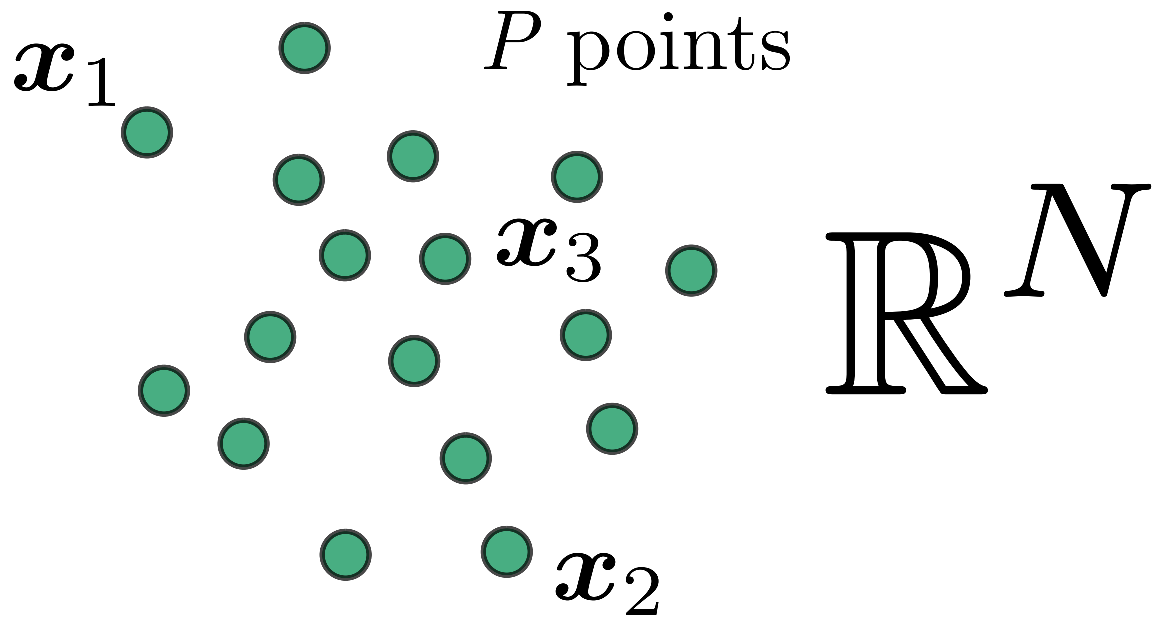
When Buffon's needle problem helps in quantizing the Johnson-Lindenstrauss Lemma

Laurent Jacques - ISPGroup
University of Louvain (UCL), Belgium

ICCHA5, Vanderbilt University
May 19, 2014

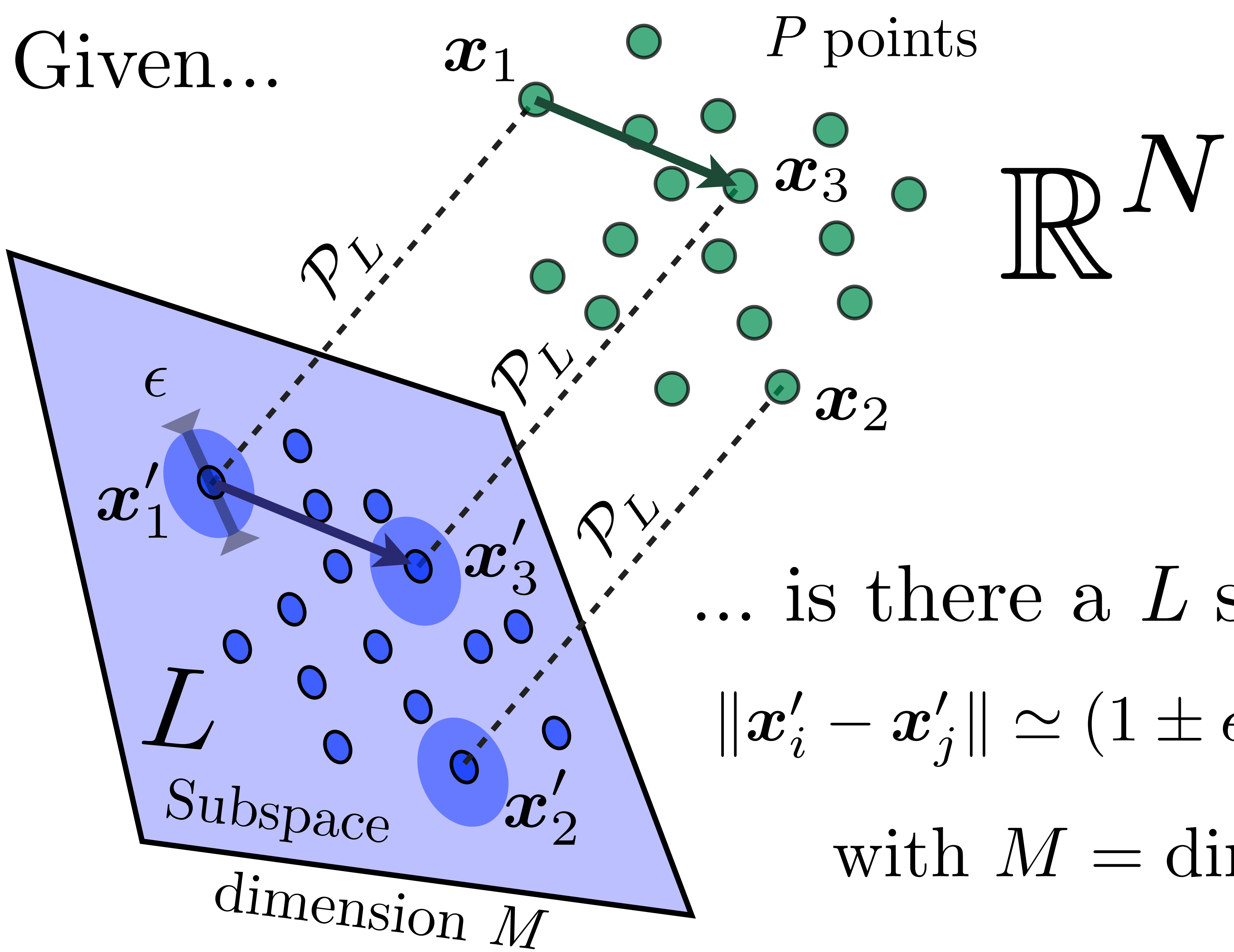
1. Linear dimensionality reduction

Linear Dimensionality Reduction



Linear Dimensionality Reduction

Given...



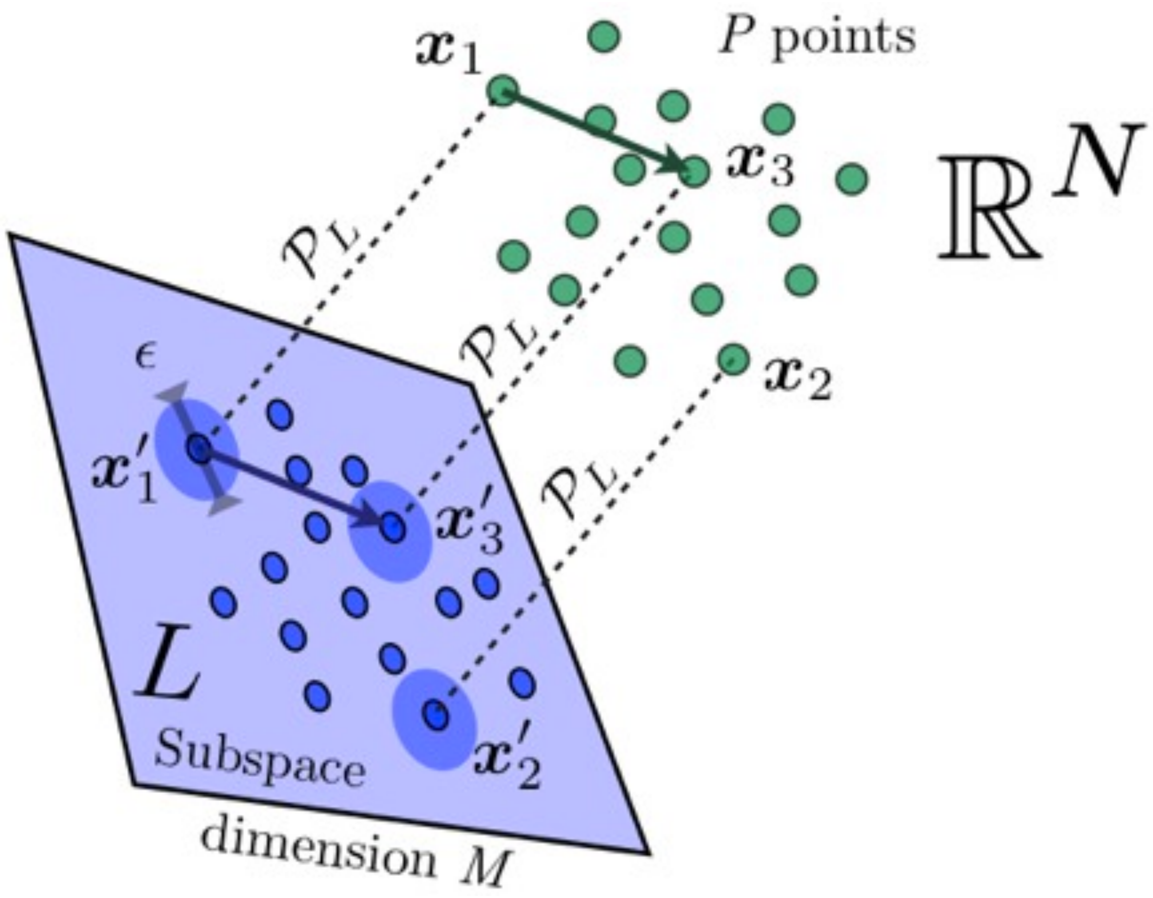
... is there a L such that

$$\|\boldsymbol{x}'_i - \boldsymbol{x}'_j\| \simeq (1 \pm \epsilon)\|\boldsymbol{x}_i - \boldsymbol{x}_j\|$$

with $M = \dim L \ll N$?

Linear Dimensionality Reduction

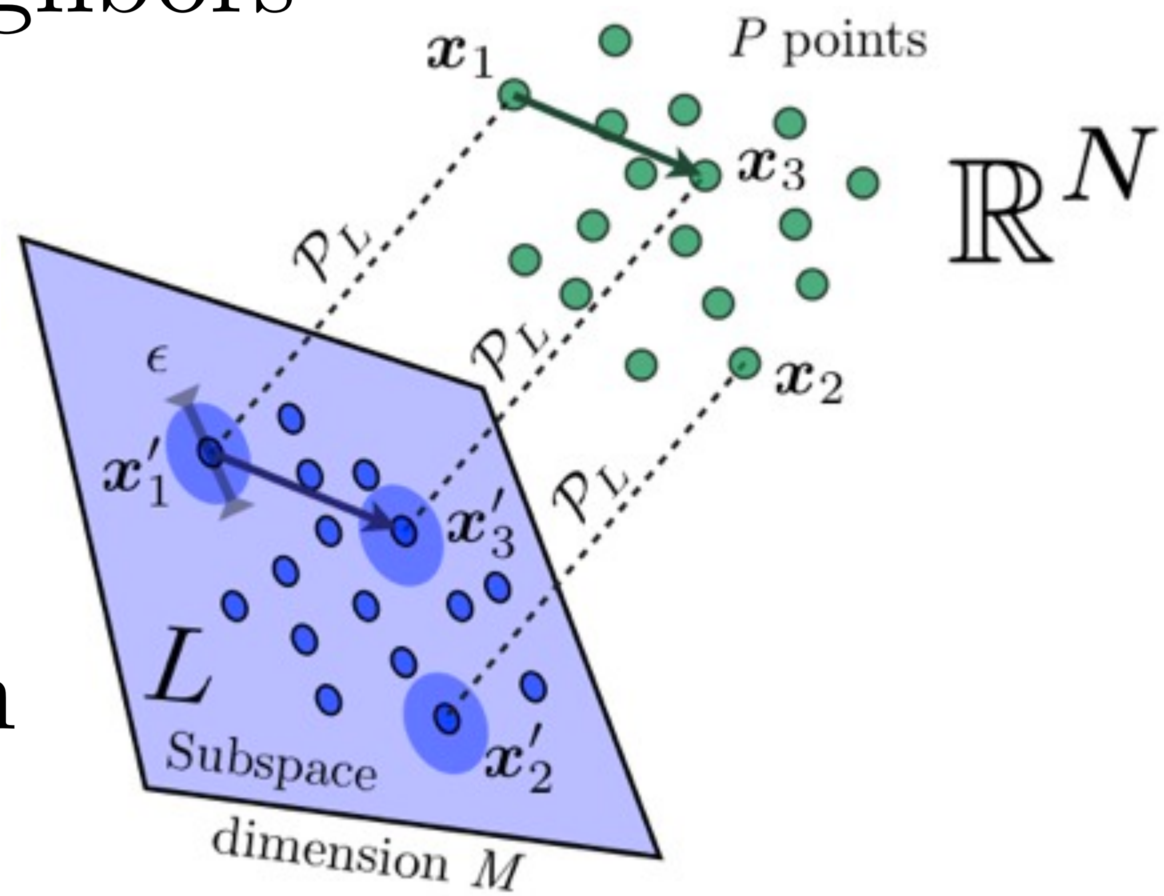
Applications of such a problem? Many!



Linear Dimensionality Reduction

Applications of such a problem? Many!

- Approximate Nearest Neighbors
- Query in Big Databases
- Machine Learning
- Signal Processing in a (easy) compressed domain
- Randomized algorithms
- ...



Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

Lemma 1 *Given an error* $0 < \epsilon < 1$, *and a point set* $\mathcal{S} \subset \mathbb{R}^N$. *If* M *is such that*

$$M > M_0 = O(\epsilon^{-2} \log |\mathcal{S}|),$$

then, there exists a (Lipschitz) mapping $\boldsymbol{f} : \mathbb{R}^N \to \mathbb{R}^M$ *such that*

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| \;\leqslant\; \|\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{v})\| \;\leqslant\; (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\|,$$

for all $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{S}$.

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

Lemma 1 *Given an error* $0 < \epsilon < 1$, *and a point set* $\mathcal{S} \subset \mathbb{R}^N$.
If M *is such that*

$$M > M_0 = O(\epsilon^{-2} \log |\mathcal{S}|),$$

then, there exists a (Lipschitz) mapping $\boldsymbol{f} : \mathbb{R}^N \to \mathbb{R}^M$ *such that*

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| \leqslant \|\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{v})\| \leqslant (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\|,$$

for all $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{S}$.

