

# Anti-concentration, from small-ball estimates to Fourier decay

---

Dan Mikulincer

UW

# Concentration

**Concentration:** Functions have small **effective** range.

- Example 1:  $f : [-1/2, 1/2]^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i$ .

Standard:  $\text{Vol}(\{x \in [-1/2, 1/2]^n \mid |f(x)| \geq t\}) \leq e^{-10t}$ .

- Example 2:  $\sqrt{n}B_2 \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=1}^n x_i^2$ .

Also standard:  $\text{Vol}(\{x \in \sqrt{n}B_2 \mid f(x) \leq \sqrt{n} - t\}) \leq e^{-10t}$ .

**Anti-concentration:** But why? How do things look like when they concentrate?

# Concentration

**Concentration:** Functions have small **effective** range.

- Example 1:  $f : [-1/2, 1/2]^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i$ .

Standard:  $\text{Vol}(\{x \in [-1/2, 1/2]^n \mid |f(x)| \geq t\}) \leq e^{-10t}$ .

- Example 2:  $\sqrt{n}B_2 \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=1}^n x_i^2$ .

Also standard:  $\text{Vol}(\{x \in \sqrt{n}B_2 \mid f(x) \leq \sqrt{n} - t\}) \leq e^{-10t}$ .

**Anti-concentration:** But why? How do things look like when they concentrate?

# Concentration

**Concentration:** Functions have small **effective** range.

- Example 1:  $f : [-1/2, 1/2]^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i$ .

Standard:  $\text{Vol}(\{x \in [-1/2, 1/2]^n \mid |f(x)| \geq t\}) \leq e^{-10t}$ .

- Example 2:  $\sqrt{n}B_2 \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=1}^n x_i^2$ .

Also standard:  $\text{Vol}(\{x \in \sqrt{n}B_2 \mid f(x) \leq \sqrt{n} - t\}) \leq e^{-10t}$ .

**Anti-concentration:** But why? How do things look like when they concentrate?

# Concentration

**Concentration:** Functions have small **effective** range.

- Example 1:  $f : [-1/2, 1/2]^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i$ .

Standard:  $\text{Vol}(\{x \in [-1/2, 1/2]^n \mid |f(x)| \geq t\}) \leq e^{-10t}$ .

- Example 2:  $\sqrt{n}B_2 \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=1}^n x_i^2$ .

Also standard:  $\text{Vol}(\{x \in \sqrt{n}B_2 \mid f(x) \leq \sqrt{n} - t\}) \leq e^{-10t}$ .

**Anti-concentration:** But why? How do things look like when they concentrate?

# Concentration

**Concentration:** Functions have small **effective** range.

- Example 1:  $f : [-1/2, 1/2]^n \rightarrow \mathbb{R}$ ,  $f(x) = \frac{1}{\sqrt{n}} \sum_{i=1}^n x_i$ .

Standard:  $\text{Vol}(\{x \in [-1/2, 1/2]^n \mid |f(x)| \geq t\}) \leq e^{-10t}$ .

- Example 2:  $\sqrt{n}B_2 \rightarrow \mathbb{R}$ ,  $f(x) = \sum_{i=1}^n x_i^2$ .

Also standard:  $\text{Vol}(\{x \in \sqrt{n}B_2 \mid f(x) \leq \sqrt{n} - t\}) \leq e^{-10t}$ .

**Anti-concentration:** But why? How do things look like when they concentrate?

**Variance:** For  $\mu$  a probability measure on  $\mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{Var}_\mu(f) = \int f^2 d\mu - (f d\mu)^2.$$

- $\text{Var}_\mu(f) = 0$  iff  $f = \text{constant}$  almost surely.
- How about  $\text{Var}_\mu(f) = \frac{1}{2}$ ? Small support? Some outliers?
- Maybe not suitable for a good quantitative description.
- ???

**Variance:** For  $\mu$  a probability measure on  $\mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{Var}_\mu(f) = \int f^2 d\mu - (fd\mu)^2.$$

- $\text{Var}_\mu(f) = 0$  iff  $f = \text{constant}$  almost surely.
- How about  $\text{Var}_\mu(f) = \frac{1}{2}$ ? Small support? Some outliers?
- Maybe not suitable for a good quantitative description.
- ???

**Variance:** For  $\mu$  a probability measure on  $\mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{Var}_\mu(f) = \int f^2 d\mu - (fd\mu)^2.$$

- $\text{Var}_\mu(f) = 0$  iff  $f = \text{constant}$  almost surely.
- How about  $\text{Var}_\mu(f) = \frac{1}{2}$ ? Small support? Some outliers?
- Maybe not suitable for a good quantitative description.
- ???

**Variance:** For  $\mu$  a probability measure on  $\mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{Var}_\mu(f) = \int f^2 d\mu - (fd\mu)^2.$$

- $\text{Var}_\mu(f) = 0$  iff  $f = \text{constant}$  almost surely.
- How about  $\text{Var}_\mu(f) = \frac{1}{2}$ ? Small support? Some outliers?
- Maybe **not suitable** for a good quantitative description.
- ???

