Discrepancy and Small Ball Inequalities

Dmitriy Bilyk & Michael Lacey & Armen Vagharshakyan

February 20, 2008

Quantitative Estimates of Uniform Distribution

$$[0, (x_1, \ldots, x_d)] = \prod_{t=1}^{d} [0, x_t]$$

Quantitative Estimates of Uniform Distribution

$$[0,(x_1,\ldots,x_d)] = \prod_{t=1}^{d} [0,x_t]$$

Definition

The *Discrepancy Function* of $\mathcal{P}_N = \{x_1,\ldots,x_N\} \subset [0,1]^d$ is

$$D_N(x) = \sharp(\mathcal{P} \cap [0,x]) - N[0,x]\,.$$

Quantitative Estimates of Uniform Distribution

$$[0,(x_1,\dots,x_d)] = \prod_{t=1}^{d}[0,x_t]$$

Definition

The *Discrepancy Function* of $\mathcal{P}_N = \{x_1,\dots,x_N\} \subset [0,1]^d$ is

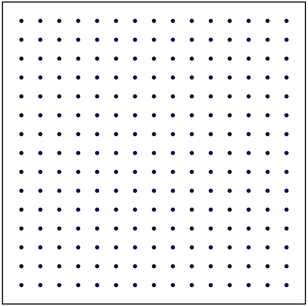
$$D_N(x) = \sharp(\mathcal{P} \cap [0,x]) - N[0,x]\,.$$

Koksma-Hlawka Inequality

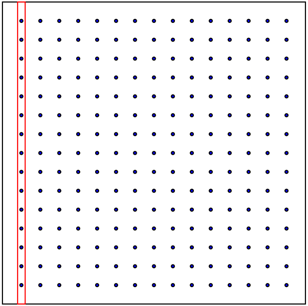
For any function $f\ :\ [0,1]^d \longrightarrow \mathbb{R}$ of bounded variation $\mathrm{V}(f)$ in the sense of Hardy, then

$$\left|\int_{[0,1]^d} f(y)\,dy - N^{-1}\sum_{j=1}^{N} f(x_j)\right| \le \mathrm{V}(f)\cdot \frac{\|D_N\|_\infty}{N}\,.$$

Lattices are *Not* Extremal Point Distributions



Lattices are *Not* Extremal Point Distributions



The area of the rectangle is tiny, but contains 1/15 of all the rectangles.

Random Selection is Bad

CLT: For measurable f, random X_n,

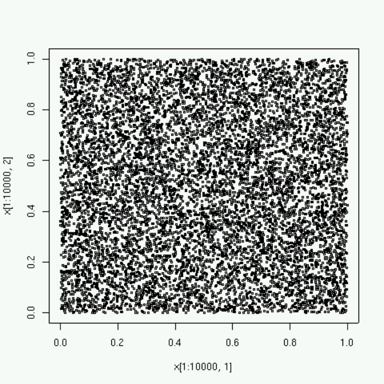
$$\frac{1}{N}\sum_{n=1}^{N} f(X_n) = \int_{[0,1]^d} f(x)\,dx + O(N^{-1/2}).$$

Random Selection is Bad

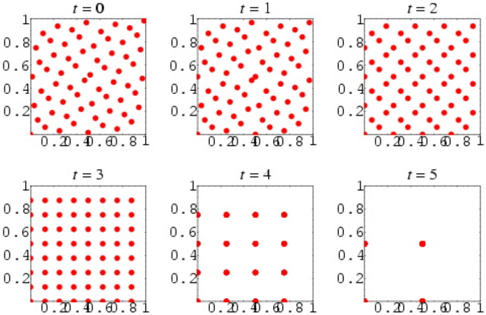
CLT: For measurable f, random X_n,

$$\frac{1}{N}\sum_{n=1}^{N} f(X_n) = \int_{[0,1]^d} f(x)\,dx + O(N^{-1/2})\,.$$

They cluster, and have gaps.