$\Rightarrow$ isometry between $(\mathcal{S}, \ell_2)$ and $(\boldsymbol{f}(\mathcal{S}), \ell_2)$

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

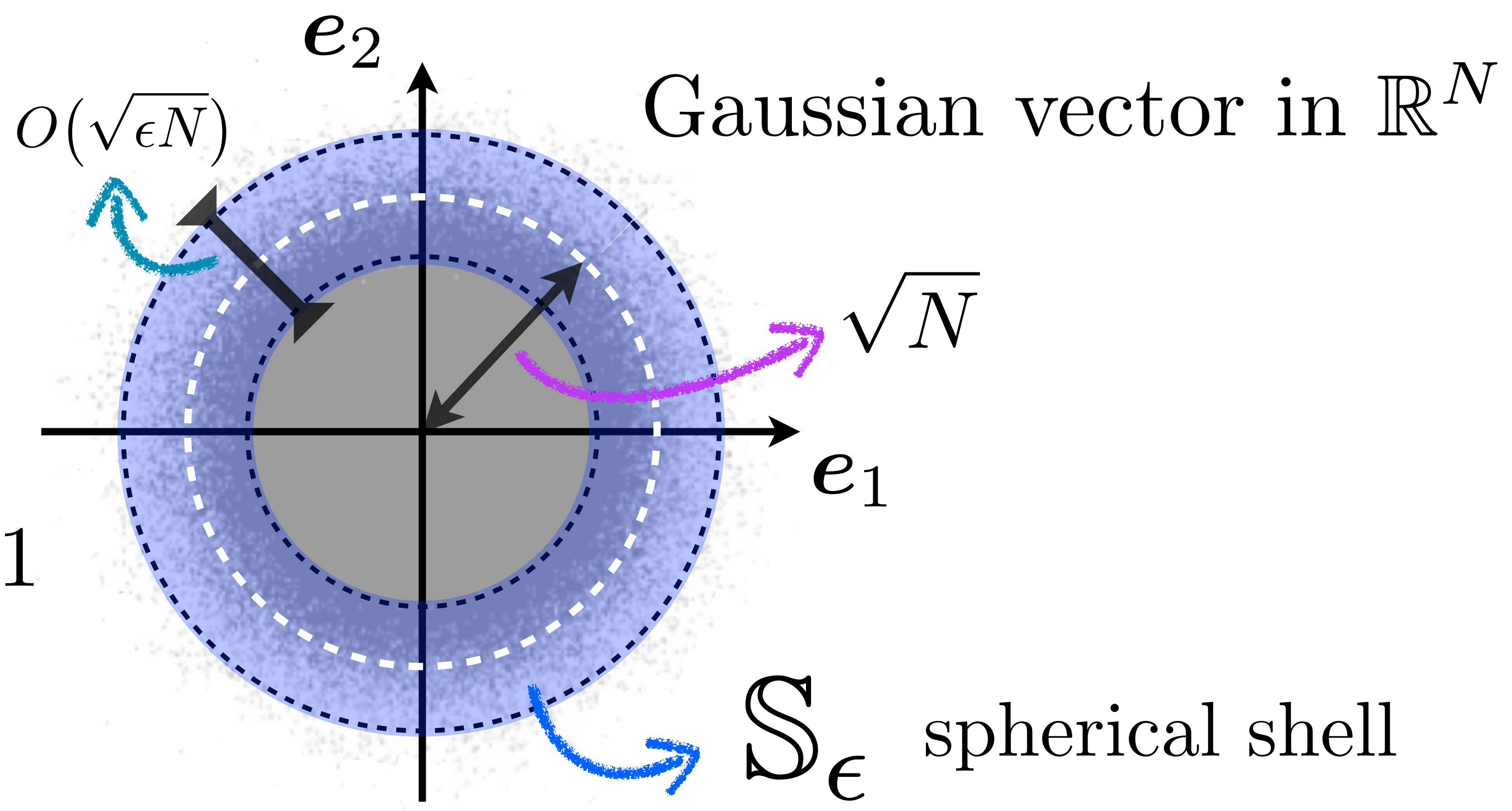
proof sketch:

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)

Weird things happens in high dimension!

$\mathbb{P}[\text{ vector} \in \mathbb{S}_\epsilon \,] \to_N 1$

and exponentially!



Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

proof sketch:

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)

Let $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0, 1/M)$, then, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N$,

$$\mathbb{P}\Big[\big|\|\underbrace{\boldsymbol{\Phi}(\boldsymbol{u}-\boldsymbol{v})}_{\text{Gaussian vector in } \mathbb{R}^M}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\big| \geqslant \epsilon\|\boldsymbol{u}-\boldsymbol{v}\|^2\Big] \;\leqslant\; 2e^{-M\epsilon^2/3},$$

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

<u>*proof sketch:*</u>

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)

Let $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0, 1/M)$, then, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N$,

$$\mathbb{P}\big[\big|\|\boldsymbol{\Phi}(\boldsymbol{u}-\boldsymbol{v})\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\big| \geqslant \epsilon\|\boldsymbol{u}-\boldsymbol{v}\|^2\big] \;\leqslant\; 2e^{-M\epsilon^2/3},$$

- Union bound on $\binom{|\mathcal{S}|}{2} = O(|\mathcal{S}|^2)$ pairs in $\mathcal{S}$: $\mathbb{P}[\cup_j \mathcal{E}_j\,] \;\leqslant\; \sum_j \mathbb{P}[\mathcal{E}_j]$

$$\Rightarrow \mathbb{P}(\exists \text{ failure for one pair in } \mathcal{S}) \;\leqslant\; 2e^{2\log|\mathcal{S}| - M\epsilon^2/3} \underset{e.g.}{<} 2/3$$

if $M \geqslant M_0 = O(\epsilon^{-2}\log|\mathcal{S}|)$

- $\exists \boldsymbol{f}$ for JL Lemma!

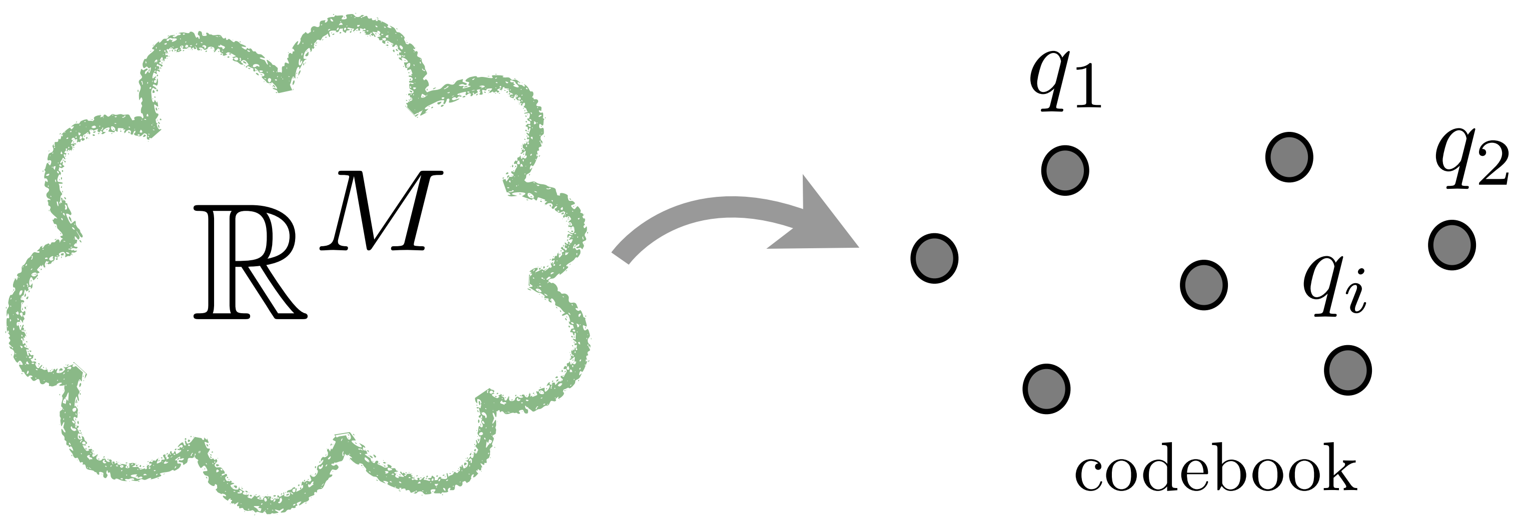
(boost $1-\mathbb{P}$ asking that $\exists$ good $\boldsymbol{\Phi}$ over many trials) □

2. Quantizing the J-L Lemma
-- prologue --

What is quantization?

- <u>Generality</u>:

  Intuitively: "*Quantization maps a continuous (bounded) domain to a set of finite elements (or codebook)*"



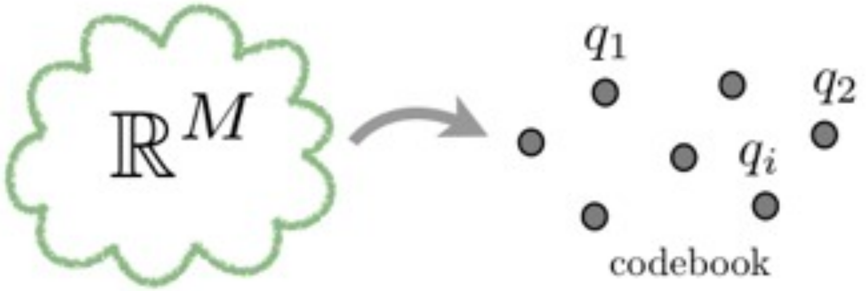
$$\mathcal{Q}[x] \in \{q_1, q_2, \cdots\}$$

- Oldest example: rounding off $\lfloor x \rfloor, \lceil x \rceil, \ldots \quad \mathbb{R} \to \mathbb{Z}$

What is quantization?

- Generality:

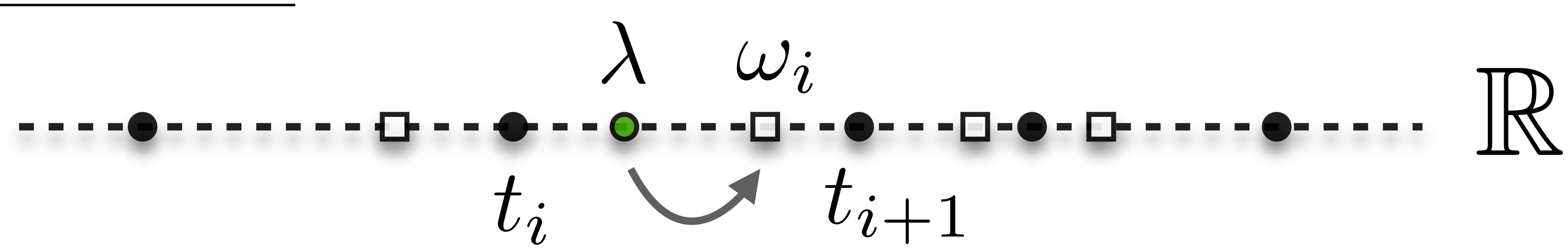
  Intuitively: "*Quantization maps a continuous (bounded) domain to a set of finite elements (or codebook)*"



- Needed for:
  - storing/computing/transmitting information
  - turning continuous values in bits (digitization)
  - quantifying/measuring information

Scalar quantization

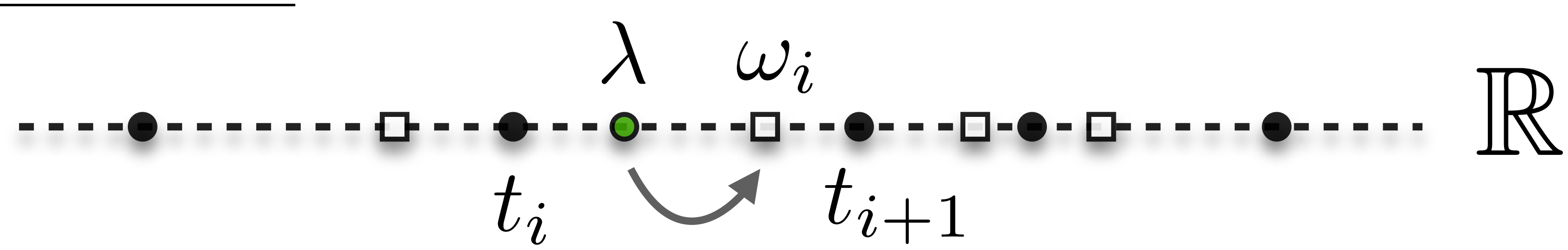
Principle in 1-D:



$$\mathcal{Q}[\lambda] = \omega_i \quad \Leftrightarrow \quad \lambda \in [t_i, t_{i+1}]$$

Scalar quantization

<u>Principle in 1-D:</u>

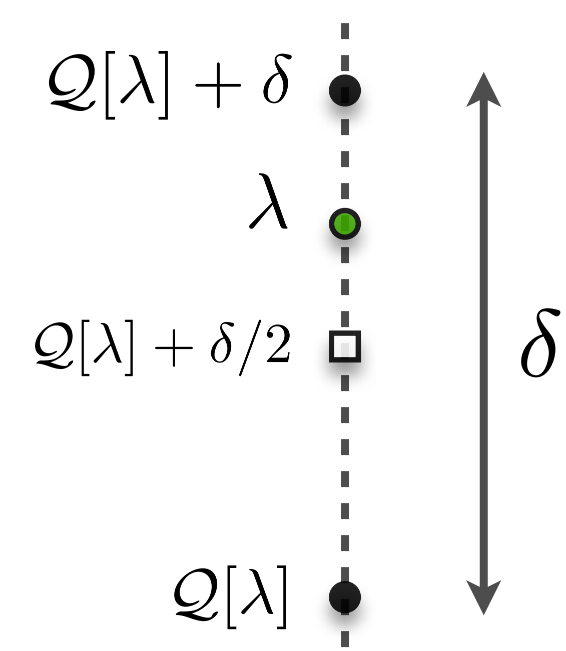


$$\mathcal{Q}[\lambda] = \omega_i \quad \Leftrightarrow \quad \lambda \in [t_i, t_{i+1}]$$

From now on: Given a resolution $\delta > 0$,

$$\mathcal{Q}[\lambda] = \delta \lfloor \lambda / \delta \rfloor \quad \in \; \mathbb{Z}_\delta := \delta \mathbb{Z}$$

and $(\mathcal{Q}[\boldsymbol{v}])_j = \mathcal{Q}[v_j]$ for vectors.



