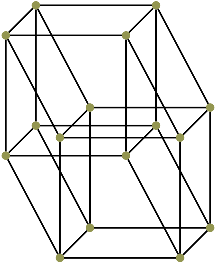
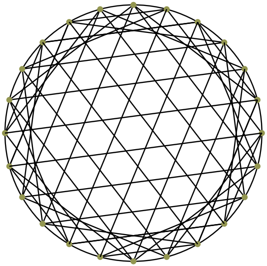
Vector-valued concentration inequalities on discrete spaces



Mira Gordin

Princeton University

November 6, 2024



Two perspectives on vector-valued concentration inequalities

Vector-valued analogs of classical concentration inequalities that help us understand high-dimensional random structures

Poincaré inequality

$$\mathrm{Var}(f(X)) \leq C_P \mathbb{E}|\nabla f(X)|^2$$

log-Sobolev inequality

$$\mathrm{Ent}(f^2(X)) \leq C \mathbb{E}|\nabla f(X)|^2$$

Vector-valued functional inequalities that describe phenomena in functional analysis and metric geometry

Example: consequences for metric embeddings of graphs in Banach spaces

$\downarrow$

algorithmic applications for embeddings

Pisier's inequalities

Let $(X, \|\cdot\|)$ be a Banach space.

Theorem (Pisier, 1985)

For $f : \mathbb{R}^n \to X$ locally Lipschitz, $G, G' \sim N(0, I_n)$ independent, and $1 \le p < \infty$,

$$\mathbb{E}\,\|f(G) - \mathbb{E}f(G)\|^p \le \left(\frac{\pi}{2}\right)^p \mathbb{E}\left\|\sum_{j=1}^n G_j' \frac{\partial f}{\partial x_j}(G)\right\|^p .$$

Pisier's inequalities

Let $(X, \|\cdot\|)$ be a Banach space.

Theorem (Pisier, 1985)

For $f : \mathbb{R}^n \to X$ locally Lipschitz, $G, G' \sim N(0, I_n)$ independent, and $1 \le p < \infty$,

$$\mathbb{E}\,\|f(G) - \mathbb{E}f(G)\|^p \le \left(\frac{\pi}{2}\right)^p \mathbb{E}\left\|\sum_{j=1}^n G'_j \frac{\partial f}{\partial x_j}(G)\right\|^p .$$

Theorem (Pisier, 1985)

For $f : \{-1,1\}^n \to X$, $\varepsilon, \varepsilon' \sim \mathrm{Unif}(\{-1,1\}^n)$ independent,

$$\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p \le C(n)^p \mathbb{E}\left\|\sum_{j=1}^n \varepsilon'_j D_j f(\varepsilon)\right\|^p \tag{1}$$

Pisier's inequalities

Let $(X, \|\cdot\|)$ be a Banach space.

Theorem (Pisier, 1985)

For $f : \{-1,1\}^n \to X$, $\varepsilon, \varepsilon' \sim \mathrm{Unif}(\{-1,1\}^n)$ independent,

$$\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p \le C(n)^p \mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon'_j D_j f(\varepsilon)\right\|^p \tag{1}$$

Pisier's inequalities

Let $(X, \|\cdot\|)$ be a Banach space.

Theorem (Pisier, 1985)

For $f : \{-1, 1\}^n \to X$, $\varepsilon, \varepsilon' \sim \mathrm{Unif}(\{-1, 1\}^n)$ independent,

$$\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p \leq C(n)^p \mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon'_j D_j f(\varepsilon)\right\|^p \tag{1}$$

- Pisier showed $C \sim \log n$
- Talagrand (1993) proved sharpness
- Naor and Schechtman (2002) dimension-free constant for UMD Banach spaces

Pisier's inequalities

Let $(X, \|\cdot\|)$ be a Banach space.