Examples of Low Discrepancy Set

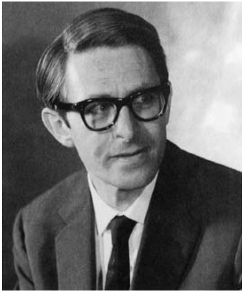


Roth's Theorem

For any choice of $\mathcal{P}_N$ we have

$$\|D_N\|_2 \gtrsim (\log N)^{(d-1)/2}$$

Two Giants: Klaus Roth and Wolfgang Schmidt



Roth Heuristic

For extremal distributions, one expects that each dyadic rectangle with volume N^{-1} has one point in it.

Roth Heuristic

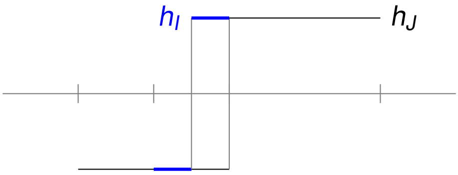
For extremal distributions, one expects that each dyadic rectangle with volume N^{-1} has one point in it.

Hyperbolic Haar Reduction

Consider dyadic rectangles of volume $(2N)^{-1}$; at least one-half of these must not contain *any* point in $\mathcal{P}_N$. Call these the *good* rectangles. And consider the Haar function associated to these dyadic rectangles.

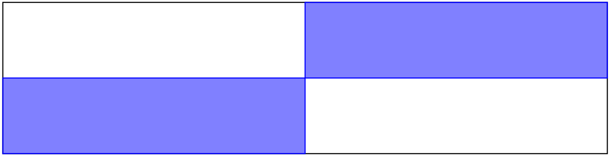
$$h_{I_1\times\cdots\times I_d}(x_1,\ldots,x_d) = \prod_{j=1}^{d}\left\{-\mathbf{1}_{I_{j,\text{left}}}(x_j) + \mathbf{1}_{I_{j,\text{right}}}(x_j)\right\}$$

One Dimensional Haar Functions



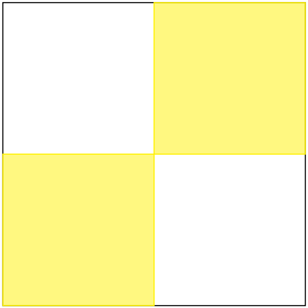
Two Dimensional Haar Functions

h_R

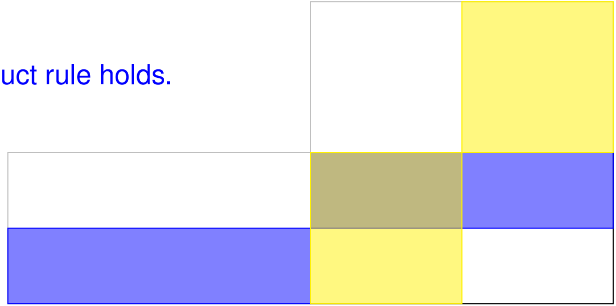


Two Dimensional Haar Functions

h_S



A product rule holds.



$$h_R \cdot h_S = -h_{R \cap S}$$

Proof of Roth Theorem

Lemma

If $R \cap \mathcal{P}_N = \emptyset$, *then* $\langle h_R, D_N \rangle < -cN|R|^2$.

Lemma

If $R \cap \mathcal{P}_N = \emptyset$, then $\langle h_R, D_N \rangle < -cN|R|^2$.

Proof.

$$\|D_N\|_2^2 \geq \sum_{R \text{ good}} |R|^{-1} |D_N, h_R|^2$$

$$\gtrsim N^2 \sum_{R \text{ good}} |R|^3 \gtrsim (\log N)^{d-1} .$$

□

Theorem

For any choice of point distribution $\mathcal{P}_N$ we have

$$\|D_N\|_p \gtrsim (\log N)^{(d-1)/2}\,, \qquad 1 < p < \infty\,.$$

Theorem

For any choice of point distribution $\mathcal{P}_N$ we have

$$\|D_N\|_p \gtrsim (\log N)^{(d-1)/2}, \qquad 1 < p < \infty .$$

There is however a ‘kink’ at L^∞ in Dimension $d = 2$.