Remark: $|\lambda - \mathcal{Q}[\lambda] - \frac{1}{2}\delta| \leqslant \frac{1}{2}\delta$ for all λ

$\Rightarrow$ Quant. error $= \frac{1}{2}\delta$

Quantizing JL (first attempt)

Given a mapping $\boldsymbol{f} : \mathbb{R}^N \to \mathbb{R}^M$ s.t. $\frac{1}{\sqrt{M}}\boldsymbol{f}$ is JL

e.g., $\boldsymbol{f}(\cdot) = \boldsymbol{\Phi}\,\cdot$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ $\Rightarrow$ constant dynamic for $f_j(\cdot)$
(important for quantizing)

Quantizing JL (first attempt)

Given a mapping $\boldsymbol{f}: \mathbb{R}^N \to \mathbb{R}^M$ s.t. $\frac{1}{\sqrt{M}}\boldsymbol{f}$ is JL

e.g., $\boldsymbol{f}(\cdot) = \boldsymbol{\Phi}\cdot$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ $\Rightarrow$ constant dynamic for $f_j(\cdot)$

Form $\boldsymbol{\psi} := \mathcal{Q} \circ \boldsymbol{f}: \ \mathbb{R}^N \ \to \ \mathbb{Z}_\delta^M$

Then, with $M \geqslant M_0 = O(\epsilon^{-2} \log|\mathcal{S}|)$, and $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathcal{S}$,

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \ \leqslant\ \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \ \leqslant\ (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

$\Rightarrow$ **quasi**-isometry between $(\mathcal{S}, \ell_2)$ and $(\boldsymbol{f}(\mathcal{S}), \ell_2)$

Quantizing JL (first attempt)

Given a mapping $\boldsymbol{f} : \mathbb{R}^N \to \mathbb{R}^M$ s.t. $\frac{1}{\sqrt{M}}\boldsymbol{f}$ is JL

e.g., $\boldsymbol{f}(\cdot) = \boldsymbol{\Phi}\,\cdot$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ $\Rightarrow$ constant dynamic for $f_j(\cdot)$

Form $\boldsymbol{\psi} := \mathcal{Q} \circ \boldsymbol{f} : \ \mathbb{R}^N \ \to \ \mathbb{Z}_\delta^M$

Then, with $M \geqslant M_0 = O(\epsilon^{-2} \log |\mathcal{S}|)$,

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \ \leqslant\ \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \ \leqslant\ (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

both smaller than $\delta/2$

Proof (easy): $|\mathcal{Q}(a) - \mathcal{Q}(b)| = |b - \mathcal{Q}(b) - \tfrac{1}{2}\delta - (a - \mathcal{Q}(a) - \tfrac{1}{2}\delta) + (a-b)|$

$$\begin{cases} \leqslant |a-b| + \delta \\ \geqslant |a-b| - \delta \end{cases}$$

Then, with 2 more lines,

$$\tfrac{1}{\sqrt{M}}\|\mathcal{Q}(\boldsymbol{f}(\boldsymbol{u})) - \mathcal{Q}(\boldsymbol{f}(\boldsymbol{v}))\| \leqslant \underbrace{\tfrac{1}{\sqrt{M}}\|\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{v})\|}_{\leqslant (1+\epsilon)\|\boldsymbol{u}-\boldsymbol{v}\| \ \text{(by JL)}} + \delta \text{ and}$$

$$\tfrac{1}{\sqrt{M}}\|\mathcal{Q}(\boldsymbol{f}(\boldsymbol{u})) - \mathcal{Q}(\boldsymbol{f}(\boldsymbol{v}))\| \geqslant \underbrace{\tfrac{1}{\sqrt{M}}\|\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{v})\|}_{\geqslant (1-\epsilon)\|\boldsymbol{u}-\boldsymbol{v}\|} - \delta. \qquad \square$$

Quantizing JL (first attempt)

Given a mapping $\boldsymbol{f} : \mathbb{R}^N \to \mathbb{R}^M$ s.t. $\frac{1}{\sqrt{M}}\boldsymbol{f}$ is JL

e.g., $\boldsymbol{f}(\cdot) = \boldsymbol{\Phi}\,\cdot$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ $\Rightarrow$ constant dynamic for $f_j(\cdot)$

Form $\boldsymbol{\psi} := \mathcal{Q} \circ \boldsymbol{f} : \; \mathbb{R}^N \; \to \; \mathbb{Z}_\delta^M$

Then, with $M \geqslant M_0 = O(\epsilon^{-2} \log |\mathcal{S}|)$,

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \leqslant \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \leqslant (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

multiplicative error — **additive error**

Quantizing JL (first attempt)

Given a mapping $\boldsymbol{f}: \mathbb{R}^N \to \mathbb{R}^M$ s.t. $\frac{1}{\sqrt{M}}\boldsymbol{f}$ is JL

e.g., $\boldsymbol{f}(\cdot) = \boldsymbol{\Phi}\,\cdot$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0,1)$ $\Rightarrow$ constant dynamic for $f_j(\cdot)$

Form $\boldsymbol{\psi} := \mathcal{Q} \circ \boldsymbol{f}: \ \mathbb{R}^N \ \to \ \mathbb{Z}_\delta^M$

Then, with $M \geqslant M_0 = O(\epsilon^{-2} \log|\mathcal{S}|)$,

$$(1-\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| - \delta \leqslant \tfrac{1}{\sqrt{M}}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\| \leqslant (1+\epsilon)\,\|\boldsymbol{u}-\boldsymbol{v}\| + \delta,$$

(decaying, good!) **multiplicative error**

(constant, weird!?) **additive error**

Problem: $\epsilon = O\big(\sqrt{\log|\mathcal{S}|/M_0}\,\big)$ but δ is constant!

Can we hope better?

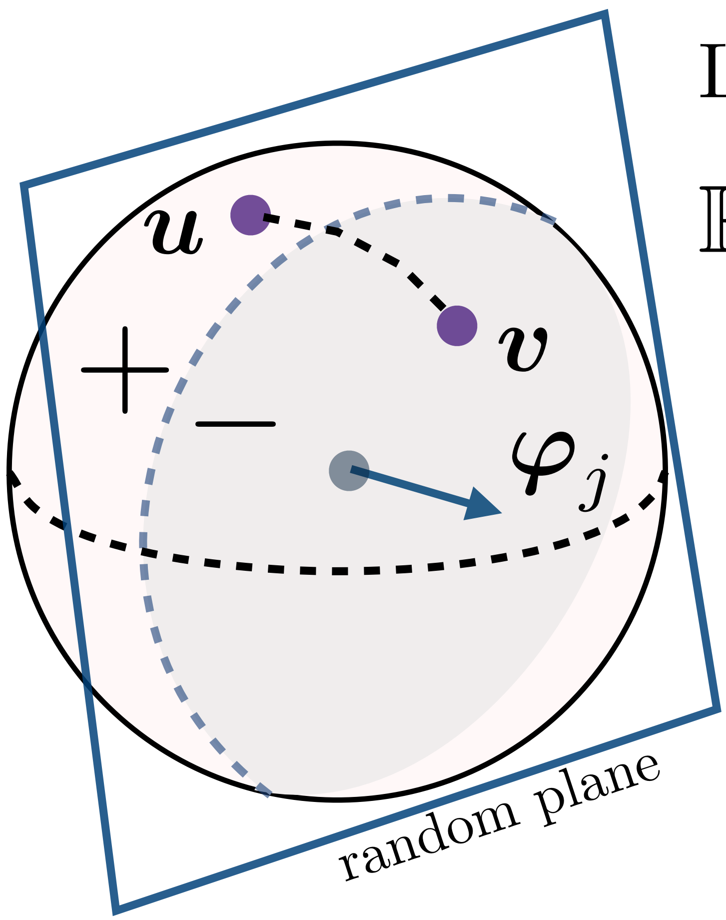
What's known? Binary Quantization

(equiv. to $\delta \gg \mathrm{diam}\mathcal{S}$)

- Let's define

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathrm{sign}\,(\boldsymbol{\Phi}\boldsymbol{u}) \;\Leftrightarrow\; \psi_j(\boldsymbol{u}) = \mathrm{sign}\,(\boldsymbol{\varphi}_j \cdot \boldsymbol{u}) \;\in \{\pm 1\}$$

j^{th} row of $\boldsymbol{\Phi}$



Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{S}^{N-1}$ (wlog)

$$\mathbb{P}[\psi_j(\boldsymbol{u}) \neq \psi_j(\boldsymbol{v})] = \;?$$

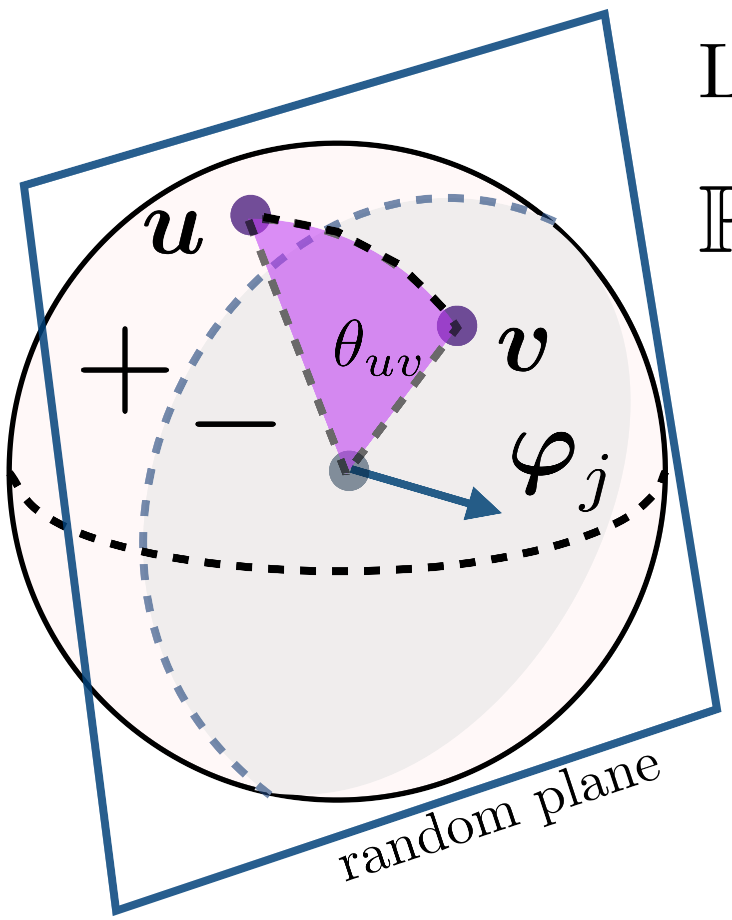
What's known? Binary Quantization

(equiv. to $\delta \gg \text{diam}\mathcal{S}$)

- Let's define

$$\boldsymbol{\psi}(\boldsymbol{u}) := \text{sign}\,(\boldsymbol{\Phi}\boldsymbol{u}) \;\Leftrightarrow\; \psi_j(\boldsymbol{u}) = \text{sign}\,(\boldsymbol{\varphi}_j \cdot \boldsymbol{u}) \;\in \{\pm 1\}$$

j^{th} row of $\boldsymbol{\Phi}$



Let $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{S}^{N-1}$ (wlog)

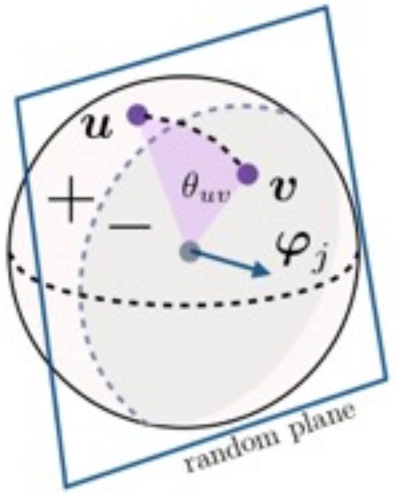
$$\mathbb{P}[\psi_j(\boldsymbol{u}) \neq \psi_j(\boldsymbol{v})] = \tfrac{1}{\pi}\,\text{angle}(\boldsymbol{u}, \boldsymbol{v}) = \theta_{uv}/\pi$$

What's known? Binary Quantization

(equiv. to $\delta \gg \mathrm{diam}\mathcal{S}$)

- Let's define

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathrm{sign}\,(\boldsymbol{\Phi}\boldsymbol{u}) \;\Leftrightarrow\; \psi_j(\boldsymbol{u}) = \mathrm{sign}\,(\boldsymbol{\varphi}_j \cdot \boldsymbol{u}) \;\in \{\pm 1\}$$



$$\mathbb{P}[\psi_j(\boldsymbol{u}) \neq \psi_j(\boldsymbol{v})] = \tfrac{1}{\pi}\,\mathrm{angle}(\boldsymbol{u}, \boldsymbol{v}) = \theta_{uv}/\pi$$

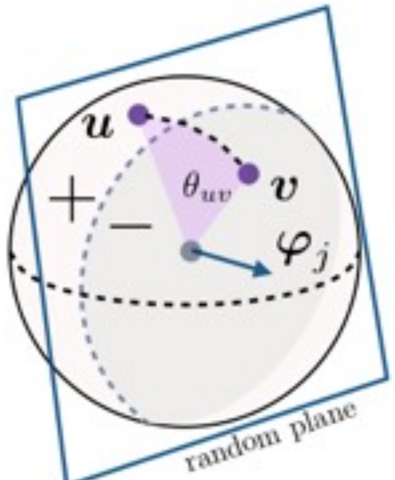
$$\Rightarrow X_j = \tfrac{1}{2}|\psi_j(\boldsymbol{u}) - \psi_j(\boldsymbol{v})| \sim \mathrm{Bernoulli}(\tfrac{\theta_{uv}}{\pi}) \in \{0, 1\}$$

What's known? Binary Quantization

(equiv. to $\delta \gg \mathrm{diam}\mathcal{S}$)

- Let's define

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathrm{sign}\,(\boldsymbol{\Phi}\boldsymbol{u}) \;\Leftrightarrow\; \psi_j(\boldsymbol{u}) = \mathrm{sign}\,(\boldsymbol{\varphi}_j \cdot \boldsymbol{u}) \;\in \{\pm 1\}$$



$$\mathbb{P}[\psi_j(\boldsymbol{u}) \neq \psi_j(\boldsymbol{v})] = \tfrac{1}{\pi}\,\mathrm{angle}(\boldsymbol{u},\boldsymbol{v}) = \theta_{uv}/\pi$$

$$\Rightarrow X_j = \tfrac{1}{2}|\psi_j(\boldsymbol{u}) - \psi_j(\boldsymbol{v})| \sim \mathrm{Bernoulli}(\tfrac{\theta_{uv}}{\pi}) \in \{0,1\}$$

From [Goemans, Williamson 1995], [LJ et al. 2011], [Plan 2011]

concentration around mean

For $M \geqslant M_0 = O(\epsilon^{-2}\log|\mathcal{S}|)$,

$$\theta_{uv} - \epsilon \;\leqslant\; \tfrac{1}{2M}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\|_1 \;\leqslant\; \theta_{uv} + \epsilon,$$

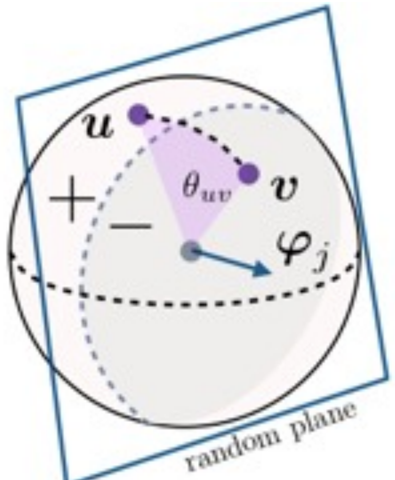
for all $\boldsymbol{u},\boldsymbol{v} \in \mathcal{S}$.

