volume of the ball in this space depends on its radius exponentially while such a dependence for the Riemannian manifolds is polynomial.

1. INTRODUCTION

1.1. Basic definitions. Let $\mathcal{X}_n = \{0, 1\}^n$ be the binary Hamming space which can be also thought of as a linear space $\mathbb{F}_2^n$ over the finite field $\mathbb{F}_2$. The cardinality $|\mathcal{X}_n| = 2^n$. Denote by $B(x, t)$ the ball with center at $x \in \mathcal{X}_n$ and radius $t \geq 0$, i.e., the set of all points $y \in \mathcal{X}_n$ with $d(x, y) \leq t$, where $d(x, y)$ is the Hamming distance. The volume of the ball $v(t) := |B(x, t)| = \sum_{i=0}^{t} \binom{n}{i}$ is independent of $x \in \mathcal{X}_n$. It is convenient to assume that $B(x, t) = \emptyset$ and $v(t) = 0$ for $t < 0$, and $B(x, t) = \mathcal{X}_n$ and $v(t) = 2^n$ for $t > n$.

For an $N$-point subset $Z_N \subset \mathcal{X}_n$ and a ball $B(y, t)$ define the local discrepancy as follows:

$$D(Z_N, y, t) = |B(y, t) \cap Z_N| - N 2^{-n}v(t).$$

(1)

We note that $D(Z_N, y, n) = 0$ for any $Z_N, y$, and thus below we limit ourselves to the values $0 \leq t \leq n - 1$. Define the weighted $L_p$-discrepancy by

$$D_p(G, Z_N) = \left( \sum_{t=0}^{n-1} g_t \sum_{y \in \mathcal{X}_n} 2^{-n}|D(Z_N, y, t)|^p \right)^{1/p}, \quad 0 < p < \infty,$$

(2)

where $G = (g_0, \ldots, g_{n-1})$ is a vector of nonnegative weights normalized by

$$\sum_{t=0}^{n-1} g_t = 1.$$

(3)

With such a normalization, we have

$$D_p(G, Z_N) \leq D_q(G, Z_N) \quad 0 < p < q < \infty.$$

(4)

The $L_\infty$-discrepancy is defined by

$$D_\infty(I, Z_N) = \max_{t \in I} \max_{y \in \mathcal{X}_n} |D(Z_N, y, t)|,$$

(5)

where $I \subseteq \{0, \ldots, n-1\}$ is a subset of the set of the radii.

We also introduce the following extremal discrepancies

$$D_p(G, n, N) = \min_{Z_N \subset \mathcal{X}_n} D_p(G, Z_N), \quad 0 < p < \infty,$$

1 Department of ECE and Institute for Systems Research, University of Maryland, College Park, MD 20742, USA and Inst. for Probl. Inform. Trans., Moscow, Russia. Email: abarg@umd.edu. Research of this author was partially supported by NSF grants CCF1618603 and CCF1814487.

2 St. Petersburg Department of Steklov Institute of Mathematics, Russian Academy of Sciences, nab. Fontanki 27, St. Petersburg, 191023, Russia. Email: maksim88138813@mail.ru.
These quantities can be thought of as geometric characteristics of the Hamming space.

It is useful to keep in mind the following simple observations:

(i) If $Z_N^c = X_n \setminus Z_N$ is the complement of $Z_N \subseteq X_n$, then $D(Z_N, y, t) = -D(Z_N^c, y, t)$, and we have

$$D_p(G, Z_N) = D_p(G, Z_N^c) \quad \text{and} \quad D_p(G, n, N) = D_p(G, n, 2^n - N),$$

for all $0 < p \leq \infty$. Hence, generally it suffices to consider only subsets $Z_N$ with $N \leq 2^{n-1}$. Together with results of [1] on quadratic discrepancies this gives rise to the next claim: Let $Z_N$ be a perfect code in $X_n$, then the set $Z_N^c$ attains the minimum value $D_2(G_1, n, 2^n - N)$, where $G_1 = (1/n, 1/n, \ldots, 1/n)$. For instance, for $n = 2^m - 1$ and $N = 2^n(1 - 2^{-m})$, $m \geq 2$ the code $Z_N$ formed of spheres of radius one around the codewords of the Hamming code (i.e., the union of the $n$ cosets of the Hamming code) is a minimizer of quadratic discrepancy. Another family of minimizers is given by $X_n \setminus \{y, \bar{y}\}$ for any $y \in X_n$, where $\bar{y} := 1^n + y$ is a point antipodal to $y$ and $1^n \in X_n$ denotes the all-one vector. Some other examples can be also given; see [1]. For the reader’s convenience, we emphasize that the quadratic discrepancy $D_{L^2}(Z_N)$ in [1] is related with our definition (2) by $D_{L^2}(Z_N) = 2^n n^{-2} (D(G_1, Z_N))^2$.

(ii) Without loss of generality we can restrict the range of summation on $t$ in (2) from $\{0, \ldots, n\}$ to $\{0, \ldots, \nu\}$, where $\nu = \lceil (n - 1)/2 \rceil$, limiting ourselves to a half of the full range. More precisely, we have

$$D_p(G, Z_N) = D_p(G^*, Z_N) \quad \text{and} \quad D_p(G, n, N) = D_p(G^*, n, N),$$

where $G^* = (g^*_1, \ldots, g^*_n)$ with $g^*_t = g_t + g_{n-t+1}$.