Theorem (Pisier, 1985)

For $f : \{-1,1\}^n \to X$, $\varepsilon, \varepsilon' \sim \mathrm{Unif}(\{-1,1\}^n)$ independent,

$$\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p \le C(n)^p \mathbb{E}\left\|\sum_{j=1}^n \varepsilon_j' D_j f(\varepsilon)\right\|^p \tag{1}$$

- Pisier showed $C \sim \log n$
- Talagrand (1993) proved sharpness
- Naor and Schechtman (2002) dimension-free constant for UMD Banach spaces

Is there another way to think about vector-valued concentration to get the "right" dimensional dependence?

Dimension-free constant on the discrete hypercube

Theorem (Ivanisvili, van Handel, Volberg 2020)

For $f : \{-1,1\}^n \to X$, $\varepsilon \sim \text{Unif}(\{-1,1\}^n)$, and $p \geq 1$,

$$(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p)^{\frac{1}{p}} \leq \frac{\pi}{2} \int \left(\mathbb{E} \left\| \sum_{j=1}^{n} \delta_j(t) D_j f(\varepsilon) \right\|^p \right)^{\frac{1}{p}} \mu(dt)$$

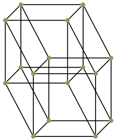
where $\mu(dt) := \frac{2}{\pi} \frac{1}{\sqrt{e^{2t}-1}}\, dt$ and $\delta_j(t)$ are appropriately renormalized *biased* Rademacher random variables.

$$D_j f(\varepsilon) := \frac{f(\varepsilon_1, \ldots, \varepsilon_j, \ldots, \varepsilon_n) - f(\varepsilon_1, \ldots, -\varepsilon_j, \ldots, \varepsilon_n)}{2}.$$

Biased measure on the discrete cube

Consider the *biased* product measure on $\{-1,1\}^n$ with parameter $\alpha \in [0,1]$

$$\mu^\alpha = \bigotimes_{i=1}^{n} \mu_i^\alpha$$

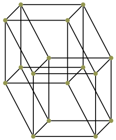


where $\mu_i^\alpha(+1) = \alpha$ and $\mu_i^\alpha(-1) = 1 - \alpha$.

Biased measure on the discrete cube

Consider the *biased* product measure on $\{-1,1\}^n$ with parameter $\alpha \in [0,1]$

$$\mu^\alpha = \bigotimes_{i=1}^{n} \mu_i^\alpha$$



where $\mu_i^\alpha(+1) = \alpha$ and $\mu_i^\alpha(-1) = 1-\alpha$.

Theorem (G. '24)

For any Banach space $(X, \|\cdot\|)$, function $f : \{-1,1\}^n \to X$, and $p \geq 1$, we have

$$\left(\mathbb{E}^\alpha \|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p\right)^{\frac{1}{p}} \leq \int_0^\infty 4\alpha(1-\alpha)\left(\mathbb{E}^\alpha \left\|\sum_{i=1}^n \delta_i(t) D_i f(\varepsilon)\right\|^p\right)^{\frac{1}{p}} dt.$$

Biased measure on the discrete cube

Theorem (G. '24)

For any Banach space $(X, \|\cdot\|)$, function $f : \{-1,1\}^n \to X$, and $p \geq 1$, we have

$$\left(\mathbb{E}^\alpha \|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p\right)^{\frac{1}{p}} \leq \int_0^\infty 4\alpha(1-\alpha) \left(\mathbb{E}^\alpha \left\|\sum_{i=1}^n \delta_i(t) D_i f(\varepsilon)\right\|^p\right)^{\frac{1}{p}} dt.$$

For $\alpha \leq \frac{1}{2}$ and $p \geq 1$,

$$\sup_{x \in \{-1,1\}^n} \mathbb{E}\left[\left|\delta_i(t)\right|^p \mid X_i(0) = x_i\right] \leq e^{-tp}(2\alpha t)^{1-p}.$$

Definition: Rademacher type

We say that Banach space $(X, \|\cdot\|)$ has Rademacher type $p \in [1, 2]$ if there exists a $C \in (0, \infty)$ so that for all $n \geq 1$,

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq C^p \sum_{j=1}^{n} \|x_j\|^p ,$$

where $x_1, \ldots, x_n \in X$ and $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. Rademacher random variables.