**Variance:** For  $\mu$  a probability measure on  $\mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\text{Var}_\mu(f) = \int f^2 d\mu - (fd\mu)^2.$$

- $\text{Var}_\mu(f) = 0$  iff  $f = \text{constant}$  almost surely.
- How about  $\text{Var}_\mu(f) = \frac{1}{2}$ ? Small support? Some outliers?
- Maybe **not suitable** for a good quantitative description.
- ???

## Variance bounds - Polynomial Regression

Given  $Y_i = h(X_i)$ ,  $h$  polynomial, find the polynomial  $h$ . In matrix notation

$$\begin{bmatrix} Y_1 \\ \dots \\ Y_m \end{bmatrix} = \begin{bmatrix} a_d X_1^d, \dots, a_1 X_1, a_0 \\ \dots \\ a_d X_m^d, \dots, a_1 X_m, a_0 \end{bmatrix} := M \begin{bmatrix} a_d \\ \dots \\ a_0 \end{bmatrix}.$$

Want to invert  $M$ . Need  $MM^T \geq cI_d$ , on  $\text{span}(Y)$ .

When  $X_i \sim \mu$ ,  $c \simeq \text{Var}_\mu(h)$ . (Emschwiller, Gamarnik, Kızıldağ, Zadik)

## Variance bounds - Polynomial Regression

Given  $Y_i = h(X_i)$ ,  $h$  polynomial, find the polynomial  $h$ . In matrix notation

$$\begin{bmatrix} Y_1 \\ \dots \\ Y_m \end{bmatrix} = \begin{bmatrix} a_d X_1^d, \dots, a_1 X_1, a_0 \\ \dots \\ a_d X_m^d, \dots, a_1 X_m, a_0 \end{bmatrix} := M \begin{bmatrix} a_d \\ \dots \\ a_0 \end{bmatrix}.$$

Want to invert  $M$ . Need  $MM^T \geq cI_d$ , on  $\text{span}(Y)$ .

When  $X_i \sim \mu$ ,  $c \simeq \text{Var}_\mu(h)$ . (Emschwiller, Gamarnik, Kızıldağ, Zadik)

**Small-balls:** For  $\mu$  a measure on  $\mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and  $\varepsilon > 0$

$$f_*\mu([-\varepsilon, \varepsilon]) = \mu(\{x \in \mathbb{R}^n \mid |f(x)| \leq \varepsilon\}).$$

- $f_*\mu([-\varepsilon, \varepsilon]) \xrightarrow{\varepsilon \rightarrow 0} 0$  iff  $f_*\mu$  non atomic.
- Multi-scaled estimate of anti-concentration.
- Strictly stronger than variance  $\text{Var}_\mu(f) \geq \varepsilon^2(1 - f_*\mu([-\varepsilon, \varepsilon]))$ .
- Much harder to compute/estimate.

# Small-Ball Estimates

**Small-balls:** For  $\mu$  a measure on  $\mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and  $\varepsilon > 0$

$$f_*\mu([- \varepsilon, \varepsilon]) = \mu(\{x \in \mathbb{R}^n \mid |f(x)| \leq \varepsilon\}).$$

- $f_*\mu([- \varepsilon, \varepsilon]) \xrightarrow{\varepsilon \rightarrow 0} 0$  iff  $f_*\mu$  non atomic.
- Multi-scaled estimate of anti-concentration.
- Strictly stronger than variance  $\text{Var}_\mu(f) \geq \varepsilon^2(1 - f_*\mu([- \varepsilon, \varepsilon]))$ .
- Much harder to compute/estimate.

# Small-Ball Estimates

**Small-balls:** For  $\mu$  a measure on  $\mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and  $\varepsilon > 0$

$$f_*\mu([- \varepsilon, \varepsilon]) = \mu(\{x \in \mathbb{R}^n \mid |f(x)| \leq \varepsilon\}).$$

- $f_*\mu([- \varepsilon, \varepsilon]) \xrightarrow{\varepsilon \rightarrow 0} 0$  iff  $f_*\mu$  non atomic.
- Multi-scaled estimate of anti-concentration.
- Strictly stronger than variance  $\text{Var}_\mu(f) \geq \varepsilon^2(1 - f_*\mu([- \varepsilon, \varepsilon]))$ .
- Much harder to compute/estimate.

# Small-Ball Estimates

**Small-balls:** For  $\mu$  a measure on  $\mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and  $\varepsilon > 0$

$$f_*\mu([- \varepsilon, \varepsilon]) = \mu(\{x \in \mathbb{R}^n \mid |f(x)| \leq \varepsilon\}).$$

- $f_*\mu([- \varepsilon, \varepsilon]) \xrightarrow{\varepsilon \rightarrow 0} 0$  iff  $f_*\mu$  non atomic.
- Multi-scaled estimate of anti-concentration.
- Strictly stronger than variance  $\text{Var}_\mu(f) \geq \varepsilon^2(1 - f_*\mu([- \varepsilon, \varepsilon]))$ .
- Much harder to compute/estimate.