Indeed, notice that $B(y, t) = X_n \setminus B(\bar{y}, n - 1 - t)$, and therefore $D(Z_N, y, t) = D(Z_N, \bar{y}, n - 1 - t)$. Also, obviously,

$$\sum_{y \in X_n} |D(Z_N, \bar{y}, t)|^p = \sum_{y \in X_n} |D(Z_N, y, t)|^p,$$

and thus

$$D_p(G, Z_N) = \left( \sum_{t=0}^{\nu} \left( g_t 2^{-n} \sum_{y \in X_n} |D(Z_N, y, t)|^p + g_{n-1-t} 2^{-n} \sum_{y \in X_n} |D(Z_N, \bar{y}, t)|^p \right) \right)^{1/p}$$

$$= \left( 2^{-n} \sum_{t=0}^{\nu} \left( g_t + g_{n-1-t} \right) \sum_{y \in X_n} |D(Z_N, y, t)|^p \right)^{1/p}.$$

We conclude that limiting the summation range of $t$ amounts to changing the weights in definition (2). Similar arguments hold true for the $L_\infty$-discrepancy [5].

1.2. Earlier results. Discrepancies in compact metric measure spaces have been studied for a long time, starting with basic results in the theory of uniform distributions [2, 3, 14]. In particular, quadratic discrepancy of finite subsets of the Euclidean sphere is related to the structure of the distances in the subset through a well-known identity called Stolarsky’s invariance principle [19]. Stolarsky’s identity expresses the $L_2$-discrepancy of a spherical set as a difference between the average distance on the sphere and the average distance in the set. Recently it has been a subject of renewed attention in the literature. In particular, papers [2, 15, 4] gave new, simplified proofs of Stolarsky’s invariance, while [18] extended Stolarsky’s principle to projective spaces and derived asymptotically tight estimates of discrepancy. Sharp bounds on quadratic discrepancy were obtained in [6, 8, 15, 16]. Finally, paper [17] introduced new asymptotic upper bounds on $L_p$-discrepancies of finite sets in compact metric measure spaces.

A recent paper [11] initiated the study of Stolarsky’s invariance in finite metric spaces, deriving an explicit form of the invariance principle in the Hamming space $X_n$ as well as bounds on the quadratic discrepancy of
subsets (codes) in $X_n$. Explicit formulas were obtained for the uniform weights $G_1 = (1/n, 1/n, \ldots, 1/n)$. Namely, let $x, y \in X_n$ be two points with $d(x, y) = w$. Define
\[
\lambda(x, y) = \lambda(w) := 2^{n-w} w \left( \left\lfloor \frac{w}{2} \right\rfloor - 1 \right), \quad w = 0, \ldots, n.
\]
As shown in [1] Eq. (23), Stolarsky’s identity for $Z_N \subset X_n$ can be written in the following form:
\[
2^n n D_2(G_1, Z_N)^2 = \frac{nN^2}{2n+1} \left( \frac{2n}{n} \right) - \sum_{i,j=1}^{N} \lambda(d(z_i, z_j)). \tag{6}
\]
Using this representation, [1] Cor.5.3, Thm.5.5] further showed that
\[
c n^{-3/4} N^{1/2} \left( 1 - \frac{N}{2^n} \right)^{1/2} \leq D_2(G_1, n, N) \leq C n^{-1/4} N^{1/2},
\]
where $c, C$ are some universal constants. Here the upper bound is proved by random choice and the lower bound by linear programming. The method of linear programming, well known in coding theory [11, 12], is applicable to the problem of bounding the quadratic discrepancy because it can be expressed as an energy functional on the code with potential given by $\lambda$. Moreover, there exist sequences of subsets (codes) $Z_N \subset X_n, n = 2^m - 1$ whose quadratic discrepancy meets the lower bound. Observe also that if $N = o(2^n)$, then the bounds differ only by a factor of $n$: for example, if $N \simeq 2^αn, 0 < α < 1$, then
\[
N^{1/2} (\log N)^{-3/4} \lesssim D_2(G_1, n, N) \lesssim N^{1/2} (\log N)^{-1/4}. \tag{7}
\]
In this short paper we develop the results of [1], proving bounds on $D_p(G, n, N), p \in (0, \infty]$. We also consider a restricted version of the discrepancy $D_p(G, Z_N)$, limiting ourselves to the case of hemispheres in $X_n$. In other words, we take local discrepancy for $t = (n - 1)/2$ in [1] (n odd) and average its value over the centers of the balls. For the case of the Euclidean sphere, quadratic discrepancy for hemispheres was previously studied in [4, 16], which established a version of Stolarsky’s invariance for this case.

2. Bounds on $D_p(G, n, N)$

We are interested in universal bounds for discrepancies [4–5] for given $n, N$ and $p \in (a, \infty]$ without accounting for the structure of the subset. For the case of finite subsets in compact Riemannian manifolds this problem was recently studied in [17], and we draw on the approach of this paper in the derivations below.

2.1. The case $0 < p < \infty$. We shall consider random subsets $Z_N \subset X_n$, using the following standard result to handle discrepancies of such subsets.