Schmidt’s Theorem ($d = 2$!)

$$\|D_N\|_{L^\infty([0,1]^2)} \gtrsim \log N$$

A gain of $\sqrt{\log N}$ *over the average case bound.*

Conjecture: Discrepancy Function in L^∞

For $d \geq 3$,

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}$$

Conjecture: Discrepancy Function in L^∞

For $d \geq 3$,

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}$$

Conjecture: Small Ball Inequality

For $d \geq 3$, and generic choices of coefficients $a_R \in \{-1, 0, 1\}$,

$$\Big\| \sum_{|R|=2^{-n}} a_R h_R(x) \Big\|_\infty \gtrsim n^{d/2} .$$

Conjecture: Discrepancy Function in L^∞

For $d \geq 3$,

$$\|D_N\|_\infty \gtrsim (\log N)^{d/2}$$

Conjecture: Small Ball Inequality

For $d \geq 3$, and generic choices of coefficients $a_R \in \{-1, 0, 1\}$,

$$\Big\| \sum_{|R|=2^{-n}} a_R h_R(x) \Big\|_\infty \gtrsim n^{d/2} .$$

- $d = 2$ is a Theorem of Talagrand.
- Both conjectures are a 'gain of a square root' over the average case bounds.

Theorem (Bilyk & L & Vagharshakyan)

In dimension $d \geq 3$ there is a $\eta = \eta(d) \geq c/d^2 > 0$ for which we have

$$\Big\| \sum_{|R|=2^{-n}} a_R h_R \Big\|_\infty \gtrsim n^{(d-1)/2+\eta} . \qquad (1)$$

Beck established a version of this Theorem with $d = 3$ and

$$n^\eta \leftarrow (\log n)^{1/8} .$$

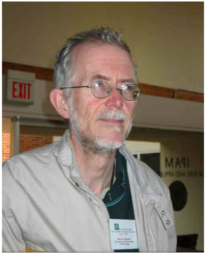
József Beck, *A two-dimensional van Aardenne-Ehrenfest theorem in irregularities of distribution* Compositio Math. **72** (1989) 269—339

Other Applications of the Small Ball Inequality

- Lower bounds on Packing Numbers of Unit Balls of certain Mixed Derivative Sobolev Spaces.
- For the Brownian Sheet B, upper bounds on

$$\mathbb{P}(\|B\|_{C([0,1]^d)} < \epsilon) , \qquad \epsilon \downarrow 0 .$$

Talagrand's Theorem–aprés Halasz, & Temlyakov



Talagrand's Theorem

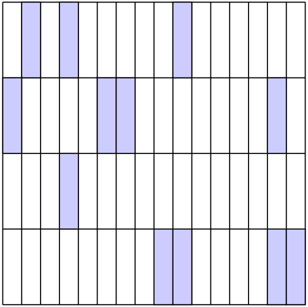
In dimension $d = 2$, for generic choices of coefficients $a_R \in \{-1, 0, 1\}$

$$\left\| \sum_{|R|=2^n} a_R \, h_R \right\|_\infty \gtrsim n \, .$$

$$H = \sum_{|R|=2^{-n}} a_R \, h_R$$

$$H = \sum_{|R|=2^{-n}} a_R \, h_R$$

$$f_{(k,n-k)} = \sum_{|R_1|=2^{-k}\,,|R_2|=2^{-n+k}} \operatorname{sgn}(a_R)\, h_R\,, \qquad 0 \le k \le n\,,$$



$$F = \prod_{k=1}^{n} \left(1 + f_{(k,n-k)}\right) \qquad F \geq 0\,, \qquad \mathbb{E}F = 1.$$