What's known? Binary Quantization

(equiv. to $\delta \gg \mathrm{diam}\mathcal{S}$)

- Let's define

$$\boldsymbol{\psi}(\boldsymbol{u}) := \mathrm{sign}\,(\boldsymbol{\Phi u}) \;\Leftrightarrow\; \psi_j(\boldsymbol{u}) = \mathrm{sign}\,(\boldsymbol{\varphi}_j \cdot \boldsymbol{u}) \;\in \{\pm 1\}$$



$$\mathbb{P}[\psi_j(\boldsymbol{u}) \neq \psi_j(\boldsymbol{v})] = \tfrac{1}{\pi}\,\mathrm{angle}(\boldsymbol{u},\boldsymbol{v}) = \theta_{uv}/\pi$$

$$\Rightarrow X_j = \tfrac{1}{2}|\psi_j(\boldsymbol{u}) - \psi_j(\boldsymbol{v})| \sim \mathrm{Bernoulli}(\tfrac{\theta_{uv}}{\pi}) \in \{0,1\}$$

From [Goemans, Williamson 1995], [LJ et al. 2011], [Plan 2011]

For $M \geqslant M_0 = O(\epsilon^{-2}\log|\mathcal{S}|)$,

$$\theta_{uv} - \epsilon \leqslant \frac{1}{2M}\,\|\boldsymbol{\psi}(\boldsymbol{u}) - \boldsymbol{\psi}(\boldsymbol{v})\|_1 \leqslant \theta_{uv} + \epsilon,$$

for all $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{S}$.

Here, we do see a decaying additive error! $\epsilon = O\big(\sqrt{\log|\mathcal{S}|/M_0}\,\big)$

3. The finding of Buffon's needle

Comte de Buffon

- Georges-Louis Leclerc, Comte de Buffon
- French Naturalist: 1707-1788
- Published 36 volumes of "L'Histoire Naturelle"
- Father of the field of "Geometrical Probability"

http://www.buffon.cnrs.fr

Buffon's needle problem

http://www.buffon.cnrs.fr

[Buffon's problem 1733, Buffon's solution 1777]

> *"I suppose that in a room where the floor is simply divided by parallel joints one throws a stick ("needle") in the air,*
>
> *and that one of the players bets that the stick will not cross any of the parallels on the floor,*
>
> *and that the other in contrast bets that the stick will cross some of these parallels;*
>
> ***one asks for the chances of these two players."***

http://www.buffon.cnrs.fr

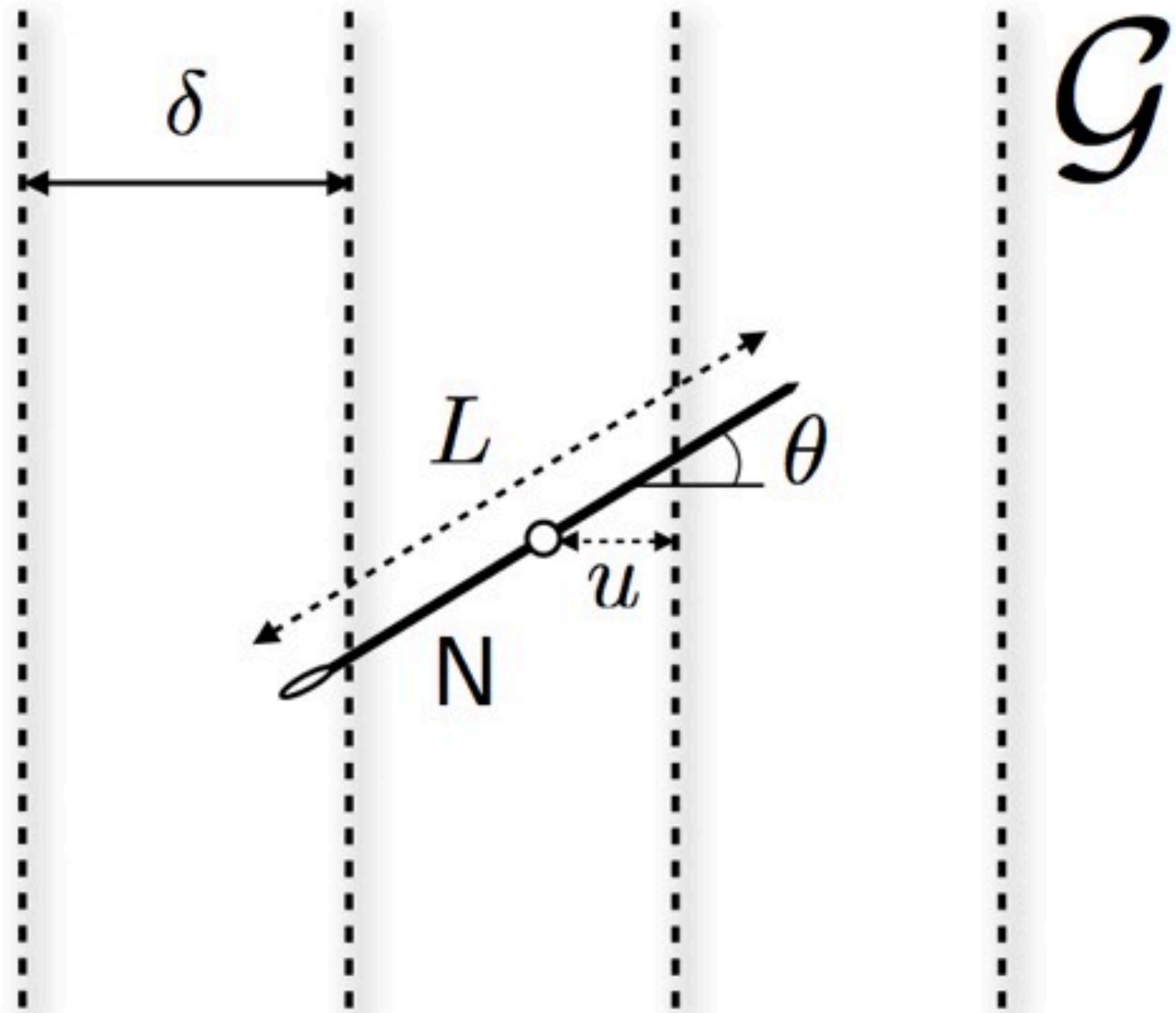
Buffon's needle problem


146 Essai

d'Arithmétique morale. 147

(Courtesy of E. Kowalski's blog)

$$\mathbb{P}[\,\mathsf{N}(u,\theta)\cap\mathcal{G}\neq\emptyset\,]$$
$$=\,?$$



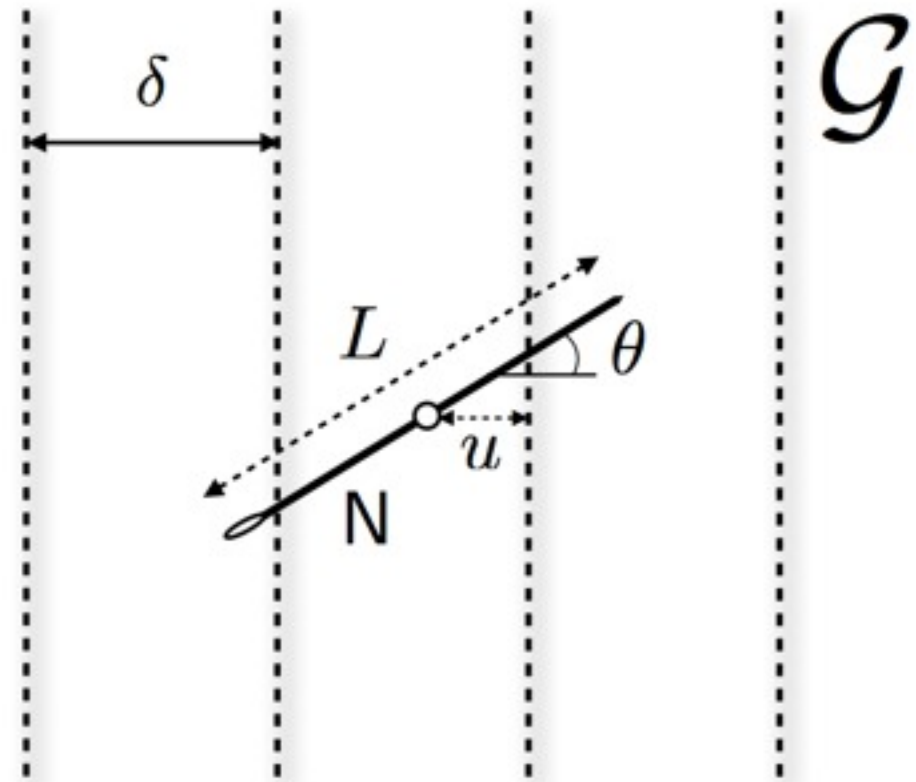
with $u \sim \mathcal{U}([0,\delta])$ and $\theta \sim \mathcal{U}([0,2\pi])$

http://www.buffon.cnrs.fr

Buffon's needle problem

Fact 1: if $L < \delta$, $\mathbb{P} = \frac{2}{\pi\delta} L$

(small integral to solve)



