CODING THEORY AND UNIFORM DISTRIBUTIONS

Alexander Barg

University of Maryland, College Park

TBSI Workshop on Learning Theory WOLT’20

July 21, 2020
Uniformly distributed point sets
U.D. point sets approximate random subsets of the metric space

Classical theory of uniform distributions (Weyl, 1916) developed to measure errors in numerical (QMC) integration on $\mathcal{X} = [0, 1]^d$

A set $Z_N = \{z_1, \ldots, z_N\} \subset \mathcal{X}$ is used to approximate $\int_{\mathcal{X}} f(x) dx \approx \frac{1}{N} \sum_{i=1}^{N} f(z_i)$

The error is measured by the discrepancy of $Z_N$. Let $A \subset \mathcal{X}$ and define

$$D(Z_N, A) = |Z_N \cap A| - N \cdot \text{vol}(A)$$

$$D_\infty(Z_N, A) = \sup_{A \in \mathcal{A}} D(Z_N, A)$$
Uniformly distributed point sets

U.D. point sets approximate random subsets of the metric space

**Classical theory** of uniform distributions (Weyl, 1916) developed to measure errors in numerical (QMC) integration on $X = [0, 1]^d$

A set $Z_N = \{z_1, \ldots, z_N\} \subset X$ is used to approximate $\int_X f(x) dx \approx \frac{1}{N} \sum_{i=1}^{N} f(z_i)$

The error is measured by the **discrepancy** of $Z_N$. Let $A \subset X$ and define

$$D(Z_N, A) = |Z_N \cap A| - N \cdot \text{vol}(A)$$

$$D_{\infty}(Z_N, \mathcal{A}) = \sup_{A \in \mathcal{A}} D(Z_N, A)$$

**Roth (1954):** For $\mathcal{A} = \{\text{boxes in } X\}$, any $Z_N$

$$D_{\infty}(Z_N, \mathcal{A}) \geq D_2(Z_N, \mathcal{A}) = \Omega_d((\log N)^{(d-1)/2})$$
Quadratic discrepancy on the sphere $S^d(\mathbb{R})$

\[ D_{\text{Cap}}^{L_2}(Z_N) = \int_{-1}^{1} \int_{S^d} \left( \frac{1}{N} |C(x, t) \cap Z_N| - \sigma(C(x, t)) \right)^2 d\sigma(x) dt \]

where $C(x, t) = \{ y \in S^d \mid (x, y) \geq t \}$ is a spherical cap
Quadratic discrepancy on the sphere

Quadratic discrepancy on the sphere $S^d(\mathbb{R})$

$$D^L_{Cap}(Z_N) = \int_{-1}^{1} \int_{S^d} \left( \frac{1}{N} |C(x, t) \cap Z_N| - \sigma(C(x, t)) \right)^2 d\sigma(x) dt$$

where $C(x, t) = \{y \in S^d \mid (x, y) \geq t\}$ is a spherical cap

Applications:

- Geometric method of testing correlations between gene expression and disease status, computationally simpler than permutation approximation

Quadratic discrepancy on the sphere

\[ D^L_\text{Cap} (Z_N) = \int_{-1}^1 \int_{S^d} \left( \frac{1}{N} |C(x, t) \cap Z_N| - \sigma(C(x, t)) \right)^2 d\sigma(x) dt \]

where \( C(x, t) = \{ y \in S^d \mid (x, y) \geq t \} \) is a spherical cap.

Applications:

- Geometric method of testing correlations between gene expression and disease status, computationally simpler than permutation approximation
  
  (He et al., *Annals of Statistics*, 2019)

- “Local variance” as order metric for point patterns in \( R^d \)
  
Stolarsky’s invariance principle for connected spaces

Stolarsky’s invariance principle, 1973

\[ c_d \left( D_{\text{Cap}}^{L^2} (Z_N) \right)^2 = \iint_{S^d \times S^d} \| x - y \| d\sigma(x) d\sigma(y) - \frac{1}{N^2} \sum_{i,j=1}^{N} \| z_i - z_j \|, \]

where \( c_d = \frac{\sqrt{\pi} \Gamma(d/2)}{\Gamma((d+1)/2)} \).
Stolarsky’s invariance principle, 1973

\[ c_d(D_{\text{Cap}}^{L_2}(Z_N))^2 = \int \int_{S^d \times S^d} \|x - y\|d\sigma(x)d\sigma(y) - \frac{1}{N^2} \sum_{i,j=1}^{N} \|z_i - z_j\|, \]

where \( c_d = \frac{d \sqrt{\pi} \Gamma(d/2)}{\Gamma((d+1)/2)} \).