$$F = \prod_{k=1}^{n} \left(1 + f_{(k,n-k)}\right) \qquad F \geq 0\,, \qquad \mathbb{E}F = 1.$$

$$\sum_{k=0}^{n} \langle H, f_{(k,n-k)} \rangle = 2^{-n} \sum_{|R|=2^{-n}} |a_R|\,,$$

$$F = \prod_{k=1}^{n} \Big(1 + f_{(k,n-k)}\Big) \qquad F \geq 0\,, \qquad \mathbb{E}F = 1\,.$$

$$\sum_{k=0}^{n} \langle H, f_{(k,n-k)} \rangle = 2^{-n} \sum_{|R|=2^{-n}} |a_R|\,,$$

H is orthogonal to the higher products of the f_k.

$$F = \prod_{k=1}^{n} \left(1 + f_{(k,n-k)}\right) \qquad F \geq 0\,, \qquad \mathbb{E}F = 1.$$

$$\sum_{k=0}^{n} \langle H, f_{(k,n-k)} \rangle = 2^{-n} \sum_{|R|=2^{-n}} |a_R|\,,$$

H is orthogonal to the higher products of the f_k.

Note that the Riesz Product is

$$F = \prod_{k=0}^{n} \left(1 + f_k\right) = 2^n \mathbf{1}_E\,,$$

$$E = \{x \ : \ \text{all } f_k(x) \text{ equal one}\}$$

The 'Short' Bernoulli Product

- Set $H = \sum_{|R|=2^{-n}} a_R h_R$.

The 'Short' Bernoulli Product

- Set $H = \sum_{|R|=2^{-n}} a_R h_R$.
- Define $\operatorname{sgn}(0) = 1$, and

$$f_{\vec{r}} = \sum_{R \,:\, |R_j| = 2^{-r_j} \ \forall j} \operatorname{sgn}(a_R) h_R$$

The 'Short' Bernoulli Product

- Set $H = \sum_{|R|=2^{-n}} a_R h_R$.
- Define $\mathrm{sgn}(0) = 1$, and

$$f_{\vec{r}} = \sum_{R\,:\,|R_j|=2^{-r_j}\ \forall j} \mathrm{sgn}(a_R) h_R$$

- Set $q = n^{\epsilon}$, this will be the length of our product.

The 'Short' Bernoulli Product

- Set $H = \sum_{|R|=2^{-n}} a_R h_R$.
- Define $\operatorname{sgn}(0) = 1$, and

$$f_{\vec{r}} = \sum_{R\,:\,|R_j|=2^{-r_j}\ \forall j} \operatorname{sgn}(a_R) h_R$$

- Set $q = n^{\epsilon}$, this will be the length of our product.
- Divide the integers $\{1, 2, \ldots, n\}$ into q disjoint intervals $I_1, \ldots, I_q$, and let $\mathbb{I}_t \stackrel{\text{def}}{=} \{\vec{r} \in \mathbb{N}_n \,:\, r_1 \in I_t\}$.

The 'Short' Bernoulli Product

- Set $H = \sum_{|R|=2^{-n}} a_R h_R$.
- Define $\operatorname{sgn}(0) = 1$, and

$$f_{\vec{r}} = \sum_{R\,:\,|R_j|=2^{-r_j}\ \forall j} \operatorname{sgn}(a_R) h_R$$

- Set $q = n^{\epsilon}$, this will be the length of our product.
- Divide the integers $\{1, 2, \ldots, n\}$ into q disjoint intervals $I_1, \ldots, I_q$, and let $\mathbb{I}_t \stackrel{\text{def}}{=} \{\vec{r} \in \mathbb{N}_n \ :\ r_1 \in I_t\}$.
- We will define F_t as a poor man's $\operatorname{sgn}\Big(\sum_{\vec{r}\in\mathbb{I}_t} f_{\vec{r}}\Big)$.