with $u \sim \mathcal{U}([0, \delta])$ and $\theta \sim \mathcal{U}([0, 2\pi])$

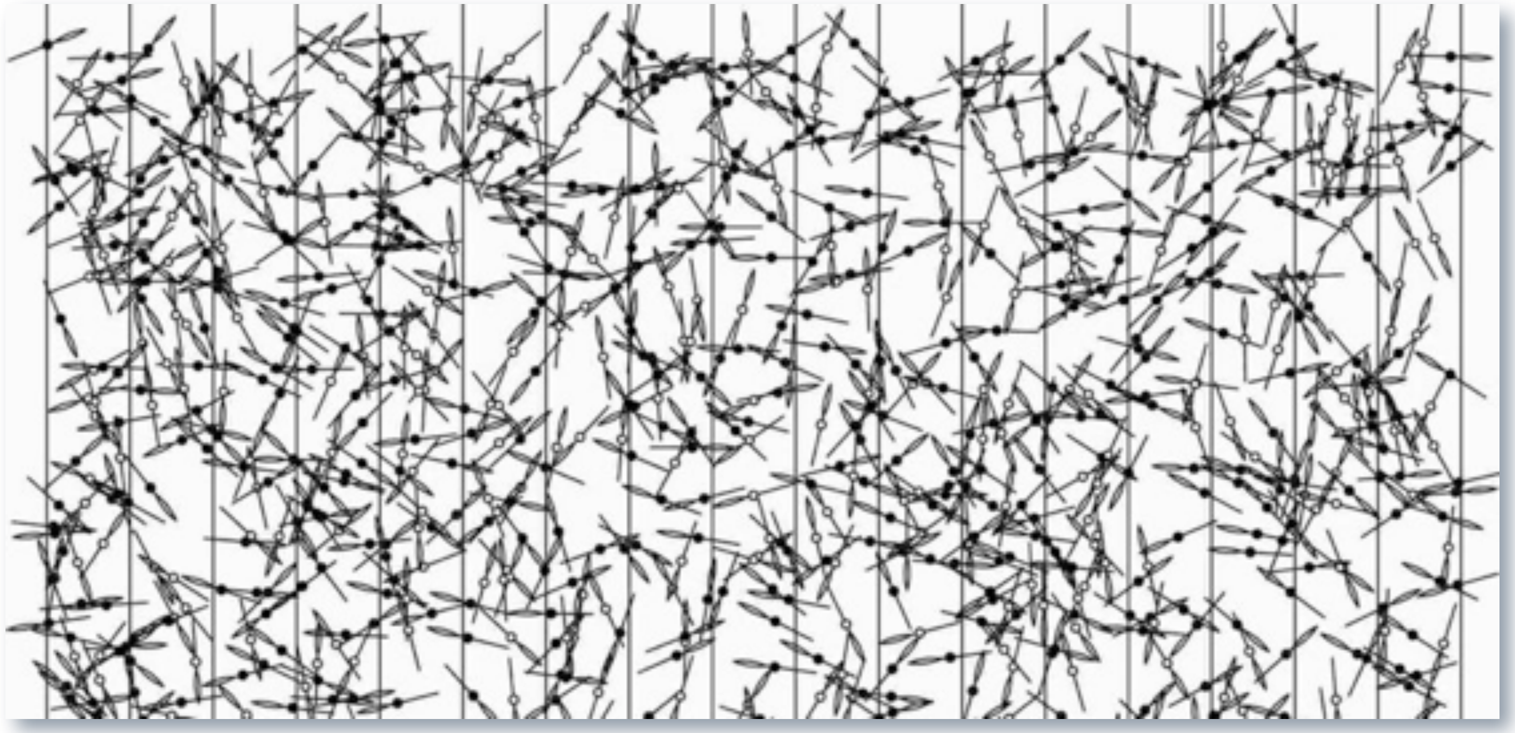
http://www.buffon.cnrs.fr

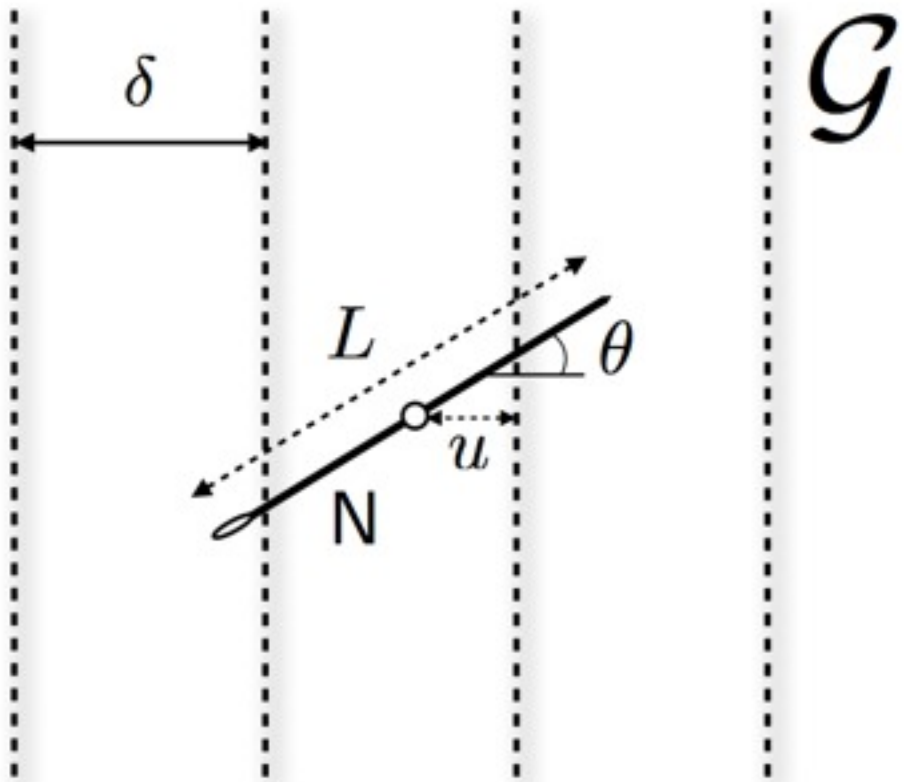
Buffon's needle problem

Fact 1: if $L < \delta$, $\mathbb{P} = \frac{2}{\pi\delta} L$

(small integral to solve)

Has been used for estimating π !
(first "Monte Carlo" method)





with $u \sim \mathcal{U}([0, \delta])$ and $\theta \sim \mathcal{U}([0, 2\pi])$

http://www.buffon.cnrs.fr

Buffon's needle problem

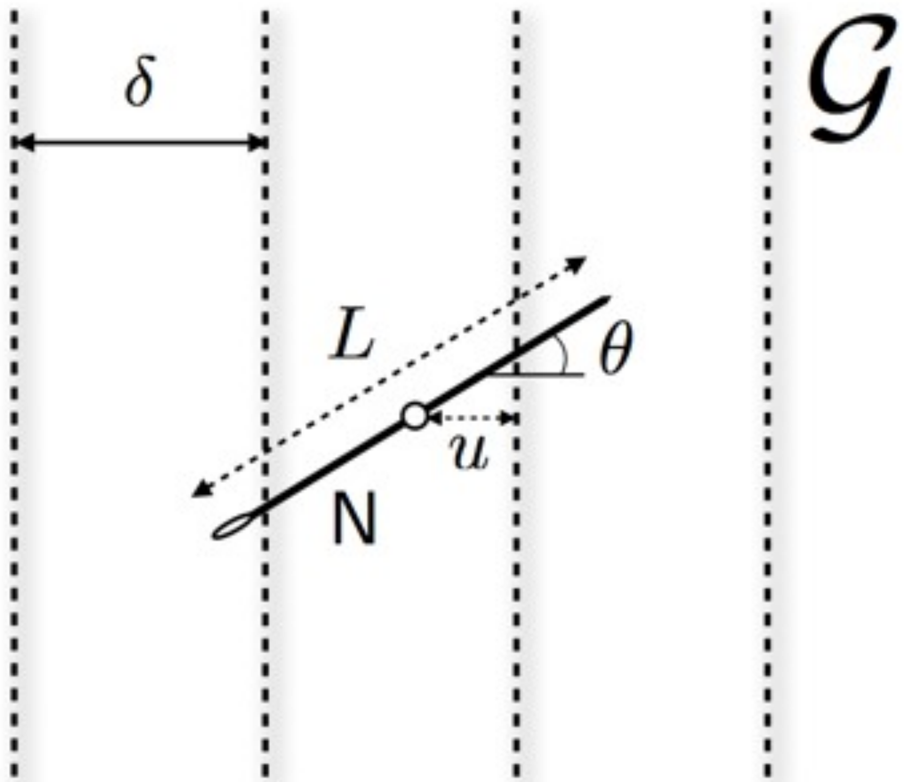
Fact 1: if $L < \delta$, $\mathbb{P} = \frac{2}{\pi\delta} L$

Fact 2: if $L \geqslant \delta$, $\mathbb{P} \neq \frac{2}{\pi\delta} L$ but

$$\mathbb{E}X = \frac{2}{\pi\delta} L,$$

with $X = \#\{\, \mathsf{N}(u,\theta) \cap \mathcal{G} \,\}$.

Proof: cut N in parts smaller than δ and sum expectations!



with $u \sim \mathcal{U}([0,\delta])$ and $\theta \sim \mathcal{U}([0,2\pi])$

Buffon's needle problem

http://www.buffon.cnrs.fr

Fact 1: if $L < \delta$, $\mathbb{P} = \frac{2}{\pi\delta} L$

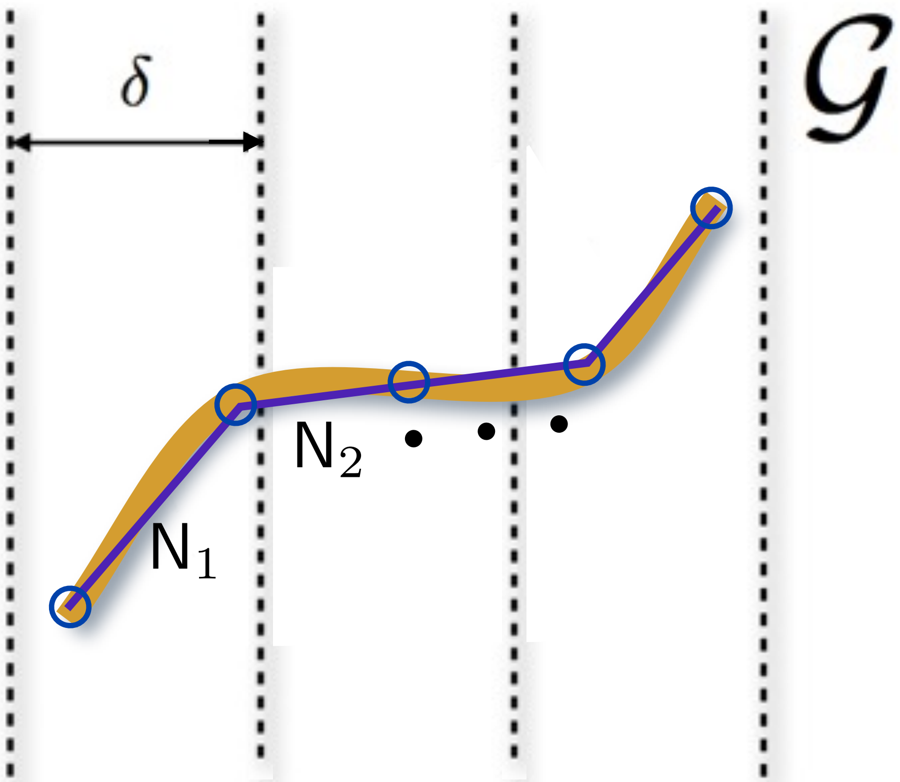
Fact 2: if $L \geqslant \delta$, $\mathbb{P} \neq \frac{2}{\pi\delta} L$ but

$$\mathbb{E}X = \frac{2}{\pi\delta} L,$$

with $X = \#\{\,\mathsf{N}(u,\theta) \cap \mathcal{G}\,\}$.

Fact 3: It works for "noodles" (smooth curves)!

For information only.



with $u \sim \mathcal{U}([0,\delta])$ and $\theta \sim \mathcal{U}([0,2\pi])$

Buffon in N-D? [LJ, 2013]

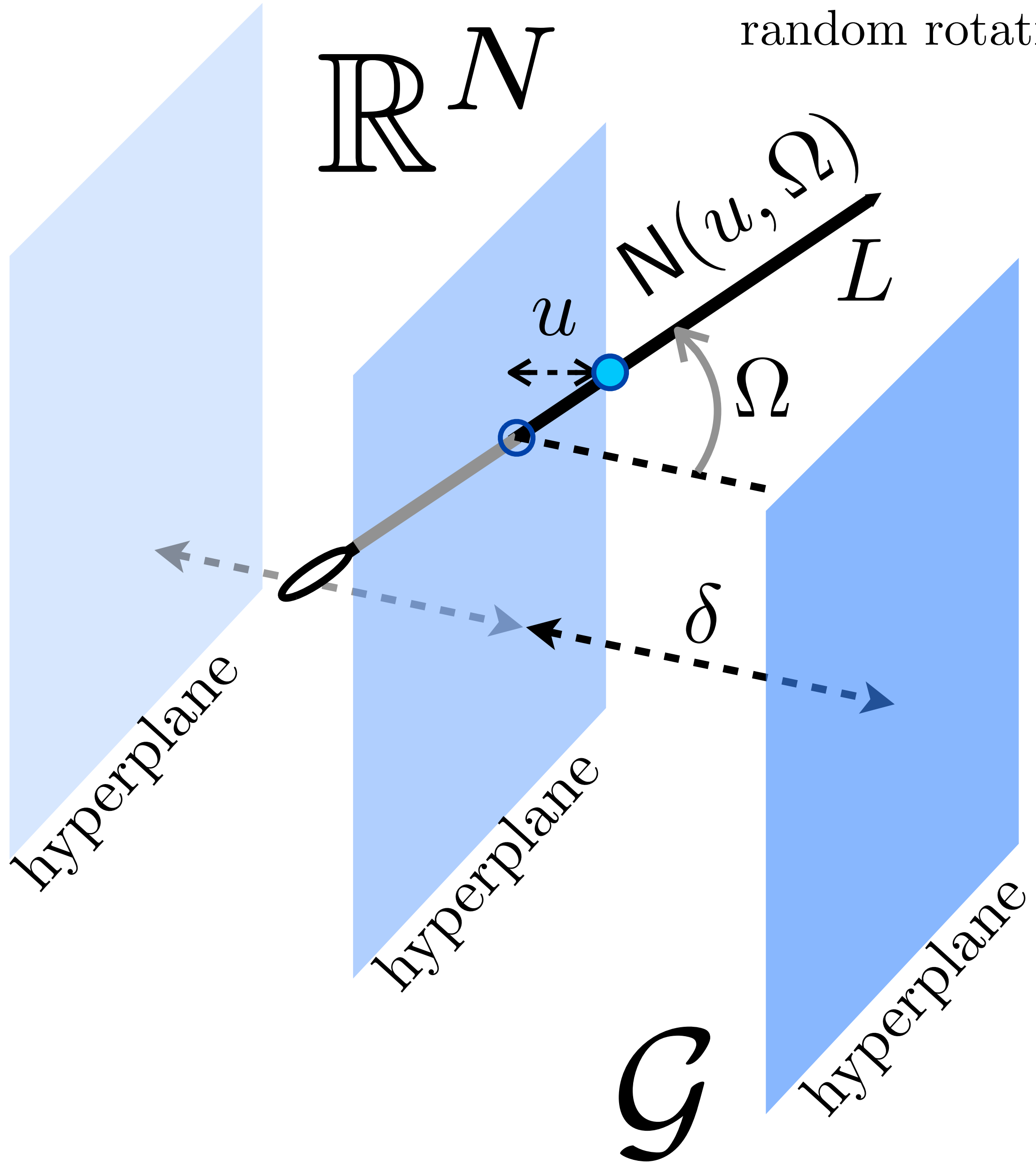
http://www.buffon.cnrs.fr

Buffon in N-D? [LJ, 2013]

(discr. r.v.) Let $X = \#\{\, \mathsf{N}(u,\Omega) \cap \mathcal{G}\,\}$,

with $\Omega \sim \mathcal{U}(SO(N))$, $u \sim \mathcal{U}([0,\delta])$.

random rotation in $\mathbb{R}^N$



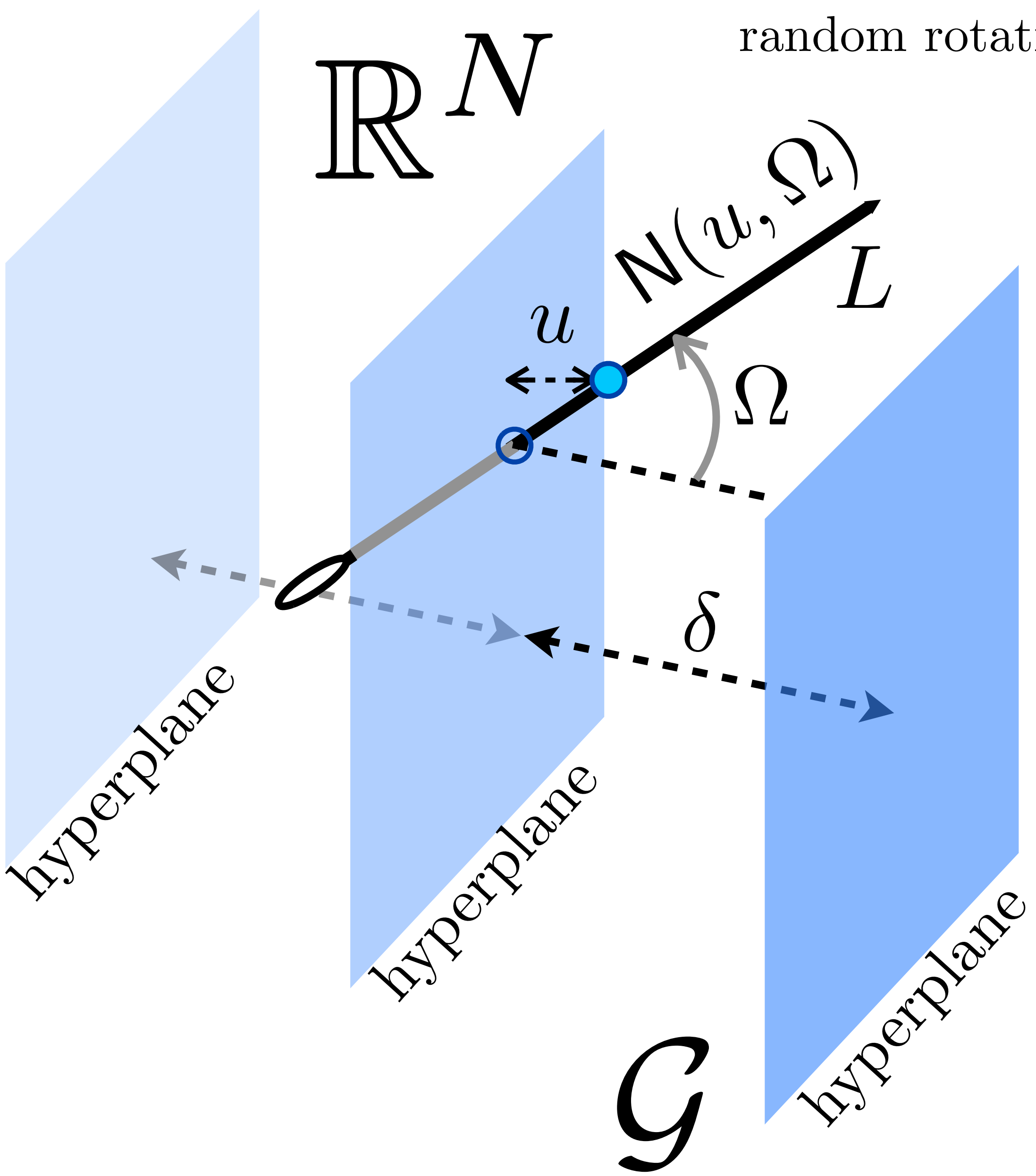
http://www.buffon.cnrs.fr

Buffon in N-D? [LJ, 2013]