Minimum quadratic discrepancy is equivalent to maximum average distance in \( Z_N \).
Stolarsky’s invariance principle for connected spaces

Stolarsky’s invariance principle, 1973

\[ c_d \left( D_{\text{Cap}}^{L_2}(Z_N) \right)^2 = \int_{S^d \times S^d} \|x - y\| d\sigma(x) d\sigma(y) - \frac{1}{N^2} \sum_{i,j=1}^{N} \|z_i - z_j\|, \]

where \( c_d = \frac{d \sqrt{\pi} \Gamma(d/2)}{\Gamma((d+1)/2)}. \)

**Minimum quadratic discrepancy** is equivalent to **maximum average distance** in \( Z_N \)

General results regarding *universally optimal spherical codes* (COHN-KUMAR, 2007) imply that they minimize quadratic discrepancy

See also BORODACHOV, HARDIN, SAFF, Discrete Energy on Rectifiable Sets. Springer, 2019
Bounds on $D_{L^2}$ (the case of $S^d$)

Classical results of BECK (1987) and ALEXANDER (1972) imply that for any $Z_N$ of size $N$

$$D_{L^2, \text{Cap}} > cN^{-1/2(1+1/d)},$$

and that there exist point distributions such that

$$D_{L^2, \text{Cap}} < CN^{-1/2(1+1/d)}.$$
Bounds on $D_{L^2}^{}$ (the case of $S^d$)

Classical results of Beck (1987) and Alexander (1972) imply that for any $Z_N$ of size $N$

$$D_{L^2}^{\text{Cap}} > cN^{-1/2(1+1/d)},$$

and that there exist point distributions such that

$$D_{L^2}^{\text{Cap}} < CN^{-1/2(1+1/d)}.$$ 

Optimal spherical designs of Bondarenko-Radchenko-Viazovska (2013) meet the lower bound:

$$cN^{-1/2(1+1/d)} < D_{L^2}^{\text{Cap}} < CN^{-1/2(1+1/d)}$$

for sufficiently large $N$, absolute constants $c$ and $C$ (Skiganov, 2019)
References

Books on discrepancy:

MATOUŠEK (2010)
References

Books on discrepancy:

MATOUŠEK (2010)

K. STOLARSKY, Sums of distances between points on a sphere II, Proc AMS, 1973
References

Books on discrepancy:

**MATOUŠEK** (2010)

**K. STOLARSKY**, *Sums of distances between points on a sphere II*, Proc AMS, 1973

*Recent literature:*

References

Books on discrepancy:

MATOUŠEK (2010)

K. STOLARSKY, Sums of distances between points on a sphere II, Proc AMS, 1973

Recent literature:


♦ D. BILYK, F. DAI, AND R. MATZKE, The Stolarsky principle and energy optimization on the sphere, Constr. Approx., 48, 2018

♦ M. M. SKRIGANOV, Point distributions in two-point homogeneous spaces, Mathematika, 65, 2019

This talk:


Discrepancy

Let $\mathcal{X}$ be a finite metric space; distances $d \in \{0, 1, \ldots, n\}$

$Z_N = \{z_1, \ldots, z_N\} \subset \mathcal{X}$

- **Problem:** Is $Z_N$ “uniformly distributed” in $\mathcal{X}$?
Discrepancy

Let $\mathcal{X}$ be a finite metric space; distances $d \in \{0, 1, \ldots, n\}$

$Z_N = \{z_1, \ldots, z_N\} \subset \mathcal{X}$

- Problem: Is $Z_N$ “uniformly distributed” in $\mathcal{X}$?

A subset $Z_N$ is u.d. if for all $x \in \mathcal{X}, t \in \{0, 1, \ldots, n\}$

$$\frac{|Z_N \cap B(x, t)|}{N} = \text{vol}(B(x, t))$$

where $B(x, t) = \{z \in \mathcal{X} | d(z, x) \leq t\}$ is a metric ball in $\mathcal{X}$ centered at $x$
Discrepancy

Let $\mathcal{X}$ be a finite metric space; distances $d \in \{0, 1, \ldots, n\}$

$Z_N = \{z_1, \ldots, z_N\} \subset \mathcal{X}$

- Problem: Is $Z_N$ “uniformly distributed” in $\mathcal{X}$?