# Small-Ball Estimates

**Small-balls:** For  $\mu$  a measure on  $\mathbb{R}^n$ ,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , and  $\varepsilon > 0$

$$f_*\mu([- \varepsilon, \varepsilon]) = \mu(\{x \in \mathbb{R}^n \mid |f(x)| \leq \varepsilon\}).$$

- $f_*\mu([- \varepsilon, \varepsilon]) \xrightarrow{\varepsilon \rightarrow 0} 0$  iff  $f_*\mu$  non atomic.
- Multi-scaled estimate of anti-concentration.
- Strictly stronger than variance  $\text{Var}_\mu(f) \geq \varepsilon^2(1 - f_*\mu([- \varepsilon, \varepsilon]))$ .
- Much harder to compute/estimate.

# Anti-concentration - Small Ball Estimates

In general, can't expect to have meaningful small-ball estimates. Both measure and function have to be 'nice'.

Today, nice  $\mu = e^{-\varphi} dx$  is **log-concave** ( $\varphi$  is convex).

**Examples:** Gaussian, uniform on convex sets...

## Theorem (Carbery-Wright 2001)

*If  $\mu$  is log-concave and  $h$  degree- $d$  polynomial*

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}.$$

## Anti-concentration - Small Ball Estimates

In general, can't expect to have meaningful small-ball estimates. Both measure and function have to be 'nice'.

Today, nice  $\mu = e^{-\varphi} dx$  is **log-concave** ( $\varphi$  is convex).

**Examples:** Gaussian, uniform on convex sets...

Theorem (Carbery-Wright 2001)

*If  $\mu$  is log-concave and  $h$  degree- $d$  polynomial*

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}.$$

## Anti-concentration - Small Ball Estimates

In general, can't expect to have meaningful small-ball estimates. Both measure and function have to be 'nice'.

Today, nice  $\mu = e^{-\varphi} dx$  is **log-concave** ( $\varphi$  is convex).

**Examples:** Gaussian, uniform on convex sets...

### Theorem (Carbery-Wright 2001)

*If  $\mu$  is log-concave and  $h$  degree- $d$  polynomial*

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}.$$

## Small Ball Estimates - Beyond Polynomials

**Carbery-Wright:**  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$ .  $h$  polynomial.

What if  $h$  is not a polynomial?

Theorem (Cornacchia, M., Mossel)

*If  $f$  is Lipschitz, analytic, normalized*

$$f_*\mu([- \varepsilon, \varepsilon]) \leq \exp(-\log(1/\varepsilon)^{2/3}).$$

1. Extend Carbery-Wright:  $d \rightarrow$  effective degree.
2. In practice need to have a lower bound on  $\text{Var}_\mu(f)$ .

## Small Ball Estimates - Beyond Polynomials

**Carbery-Wright:**  $h_*\mu([-ε, ε]) \leq C \left( \frac{ε}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$ .  $h$  polynomial.

What if  $h$  is not a polynomial?

Theorem (Cornacchia, M., Mossel)

*If  $f$  is Lipschitz, analytic, normalized*

$$f_*\mu([-ε, ε]) \leq \exp(-\log(1/ε)^{2/3}).$$

1. Extend Carbery-Wright:  $d \rightarrow$  effective degree.
2. In practice need to have a lower bound on  $\text{Var}_\mu(f)$ .

## Small Ball Estimates - Beyond Polynomials

**Carbery-Wright:**  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$ .  $h$  polynomial.

What if  $h$  is not a polynomial?

### Theorem (Cornacchia, M., Mossel)

*If  $f$  is Lipschitz, analytic, normalized*

$$f_*\mu([- \varepsilon, \varepsilon]) \leq \exp(-\log(1/\varepsilon)^{2/3}).$$

1. Extend Carbery-Wright:  $d \rightarrow$  effective degree.
2. In practice need to have a lower bound on  $\text{Var}_\mu(f)$ .

## Anti-Concentration - Variance Bounds

**Carbery-Wright:**  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$ .  $h$  polynomial.

Two forms of anti-concentration:

- **Small-ball:**  $\forall \varepsilon > 0, \quad h_*\mu([- \varepsilon, \varepsilon]) \leq p(\varepsilon)$ .
- **Variance bound:**  $\text{Var}_\mu(h) \geq c$ .

In general **small-ball**  $\implies$  **variance bound**.

Carbery-Wright says: **variance bound**  $\implies$  **small-ball**. (for polynomials and log-concave measures)

## Anti-Concentration - Variance Bounds

**Carbery-Wright:**  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$ .  $h$  polynomial.

Two forms of anti-concentration:

- **Small-ball:**  $\forall \varepsilon > 0, \quad h_*\mu([- \varepsilon, \varepsilon]) \leq p(\varepsilon)$ .
- **Variance bound:**  $\text{Var}_\mu(h) \geq c$ .

In general **small-ball**  $\implies$  **variance bound**.

Carbery-Wright says: **variance bound**  $\implies$  **small-ball**. (for polynomials and log-concave measures)

# Variance Bounds

However,  $C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$  is unsuitable for a variance bound.