**Lemma 2.1** (Marcinkiewicz–Zygmund inequality; [10], Sec.10.3). Let $ζ_j, j \in J, |J| < \infty$, be a finite collection of real-valued independent random variables with expectations $\mathbb{E} \zeta_j = 0, j \in J$. Then, we have
\[
\mathbb{E} \left| \sum_{j \in J} ζ_j \right|^p \leq 2^p (p + 1)^{p/2} \mathbb{E} \left( \sum_{j \in J} ζ_j^2 \right)^{p/2}, \quad 1 \leq p < \infty.
\]

In our first result we construct a random subset $Z_N$ by uniform random choice. Later we will refine this procedure, obtaining a more precise bound on $D_p$.

**Theorem 2.2.** For all $N \leq 2^{n-1}$, we have
\[
D_p(G, n, N) \leq 2(p + 1)^{1/2} N^{1/2}
\]
for $1 \leq p < \infty$, and $D_p(G, n, N) \leq 2^{3/2} N^{1/2}$ for $0 < p < 1$. 

Remark 2.1. Bounds of the type \(8\) hold true for arbitrary compact metric measure spaces. Theorem 2.2 is given here to compare it with Theorem 2.3 below. Notice also that the upper bound \(7\) is better than \(8\) with \(p = 2\) by a logarithmic factor. Such an improvement is obtained in \([11]\) because of the explicit formula \(6\) for the quadratic discrepancy with the uniform weights \(G_1\).

Proof. Choose a subset \(Z_N\) by selecting the points \(\{z_i\}_1^N\) independently and uniformly in \(X_n\). The probability that such a point falls into a subset \(E \subseteq X_n\) equals to \(|E|/|X_n|\). Therefore, for the local discrepancy \(1\) of this random subset \(Z_N\) we have

\[
D(Z_N, y, t) = \sum_{i=1}^N \zeta_i(y, t),
\]

where

\[
\zeta_i(y, t) = 1_{B(y, t)}(z_i) - \frac{v(t)}{|X_n|},
\]

where \(1_E\) is the indicator function of a subset \(E \subseteq X_n\). The quantities \(\zeta_i(y, t)\) are independent random variables that satisfy \(|\zeta_i(y, t)| \leq 1\) and \(\mathbb{E}\zeta_i(y, t) = 0\).

Applying the Marcinkiewicz–Zygmund inequality to the sum \(9\), we obtain

\[
\mathbb{E} |D(Z_N, y, t)|^p \leq 2^p (p + 1)^{p/2} N^{p/2}, \quad 1 \leq p < \infty,
\]

and, therefore, in view of \(3\),

\[
\mathbb{E} D(G, Z_N)^p \leq 2^p (p + 1)^{p/2} N^{p/2}, \quad 1 \leq p < \infty.
\]

Thus, there exists a subset \(Z_N = Z_N(p) \subseteq X_n\), \(1 \leq p < \infty\), whose discrepancy is bounded above as in this inequality. For \(0 < p < 1\), in view of \(4\), we can put \(Z(p) = Z(1)\) to complete the proof.

In some situations the bound of this theorem can be improved relying on the method of jittered (or stratified) sampling, which uses a partition of the metric space into subsets of small diameter and equal volume. This idea goes back to classical works on discrepancy theory \([2,3]\) pp.237-240 and it was used more recently in \([5,6,7]\) for the case of the Euclidean sphere and in \([17]\) for general metric spaces. Below we follow the approach of \([17]\). In the case of the Hamming space the natural way to proceed is to partition \(X_n\) into sub-hypercubes of a fixed dimension.

In our analysis bounds on the volume of ball \(v(t)\) are crucial. For large \(n\) and \(t = \lambda n, 0 \leq \lambda \leq 1\), the well-known bound on \(v(t)\) (cf. \([13]\) p. 310]), can be written in the form

\[
v(\lambda n) \leq 2^{nH(\lambda)},
\]

where

\[
H(\lambda) = \begin{cases} h(\lambda), & \text{if } 0 \leq \lambda \leq 1/2, \\ 1, & \text{if } 1/2 < \lambda \leq 1, \end{cases}
\]

and \(h(\lambda) = -\lambda \log_2 \lambda - (1 - \lambda) \log_2 (1 - \lambda)\) is the standard binary entropy, and in general, the bound \(10\) cannot be improved. Formally speaking, the statement \(10\) requires \(\lambda n\) be integer, but this does not matter for the asymptotic arguments that we employ.