- $$F_t = \widetilde{\rho} \sum_{\vec{r} \in \mathbb{I}_t} f_{\vec{r}} \,.$$

  $$\rho = \frac{q^{1/2}}{n^{(d-1)/2}}\,, \qquad \widetilde{\rho} = \frac{a q^{1/4}}{n^{(d-1)/2}}\,.$$

- $$F_t = \widetilde{\rho} \sum_{\vec{r} \in \mathbb{I}_t} f_{\vec{r}} \,.$$

  $$\rho = \frac{q^{1/2}}{n^{(d-1)/2}} \,, \qquad \widetilde{\rho} = \frac{a q^{1/4}}{n^{(d-1)/2}} \,.$$

- $$\Psi \stackrel{\text{def}}{=} \prod_{t=1}^{q} (1 + F_t) \,.$$

  This tends to be large when the F_t are all positive.

- $$F_t = \widetilde{\rho} \sum_{\vec{r} \in \mathbb{I}_t} f_{\vec{r}} \, .$$

  $$\rho = \frac{q^{1/2}}{n^{(d-1)/2}} \, , \qquad \widetilde{\rho} = \frac{a q^{1/4}}{n^{(d-1)/2}} \, .$$

- $$\Psi \stackrel{\text{def}}{=} \prod_{t=1}^{q} (1 + F_t) \, .$$

  This tends to be large when the F_t are all positive.

- $\mathbb{E}\Psi = 1$, as an easy conditional expectation argument shows, but Ψ takes negative values.

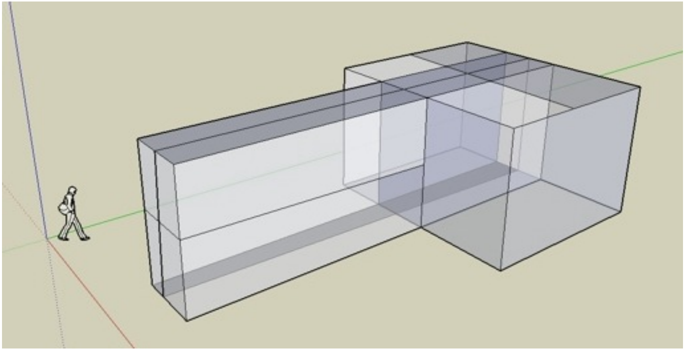
- $$F_t = \widetilde{\rho} \sum_{\vec{r} \in \mathbb{I}_t} f_{\vec{r}} \, .$$

  $$\rho = \frac{q^{1/2}}{n^{(d-1)/2}} \, , \qquad \widetilde{\rho} = \frac{a q^{1/4}}{n^{(d-1)/2}} \, .$$

- $$\Psi \stackrel{\mathrm{def}}{=} \prod_{t=1}^{q} (1 + F_t) \, .$$

  This tends to be large when the F_t are all positive.
- $\mathbb{E}\Psi = 1$, as an easy conditional expectation argument shows, but Ψ takes negative values.
- And, Ψ can not be the test function since...

Product Rule Fails in Three Dimensions



Coincidences

Definition

Say that $\vec{r}$, $\vec{s}$ have a *coincidence* if $\vec{r}$ and $\vec{s}$ agree in any one coordinate.

Coincidences

Definition

Say that $\vec{r}, \vec{s}$ have a *coincidence* if $\vec{r}$ and $\vec{s}$ agree in any one coordinate.

- Write $\Psi = 1 + \Psi^{\mathrm{NoCoin}} + \Psi^{\mathrm{Coin}}$

Coincidences

Definition

Say that $\vec{r}, \vec{s}$ have a *coincidence* if $\vec{r}$ and $\vec{s}$ agree in any one coordinate.