A subset $Z_N$ is u.d. if for all $x \in \mathcal{X}, t \in \{0, 1, \ldots, n\}$

$$\frac{|Z_N \cap B(x, t)|}{N} = \text{vol}(B(x, t))$$

where $B(x, t) = \{z \in \mathcal{X} | d(z, x) \leq t\}$ is a metric ball in $\mathcal{X}$ centered at $x$

Quadratic discrepancy of $Z_N$:

$$D^{L2}(Z_N) = \sum_{t=0}^{n} (D_t(Z_N))^2$$

where

$$D_t(Z_N) := \left( \sum_{x \in \mathcal{X}} \left( \frac{|B(x, t) \cap Z_N|}{N} - \frac{1}{|\mathcal{X}|} |B(x, t)| \right)^2 \right)^{1/2}.$$
Finite metric spaces

Theorem (STOLARSKY’S INVARIANCE PRINCIPLE)

Let $Z_N = \{z_1, \ldots, z_N\}$ be a subset of a finite metric space $\mathcal{X}$. Then

$$D^{L^2}(Z_N) = \frac{1}{2} \left( \frac{1}{|\mathcal{X}|^2} \sum_{x,y \in \mathcal{X}} \sum_{u \in \mathcal{X}} |d(x, u) - d(y, u)| - \frac{1}{N^2} \sum_{i,j=1}^{N} \sum_{u \in \mathcal{X}} |d(z_i, u) - d(z_j, u)| \right).$$
Theorem (Stolarsky’s Invariance Principle)

Let $Z_N = \{z_1, \ldots, z_N\}$ be a subset of a finite metric space $X$. Then

$$D^{L^2}(Z_N) = \frac{1}{2}\left(\frac{1}{|X|^2} \sum_{x,y \in X} \sum_{u \in X} |d(x,u) - d(y,u)| - \frac{1}{N^2} \sum_{i,j=1}^{N} \sum_{u \in X} |d(z_i,u) - d(z_j,u)|\right).$$

Rephrasing:

$$\lambda(x,y) := \frac{1}{2} \sum_{u \in X} |d(x,u) - d(y,u)|$$

Then

$$D^{L^2}(Z_N) = \frac{1}{|X|^2} \sum_{x,y \in X} \lambda(x,y) - \frac{1}{N^2} \sum_{i,j=1}^{N} \lambda(z_i,z_j)$$

$$= \langle \lambda \rangle_X - \langle \lambda \rangle_{Z_N}$$
Proof outline

\[ D_t(Z_N)^2 = \sum_{x \in \mathcal{X}} \left( \frac{1}{N} \sum_{j=1}^{N} \mathbb{1}_{B(x, t)}(z_j) - \frac{1}{|\mathcal{X}|} |B(x, t)| \right)^2 \]

\[ = \frac{1}{N^2} \sum_{i,j=1}^{N} |B(z_i, t) \cap B(z_j, t)| - \frac{|B(u, t)|^2}{|\mathcal{X}|}, \]

\[ \sum_{t=0}^{n} |B(x, t) \cap B(y, t)| = |\mathcal{X}|(n + 1) - \sum_{z \in \mathcal{X}} d(z, u) - \frac{1}{2} \sum_{z \in \mathcal{X}} |d(z, x) - d(z, y)| \,
\]
Proof outline

\[ D_t(Z_N)^2 = \sum_{x \in \mathcal{X}} \left( \frac{1}{N} \sum_{j=1}^{N} \mathbf{1}_{B(x, t)}(z_j) - \frac{1}{|\mathcal{X}|} |B(x, t)| \right)^2 \]

\[ = \frac{1}{N^2} \sum_{i,j=1}^{N} |B(z_i, t) \cap B(z_j, t)| - \frac{|B(u, t)|^2}{|\mathcal{X}|}, \]

\[ \sum_{t=0}^{n} |B(x, t) \cap B(y, t)| = |\mathcal{X}|(n + 1) - \sum_{z \in \mathcal{X}} d(z, u) - \frac{1}{2} \sum_{z \in \mathcal{X}} |d(z, x) - d(z, y)| \]

Another form of the invariance principle: Define

\[ \mu(x, y) = \sum_{t=0}^{\infty} \mu_t(x, y), \text{ where } \mu_t(x, y) := |B(x, t) \cap B(y, t)| \]

Then

\[ D(Z_N)^2 = \langle \mu \rangle_{Z_N} - \langle \mu \rangle_{\mathcal{X}} \]
Kernel $\lambda(x, y)$ in the Hamming space