Theorem 2.3. Let \(0 < p < \infty\), \(N = \Theta(2^{\alpha n})\), \(0 < \alpha < 1\). Suppose that the weights \(g_t = 0\) for \(t > \beta n\), \(0 < \beta < 1/2\). Then

\[
D_p(G, n, N) \leq 2(p + 1)^{1/2} N^{(1-\kappa)/2}
\]

for \(1 \leq p < \infty\), and \(D_p(G, n, N) \leq 2^{3/2} N^{(1-\kappa)/2}\) for \(0 < p < 1\). Here

\[
\kappa = \kappa(\alpha, \beta) = \frac{1 - H(1 + \beta - \alpha)}{\alpha} \geq 0.
\]

If \(\alpha > \frac{1}{2} + \beta\), then the exponent \(\kappa(\alpha, \beta) > 0\), and the bound \(12\) is better than \(8\).
Proof. Let $V \subset X_n$ be the $k$-dimensional subspace, $k = \gamma n$, $0 < \gamma < 1$, consisting of all vectors $(x_1, \ldots, x_n)$ with $x_i = 0$ if $i > k$. Let $N = 2^{n-k} = 2^\alpha n$, $\alpha = 1 - \gamma$. The affine subspaces

$$V_i = V + s_i, \quad s_i \in X_n/V$$

form a partition of the Hamming space

$$X_n = \bigcup_{i=1}^N V_i, \quad V_i \cap V_j = \emptyset,$$

where $|V_i| = 2^\gamma n$, $\text{diam} V_i = \gamma n$, where $\text{diam} \, \mathcal{E} = \max\{d(x_1, x_2) : x_1, x_2 \in \mathcal{E}\}$ denotes the diameter of a subset $\mathcal{E} \subseteq X_n$.

We consider a subset $Z_N = \{z_i\}_{i=1}^N$ with $z_i \in V_i$, $i = 1, \ldots, N$. For such a subset, the local discrepancy (1) can be written as follows

$$D(Z_N, y, t) = \sum_{i=1}^N \zeta_i(y, t),$$

(14)

where

$$\zeta_i(y, t) = \mathbb{1}_{\{B(y, t) \cap V_i\}}(z_i) - N \frac{|(B(y, t) \cap V_i)|}{|X_n|}.$$  

Notice that if $V_i \subset B(y, t)$, then $\zeta_i(y, t) \equiv 0$ (recall that $x_i \in V_i$). Therefore, the sum (14) takes the form

$$D(Z_N, y, t) = \sum_{i \in J}^N \zeta_i(y, t),$$

where $J$ is a subset of indices $i$ such that $V_i \cap B(y, t) \neq \emptyset$ but $V_i \not\subset B(y, t)$ ($V_i$ is not either completely inside or completely outside $B(y, t)$). Since $\text{diam} V_i = k$, we conclude that all $V_i$, $i \in J$, are contained in the ball $B(y, t + k)$ and do not intersect the ball $B(y, t - k - 1)$. Therefore,

$$|J| |V_i| \leq v(t + k) - v(t - k - 1) \leq v(t + k).$$

Here we estimate $J$ from above by the number of sets $V_i$ such that $B(y, t) \subset V_i$. We note that discarding the term $v(t - k - 1)$ entails no significant loss in the asymptotics because this term is exponentially small compared to $v(t + k)$. For $t \leq \beta n$, using the bound (11) and $\alpha + \gamma = 1$, we obtain

$$|J| \leq 2^{nH(\beta + \gamma) - \gamma n} = 2^{\alpha n(1-\kappa)} = N^{1-\kappa},$$

where $\kappa$ is defined in (13).

Now consider a random subset $Z_N = \{z_i\}_{i=1}^N$ in which each point $z_i$ is selected independently and uniformly in $V_i$. For a subset $\mathcal{E} \subset V_i$ we have $\Pr(z_i \in \mathcal{E}) = |\mathcal{E}|/|V_i| = N|\mathcal{E}|/|X_n|$. The quantities $\zeta_i(y, t)$ are bounded independent random variables that satisfy $|\zeta_i(y, t)| \leq 1$ and $\mathbb{E} \zeta_i(y, t) = 0$. Applying the Marcinkiewicz–Zygmund inequality to the sum (14), we obtain

$$\mathbb{E} |D(Z_N, y, t)|^p \leq 2^p (p + 1)^{p/2} N^{p(1-\kappa)/2}$$

and, therefore, in view of (3),

$$\mathbb{E} |D(G, Z_N)|^p \leq 2^p (p + 1)^{p/2} N^{p(1-\kappa)/2}.$$  

(15)

Thus, there exists a subset $Z_N = Z_N(p) \subset X_n$, $1 \leq p < \infty$, whose discrepancy is bounded above as in this inequality. For $0 < p < 1$, in view of (4), we can put $Z(p) = Z(1)$ to complete the proof. \qed
2.2. The case $p = \infty$. The following statement is analogous to [17, Prop. 2.2]. For $1 \leq p < \infty$ and any subset $Z_N \subseteq X_n$, we have

$$D_{\infty}(I, Z_N) \leq |I|^{1/p} \, 2^{n/p} \, D_p(G_I, Z_N),$$

(16)

where

$$D_p(G_I, Z_N) = \left( \sum_{t=0}^{\nu} |I|^{-1} \sum_{y \in X_n} 2^{-n}|D(Z_N, y, t)|^p \right)^{1/p},$$

is a special $L_p$-discrepancy with $G_I = (g_1, \ldots, g_\nu)$, where $g_t = |I|^{-1}$ for $t \in I$ and $g_t = 0$ otherwise.