(discr. r.v.) Let $X = \#\{\, \mathsf{N}(u,\Omega) \cap \mathcal{G}\, \}$,

with $\Omega \sim \mathcal{U}(SO(N))$, $u \sim \mathcal{U}([0,\delta])$.

random rotation in $\mathbb{R}^N$



We still have: $\mathbb{E}X = \tau_N \frac{L}{\delta}$,

with $\tau_N = \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi}\,\Gamma(\frac{N+1}{2})} \simeq_N 1/\sqrt{N}$

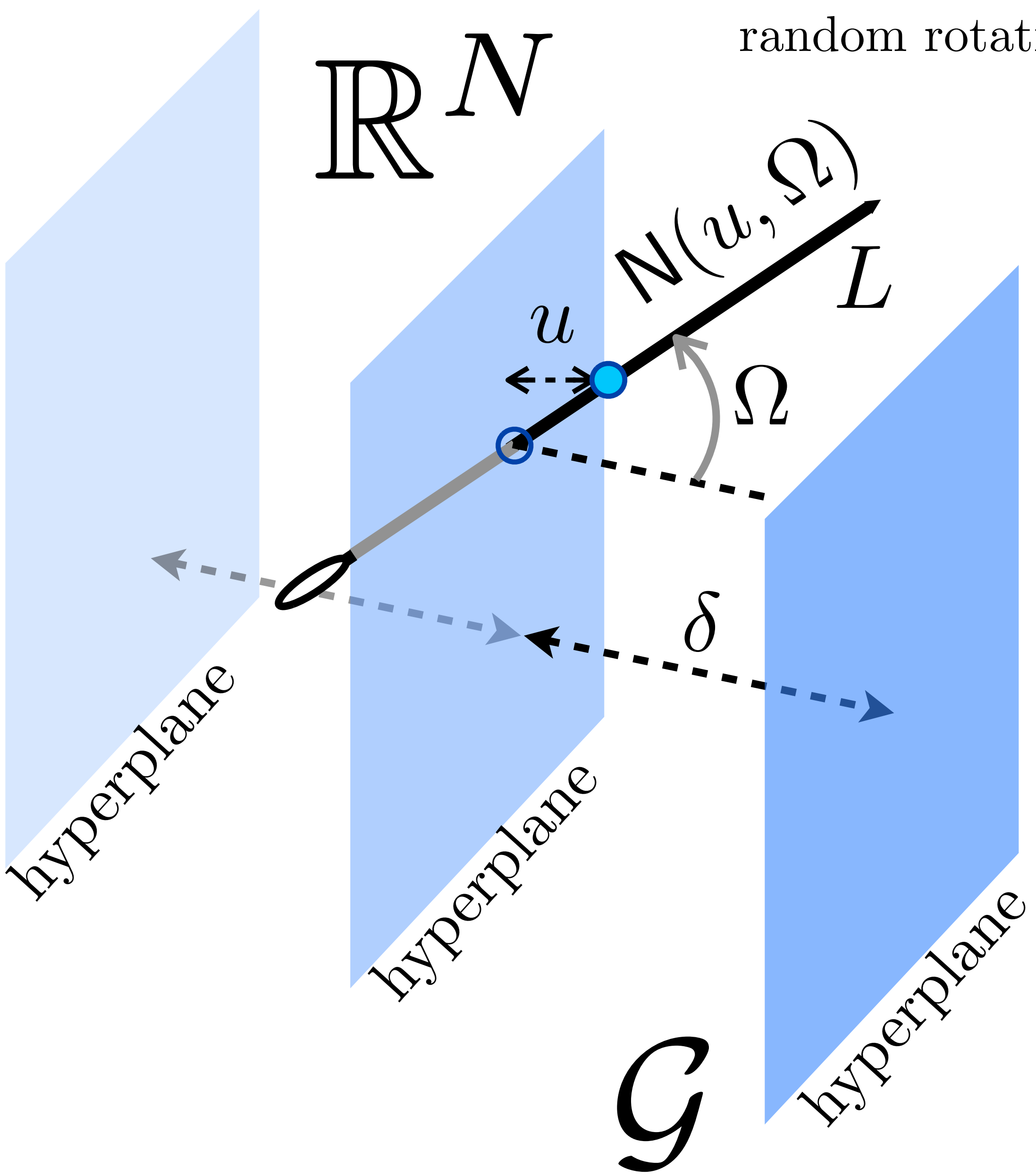
http://www.buffon.cnrs.fr

Buffon in N-D? [LJ, 2013]

(discr. r.v.) Let $X = \#\{\, \mathsf{N}(u,\Omega) \cap \mathcal{G}\,\}$,

with $\Omega \sim \mathcal{U}(SO(N))$, $u \sim \mathcal{U}([0,\delta])$.

random rotation in $\mathbb{R}^N$



We still have: $\mathbb{E}X = \tau_N \frac{L}{\delta}$,

with $\tau_N = \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi}\,\Gamma(\frac{N+1}{2})} \simeq_N 1/\sqrt{N}$

Moreover,

$p_k := \mathbb{P}[X = k]$ is computable!

(and all its moments: $\mathbb{E}X^q \leqslant c_q\,(L/\delta)^q$, $c_q > 0$)

We write:

$$X \sim \mathrm{Buffon}(L/\delta, N)$$

with $0 \leqslant X \leqslant \lceil L/\delta \rceil$.

4. Quantizing the J-L Lemma
-- epilogue --

Where's the equivalence? (what's the point?)

Let $\psi(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\varphi} \cdot \boldsymbol{x} + u)$, for $\boldsymbol{x} \in \mathbb{R}^N$

random Gaussian projection

$\boldsymbol{\varphi} \sim \mathcal{N}^{N \times 1}(0,1)$

random "dithering"

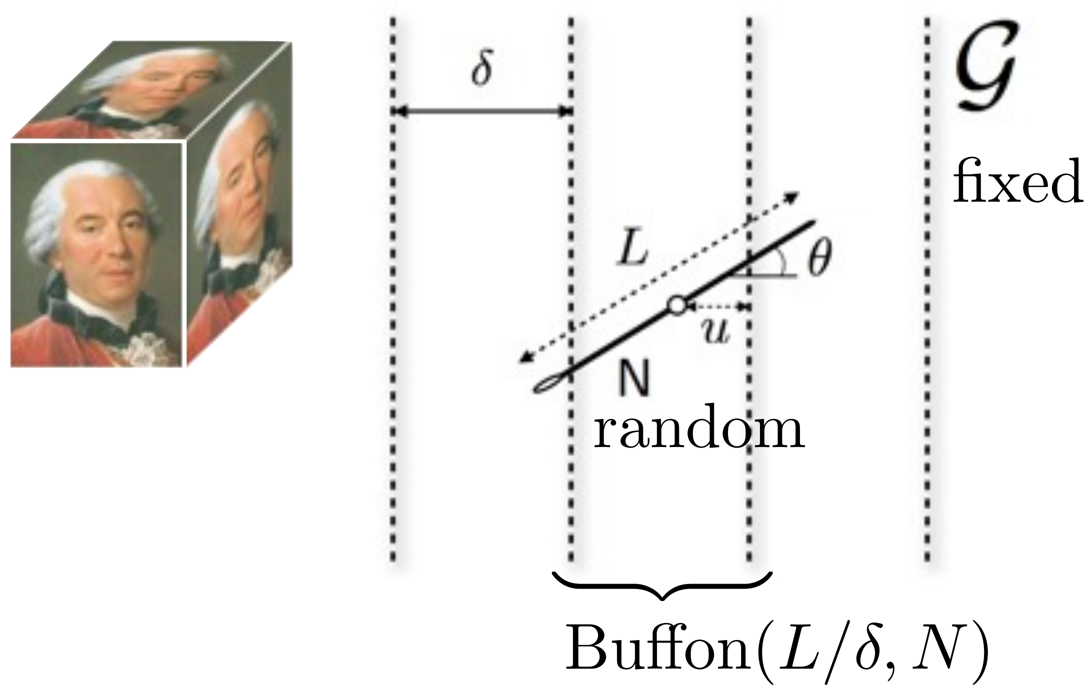
$u \sim \mathcal{U}([0,\delta])$

Scalar Quantization

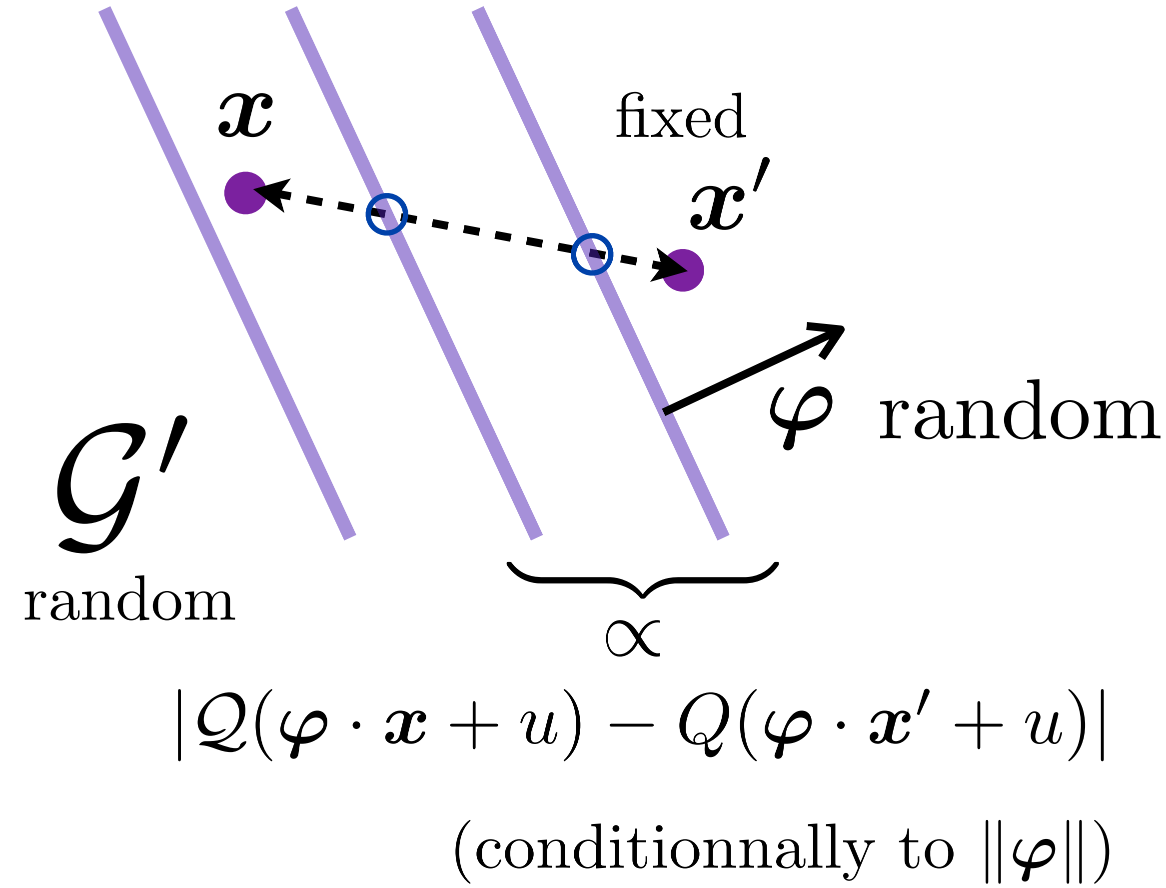
resolution $\delta > 0$

Where's the equivalence?

Let $\psi(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\varphi} \cdot \boldsymbol{x} + u)$, for $\boldsymbol{x} \in \mathbb{R}^N$



$\equiv$



Idea: in N-D

- grid $\mathcal{G} \leftrightarrow$ quant. $\mathcal{Q}$ with resol. δ
- needle N $\leftrightarrow$ segment $\overline{\boldsymbol{x}\boldsymbol{x}'}$
- fixed grid $\mathcal{G}$/random N $\leftrightarrow$ random grid $\mathcal{G}$/fixed $\overline{\boldsymbol{x}\boldsymbol{x}'}$

For M quantized projections?

Let $\boldsymbol{\psi}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{u})$ for $\boldsymbol{x} \in \mathbb{R}^N$

with $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$ and $\boldsymbol{u} \sim \mathcal{U}^M([0,\delta])$

For M quantized projections?

Let $\boldsymbol{\psi}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u})$ for $\boldsymbol{x} \in \mathbb{R}^N$

with $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$ and $\boldsymbol{u} \sim \mathcal{U}^M([0,\delta])$

Proposition *For each $1 \leqslant j \leqslant M$ and conditionally to the knowledge of $r_j = \|\boldsymbol{\varphi}_j\|$, we have*

$$X_j := \tfrac{1}{\delta}|(\boldsymbol{\psi}(\boldsymbol{x}))_j - (\boldsymbol{\psi}(\boldsymbol{x}'))_j| \sim_{\rm iid} \mathrm{Buffon}(\tfrac{r_j}{\delta}\|\boldsymbol{x} - \boldsymbol{x}'\|, N).$$

Proof: 1 page but intuition was given before.