Definition: Rademacher type

We say that Banach space $(X, \|\cdot\|)$ has Rademacher type $p \in [1,2]$ if there exists a $C \in (0,\infty)$ so that for all $n \geq 1$,

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq C^p \sum_{j=1}^{n} \|x_j\|^p ,$$

where $x_1, \ldots, x_n \in X$ and $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. Rademacher random variables.

Examples:

- all Banach spaces have type 1 ("trivial type")

Definition: Rademacher type

We say that Banach space $(X, \|\cdot\|)$ has Rademacher type $p \in [1, 2]$ if there exists a $C \in (0, \infty)$ so that for all $n \geq 1$,

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq C^p \sum_{j=1}^{n} \|x_j\|^p ,$$

where $x_1, \ldots, x_n \in X$ and $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. Rademacher random variables.

Examples:

- all Banach spaces have type 1 ("trivial type")
- Hilbert spaces have type 2

Definition: Rademacher type

We say that Banach space $(X, \|\cdot\|)$ has Rademacher type $p \in [1,2]$ if there exists a $C \in (0,\infty)$ so that for all $n \geq 1$,

$$\mathbb{E}\left\|\sum_{j=1}^{n} \varepsilon_j x_j\right\|^p \leq C^p \sum_{j=1}^{n} \|x_j\|^p ,$$

where $x_1, \ldots, x_n \in X$ and $\varepsilon_1, \ldots, \varepsilon_n$ i.i.d. Rademacher random variables.

Examples:

- all Banach spaces have type 1 ("trivial type")
- Hilbert spaces have type 2
- L^p, ℓ^p spaces have type p for $p \in [1,2]$

Vector-valued Poincaré inequality

Corollary (G. '24)

For any Banach space $(X, \|\cdot\|)$ of Rademacher type $p \in [1,2]$, function $f : \{-1,1\}^n \to X$, and $\alpha < \frac{1}{2}$, we have that

$$\left(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p\right)^{\frac{1}{p}} \lesssim T_p(X)\alpha^{\frac{1}{p}} \left(\sum_{i=1}^{n} \mathbb{E}\|D_i f(\varepsilon)\|^p\right)^{\frac{1}{p}}.$$

Vector-valued Poincaré inequality

Corollary (G. '24)

For any Banach space $(X, \|\cdot\|)$ of Rademacher type $p \in [1,2]$, function $f : \{-1,1\}^n \to X$, and $\alpha < \frac{1}{2}$, we have that

$$(\mathbb{E}\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\|^p)^{\frac{1}{p}} \lesssim T_p(X)\alpha^{\frac{1}{p}} \left(\sum_{i=1}^{n} \mathbb{E}\|D_i f(\varepsilon)\|^p\right)^{\frac{1}{p}}.$$

Example: Let $(X, \|\cdot\|) = (\mathbb{R}^n, \|\cdot\|_{\ell^p})$ and $f(\varepsilon) := n^{-\frac{1}{p}} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$.

$$\|f(\varepsilon) - \mathbb{E}f(\varepsilon)\| = \left(\frac{1}{n}\sum_{i=1}^{n} |\varepsilon_i - \mathbb{E}\varepsilon_i|^p\right)^{\frac{1}{p}} \xrightarrow[n\to\infty]{} (\mathbb{E}|\varepsilon_1 - \mathbb{E}\varepsilon_1|^p)^{\frac{1}{p}} \sim \alpha^{\frac{1}{p}}.$$

Lower bounds on average distortion

We say a D-Lipschitz mapping $f : (M, d) \to (X, \|\cdot\|)$ has ν-average distortion D if

$$\mathbb{E}_{\nu\otimes\nu}\|f(\varepsilon) - f(\varepsilon')\| \geq \mathbb{E}_{\nu\otimes\nu}|d(\varepsilon, \varepsilon')|.$$

Lower bounds on average distortion

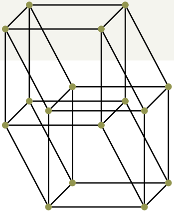
We say a D-Lipschitz mapping $f : (M, d) \to (X, \|\cdot\|)$ has ν-average distortion D if

$$\mathbb{E}_{\nu\otimes\nu}\|f(\varepsilon) - f(\varepsilon')\| \geq \mathbb{E}_{\nu\otimes\nu}|d(\varepsilon, \varepsilon')|.$$