Lemma

Let $x, y \in X_n$ be two points such that $d(x, y) = w$. Then

$$
\lambda(x, y) = \lambda(w) := 2^{n-w} w \left( \frac{w - 1}{\left\lfloor \frac{w}{2} \right\rfloor - 1} \right), \quad w = 0, 1, \ldots, n.
$$

We have

$$
\frac{\lambda(2i + 1)}{2i + 1} = \frac{\lambda(2i)}{2i}, \quad i \geq 1,
$$

and thus $\lambda(i)$ is a monotone non-decreasing function of $i$ for all $i \geq 1$. 
Average value $\langle \lambda \rangle_{\mathcal{X}_n}$

$$
\langle \lambda \rangle_{\mathcal{X}_n} = 2^{-2n} \sum_{x,y \in \mathcal{X}_n} \lambda(d(x,y))
$$

$$
= 2^{-n} \sum_{w=0}^{n} \binom{n}{w} \lambda(w)
$$

$$
= 2^{-n} \sum_{w=1}^{n} 2^{n-w} \binom{n}{w} \left( \binom{w-1}{\left\lfloor \frac{w}{2} \right\rfloor} - 1 \right)
$$

$$
= \frac{n}{2^{n+1}} \binom{2n}{n}
$$
**Average value** $\langle \lambda \rangle_{\mathcal{X}_n}$

\begin{align*}
\langle \lambda \rangle_{\mathcal{X}_n} &= 2^{-2n} \sum_{x, y \in \mathcal{X}_n} \lambda(d(x, y)) \\
&= 2^{-n} \sum_{w=0}^{n} \binom{n}{w} \lambda(w) \\
&= 2^{-n} \sum_{w=1}^{n} 2^{n-w} \binom{n}{w} \left( \binom{w - 1}{\left\lfloor \frac{w}{2} \right\rfloor} - 1 \right) \\
&= \frac{n}{2^{n+1}} \binom{2n}{n}
\end{align*}

Computed using

\begin{equation*}
\sum_{i=0}^{t} \left( \sum_{i=0}^{n} \binom{n}{i} \right)^2 = 2^{2n-1} (n + 2) - \frac{n}{2} \binom{2n}{n}
\end{equation*}

(OEIS A002457)
Stolarsky’s identity in the Hamming space

Distance distribution of the code $Z_N \subset \{0, 1\}^n$:

$$A_w = \frac{1}{N} |\{(x, y) \in Z_N \mid d(x, y) = w\}|, \quad w = 0, 1, \ldots, n$$
Stolarsky’s identity in the Hamming space

Distance distribution of the code $Z_N \subset \{0, 1\}^n$:

$$A_w = \frac{1}{N} \left| \{(x, y) \in Z_N \mid d(x, y) = w\} \right|, \quad w = 0, 1, \ldots, n$$

**Theorem** (STOLARSKY’S INVARiance FOR THE HAMMING SPACE)

Let $Z_N \subset \{0, 1\}^n$ be a subset of size $N$ with distance distribution $A(Z_N) = (1, A_1, \ldots, A_n)$. Then

$$D^2(Z_N) = \Lambda_n - \frac{1}{N} \sum_{w=1}^{n} A_w \lambda(w),$$

where $\Lambda_n := \frac{n}{2^{n+1}} \binom{2n}{n}$
Stolarsky’s identity in the Hamming space

Distance distribution of the code $Z_N \subset \{0, 1\}^n$:

$$A_w = \frac{1}{N} \left| \{(x, y) \in Z_N \mid d(x, y) = w\} \right|, \quad w = 0, 1, \ldots, n$$

Theorem (Stolarsky’s invariance for the Hamming space)

Let $Z_N \subset \{0, 1\}^n$ be a subset of size $N$ with distance distribution $A(Z_N) = (1, A_1, \ldots, A_n)$. Then

$$D^{L_2}(Z_N) = \Lambda_n - \frac{1}{N} \sum_{w=1}^{n} A_w \lambda(w),$$

where $\Lambda_n := \frac{n}{2^{n+1}} \binom{2n}{n}$

This result enables us to compute or estimate $D^{L_2}(Z_N)$ for various binary codes.
First results

**Moments of discrepancy**

Choose a subset $\mathcal{Z}_N \subset \mathcal{X}_n$ randomly and independently

**Proposition**

The expected discrepancy of a random code of size $N$ in $\{0, 1\}^n$ equals

$$E[D^2(\mathcal{Z}_N)] = \frac{n}{N2^{n+1}} \binom{2n}{n} \approx \sqrt{\frac{n}{\pi}} \frac{2^{n-1}}{N}$$
Dual view of discrepancy