## Question

Can we prove  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{n(h)} \right)^{\frac{1}{d}}$ ,  $n(h)$  independent of  $\mu$ .

**Answer:** Generally no, if  $\mu$  is uniform on ball,  $\text{Var}_\mu(\|\cdot\|^2) = \frac{1}{n}$ .  
But sometimes yes,

## Theorem (Glazer, M.)

If  $\mu$  is log-concave **product** measure and  $h$  degree- $d$  polynomial<sup>1</sup>

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{coeff}_d(h)} \right)^{\frac{1}{d}}$$

<sup>1</sup>coeff<sub>d</sub> sum of squares of coefficients

# Variance Bounds

However,  $C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$  is unsuitable for a variance bound.

## Question

Can we prove  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{n(h)} \right)^{\frac{1}{d}}$ ,  $n(h)$  independent of  $\mu$ .

**Answer:** Generally no, if  $\mu$  is uniform on ball,  $\text{Var}_\mu(\|\cdot\|^2) = \frac{1}{n}$ .

But sometimes yes,

## Theorem (Glazer, M.)

If  $\mu$  is log-concave **product** measure and  $h$  degree- $d$  polynomial<sup>1</sup>

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{coeff}_d(h)} \right)^{\frac{1}{d}}$$

---

<sup>1</sup>coeff <sub>$d$</sub>  sum of squares of coefficients

# Variance Bounds

However,  $C \left( \frac{\varepsilon}{\sqrt{\text{Var}_\mu(h)}} \right)^{\frac{1}{d}}$  is unsuitable for a variance bound.

## Question

Can we prove  $h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{n(h)} \right)^{\frac{1}{d}}$ ,  $n(h)$  independent of  $\mu$ .

**Answer:** Generally no, if  $\mu$  is uniform on ball,  $\text{Var}_\mu(\|\cdot\|^2) = \frac{1}{n}$ .  
But sometimes yes,

## Theorem (Glazer, M.)

If  $\mu$  is log-concave **product** measure and  $h$  degree- $d$  polynomial<sup>1</sup>

$$h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{coeff}_d(h)} \right)^{\frac{1}{d}}$$

<sup>1</sup>coeff <sub>$d$</sub>  sum of squares of coefficients

# Oscillatory Integrals

For  $\mu, h$ , consider  $H_\mu(t) := \int e^{ith(x)} d\mu(x)$ .

- $H_\mu(t) \equiv 1 \iff h_*\mu = \delta_0$ .
- By Riemann-Lebesgue, if  $h_*\mu$  abs. continuous  $H_\mu(t) \xrightarrow{|t| \rightarrow \infty} 0$ .

Classical result by Esseen:

$$h_*\mu([- \varepsilon, \varepsilon]) \leq \varepsilon \int_{-2\pi/\varepsilon}^{2\pi/\varepsilon} |H_\mu(t)| dt.$$

Fourier decay  $\implies$  small-ball estimates.

# Oscillatory Integrals

For  $\mu, h$ , consider  $H_\mu(t) := \int e^{ith(x)} d\mu(x)$ .

- $H_\mu(t) \equiv 1 \iff h_*\mu = \delta_0$ .
- By Riemann-Lebesgue, if  $h_*\mu$  abs. continuous  $H_\mu(t) \xrightarrow{|t| \rightarrow \infty} 0$ .

Classical result by Esseen:

$$h_*\mu([- \varepsilon, \varepsilon]) \leq \varepsilon \int_{-2\pi/\varepsilon}^{2\pi/\varepsilon} |H_\mu(t)| dt.$$

Fourier decay  $\implies$  small-ball estimates.

# Oscillatory Integrals

For  $\mu, h$ , consider  $H_\mu(t) := \int e^{ith(x)} d\mu(x)$ .

- $H_\mu(t) \equiv 1 \iff h_*\mu = \delta_0$ .
- By Riemann-Lebesgue, if  $h_*\mu$  abs. continuous  $H_\mu(t) \xrightarrow{|t| \rightarrow \infty} 0$ .

Classical result by Esseen:

$$h_*\mu([- \varepsilon, \varepsilon]) \leq \varepsilon \int_{-2\pi/\varepsilon}^{2\pi/\varepsilon} |H_\mu(t)| dt.$$

Fourier decay  $\implies$  small-ball estimates.

Another classical result by van der Corput:

## Lemma

If  $\mu = 1_{0 < x < 1} dx$ , and  $h$  degree- $d$  monic polynomial

$$|H_\mu(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

Set  $A_\delta = \{x \in [0, 1] \mid |h'(x)| \leq \delta\}$ .

$$|H_\mu(t)| \leq \left| \int_{A_\delta} e^{-ith(x)} dx \right| + \left| \int_{A_\delta^c} e^{-ith(x)} dx \right|.$$

First integral bounded by  $h'_* \mu([-\delta, \delta])$ . Integrate by parts second.