Indeed, for $y_1 \in X_n$ and $t \in I$ we have

$$|D(Z_N, y, t)| \leq \left( \sum_{t=1}^{\nu} \sum_{y \in X_n} |D(Z_N, y, t)|^p \right)^{1/p} = |I|^{1/p} \, 2^{n/p} \left( \sum_{t=1}^{\nu} |I|^{-1} \sum_{y \in X_n} 2^{-n}|D(Z_N, y, t)|^p \right)^{1/p}.$$

Theorem 2.4. (i) Let $I \subseteq \{0, 1, \ldots, n\}$ be an arbitrary subset of the set of radii, and $N \leq 2^{n-1}$. Then

$$D_{\infty}(I, Z_N) \leq 8 \, (1 + n)^{1/2} \, N^{1/2}.$$  

(17)

If $N$ increases exponentially, $N \geq 2^{\alpha n}$, then $D_{\infty}(I, n, N) = O((\log_2 N)^{1/2} \, N^{1/2})$.

(ii) Let $I \subseteq \{0, 1, \ldots, \beta n\}$ be an arbitrary subset of the set of radii $t \leq \beta n$, $0 < \beta < 1/2$, and $N = 2^{\alpha n} \leq 2^{n-1}$. Then

$$D_{\infty}(I, n, N \leq 8 \left( 2 + \frac{\log_2 N}{\alpha} \right)^{1/2} \, N^{(1-\kappa)/2},$$

(18)

where the exponent $\kappa = \kappa(\alpha, \beta)$ is given in [13]. If $\alpha > \frac{1}{2} + \beta$, then the exponent $\kappa(\alpha, \beta) > 0$, and the bound (18) is better than (17).

Proof. Substituting the bounds (8) and (12) into inequality (16), we obtain

$$D_{\infty}(I, Z_N) \leq n^{1/p} \, 2^{n/p} \, 2 \, (p + 1)^{1/2} \, N^{1/2}$$

(19)

and

$$D_{\infty}(I, Z_N) \leq n^{1/p} \, 2^{n/p} \, 2 \, (p + 1)^{1/2} \, N^{(1-\kappa)/2}.$$  

(20)

Now, we put $p = n$ in (19) and (20) to obtain, respectively, (17) and (18).

3. Discrepancy for hemispheres

Let $X_n, n = 2m + 1$ be the Hamming space. In this section we consider a restricted version of discrepancy where instead of all the ball radii in (2) we consider discrepancy only with respect to the balls of radius $m$, calling them hemispheres. For any pair of antipodal points $y, \bar{y}$

$$X_n = B(y, m) \cup B(\bar{y}, m), \quad B(y, m) \cap B(\bar{y}, m) = \emptyset,$$

hence $2^{-n} v(m) = 2^{-n} |B(y, m)| = 1/2$.

For a subset $Z_N \subseteq X_n$ define

$$D^{(m)}_{\infty}(Z_N) = \left( \sum_{y \in X_n} |D(Z_N, y, m)|^p \right)^{1/p}, \quad 0 < p < \infty,$$

(21)

where

$$D(Z_N, y, m) = |B(y, m) \cap Z_N| - \frac{N}{2}.$$
is the local discrepancy defined in (1). In the previous notation \( D_p^{(m)}(Z_N) = D_p(G(m), Z_N) \), with the weights \( G(m) = (g_1, \ldots, g_n) \), where \( g_m = 1 \) and \( g_t = 0 \) if \( t \neq m \). Further, let

\[
D^{(m)}(Z_N) = \max_{y \in X_n} |D(Z_N, y)|.
\]

As before, define

\[
D_p^{(m)}(n, N) = \min_{Z_N \subseteq X_n} D_p^{(m)}(Z_N), \quad p \in (0, \infty] .
\]

First we address the question of global minimizers of discrepancy.

**Theorem 3.1.** (i) Let \( N = 2K \) be even, then for all subsets \( Z_N \subseteq X_n \) and \( p \in (0, \infty] \)

\[
D_p^{(m)}(Z_N) \geq 0
\]

with equality for subsets \( Z_N \) consisting of \( K \) pairs of antipodal points.

(ii) Let \( N = 2K + 1 \) be odd, then for all subsets \( Z_N \subseteq X_n \) and \( p \in (0, \infty] \)

\[
D_p^{(m)}(Z_N) \geq 1/2
\]

with equality for subsets \( Z_N \) consisting of \( K \) pairs of antipodal points supplemented with a single point.

In other words, for all \( p \in (0, \infty] \) the extremal discrepancies \( D_p^{(m)}(n, N) = 0 \) if \( N \) is even and \( D_p^{(m)}(n, N) = 1/2 \) if \( N \) is odd.

**Proof.** From (21) we conclude that

\[
N = |B(y, m) \cap Z_N| + |B(\bar{y}, m) \cap Z_N|,
\]

and for any \( y \in X_n \), the local discrepancy can be written as

\[
2|D(Z_N, y, m)| = 2|B(y, m) \cap Z_N| - N = |B(y, m) \cap Z_N| - |B(\bar{y}, m) \cap Z_N| .
\]

Let \( N = 2K \). Inequality (22) holds for all subsets \( Z_N \). If \( Z_N \) is formed of \( K \) pairs of antipodal points, then \( |D(Z_N, y, m)| = 0 \) for all \( y \in X_n \). This proves part (i).

Let \( N = 2K + 1 \). It follows from (24) that \( |D(Z_N, y, m)| \geq 1 \), since \( N \) is odd and \( 2|B(y, m) \cap Z_N| \) is even. This implies inequality (23). Furthermore, it also follows from (24) that \( |D(Z_N, y, m)| = 1 \) for all \( y \in X_n \) if \( Z_N \) consists of \( K \) pairs of antipodal points supplemented with a single point. This proves part (ii).