Define the dual distance distribution $A^\perp (Z_N) = (A_0^\perp, \ldots, A_n^\perp)$ of the code $Z_N$:

$$A_w^\perp = \frac{1}{N} \sum_{i=0}^{n} K_w^{(n)}(i) A_i, \quad w = 0, 1, \ldots, n$$

where

$$K_k^{(n)}(x) = \binom{n}{k} {}_2F_1(-k, -x; -n; 2).$$

are the Krawtchouk polynomials.
Dual view of discrepancy

Define the dual distance distribution \( A^\perp(Z_N) = (A^\perp_0, \ldots, A^\perp_n) \) of the code \( Z_N \):

\[
A^\perp_w = \frac{1}{N} \sum_{i=0}^{n} K_w^{(n)}(i)A_i, \quad w = 0, 1, \ldots, n
\]

where

\[
K_k^{(n)}(x) = \binom{n}{k} {}_2F_1(-k, -x; -n; 2).
\]

are the Krawtchouk polynomials.

We obtain

\[
D^{L2}(Z_N) = \Lambda_n - \frac{1}{2^n} \sum_{i=0}^{n} A_i^\perp \sum_{w=0}^{n} K_w^{(n)}(i)\lambda(w)
\]

\[
= -\frac{1}{2^n} \sum_{i=1}^{n} A_i^\perp \sum_{w=0}^{n} K_w^{(n)}(i)\lambda(w)
\]
Hamming codes

**Theorem**

The quadratic discrepancy of the Hamming code $Z_N = \mathcal{H}_m$ of length $n = 2^m - 1, m \geq 2$ equals

$$D^{L^2} (\mathcal{H}_m) = \frac{n}{2^n} \left( \frac{n - 1}{n - \frac{1}{2}} \right)$$

For large $n$ the discrepancy $D^{L^2} (\mathcal{H}_m) = \sqrt{n/4\pi} (1 - o(1))$.

### Discrepancy of the Hamming Codes and their Duals

<table>
<thead>
<tr>
<th>$m$</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D^{L^2} (\mathcal{H}_m)$</td>
<td>1.571</td>
<td>2.239</td>
<td>3.179</td>
<td>4.50471</td>
<td>6.377</td>
<td>9.027</td>
<td>12.763</td>
</tr>
<tr>
<td>$ED^{L^2} (N)$</td>
<td>17.336</td>
<td>50.058</td>
<td>143.016</td>
<td>406.518</td>
<td>1152.64</td>
<td>3264.14</td>
<td>9238.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^{-n} D^{L^2} (\mathcal{H}^\perp_m)$</td>
<td>0.058</td>
<td>0.042</td>
<td>0.030</td>
<td>0.021</td>
<td>0.015</td>
<td>0.011</td>
<td>0.008</td>
</tr>
<tr>
<td>$2^{-n} ED^{L^2} (N)$</td>
<td>0.068</td>
<td>0.049</td>
<td>0.035</td>
<td>0.025</td>
<td>0.018</td>
<td>0.012</td>
<td>0.009</td>
</tr>
</tbody>
</table>
Fourier-Krawtchouk expansions, I

Our next goal is to identify binary codes that have the smallest discrepancy, using tools from harmonic analysis.
Our next goal is to identify binary codes that have the smallest discrepancy, using tools from harmonic analysis.

**First step:** Compute a Krawtchouk expansion of $\lambda(x, y) = \lambda(w)$. Start with

$$\mu_t(x, y) := |B(x, t) \cap B(y, t)| = \sum_{z \in X_n} \phi_t(d(x, z)) \phi_t(d(z, y))$$

where $\phi_t = 1_{\{0, 1, \ldots, t\}}$
Our next goal is to identify binary codes that have the smallest discrepancy, using tools from harmonic analysis.