For M quantized projections?

Let $\boldsymbol{\psi}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u})$ for $\boldsymbol{x} \in \mathbb{R}^N$

with $\boldsymbol{\Phi} \sim \mathcal{N}^{M\times N}(0,1)$ and $\boldsymbol{u} \sim \mathcal{U}^M([0,\delta])$

Proposition *For each* $1 \leqslant j \leqslant M$ *and conditionally to the knowledge of* $r_j = \|\boldsymbol{\varphi}_j\|$, *we have*

$$X_j := \tfrac{1}{\delta}|(\boldsymbol{\psi}(\boldsymbol{x}))_j - (\boldsymbol{\psi}(\boldsymbol{x}'))_j| \sim_{\rm iid} \mathrm{Buffon}(\tfrac{r_j}{\delta}\|\boldsymbol{x}-\boldsymbol{x}'\|, N).$$

Proof: 1 page but intuition was given before.

Moreover,

Buffon expect.

$$\mathbb{E}X_j = \mathbb{E}_{r_j}\big(\mathbb{E}(X_j|r_j)\big) = \tau_N\,\mathbb{E}_{r_j}(\tfrac{r_j}{\delta}\,\|\boldsymbol{x}-\boldsymbol{x}'\|) = \tfrac{\sqrt{2}}{\sqrt{\pi}}\,\tfrac{1}{\delta}\,\|\boldsymbol{x}-\boldsymbol{x}'\|,$$

since $\mathbb{E}\|\boldsymbol{\varphi}_j\| = \frac{\sqrt{2}}{\tau_N\sqrt{\pi}}$! (Chi($N$) distr.)

coincidences happen!

and finally ...

(reminder: $\boldsymbol{\psi}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi}\boldsymbol{x} + \boldsymbol{u})$)

- Knowing/bounding the expectation/moments of X_j
- and using measure concentration (by Bernstein) for

$$\tfrac{1}{M}\sum_j X_j = \tfrac{1}{\delta M}\|\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}(\boldsymbol{x}')\|_1$$

and finally ...

(reminder: $\psi(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u})$)

- Knowing/bounding the expectation/moments of X_j
- and using measure concentration (by Bernstein) for

$$\tfrac{1}{M}\textstyle\sum_j X_j = \tfrac{1}{\delta M}\|\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}(\boldsymbol{x}')\|_1$$

Quasi-isometry!

Lemma 1 *Given an error* $0 < \epsilon < 1$, *and a point set* $\mathcal{S} \subset \mathbb{R}^N$.
If M *is such that*

$$M \geqslant M_0 = O(\epsilon^{-2} \log |\mathcal{S}|),$$

then, for $c > 0$ *and with high probability, we have*

$$(1-\epsilon)\,\|\boldsymbol{x}-\boldsymbol{x}'\| - c\delta\epsilon \;\leqslant\; \frac{\sqrt{\pi}}{M\sqrt{2}}\,\|\boldsymbol{\psi}(\boldsymbol{x})-\boldsymbol{\psi}(\boldsymbol{x}')\|_1 \;\leqslant\; (1+\epsilon)\,\|\boldsymbol{x}-\boldsymbol{x}'\| + c\delta\epsilon,$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{S}$.

and finally ...

(reminder: $\boldsymbol{\psi}(\boldsymbol{x}) = \mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u})$)

- Knowing/bounding the expectation/moments of X_j
- and using measure concentration (by Bernstein) for

$$\tfrac{1}{M}\sum_j X_j = \tfrac{1}{\delta M}\|\boldsymbol{\psi}(\boldsymbol{x}) - \boldsymbol{\psi}(\boldsymbol{x}')\|_1$$

Quasi-isometry!

Lemma 1 *Given an error* $0 < \epsilon < 1$, *and a point set* $\mathcal{S} \subset \mathbb{R}^N$. *If* M *is such that*

$$M \geqslant M_0 = O(\epsilon^{-2} \log |\mathcal{S}|),$$

$\epsilon = O\big(\sqrt{\log |\mathcal{S}|/M_0}\big)$

then, for $c > 0$ *and with high probability, we have*

$$(1-\epsilon)\,\|\boldsymbol{x}-\boldsymbol{x}'\| - c\delta\epsilon \leqslant \frac{\sqrt{\pi}}{M\sqrt{2}}\,\|\boldsymbol{\psi}(\boldsymbol{x})-\boldsymbol{\psi}(\boldsymbol{x}')\|_1 \leqslant (1+\epsilon)\,\|\boldsymbol{x}-\boldsymbol{x}'\| + c\delta\epsilon,$$

for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{S}$.

decaying multiplicative and additive errors!

5. A few numerical tests

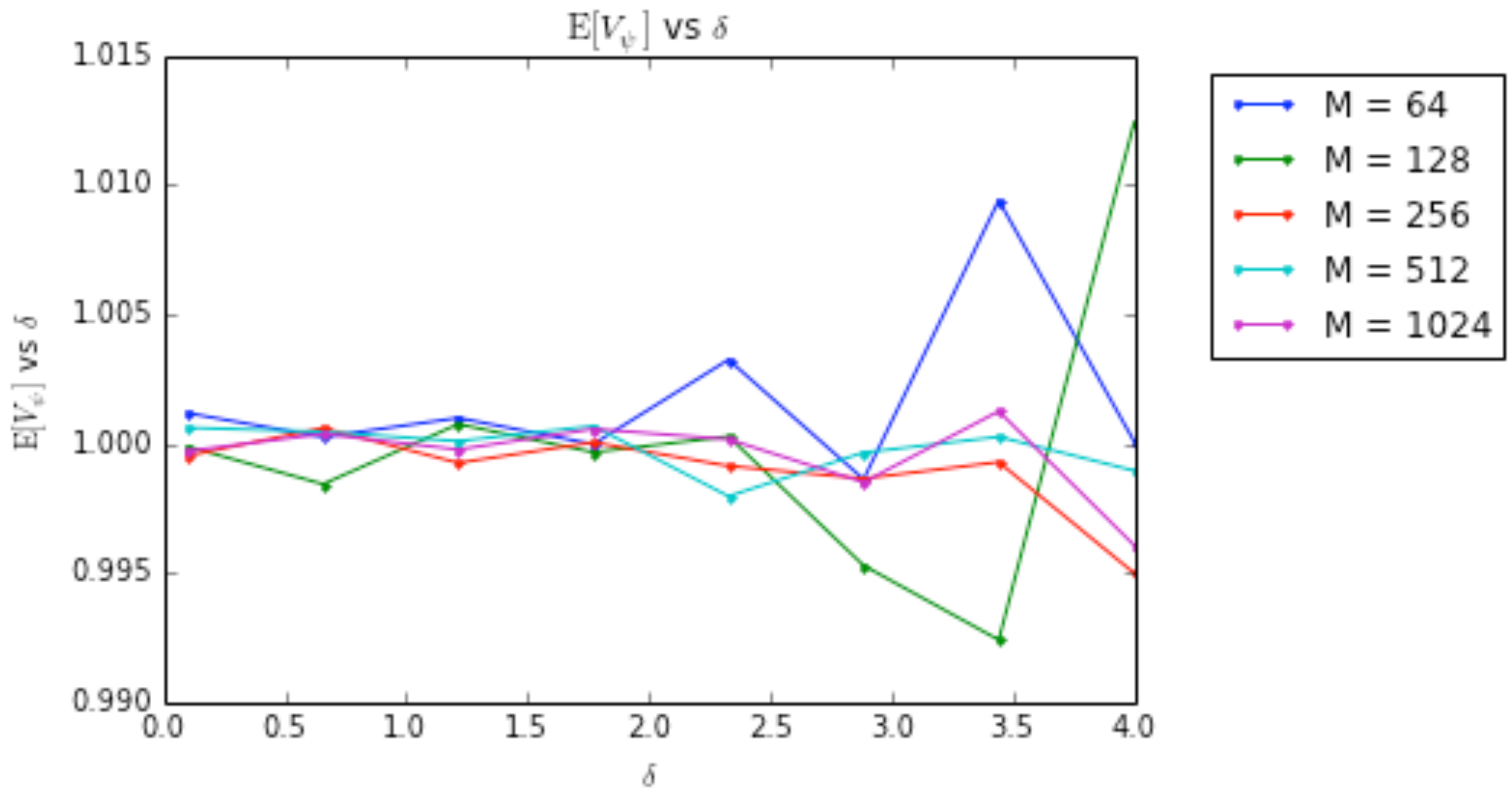
Simulations

(demo available on http://tinyurl.com/quantJL)

Idea: testing $V_\psi = \frac{\sqrt{\pi}}{M\sqrt{2}} \|\mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u}) - \mathcal{Q}(\boldsymbol{\Phi x}' + \boldsymbol{u})\|_1$

- $N = 256$, $M \in \{64, 128, \cdots, 1024\}$ and $\delta \in [0.1, 4]$.
- For each (M, N, δ), 100 trials on $(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{\Phi})$ with $\|\boldsymbol{x} - \boldsymbol{x}'\| = 1$ (WLOG)

Expectation of V_ψ?



Simulations

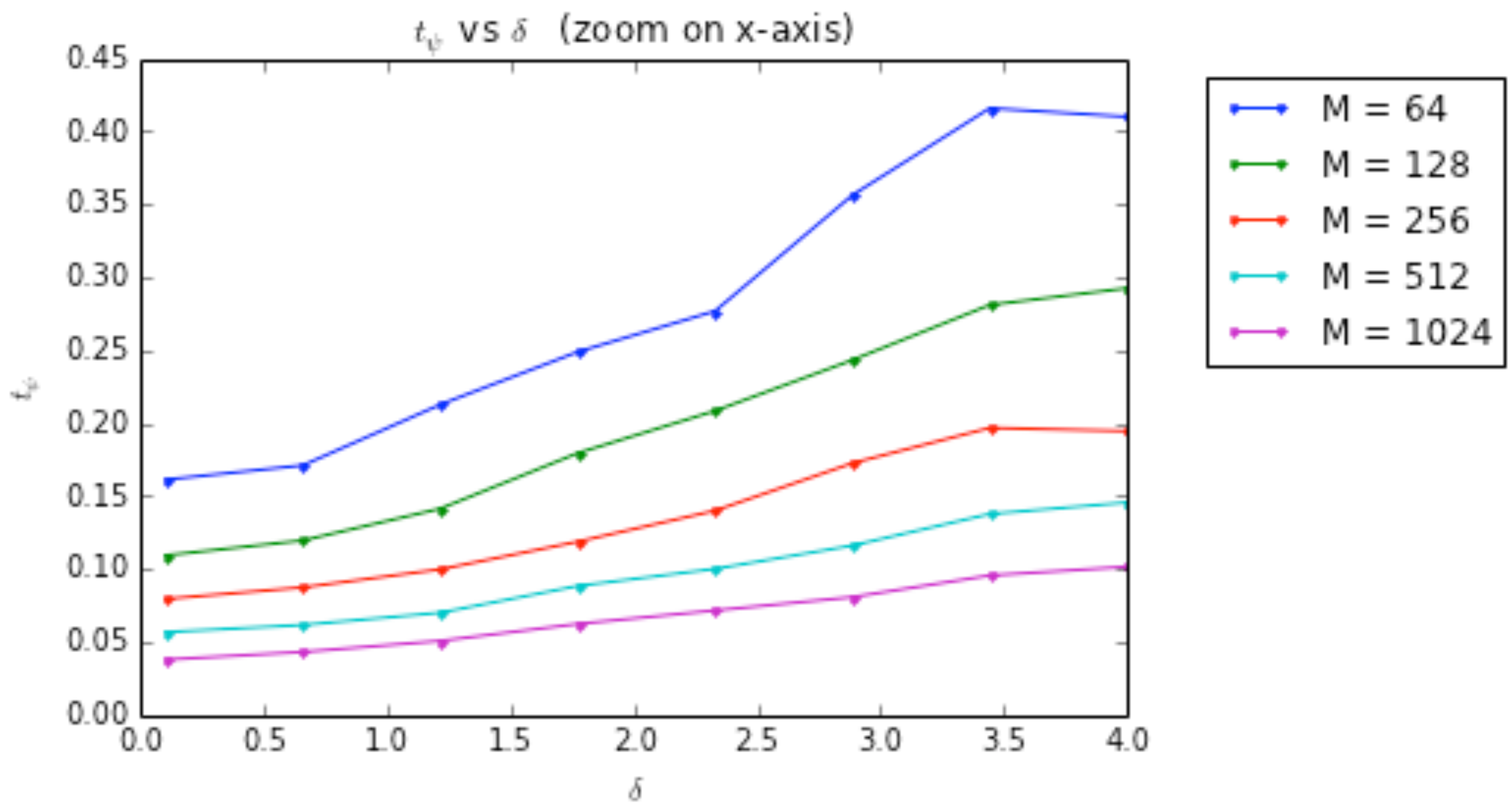
(demo available on http://tinyurl.com/quantJL)

Idea: testing $V_\psi = \frac{\sqrt{\pi}}{M\sqrt{2}} \|\mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u}) - \mathcal{Q}(\boldsymbol{\Phi x}' + \boldsymbol{u})\|_1$

- $N = 256$, $M \in \{64, 128, \cdots, 1024\}$ and $\delta \in [0.1, 4]$.
- For each (M, N, δ), 100 trials on $(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{\Phi})$ with $\|\boldsymbol{x} - \boldsymbol{x}'\| = 1$ (WLOG)

t_ψ s.t.