Thus in particular, any linear code \( Z_N \subset X_n \) that contains the all-ones vector has discrepancy zero (such codes are called self-complementary). Many well-known families of binary linear codes such as the Hamming codes, BCH codes, etc. possess this property.

A minor generalization of the above proof implies the following useful relation. Let \( Z_N = Z'_N \cup Z''_N \) be a union of two subsets, where \( Z'_N \) contains all pairs of antipodal points in \( Z_N \) then

\[
D_p^{(m)}(Z_N) = D_p^{(m)}(Z'_N) , \quad p \in (0, \infty] .
\]

### 3.1. Quadratic discrepancy for hemispheres.

In this section we consider the discrepancy \( D_p^{(m)}(Z_N) \) defined in (24) for the special case \( p = 2 \). Let \( Z_N \subset X_n \) be a code, where \( n = 2m + 1 \). For a pair of points \( x, y \in X_n \) such that \( d(x, y) = w \) let \( \mu_m(x, y) = \mu_m(w) = |B(x) \cap B(y)| \) be the size of the intersection of the balls of radius \( t \) with centers at \( x \) and \( y \). By abuse of notation we write \( \mu_m \) both as a kernel on \( X_n \times X_n \) and as a function on \{0, 1, \ldots, n\}. This is possible because \( \mu_m(x, y) \) depends only on the distance between \( x \) and \( y \). Note that \( \mu_m(0) = v(m) = 2^{n-1} \) and \( \mu_m(n) = 0 \).
In this subsection we use some more specific facts of coding theory. We refer to [13] for details. For a code \( Z_N \subset X_n \) let
\[
A_w = A_w(Z_N) = \frac{1}{N} | \{(z_i, z_j) \in Z_N^2 \mid d(z_i, z_j) = w\} |, \quad w = 0, 1, \ldots, n
\]
be the normalized number of ordered pairs of points at distance \( w \) (the distance distribution of \( Z_N \)). Before stating it, recall the dual distance distribution of the code \( Z_N \), given by
\[
A_i^\perp = \frac{1}{N} \sum_{w=0}^n A_w K_i^{(n)}(w), \quad i = 0, 1, \ldots, n, \tag{25}
\]
where \( K_i^{(n)}(x) \) be the binary Krawtchouk polynomial of degree \( k = 0, \ldots, n \), defined as follows:
\[
K_i^{(n)}(x) = \sum_{j=0}^i (-1)^j \binom{x}{j} \binom{n+x}{i-j}. \tag{26}
\]
The vector \( (A_i^\perp) \) forms the MacWilliams transform of the distance distribution of the code \( Z_N \), and if \( Z_N \) is a linear code, it coincides with the weight distribution of the dual code \( Z_N^\perp \) [13, pp. 129,138]. The MacWilliams transform is an involuton [11, Thm. 3], which enables us to invert relations (24).
\[
A_i = \frac{2^n}{N} \sum_{w=0}^n A_w^\perp K_i^{(n)}(w), \quad i = 0, 1, \ldots, n. \tag{27}
\]
The following result is implied by [11], Lemma 4.1.

**Lemma 3.2.** The Krawtchouk expansion of the function \( \mu_m(w), w = 0, 1, \ldots, n \) has the form
\[
\mu_m(w) = \hat{\mu}_0 + \sum_{k=1}^n \hat{\mu}_k K_k^{(n)}(w)
\]
where \( \hat{\mu}_0 = 2^{n-2} \) and for all \( k = 1, 3, \ldots, n \)
\[
\hat{\mu}_k = 2^{-n} \binom{2m}{m} \frac{(2m)^2}{(k-1)^2}. \]

In the next proposition we establish a version of Stolarsky’s invariance principle for the quadratic discrepancy \( D_2^{(m)}(Z_N) \) defined above in (21).

**Proposition 3.3.** We have
\[
2^n N^{-2} D_2^{(m)}(Z_N)^2 = \frac{1}{N} \sum_{w=0}^n A_w \mu_m(w) - 2^{n-2} \tag{28}
= \sum_{k=1}^n \hat{\mu}_k A_k. \tag{29}
\]

**Proof.** Starting with (21), we compute
\[
2^n D_2^{(m)}(Z_N)^2 = \sum_{y \in X_n} \left( \sum_{j=1}^N \mathbb{1}_{B(y,m)}(z_j) - N \frac{N}{2} \right)^2 = \sum_{y \in X_n} \left( \sum_{j=1}^N \mathbb{1}_{B(z_j,m)}(y) - N \frac{N}{2} \right)^2
\]
\[
= \sum_{y \in X_n} \left( \sum_{i,j=1}^N \mathbb{1}_{B(z_i,m)}(y) \mathbb{1}_{B(z_j,m)}(y) - N \sum_{j=1}^N \mathbb{1}_{B(z_j,m)}(y) + N^2 \frac{N}{4} \right)
\]
\[
= \sum_{i,j=1}^N \sum_{y \in X_n} \mathbb{1}_{B(z_i,m)}(y) \mathbb{1}_{B(z_j,m)}(y) - 2^{n-2} N^2
\]
\[
= \sum_{i,j=1}^N \mu_m(z_i, z_j) - 2^{n-2} N^2 = N \sum_{w=0}^n A_w \mu_m(w) - 2^{n-2} N^2,
\]
where the last equality uses the definition of $A_w$. This proves (28).