First step: Compute a Krawtchouk expansion of \( \lambda(x, y) = \lambda(w) \). Start with

\[
\mu_t(x, y) := |B(x, t) \cap B(y, t)| = \sum_{z \in X_n} \phi_t(d(x, z)) \phi_t(d(z, y))
\]

where \( \phi_t = 1_{\{0, 1, \ldots, t\}} \)

For the indicator function \( f \) we compute

\[
\phi_t(l) = 2^{-n} \sum_{k=0}^{n} c_k(t) K_k^{(n)}(l), \quad l = 0, 1, \ldots, n
\]

where

\[
c_0(t) = \sum_{i=0}^{t} \binom{n}{i}; \quad c_k(t) = \frac{1}{\binom{n}{k}} \sum_{i=0}^{t} \binom{n}{i} K_k^{(n)}(i) = K_t^{(n-1)}(k - 1), \quad k \geq 1
\]
Fourier-Krawtchouk expansions, II

Lemma

Let \( x, y \in X_n \) be such that \( d(x, y) = w \). The Krawtchouk expansion of the kernel \( \mu_t(x, y), t = 0, \ldots, n \) has the following form:

\[
\mu_t(x, y) = 2^{-n} \sum_{k=0}^{n} c_k(t)^2 K_k^{(n)}(w),
\]
Lemma

Let \( x, y \in \mathcal{X}_n \) be such that \( d(x, y) = w \). The Krawtchouk expansion of the kernel \( \mu_t(x, y), t = 0, \ldots, n \) has the following form:

\[
\mu_t(x, y) = 2^{-n} \sum_{k=0}^{n} c_k(t)^2 K_k^{(n)}(w),
\]

Corollary

Let \( x, y \in \mathcal{X}_n \) be such that \( d(x, y) = w \). We have

\[
\lambda(x, y) = \lambda(w) = \sum_{k=0}^{n} \hat{\lambda}_k K_k^{(n)}(w)
\]

\[
\hat{\lambda}_0 = \Lambda_n, \quad \hat{\lambda}_k = -2^{-n} \frac{(2n-2k)(2k-2)}{n-k} \binom{n-1}{k-1}, \quad k = 1, 2, \ldots, n,
\]

and thus the kernel \((-\lambda(x, y))\) is positive definite up to an additive constant.
FIG. 1: The plots show $\lambda(w)$ for $n = 20$ (left figure) and $n = 6$ (right figure). In the right plot we also show the Krawtchouk expansion, that is equal to $\lambda(w)$ at integer values of $w$. The plots are scaled by $2^{-n}$. 
Transform domain representation of $D_{L^2}(Z_N)$

We obtain an expansion of the discrepancy of the code $Z_N$

$$D_{L^2}(Z) = 2^{-n} \sum_{k=1}^{n} \frac{(2n-2k)(2k-2)}{n-k} \binom{n-1}{k-1} A_k$$
Transform domain representation of $D^{L_2}(Z_N)$

We obtain an expansion of the discrepancy of the code $Z_N$

\[ D^{L_2}(Z) = 2^{-n} \sum_{k=1}^{n} \frac{(2n-2k+1)(2k-1)}{(n-k)(k-1)} A_k^{n} \]

For instance, let $Z_N = \mathcal{H}_m$ be the Hamming code. We have $A_{\frac{n+1}{2}} = n$ and $A_k^{n} = 0$ o/w $(k \geq 1)$. Thus

\[ D^{L_2}(\mathcal{H}_m) = -n \lambda_{\frac{n+1}{2}} \]

\[ \lambda_{\frac{n+1}{2}} = -2^{-n} \binom{n-1}{\frac{n-1}{2}} \]
Define the “potential energy” of the code $Z_N \subset \mathcal{X}_n$

$$E_\lambda(Z_N) = \frac{1}{N} \sum_{i,j=1}^{N} \lambda(d(z_i, z_j))$$
Discrepancy and energy minimization

Define the “potential energy” of the code $Z_N \subset \mathcal{X}_n$

$$E_{\lambda}(Z_N) = \frac{1}{N} \sum_{i,j=1}^{N} \lambda(d(z_i, z_j))$$

Minimizing $D^2(Z_N)$ is equivalent to maximizing $E_{\lambda}(Z_N)$. This problem can be addressed by linear programming using the Delsarte conditions

$$\sum_{k=1}^{n} A_k K_i^{(n)}(k) \geq -\binom{n}{i}, \quad i = 1, \ldots, n$$

and $\sum_{k=1}^{n} A_k = N - 1$. 
Discrepancy and energy minimization

Define the “potential energy” of the code $Z_N \subset \mathcal{X}_n$

$$E_\lambda(Z_N) = \frac{1}{N} \sum_{i,j=1}^{N} \lambda(d(z_i, z_j))$$

Minimizing $D^2(Z_N)$ is equivalent to maximizing $E_\lambda(Z_N)$. This problem can be addressed by linear programming using the Delsarte conditions

$$\sum_{k=1}^{n} A_k K_i^{(n)}(k) \geq -\binom{n}{i}, i = 1, \ldots, n$$

and $\sum_{k=1}^{n} A_k = N - 1$.