- Write $\Psi = 1 + \Psi^{\text{NoCoin}} + \Psi^{\text{Coin}}$
- $$\Psi^{\text{NoCoin}} \stackrel{\text{def}}{=} \text{All sums of products of } \vec{r} \text{ function in the expansion of } \Psi \text{ without a coincidence.}$$

Coincidences

Definition

Say that $\vec{r}, \vec{s}$ have a *coincidence* if $\vec{r}$ and $\vec{s}$ agree in any one coordinate.

- Write $\Psi = 1 + \Psi^{\text{NoCoin}} + \Psi^{\text{Coin}}$
-

$$\Psi^{\text{NoCoin}} \stackrel{\text{def}}{=} \text{All sums of products of } \vec{r} \text{ function in the expansion of } \Psi \text{ without a coincidence.}$$

- Most of the analysis takes place on Ψ^{Coin}.

The Crucial Lemma of Beck—In the Simplest Case

Lemma

We have the estimate

$$\Big\| \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{N}_n \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}} \Big\|_p \lesssim p^{2d\text{-}3/2+1} n^{2d\text{-}3/2}$$

The Crucial Lemma of Beck—In the Simplest Case

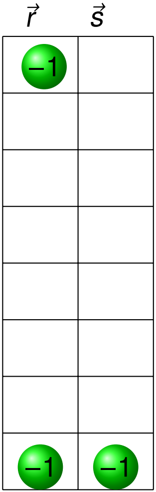
Lemma

We have the estimate

$$\Big\| \sum_{\substack{\vec{r} \neq \vec{s} \in \mathbb{N}_n \\ r_1 = s_1}} f_{\vec{r}} \cdot f_{\vec{s}} \Big\|_p \lesssim p^{(2d-3)/2+1} n^{(2d-3)/2}$$

There are 2d-3 'free' parameters.

The Number of Parameters

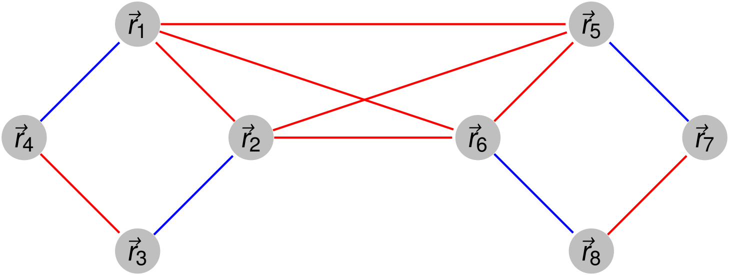


No of Parameters

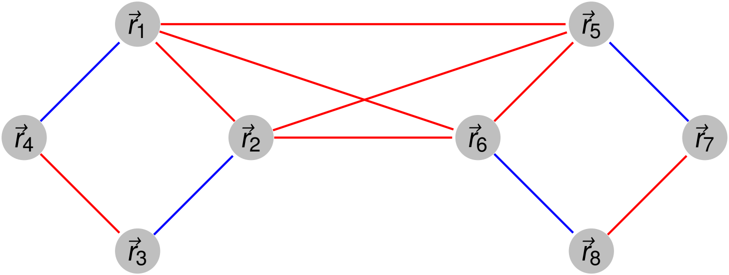
$$= 2d - 1 - 1 - 1$$

Longer Products: Graphs as Bookkeeping Device

Longer Products: Graphs as Bookkeeping Device



Longer Products: Graphs as Bookkeeping Device



- A graph on eight vertices, with two different colors.
- An edge means equality between the different vectors.
- So the number of parameters decreases with the length of spanning tree of the graph.
- The Beck Gain reflects a full proportion of the loss of parameters.

An Example Inequality, using previous graph:

For absolute $\zeta > 0$,

$$\Big\| \sum_{\substack{\vec{r}_1,\dots,\vec{r}_8 \\ \text{satisfy 'graph conditions'}}} \prod_{j=1}^{8} f_{\vec{r}_j} \Big\|_p \lesssim p^{4(d-1)} n^{4(d-1)-8\zeta},$$