To obtain (29), we note that, from (27), we have $\sum_{w=0}^n A_w \mu_m(w) = N \sum_{k=0}^n A_k^\perp \hat{\mu}_k$. Substituting this into (28) and using the obvious equality $\hat{\mu}_k = 2^{-n} \sum_{w=0}^n K_w^{(n)} \mu_m(w)$, we obtain (29).

The size of the intersection of the balls can be written in a more explicit form:

$$\mu_m(w) = \sum_{i,j} \binom{w}{i} \binom{n-w}{j}, \quad w = 0, \ldots, n,$$

where $i + j \leq m$, $0 \leq w - i + j \leq m$; in particular, $\mu_m(0) = 2^{n-1}$. It is not difficult to show that for any $l = 1, 2, \ldots, \lfloor n/2 \rfloor$ we have $\mu_m(2l - 1) = \mu_m(2l)$ and otherwise $\mu_m(w)$ is a decreasing function of $w$.

Let $(\mu_m)_{x\in G}$ be the average value of the kernel $\mu_m(x, y)$ over the subset $E \subset X_n$. Since $(\mu_m)_{x\in \cal{X}_n} = \hat{\mu}_0$, we can write (28) in the following form:

$$2^n N^{-2} D_2^{(m)}(Z_N)^2 = \langle \mu_m \rangle_{Z_N} - \langle \mu_m \rangle_{\cal{X}_n}. \quad (30)$$

Relations (30), (28) are similar to the invariance principle for hemispheres in the case of the Euclidean sphere, [4] Thm. 3.1. At the same time, the concrete forms of the results for the Hamming space and the sphere are different: while for the sphere the quadratic discrepancy is expressed via the average geodesic distance in $Z_N$, in the Hamming case it is related to the average of the kernel $\mu_m$ and is not immediately connected to the average distance. Note that for quadratic discrepancy $D_2(G, Z_N)$ for the Hamming space defined above in (2), results of this form were previously established in [1].

Our final result in this section concerns a characterization of codes with zero discrepancy for hemispheres for the case of even $N$.

**Theorem 3.4.** Let $Z_N$ be a code of even size $N$. Then $D_2^{(m)}(Z_N) = 0$ if and only if the code $Z_N$ is formed of $N/2$ antipodal pairs of points.

**Proof.** The sufficiency part has been proved in Theorem 3.1. The proof in the other direction is a combination of the following steps.

**Step 1.** Since $\hat{\mu}_k > 0$ for all $k$, expression (29) implies that a code $Z_N \subset X_n$ has zero quadratic discrepancy for hemispheres if and only if its dual distance coefficients $A_k^\perp \neq 0$ only if $k$ is even.

**Step 2.** A code $Z_N$ is formed of antipodal pairs if and only if its distance distribution is symmetric, i.e., $A_w = A_{n-w}$ for all $w = 0, \ldots, n$.

Indeed, the distance distribution coefficients $A_w, w = 0, \ldots, n$ can be written as

$$A_w = \sum_{z \in Z_N} A_w(z), \quad (31)$$

where $A_w(z) = \frac{1}{N} \{ y : d(z, y) = w \}$ is the local distance distribution at the point $z \in Z_N$.

Suppose the code is formed of antipodal pairs. For every $y \in Z_N$ such that $d(z, y) = w$, the opposite point $\bar{y}$ satisfies $d(z, \bar{y}) = n - w$, and thus, the pair $(y, \bar{y})$ contributes to $A_w(z)$ and $A_{n-w}(z)$ in equal amounts. Therefore, from (31) also $A_w = A_{n-w}$.

Now suppose that the distance distribution is symmetric. For any code $A_0 = 1$, and then also $A_n = 1$, but this means that every code point has a diametrically opposite one, or otherwise (31) cannot be satisfied for $w = n$.

**Step 3.** The matrix

$$\Phi_m = \begin{pmatrix}
K_1^{(n)}(0) & K_1^{(n)}(1) & \cdots & K_1^{(n)}(m) \\
K_2^{(n)}(0) & K_2^{(n)}(1) & \cdots & K_2^{(n)}(m) \\
\vdots & \vdots & \ddots & \vdots \\
K_{2m+1}^{(n)}(0) & K_{2m+1}^{(n)}(1) & \cdots & K_{2m+1}^{(n)}(m)
\end{pmatrix}$$
has rank $m + 1$. This is shown as follows. Orthogonality of Krawtchouk polynomials [11, Thm 5.16] implies that
\[
\binom{n}{k} 2^n \delta_{j,k} = \sum_{w=0}^{2m+1} K_k^{(n)}(w) K_j^{(n)}(w) \binom{n}{w}
\]
\[
= \sum_{w=0}^{m} K_k^{(n)}(w) K_j^{(n)}(w) \binom{n}{w} + \sum_{w=m+1}^{2m+1} (-1)^{j+k} K_k^{(n)}(n-w) K_j^{(n)}(n-w) \binom{n}{n-w}
\]
\[
= 2 \sum_{w=0}^{m} K_k^{(n)}(w) K_j^{(n)}(w) \binom{n}{w}.
\]