Another classical result by van der Corput:

## Lemma

If  $\mu = 1_{0 < x < 1} dx$ , and  $h$  degree- $d$  monic polynomial

$$|H_\mu(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

Set  $A_\delta = \{x \in [0, 1] \mid |h'(x)| \leq \delta\}$ .

$$|H_\mu(t)| \leq \left| \int_{A_\delta} e^{-ith(x)} dx \right| + \left| \int_{A_\delta^c} e^{-ith(x)} dx \right|.$$

First integral bounded by  $h'_* \mu([-\delta, \delta])$ . Integrate by parts second.

**Van der Corput:** On  $\mathbb{R}$ , **small-ball estimates**  $\implies$  **Fourier decay**.

In fact can say more, building on Kosov 21'.

## Theorem

Let  $\mu$  be log-concave on  $\mathbb{R}^n$  and let  $h$  be degree- $d$  homogeneous polynomial. The following are equivalent:

1.  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .
2.  $\forall \varepsilon > 0, h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{coeff}_d(h)} \right)^{\frac{1}{d}}$ .
3.  $\forall t \in \mathbb{R}, |H_\mu(t)| \leq \frac{C}{|\text{coeff}_d(h)t|^{\frac{1}{d}}}$ .

**Van der Corput:** On  $\mathbb{R}$ , **small-ball estimates**  $\implies$  **Fourier decay**.  
In fact can say more, building on Kosov 21'.

## Theorem

Let  $\mu$  be log-concave on  $\mathbb{R}^n$  and let  $h$  be degree- $d$  homogeneous polynomial. The following are equivalent:

1.  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .
2.  $\forall \varepsilon > 0, h_*\mu([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{coeff}_d(h)} \right)^{\frac{1}{d}}$ .
3.  $\forall t \in \mathbb{R}, |H_\mu(t)| \leq \frac{C}{|\text{coeff}_d(h)t|^{\frac{1}{d}}}$ .

## Sample Application

Recall: if  $\mu$  is product, then  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .

### Corrolary

For  $\gamma$  the standard Gaussian in  $\mathbb{R}^n$  and  $h$  degree- $d$  normalized

$$|H_\gamma(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

In particular:

### Corrolary

After  $d$  convolutions  $h_*\gamma$  has a bounded density. Bound does not depend on the dimension.

## Sample Application

Recall: if  $\mu$  is product, then  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .

### Corrolary

For  $\gamma$  the standard Gaussian in  $\mathbb{R}^n$  and  $h$  degree- $d$  normalized

$$|H_\gamma(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

In particular:

### Corrolary

After  $d$  convolutions  $h_*\gamma$  has a bounded density. Bound does not depend on the dimension.

### Question

What happens when  $\mu$  is log-concave but not product?

When  $\mu$  is uniform on a ball,  $h(x) = \frac{1}{\sqrt{n}}\|x\|^2$  is a 'bad polynomial'.

**Can somewhat generalize:** Let  $B_{p,n} = \{x \in \mathbb{R}^n \mid \|x\|_p \leq R\}$ , and  $\mu_{p,n}$  uniform on  $B_{p,n}$ . Choose  $R$  so that  $\text{Var}_{\mu_{p,n}}(x_1) = 1$ .

Then,

$$\text{Var}_{\mu_{p,n}} \left( \frac{1}{\sqrt{n}} \|x\|_p^p \right) = o(1).$$

In particular, when  $p$  is even,  $h(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$  is bad.

### Question

What happens when  $\mu$  is log-concave but not product?

When  $\mu$  is uniform on a ball,  $h(x) = \frac{1}{\sqrt{n}}\|x\|^2$  is a 'bad polynomial'.

**Can somewhat generalize:** Let  $B_{p,n} = \{x \in \mathbb{R}^n \mid \|x\|_p \leq R\}$ , and  $\mu_{p,n}$  uniform on  $B_{p,n}$ . Choose  $R$  so that  $\text{Var}_{\mu_{p,n}}(x_1) = 1$ .

Then,

$$\text{Var}_{\mu_{p,n}} \left( \frac{1}{\sqrt{n}} \|x\|_p^p \right) = o(1).$$

In particular, when  $p$  is even,  $h(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$  is bad.

### Question

What happens when  $\mu$  is log-concave but not product?

When  $\mu$  is uniform on a ball,  $h(x) = \frac{1}{\sqrt{n}}\|x\|^2$  is a 'bad polynomial'.

**Can somewhat generalize:** Let  $B_{p,n} = \{x \in \mathbb{R}^n \mid \|x\|_p \leq R\}$ , and  $\mu_{p,n}$  uniform on  $B_{p,n}$ . Choose  $R$  so that  $\text{Var}_{\mu_{p,n}}(x_1) = 1$ .

Then,

$$\text{Var}_{\mu_{p,n}} \left( \frac{1}{\sqrt{n}} \|x\|_p^p \right) = o(1).$$

In particular, when  $p$  is even,  $h(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$  is bad.

### Question

What happens when  $\mu$  is log-concave but not product?

When  $\mu$  is uniform on a ball,  $h(x) = \frac{1}{\sqrt{n}}\|x\|^2$  is a 'bad polynomial'.

**Can somewhat generalize:** Let  $B_{p,n} = \{x \in \mathbb{R}^n \mid \|x\|_p \leq R\}$ , and  $\mu_{p,n}$  uniform on  $B_{p,n}$ . Choose  $R$  so that  $\text{Var}_{\mu_{p,n}}(x_1) = 1$ .