Corollary (G. '24)

For any $f : \{-1, 1\}^n \to (X, \|\cdot\|)$ with X of type p and f having μ-average distortion D,

$$D \gtrsim (\alpha n)^{1-\frac{1}{p}}.$$



Lower bounds on average distortion

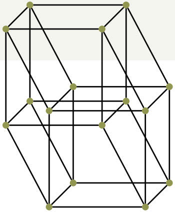
We say a D-Lipschitz mapping $f : (M, d) \to (X, \|\cdot\|)$ has ν-average distortion D if

$$\mathbb{E}_{\nu\otimes\nu}\|f(\varepsilon) - f(\varepsilon')\| \geq \mathbb{E}_{\nu\otimes\nu}|d(\varepsilon, \varepsilon')|.$$

Corollary (G. '24)

For any $f : \{-1, 1\}^n \to (X, \|\cdot\|)$ with X of type p and f having μ-average distortion D,

$$D \gtrsim (\alpha n)^{1-\frac{1}{p}}.$$



Nonembeddability even for sparse ε, e.g. $\alpha = \frac{\log n}{n}$.

Scaling limit: Poisson distribution

Theorem

For any Banach space $(X, \|\cdot\|)$, bounded function $f : \mathbb{N}^m \to X$, $1 \le p < \infty$, and ,

$$(\mathbb{E}\|f(N) - \mathbb{E}f(N)\|^p)^{\frac{1}{p}} \le \int_0^\infty \left(\mathbb{E}\left\|\sum_{i=1}^m \tilde{\eta}_i(t) D_i^{\mathbb{Z}} f(N)\right\|^p\right)^{\frac{1}{p}} dt,$$

where

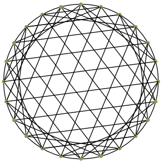
$$D_i^{\mathbb{Z}} := f(x_1, \ldots, x_i + 1, \ldots, x_m) - f(x_1, \ldots, x_i, \ldots, x_m)$$

and

$$\tilde{\eta}_i(t) := e^{-t} - \frac{e^{-t}}{(1 - e^{-t})} \eta_i(t),$$

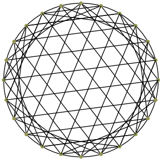
for $\eta_i(t)$ independent $\mathrm{Poisson}(1 - e^{-t})$.

Beyond product spaces: finite groups



Let (G, S) be a finite group with a symmetric set of generators S.

Beyond product spaces: finite groups



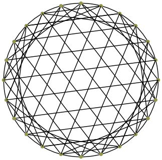
Let (G, S) be a finite group with a symmetric set of generators S. Denote

$$D_s f(x) = f(x) - f(sx).$$

The Laplacian on G is given by

$$\Delta f = -\sum_{s \in S} D_s f$$

Beyond product spaces: finite groups



Let (G, S) be a finite group with a symmetric set of generators S. Denote

$$D_s f(x) = f(x) - f(sx).$$

The Laplacian on G is given by

$$\Delta f = -\sum_{s \in S} D_s f$$

We denote by $P_t := e^{t\Delta}$ the standard heat semigroup on G.

Heat semigroup on (G, S)

Let $\{X_t\}$ be a continuous-time random walk on G
with stationary measure $\mu = \mathrm{Unif}(G)$.

heat kernel of the random walk $p_t(x, y) := \mathbb{P}_x(X_t = y)$

Heat semigroup on (G, S)

Let $\{X_t\}$ be a continuous-time random walk on G
with stationary measure $\mu = \mathrm{Unif}(G)$.

heat kernel of the random walk $p_t(x, y) := \mathbb{P}_x(X_t = y)$

For every $s \in S$,

$$\delta_s(t) = \frac{p_t(x, X_t) - p_t(sx, X_t)}{p_t(x, X_t)} = \frac{D_s p_t(x, X_t)}{p_t(x, X_t)}.$$

Heat semigroup on (G, S)