Here on the third line we used the relation
\[
K_k^{(n)}(w) = (-1)^k K_k^{(n)}(n-w), \quad 0 \leq k, w \leq n.
\]
which is immediate from (26). In other words, for odd $j, k$ we have
\[
\sum_{w=0}^{m} K_k^{(n)}(w) K_j^{(n)}(w) \binom{n}{w} = \delta_{k,j} 2^{n-1} \binom{n}{k}.
\]
Rephrasing this relation, we obtain
\[
\Phi_m^T B \Phi_m = 2^{n-1} \text{diag}(\binom{n}{1}, \binom{n}{3}, \ldots, \binom{n}{2m+1}),
\]
where $B = \text{diag}(\binom{n}{w}, w = 0, 1, \ldots, m)$. This implies that $\text{rank}(\Phi_m) = m$.

**Step 4.** To complete the proof, suppose that $B_2^{(m)}(Z_N) = 0$ and thus from Step 1 above, $A_k^T = 0$ for all odd $k$. In particular, for $k = 1, 3, \ldots, 2m + 1$ using (25) and (32) we obtain
\[
\sum_{w=0}^{2m+1} A_w K_k^{(n)}(w) = \sum_{w=0}^{m} (A_w - A_{n-w}) K_k^{(n)}(w) = 0.
\]
Define the vector
\[
\alpha_m = (A_w - A_{n-w}, w = 0, 1, \ldots, m).
\]
From (34) and the definition of $\Phi_m$ we obtain that $\Phi_m \alpha_T = 0$. From Step 3), we conclude that $\alpha = 0$ or $A_w = A_{n-w}$, $w = 0, 1, \ldots, m$. Now Step 2 implies our claim.

\[\square\]

**References**

[1] A. Barg, “Stolarsky’s invariance principle for finite metric spaces,” preprint [arXiv:2005.12995], May 2020.
[2] J. Beck, “Sums of distances between points on a sphere—an application of the theory of irregularities of distribution to discrete geometry,” *Mathematika*, vol. 31, no. 1, pp. 33–41, 1984.
[3] J. Beck and W. W. L. Chen, *Irregularities of Distribution*. Cambridge, UK: Cambridge University Press, 1987.
[4] D. Bilyk, F. Dai, and R. Matzke, “The Stolarsky principle and energy optimization on the sphere,” *Constr. Approx.*, vol. 48, pp. 31–60, 2018.
[5] D. Bilyk and M. Lacey, “One-bit sensing, discrepancy and Stolarsky’s principle,” *Sbornik: Mathematics*, vol. 208, no. 6, pp. 744–763, 2017, translation from the Russian, *Mat. Sbornik*, vol. 208, no. 6, pp.4–25.
[6] L. Brandolini, W. W. L. Chen, L. Colzani, G. Gigante, and G. Travaglini, “Discrepancy and numerical integration on metric measure spaces,” *J. Geom. Anal.*, vol. 29, no. 1, pp. 328–369, 2019.
[7] J. S. Brauchart, P. J. Grabner, W. B. Kusner, and J. Ziefle, “Hyperuniform point sets on the sphere: probabilistic aspects,” *Monatshefte für Mathematik*, vol. 192, pp. 763–781, 2020.
[8] J. S. Brauchart and P. J. Grabner, “Distributing many points on spheres: minimal energy and designs,” *J. Complexity*, vol. 31, no. 3, pp. 293–326, 2015.
[9] J. Brauchart and J. Dick, “A simple proof of Stolarsky’s invariance principle,” *Proc. Amer. Math. Soc.*, vol. 141, pp. 2085–2096, 2013.
[10] Y. Chow and H. Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 2nd ed. Springer, 1988.
[11] P. Delsarte, “Bounds for unrestricted codes, by linear programming,” *Philips Res. Rep.*, vol. 27, pp. 272–289, 1972.
[12] V. I. Levenshtein, “Universal bounds for codes and designs,” in *Handbook of Coding Theory* (V. S. Pless and W. C. Huffman eds.), Vol. 1, Chapter 6, pages 499–648, Elsevier, 1998.
[13] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North-Holland, 1991.
[14] J. Matoušek, *Geometric Discrepancy: An Illustrated Guide*. Berlin-Heidelberg: Springer-Verlag, 1999.
[15] M. M. Skriganov, “Point distributions in compact metric spaces,” *Mathematika*, vol. 63, no. 3, pp. 1152–1171, 2017.
[16] ——, “Point distributions in two-point homogeneous spaces,” *Mathematika*, vol. 65, no. 3, pp. 557–587, 2019.
[17] ——, “Bounds for $L_p$-discrepancies of point distributions in compact metric spaces,” *Constr. Approx.*, vol. 51, pp. 413–425, 2020.
[18] ——, “Stolarsky’s invariance principle for projective spaces,” *J. Complexity*, vol. 56, p. 101428, 2020.
[19] K. B. Stolarsky, “Sums of distances between points on a sphere, II,” *Proc. AMS*, vol. 41, pp. 575–582, 1973.