Then,

$$\text{Var}_{\mu_{p,n}} \left( \frac{1}{\sqrt{n}} \|x\|_p^p \right) = o(1).$$

In particular, when  $p$  is even,  $h(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$  is bad.

## Anti-Concentration in $L_p$ balls

Turns out, that in the class of  $L_p$  balls these are the only bad examples.

### Theorem

*Except at a few exceptional cases*

$$\text{Var}_{\mu_{p,n}}(h) \geq \frac{1}{C}.$$

*Consequently,*

$$|H_{\mu_{p,n}}(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

*Exceptional cases:  $p$  is an even integer,  $d = p$ , and*

$$h(x) = \frac{1}{\sqrt[n]{n}} \|x\|_p^p.$$

## Anti-Concentration in $L_p$ balls

Turns out, that in the class of  $L_p$  balls these are the only bad examples.

### Theorem

*Except at a few exceptional cases*

$$\text{Var}_{\mu_{p,n}}(h) \geq \frac{1}{C}.$$

*Consequently,*

$$|H_{\mu_{p,n}}(t)| \leq \frac{C}{|t|^{\frac{1}{d}}}.$$

*Exceptional cases:  $p$  is an even integer,  $d = p$ , and*

$$h(x) = \frac{1}{\sqrt{n}} \|x\|_p^p.$$

# Anti-Concentration with Symmetries

$\mu_{p,n}$  are examples of measures which are invariant to

- Permutation of the coordinates.
- Reflection of any coordinate.

Since  $\text{Var}$  is an “ $L_2$ -quantity”, can study with representation theory. Some partial results:

## Theorem

*If  $\mu$  is log-concave and permutation+reflection invariant,  $h$  degree- $d$  polynomial.*

- *when  $d = 2$ ,  $\text{Var}_\mu(h) \geq \frac{1}{C}$ , unless  $\mu$  is (close to)  $\mu_{2,n}$ .*
- *when  $d = 3$ ,  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .*

# Anti-Concentration with Symmetries

$\mu_{p,n}$  are examples of measures which are invariant to

- Permutation of the coordinates.
- Reflection of any coordinate.

Since  $\text{Var}$  is an “ $L_2$ -quantity”, can study with representation theory. Some partial results:

## Theorem

*If  $\mu$  is log-concave and permutation+reflection invariant,  $h$  degree- $d$  polynomial.*

- *when  $d = 2$ ,  $\text{Var}_\mu(h) \geq \frac{1}{C}$ , unless  $\mu$  is (close to)  $\mu_{2,n}$ .*
- *when  $d = 3$ ,  $\text{Var}_\mu(h) \geq \frac{1}{C}$ .*

## Further Questions

Many questions remain. First can we classify the bad measures and polynomials?

### Question

Among symmetric measures, are the only bad cases of the form  $\mu = 1_{|h(x)| \leq R} dx$ , for some **convex** polynomial?

What about no symmetries?

Second, we focused on polynomials and specific decay functions.

### Question

Can we replace the exponent  $\frac{1}{d}$  by something more refined?

For example for the Gaussian and  $h$  multilinear,

$$h_* \gamma([- \delta, \delta]) \lesssim \varepsilon.$$

## Further Questions

Many questions remain. First can we classify the bad measures and polynomials?

### Question

Among symmetric measures, are the only bad cases of the form  $\mu = 1_{|h(x)| \leq R} dx$ , for some **convex** polynomial?

What about no symmetries?

Second, we focused on polynomials and specific decay functions.

### Question

Can we replace the exponent  $\frac{1}{d}$  by something more refined?

For example for the Gaussian and  $h$  multilinear,

$$h_* \gamma([- \delta, \delta]) \lesssim \varepsilon.$$

## Further Questions

Another question deals with non-Euclidean spaces.

In general Carbery-Wright can be extended to polynomials on the sphere. That is, if  $\mu$  is uniform on the sphere and  $h : \mathbb{S}^{n-1} \rightarrow \mathbb{R}$  is a degree  $d$  polynomial.

$$\mu_* h([- \varepsilon, \varepsilon]) \leq C \left( \frac{\varepsilon}{\text{Var}_\mu(h)} \right)^{\frac{1}{2d}}.$$

### Question

Can we get the same sort of equivalence for polynomials on the sphere?

How about other spaces? Product of spheres? The orthogonal group?

*Thank You*

# Motivating Problem

**Two optimization problems:** Given  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $\theta_* \in \mathbb{S}^{d-1}$ .

---

## Classical:

- Find  $\min_x f(x)$
- Choose  $x_1, \dots, x_T$
- Query  $f(x_t)$
- $x_{t+1} = x_t - \eta \nabla f(x_t)$

## Stochastic:

- Find  $\theta_*$
- Get random  $X_1, \dots, X_T$
- Choose  $\theta_1, \dots, \theta_T$
- Query  $f(\theta_t^T X_t), f(\theta_*^T X_t)$
- $\theta_{t+1} = \theta_t - \eta \nabla_{\mathbb{S}} (f(\theta_t^T, X_t) - f(\theta_*^T, X_t))^2$

# Motivating Problem

**Two optimization problems:** Given  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $\theta_* \in \mathbb{S}^{d-1}$ .