Dualising, we obtain that any feasible solution of the linear program

$$\min \left\{ \sum_{i=0}^{n} \binom{n}{i} h_i - h_0 N \left| \sum_{i=0}^{n} h_i K_i^{(n)}(k) \leq -\lambda(k), k = 1, \ldots, n; h_i \geq 0, i = 1, \ldots, n \right\}$$

gives an upper bound on $E_\lambda(Z_N)$
Proposition (LP)

Let \( h(i) = \sum_{k=0}^{n} h_k K_k^{(n)}(i) \) be a polynomial on \( \{0, 1, \ldots, n\} \) such that (a), \( h_k \geq 0 \) for all \( k \geq 1 \) such that \( A_k > 0 \) and (b), \( h(i) \leq -\lambda(i) \) for all \( i \geq 1 \) such that \( A_i^\perp > 0 \). Then

\[
E_\lambda(Z_N) \leq h(0) - Nh_0
\]

with equality if and only if all the inequalities in the assumptions (a),(b) are satisfied with equality.

Uniform distributions and their applications
Finite metric spaces
Hamming space
Extensions

Bounds on discrepancy

Theorem

For any $N \geq 1$

$$D^{L^2}(n, N) \geq \Lambda_n - \frac{N - 1}{N} \lambda(n).$$

(1)

For $n = 2t - 1$, $N \geq 1$

$$D^{L^2}(n, N) \geq \begin{cases} 
\Lambda_n - \frac{2N-1}{2N} \lambda(t) & \text{if } t \text{ even} \\
\Lambda_n - \frac{Nn-(n-1)/2}{N(n+1)} \lambda(t) & \text{if } t \text{ odd.}
\end{cases}$$

(2)

For any $N \geq 1$

$$D^{L^2}(n, N) \geq \begin{cases} 
-(\frac{2^n}{N} - 1) \hat{\lambda}_{\frac{n}{2}} & \text{if } n \text{ even} \\
-(\frac{2^n}{N} - 1) \hat{\lambda}_{\frac{n+1}{2}} & \text{if } n \text{ odd.}
\end{cases}$$

(3)

Proof by fitting a polynomial to satisfy the conditions in Proposition (LP).

Computations are aided by knowing the Fourier coefficients $\hat{\lambda}_k$.
Bounds on discrepancy

Theorem

For any \( N \geq 1 \)

\[
D^{L^2}(n, N) \geq \Lambda_n - \frac{N - 1}{N} \lambda(n).
\]  

(1)

For \( n = 2t - 1, N \geq 1 \)

\[
D^{L^2}(n, N) \geq \begin{cases} 
\Lambda_n - \frac{2N-1}{2N} \lambda(t) & \text{t even} \\
\Lambda_n - \frac{Nn-(n-1)/2}{N(n+1)} \lambda(t) & \text{t odd.}
\end{cases}
\]  

(2)

For any \( N \geq 1 \)

\[
D^{L^2}(n, N) \geq \begin{cases} 
-(\frac{2^n}{N} - 1) \hat{\lambda}_{n/2}^2, & \text{n even} \\
-(\frac{2^n}{N} - 1) \hat{\lambda}_{n+1/2}, & \text{n odd.}
\end{cases}
\]  

(3)

▶ Proof by fitting a polynomial to satisfy the conditions in Proposition (LP).
▶ Computations are aided by knowing the Fourier coefficients \( \hat{\lambda}_k \)
▶ The bounds are obtained by using \( h(x) \) of degree 0 and \( n \)
Discrepancy minimizers

Theorem: *Binary perfect codes are discrepancy minimizers*

The following codes were found to be discrepancy minimizers by computer:

1. the Golay code with $n = 23, N = 4096$
2. the shortened Golay code
3. the twice shortened Golay code
4. the quadratic residue code with $n = 17, N = 512$
5. the 2-error-correcting BCH codes with $n = 31, N = 2^{21}$ and $n = 127, N = 2^{113}$ and their shortened codes.
In summary, we proved the following bounds on the quadratic discrepancy of binary codes in the Hamming space:

**Theorem**

*For large $n$ and $N = o(2^n)$ we have the asymptotic bounds*

$$c \frac{1}{\sqrt{n}} \frac{2^n}{N} \leq D^{L^2}(n, N) \leq C \sqrt{n} \frac{2^n}{N}$$

*for some constants $c, C$. The discrepancy $D^{L^2}(n, N)$ is bounded away from zero unless $\frac{2^n}{\sqrt{n}} = o(N)$.*