Let $\{X_t\}$ be a continuous-time random walk on G with stationary measure $\mu = \mathrm{Unif}(G)$.

heat kernel of the random walk $p_t(x, y) := \mathbb{P}_x(X_t = y)$

For every $s \in S$,

$$\delta_s(t) = \frac{p_t(x, X_t) - p_t(sx, X_t)}{p_t(x, X_t)} = \frac{D_s p_t(x, X_t)}{p_t(x, X_t)}.$$

$$D_s P_t f(x) = \mathbb{E}_x[f(X_t)\delta_s(t)]$$

Pisier-like inequality on a finite group

Proposition (G., van Handel 2024+)

For any function $f : G \to X$, and $1 \leq p < \infty$, we have

$$\left(\mathbb{E}_\mu \|f - \mathbb{E}_\mu f\|^p\right)^{\frac{1}{p}} \leq \frac{1}{2} \int_0^\infty \left(\mathbb{E}_\mu \left\| \sum_{s \in S} \delta_s(t) D_s f(X_0) \right\|^p\right)^{\frac{1}{p}} dt.$$

Pisier-like inequality on a finite group

Proposition (G., van Handel 2024+)

For any function $f : G \to X$, and $1 \leq p < \infty$, we have

$$\left(\mathbb{E}_\mu \|f - \mathbb{E}_\mu f\|^p\right)^{\frac{1}{p}} \leq \frac{1}{2} \int_0^\infty \left(\mathbb{E}_\mu \left\| \sum_{s \in S} \delta_s(t) D_s f(X_0) \right\|^p\right)^{\frac{1}{p}} dt.$$

Pisier-like inequality on the symmetric group

Theorem (G., van Handel 2024+)

Let S_n denote the symmetric group with generator set

$$S = \{(ij) : i \neq j,\ i, j \in [n]\}.$$

If $(X, \|\cdot\|)$ is a Banach space of type $p \in [1, 2]$ and $f : (S_n, S) \to (X, \|\cdot\|)$, then for $n \geq 2$,

$$\mathbb{E}_\mu \|f - \mathbb{E}_\mu f\| \lesssim \left(\frac{\log n}{n}\right)^{\frac{1}{p}} \left(\sum_{\substack{i,j=1, \\ i<j}}^{n} \mathbb{E}\|D_{ij} f(X_0)\|^p \right)^{\frac{1}{p}}.$$

Corollary: unbounded distortion in bilipschitz embedding of (S_n, S) into $(X, \|\cdot\|)$

Corollary (G., van Handel 2024+)

For any $f : (S_n, S) \to (X, \|\cdot\|)$, an embedding with μ-average distortion D with $(X, \|\cdot\|)$ of type $p \in [1, 2]$,

$$D \gtrsim n^{1-\frac{1}{p}} \left(\frac{1}{\log n}\right)^{\frac{1}{p}}.$$

Main techniques

How to "isolate" the $\delta_s(t)$s:

- decoupling and symmetrization arguments

Obtaining bounds on moments of the $\delta_s(t)$s:

- Small t: Bakry-Émery curvature and Gamma calculus
- Large t: Markov chain mixing

References

A. Eskenazis. Some geometric applications of the discrete heat flow. *arXiv preprint arXiv:2310.01868*, 2023.

M. Gordin. Vector-valued concentration inequalities on the biased discrete cube. *arXiv preprint arXiv:2410.09607*, 2024.

P. Ivanisvili, R. van Handel, and A. Volberg. Rademacher type and Enflo type coincide. *Annals of Mathematics*, 192(2):665–678, 2020.

G. Pisier. Probabilistic methods in the geometry of Banach spaces. In *Probability and Analysis: Lectures Notes in Mathematics, Como, Italy 1985*, pages 167–241. Springer, 2006.

Y. Rabinovich. On average distortion of embedding metrics into the line. *Discrete & Computational Geometry*, 39(4):720–733, 2008.

M. Talagrand. Isoperimetry, logarithmic Sobolev inequalities on the discrete cube, and Margulis' graph connectivity theorem. *Geometric & Functional Analysis GAFA*, 3(3):295–314, 1993.