---

## Classical:

- Find  $\min_x f(x)$
- Choose  $x_1, \dots, x_T$
- Query  $f(x_t)$
- $x_{t+1} = x_t - \eta \nabla f(x_t)$

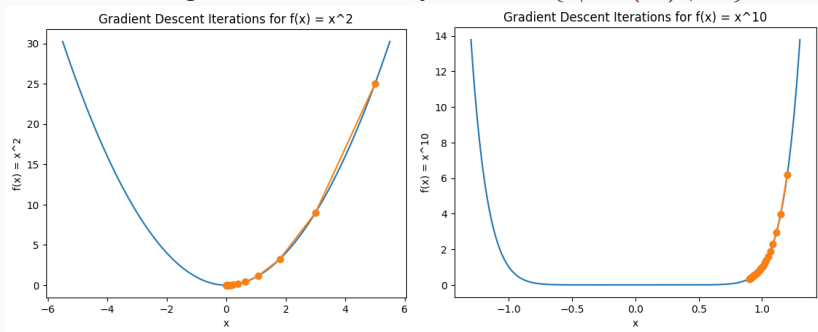
## Stochastic:

- Find  $\theta_*$
- Get random  $X_1, \dots, X_T$
- Choose  $\theta_1, \dots, \theta_T$
- Query  $f(\theta_t^T X_t), f(\theta_*^T X_t)$
- $\theta_{t+1} = \theta_t - \eta \nabla_{\mathbb{S}}(f(\theta_t^T, X_t) - f(\theta_*^T, X_t))^2$

# How Fast? - Classical

$x_{t+1} = x_t - \eta \nabla f(x_t)$ : nice  $f$  and  $x_0, x_t \rightarrow x_* = \arg \min f$ .

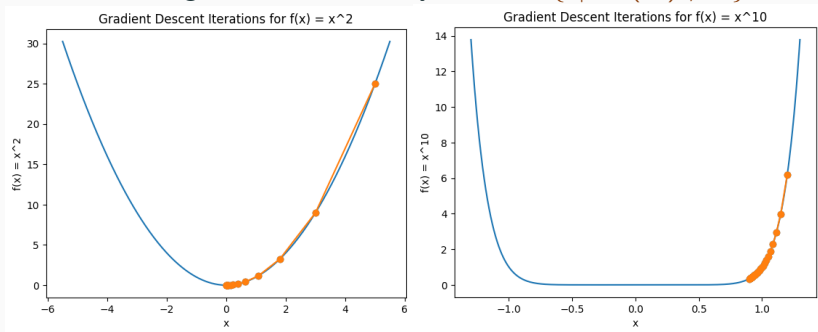
Rate of convergence determined by  $k := \min\{\ell \mid f^{(\ell)}(x_*) \neq 0\}$ .



# How Fast? - Classical

$x_{t+1} = x_t - \eta \nabla f(x_t)$ : nice  $f$  and  $x_0, x_t \rightarrow x_* = \arg \min f$ .

Rate of convergence determined by  $k := \min\{\ell | f^{(\ell)}(x_*) \neq 0\}$ .

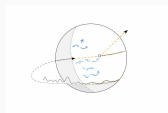


## How Fast? - Stochastic

$\theta_{t+1} = \theta_t - \eta \nabla_S (f(\theta_t^T, X_t) - f(\theta_*^T, X_t))^2$ : nice  $f$  and  $\theta_0, \theta_t \rightarrow \theta_*$ .

- If  $X_t \sim \gamma$  are Gaussian<sup>2</sup>, set  $H_\ell$  Hermite polynomials.
- Since  $H_\ell$  orthonormal for  $L_2(\gamma)$ , write  $f = \sum a_\ell H_\ell$ .
- Rate of convergence determined by  $k := \min\{\ell | a_\ell \neq 0\}$ .  
(Ben Arous, Gheissari, Jagannath 21')

Remark  $a_\ell = \frac{1}{\sqrt{2\pi}} \int f H_\ell e^{-x^2/2} = \frac{1}{\sqrt{2\pi}} \int f^{(\ell)} e^{-x^2/2}$



---

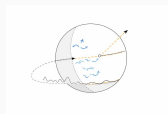
<sup>2</sup>  $\frac{1}{(2\pi)^{d/2}} e^{-\|x\|^2/2}$

## How Fast? - Stochastic

$\theta_{t+1} = \theta_t - \eta \nabla_S (f(\theta_t^T, X_t) - f(\theta_*^T, X_t))^2$ : nice  $f$  and  $\theta_0, \theta_t \rightarrow \theta_*$ .