$\mathbb{P}[V_\psi > 1 + t_\psi] = 5\%$



Simulations

(demo available on http://tinyurl.com/quantJL)

Idea: testing $V_\psi = \frac{\sqrt{\pi}}{M\sqrt{2}} \| \mathcal{Q}(\boldsymbol{\Phi x} + \boldsymbol{u}) - \mathcal{Q}(\boldsymbol{\Phi x}' + \boldsymbol{u}) \|_1$

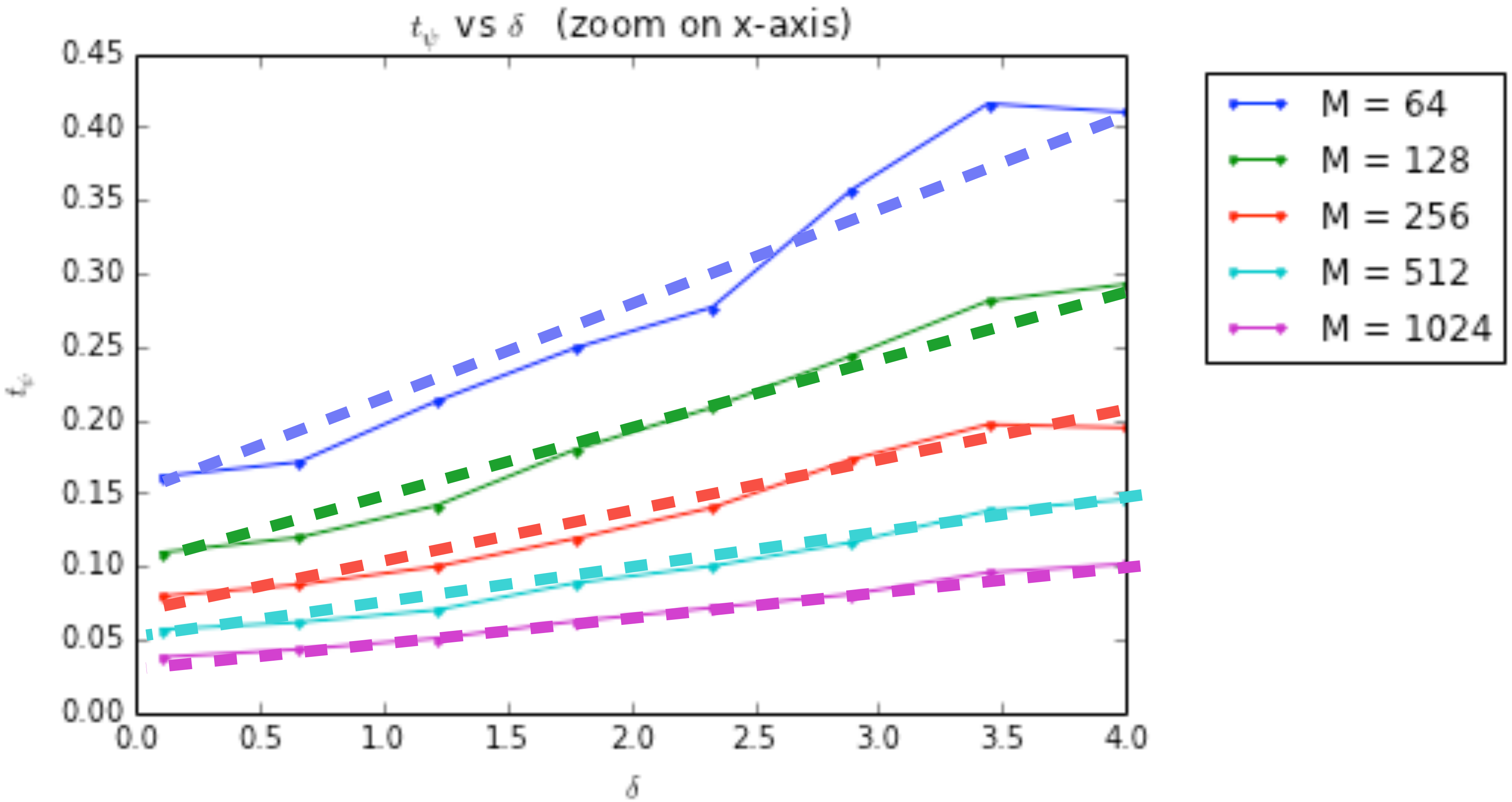
- $N = 256$, $M \in \{64, 128, \cdots, 1024\}$ and $\delta \in [0.1, 4]$.
- For each (M, N, δ), 100 trials on $(\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{\Phi})$ with $\|\boldsymbol{x} - \boldsymbol{x}'\| = 1$ (WLOG)

t_ψ s.t.

$\mathbb{P}[V_\psi > 1 + t_\psi] = 5\%$

Good match with:

$$t_\psi \simeq (a\epsilon) + (b\epsilon)\delta$$
$$\simeq (a + b\delta)\sqrt{1/M}$$



6. Conclusions

Conclusions and perspectives

- Possible to prove a Q-JL: $+$ and $\times$ distortions exist!
- Both distortions decays as $\sqrt{1/M}$
- Not shown here: almost a Q-JL with $\ell_2 \to \ell_2$

Conclusions and perspectives

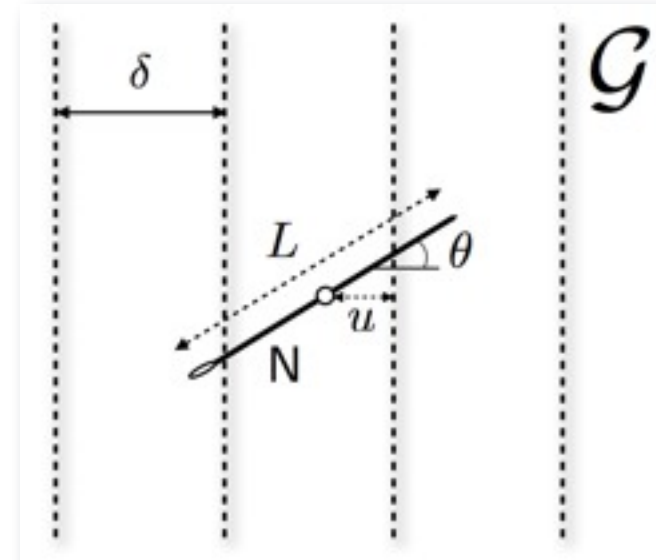
- Possible to prove a Q-JL: $+$ and $\times$ distortions exist!
- Both distortions decays as $\sqrt{1/M}$
- Not shown here: almost a Q-JL with $\ell_2 \to \ell_2$
- Future:
  - extend Q-JL to K- sparse vectors (QRIP?)
  - useful for *quantized compressed sensing* :

A K-sparse signal $\boldsymbol{x}$ is sensed by $\boldsymbol{q} = \mathcal{Q}[\boldsymbol{\Phi x}]$

How to recover $\boldsymbol{x}$?

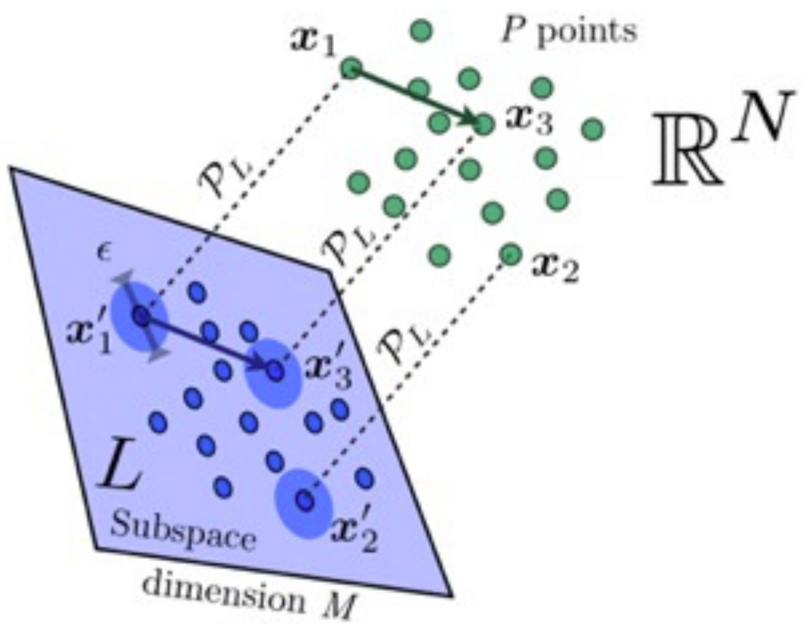
Guarantees if $\boldsymbol{x}^*$ both sparse and consistent with $\boldsymbol{q}$?

Lower bound $\|\boldsymbol{x} - \boldsymbol{x}^*\| = \Omega(K/M)$



Thank you!

Your next dinner?



A few references ...

- W.B. Johnson and J. Lindenstrauss, "*Extensions of Lipschitz mappings into a Hilbert space*," Contemporary mathematics, vol. 26, no. 189-206, pp. 1, 1984.
- S. Dasgupta and A. Gupta, "*An elementary proof of the Johnson-Lindenstrauss Lemma*," Tech. Rep. TR-99-006, Berkeley, CA, 1999.
- P. T. Boufounos, "*Universal Rate-Efficient Scalar Quantization*," Sept. 2010.
- D. Achlioptas, "*Database-friendly random projections: Johnson-Lindenstrauss with binary coins*," Journal of Computer and System Sciences, Jan. 2003.

- LJ, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, "*Robust 1-Bit Compressive Sensing via Binary Stable Embeddings of Sparse Vectors*," IEEE Transactions on Information Theory, Vol. 59(4), pp. 2082-2102, 2013.
- LJ, "A Quantized Johnson Lindenstrauss Lemma: The Finding of Buffon's Needle", http://arxiv.org/abs/1309.1507 (submitted)
- My blog: "Le petit chercheur illustré", http://yetaspblog.wordpress.com

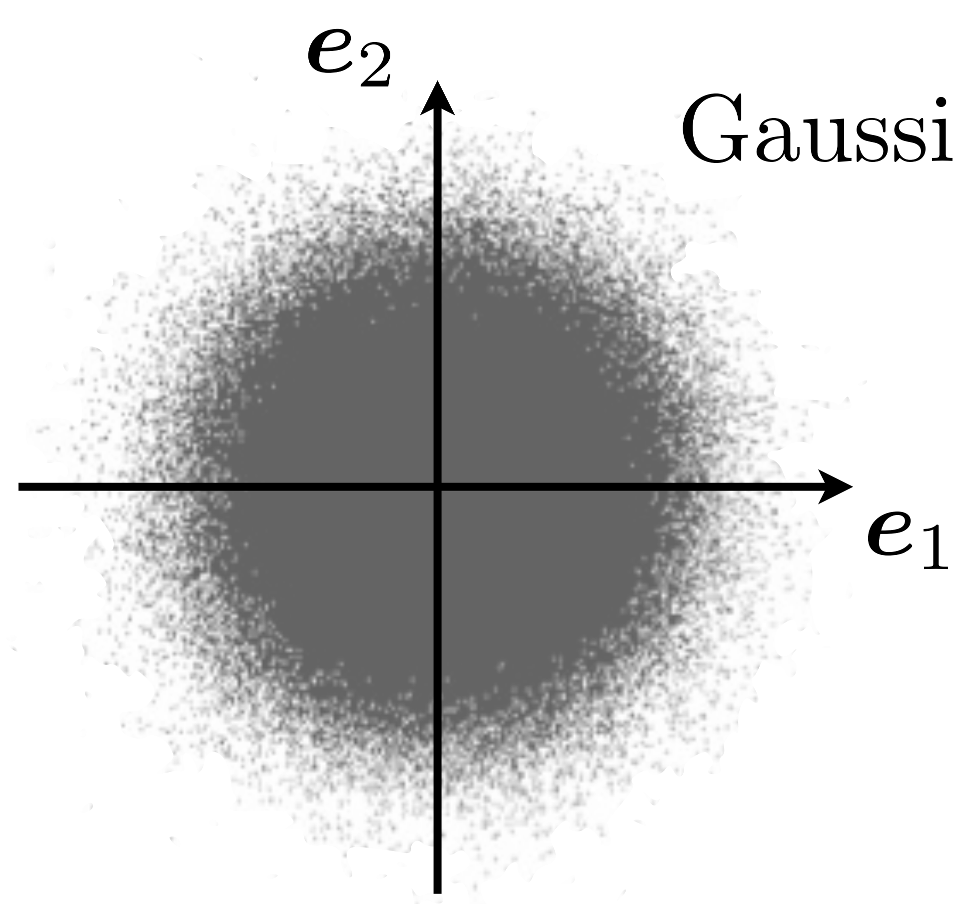
Appendix

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

<u>*proof sketch:*</u>

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)



Gaussian vector in $\mathbb{R}^N$

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

<u>*proof sketch:*</u>

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)

Let $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ with $\Phi_{ij} \sim_{\rm iid} \mathcal{N}(0, 1/M)$, then, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N$,

$$\mathbb{P}\big[\big|\|\boldsymbol{\Phi}(\boldsymbol{u}-\boldsymbol{v})\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\big| \geqslant \epsilon\|\boldsymbol{u}-\boldsymbol{v}\|^2\big] \ \leqslant\ 2e^{-M\epsilon^2/3},$$

- Union bound on $\binom{|\mathcal{S}|}{2} = O(|\mathcal{S}|^2)$ pairs in $\mathcal{S}$: $\mathbb{P}[\cup_j \mathcal{E}_j] \ \leqslant\ \sum_j \mathbb{P}[\mathcal{E}_j]$

Linear Dimensionality Reduction

- The Johnson-Lindenstrauss Lemma (1984)

proof sketch:

- Randomness helps! (Achlioptas 2003)
- and "measure concentration" (Ledoux, Talagrand, ...)

Let $\boldsymbol{\Phi} \in \mathbb{R}^{M \times N}$ with $\Phi_{ij} \sim_{\text{iid}} \mathcal{N}(0, 1/M)$, then, for $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N$,

$$\mathbb{P}\Big[\big|\|\boldsymbol{\Phi}(\boldsymbol{u}-\boldsymbol{v})\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2\big| \geqslant \epsilon\|\boldsymbol{u}-\boldsymbol{v}\|^2\Big] \;\leqslant\; 2e^{-M\epsilon^2/3},$$

- Union bound on $\binom{|\mathcal{S}|}{2} = O(|\mathcal{S}|^2)$ pairs in $\mathcal{S}$: $\mathbb{P}[\cup_j \mathcal{E}_j] \;\leqslant\; \sum_j \mathbb{P}[\mathcal{E}_j]$

$$\Rightarrow \mathbb{P}(\exists \text{ failure for one pair in } \mathcal{S}) \;\leqslant\; 2e^{2\log|\mathcal{S}| - M\epsilon^2/3} \underset{e.g.}{<} 2/3$$

if $M \geqslant M_0 = O(\epsilon^{-2}\log|\mathcal{S}|)$

Outline

1. An introduction to linear dimensionality reduction
2. Quantizing the J-L Lemma -- prologue
   - Quantization?
   - The naive way
   - What is know: binary embeddings ...
3. The finding of Buffon's needle
4. Quantizing the J-L Lemma -- epilogue
5. A few numerical tests
6. Conclusion