*If $N = 2^r n$, $0 < r < 1$, then*

$$(\log N)^{-1/2} N^{\alpha} \leq D^{L^2}(n, N) \leq (\log N)^{1/2} N^{\alpha},$$

*where $\alpha = \frac{1}{r} - 1$.***
Generalization: Weighted $L_p$ discrepancies

(This part is based on a joint work with Maxim Skriganov, arXiV:2007)

(Weighted) $L_p$ discrepancy:

$$D^{L_p}(G, Z_n) = \left( \sum_{t=0}^{n} g_t \sum_{x \in X} |D(Z_n, y, t)|^p \right)^{1/p}, \quad 0 < p < \infty$$

where

$$D(Z_n, x, t) = \frac{|B(x, t) \cap Z_N|}{N} - 2^{-n} |B(x, t)|$$

is the local discrepancy, and $G = (g_0, g_1, \ldots, g_n), g_t \geq 0, \sum_{t=0}^{n} g_t = 1$ is a vector of weights.

REMARK: This definition is analogous to weighted $L_p$ norms in functional analysis.

$L_p$ discrepancies have been earlier considered for the case of spherical sets (M.M. Skriganov, J. Complexity, 2020)
Bounds on $L_p$ discrepancy

**Idea:** Choose $N$ points randomly in $\mathcal{X}_n$, then

$$D(Z_n, x, t) = \sum_{i=1}^{N} \frac{1}{N} \zeta_i$$

where $\zeta_i(x, t) = 1_{B(x, t)}(z_i) - 2^{-n}|B(x, t)|$, $i = 1, \ldots, N$ are zero-mean random variables. Khinchine-type inequalities for the $p$th moment of the sum of independent RVs enable one to derive estimates of $D^{L_p}$. 
Bounds on $L_p$ discrepancy

Idea: Choose $N$ points randomly in $\mathcal{X}_n$, then

$$D(Z_n, x, t) = \sum_{i=1}^{N} \frac{1}{N} \zeta_i$$

where $\zeta_i(x, t) = \mathbb{1}_{B(x, t)}(z_i) - 2^{-n}|B(x, t)|$, $i = 1, \ldots, N$ are zero-mean random variables. Khinchine-type inequalities for the $p$th moment of the sum of independent RVs enable one to derive estimates of $D^{L_p}$.

**Theorem**

*For all $N \leq 2^{n-1}$, we have*

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

*for $1 \leq p < \infty$, and $D_p(G, n, N) \leq 2^{-(n/p) + 3/2} N^{-1/2}$ for $0 < p < 1$.***
Let $n = 2m + 1$ and consider only balls of radius $t = m$

For a subset $Z_N \subset \mathcal{X}_n$ define

$$D^{(m)}_p(Z_N) = \left( \sum_{x \in \mathcal{X}_n} |D(Z_N, x, m)|^p \right)^{1/p}, \quad 0 < p < \infty,$$

where

$$D(Z_N, x, m) = \frac{|B(x, m) \cap Z_N|}{N} - \frac{1}{2},$$

Let

$$D^{(m)}_\infty(Z_N) = \max_{x \in \mathcal{X}_n} |D(Z_N, x)|.$$
Main results for $D_p^{(m)}$

- Let $N = 2K$ be even, then

$$D_p^{(m)}(Z_N) \geq 0 \quad \forall Z_N \subseteq X_n, p \in (0, \infty],$$

with equality for subsets $Z_N$ consisting of $K$ pairs of antipodal points. For $p = 2$ the condition for equality is also necessary.

- Let $N = 2K + 1$ be odd, then

$$D_p^{(m)}(Z_N) \geq 2^{n/p-1}/N \quad \forall Z_N \subseteq X_n, p \in (0, \infty],$$

with equality for subsets $Z_N$ consisting of $K$ pairs of antipodal points supplemented with a single point.
Summary and open problems

- Previously the invariance principle was studied only for connected spaces such as $S^d$ and related projective spaces (Riemannian symmetric spaces of rank one). Concrete results relied on analytic methods specific to such spaces.
- Finite metric spaces require different methods (combinatorial, etc.). Relying on the structure of the Hamming space, we showed a number of results with no direct analogs in the continuous case.

A multitude of open questions:

- Classify discrepancy minimizers
- Study asymptotic behavior of discrepancy
- Structural results for other distance transitive finite (or disconnected infinite) metric spaces
- Explore applications of sets with small discrepancy