- If  $X_t \sim \gamma$  are Gaussian<sup>2</sup>, set  $H_\ell$  Hermite polynomials.
- Since  $H_\ell$  orthonormal for  $L_2(\gamma)$ , write  $f = \sum a_\ell H_\ell$ .
- Rate of convergence determined by  $k := \min\{\ell | a_\ell \neq 0\}$ .  
(Ben Arous, Gheissari, Jagannath 21')

Remark  $a_\ell = \frac{1}{\sqrt{2\pi}} \int f H_\ell e^{-x^2/2} = \frac{1}{\sqrt{2\pi}} \int f^{(\ell)} e^{-x^2/2}$



---

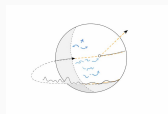
<sup>2</sup>  $\frac{1}{(2\pi)^{d/2}} e^{-\|x\|^2/2}$

## How Fast? - Stochastic

$\theta_{t+1} = \theta_t - \eta \nabla_S (f(\theta_t^T, X_t) - f(\theta_*^T, X_t))^2$ : nice  $f$  and  $\theta_0, \theta_t \rightarrow \theta_*$ .

- If  $X_t \sim \gamma$  are Gaussian<sup>2</sup>, set  $H_\ell$  Hermite polynomials.
- Since  $H_\ell$  orthonormal for  $L_2(\gamma)$ , write  $f = \sum a_\ell H_\ell$ .
- Rate of convergence determined by  $k := \min\{\ell | a_\ell \neq 0\}$ .  
(Ben Arous, Gheissari, Jagannath 21')

Remark  $a_\ell = \frac{1}{\sqrt{2\pi}} \int f H_\ell e^{-x^2/2} = \frac{1}{\sqrt{2\pi}} \int f^{(\ell)} e^{-x^2/2}$



---

<sup>2</sup>  $\frac{1}{(2\pi)^{d/2}} e^{-\|x\|^2/2}$

## Major difference - Are the Problems Stable?

For some shift  $\mu$  transform  $f(\cdot) \rightarrow f_\mu := f(\cdot + \mu)$ .

---

### Classical:

- $k := \min\{\ell | f^{(\ell)}(x_*) \neq 0\}$
- Everything is invariant:  
 $\arg \min f_\mu = x_* - \mu$
- $k = \min\{\ell | f_\mu^{(\ell)}(x_* - \mu) \neq 0\}$

### Stochastic:

- $k := \min\{\ell | \int f^{(\ell)} d\gamma \neq 0\}$
- No invariance...
- $\min\{\ell | \int f_\mu^{(\ell)} d\gamma \neq 0\} = ??$

---

**Fact:**  $F_1(\mu) := \int f'_\mu d\gamma$  is analytic  $\implies F_1(\mu) \neq 0$ .

## Major difference - Are the Problems Stable?

For some shift  $\mu$  transform  $f(\cdot) \rightarrow f_\mu := f(\cdot + \mu)$ .

---

### Classical:

- $k := \min\{\ell | f^{(\ell)}(x_*) \neq 0\}$
- Everything is invariant:  
 $\arg \min f_\mu = x_* - \mu$
- $k = \min\{\ell | f_\mu^{(\ell)}(x_* - \mu) \neq 0\}$

### Stochastic:

- $k := \min\{\ell | \int f^{(\ell)} d\gamma \neq 0\}$
- No invariance...
- $\min\{\ell | \int f_\mu^{(\ell)} d\gamma \neq 0\} = ??$

---

Fact:  $F_1(\mu) := \int f'_\mu d\gamma$  is analytic  $\implies F_1(\mu) \neq 0$ .

## Major difference - Are the Problems Stable?

For some shift  $\mu$  transform  $f(\cdot) \rightarrow f_\mu := f(\cdot + \mu)$ .

---

### Classical:

- $k := \min\{\ell | f^{(\ell)}(x_*) \neq 0\}$
- Everything is invariant:  
 $\arg \min f_\mu = x_* - \mu$
- $k = \min\{\ell | f_\mu^{(\ell)}(x_* - \mu) \neq 0\}$

### Stochastic:

- $k := \min\{\ell | \int f^{(\ell)} d\gamma \neq 0\}$
- No invariance...
- $\min\{\ell | \int f_\mu^{(\ell)} d\gamma \neq 0\} = ??$

---

Fact:  $F_1(\mu) := \int f'_\mu d\gamma$  is analytic  $\implies F_1(\mu) \neq 0$ .

## Major difference - Are the Problems Stable?

For some shift  $\mu$  transform  $f(\cdot) \rightarrow f_\mu := f(\cdot + \mu)$ .

---

### Classical:

- $k := \min\{\ell | f^{(\ell)}(x_*) \neq 0\}$
- Everything is invariant:  
 $\arg \min f_\mu = x_* - \mu$
- $k = \min\{\ell | f_\mu^{(\ell)}(x_* - \mu) \neq 0\}$

### Stochastic:

- $k := \min\{\ell | \int f^{(\ell)} d\gamma \neq 0\}$
- No invariance...
- $\min\{\ell | \int f_\mu^{(\ell)} d\gamma \neq 0\} = ??$

---

**Fact:**  $F_1(\mu) := \int f'_\mu d\gamma$  is analytic  $\implies F_1(\mu) \neq 0$ .