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

$\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $P$ points

$\mathbb{R}^N$
Given...

... is there a $L$ such that

$$\|x'_i - x'_j\| \simeq (1 \pm \epsilon)\|x_i - x_j\|$$

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


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

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

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

$$(1 - \epsilon) \|u - v\| \leq \|f(u) - f(v)\| \leq (1 + \epsilon) \|u - v\|,$$

for all $u, v \in S$. 

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

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

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

$$(1 - \epsilon) \|u - v\| \leq \|f(u) - f(v)\| \leq (1 + \epsilon) \|u - v\|,$$

for all $u, v \in S$.

$\Rightarrow$ isometry between $(S, \ell_2)$ and $(f(S), \ell_2)$
Linear Dimensionality Reduction


  *proof sketch:*
  - Randomness helps! (Achlioptas 2003)
  - and “measure concentration” (Ledoux, Talagrand, ...)

Weird things happen in high dimension!

\[
\mathbb{P} \left[ \text{vector } \in S_\epsilon \right] \rightarrow N 1 \\
\text{and exponentially!}
\]
Linear Dimensionality Reduction


**proof sketch:**

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

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

$$
\mathbb{P}[|\|\Phi(\mathbf{u} - \mathbf{v})\|^2 - \|\mathbf{u} - \mathbf{v}\|^2| \geq \epsilon\|\mathbf{u} - \mathbf{v}\|^2] \leq 2e^{-M\epsilon^2/3},
$$

Gaussian vector in $\mathbb{R}^M$
Linear Dimensionality Reduction


  **Proof sketch:**
  - Randomness helps! (Achlioptas 2003)
  - and “measure concentration” (Ledoux, Talagrand, ...)

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

$$
\mathbb{P}
\left[\left|\|\Phi(u - v)\|^2 - \|u - v\|^2\right| \geq \epsilon \|u - v\|^2\right] \leq 2e^{-M\epsilon^2/3},
$$

- Union bound on $\binom{|S|}{2} = O(|S|^2)$ pairs in $S$: $\mathbb{P}[\bigcup_j \mathcal{E}_j] \leq \sum_j \mathbb{P}[\mathcal{E}_j]$
  $\Rightarrow \mathbb{P}(\exists \text{ failure for one pair in } S) \leq 2e^{2\log |S| - M\epsilon^2/3} < 2/3$

  - $\exists f$ for JL Lemma! (boost $1 - \mathbb{P}$ asking that $\exists$ good $\Phi$ over many trials)

  if $M \geq M_0 = O(\epsilon^{-2} \log |S|)$
2. Quantizing the J-L Lemma
   -- prologue --
What is quantization?

- Generality:
  Intuitively: “Quantization maps a continuous (bounded) domain to a set of finite elements (or codebook)”

\[ \mathbb{R}^M \rightarrow \text{codebook} \]

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

- Oldest example: rounding off \[ \lfloor x \rfloor, [x], \ldots \quad \mathbb{R} \rightarrow \mathbb{Z} \]
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:

\[ Q[\lambda] = \omega_i \iff \lambda \in [t_i, t_{i+1}] \]
Scalar quantization

Principle in 1-D:

\[ Q[\lambda] = \omega_i \iff \lambda \in [t_i, t_{i+1}] \]

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

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

and \( (Q[v])_j = Q[v_j] \) for vectors.

Remark: \[ |\lambda - Q[\lambda] - \frac{1}{2} \delta| \leq \frac{1}{2} \delta \text{ for all } \lambda \]

\[ \Rightarrow \text{Quant. error} = \frac{1}{2} \delta \]
Quantizing JL (first attempt)

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

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

(important for quantizing)
Quantizing JL (first attempt)

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

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

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

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

$$(1 - \epsilon) \|u - v\| - \delta \leq \frac{1}{\sqrt{M}} \|\psi(u) - \psi(v)\| \leq (1 + \epsilon) \|u - v\| + \delta,$$

$\Rightarrow$ quasi-isometry between $(S, \ell_2)$ and $(f(S), \ell_2)$
Quantizing JL (first attempt)

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

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

Form $\psi := Q \circ f : \mathbb{R}^N \to \mathbb{Z}_M^M$

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

$$(1 - \epsilon) \|u - v\| - \delta \leq \frac{1}{\sqrt{M}} \|\psi(u) - \psi(v)\| \leq (1 + \epsilon) \|u - v\| + \delta,$$

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

\begin{align*}
&\leq |a - b| + \delta \\
&\geq |a - b| - \delta
\end{align*}

both smaller than $\delta/2$

$\Longrightarrow (1 + \epsilon)\|u - v\|$ (by JL)

Then, with 2 more lines, $\frac{1}{\sqrt{M}} \|Q(f(u)) - Q(f(v))\| \leq \frac{1}{\sqrt{M}} \|f(u) - f(v)\| + \delta$ and $\frac{1}{\sqrt{M}} \|Q(f(u)) - Q(f(v))\| \geq \frac{1}{\sqrt{M}} \|f(u) - f(v)\| - \delta$. 

$\geq (1 - \epsilon)\|u - v\|$
Quantizing JL (first attempt)

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

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

Form \( \psi := Q \circ f : \mathbb{R}^N \to \mathbb{Z}^M_\delta \)

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

\[
(1 - \epsilon) \|u - v\| - \delta \leq \frac{1}{\sqrt{M}} \|\psi(u) - \psi(v)\| \leq (1 + \epsilon) \|u - v\| + \delta,
\]

multiplicative error
additive error
Quantizing JL (first attempt)

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

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

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

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

$$(1 - \epsilon) \|u - v\| - \delta \leq \frac{1}{\sqrt{M}} \|\psi(u) - \psi(v)\| \leq (1 + \epsilon) \|u - v\| + \delta,$$

(decaying, good!) multiplicative error

(constant, weird!?) additive error

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

Can we hope better?
What’s known? Binary Quantization

Let’s define

\[ \psi(u) := \text{sign}(\Phi u) \iff \psi_j(u) = \text{sign}(\varphi_j \cdot u) \in \{\pm 1\} \]

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

\[ P[\psi_j(u) \neq \psi_j(v)] = ? \]
What's known? Binary Quantization

- Let's define

\[ \psi(u) := \text{sign} (\Phi u) \iff \psi_j(u) = \text{sign} (\varphi_j \cdot u) \in \{\pm 1\} \]

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

\[ \mathbb{P}[\psi_j(u) \neq \psi_j(v)] = \frac{1}{\pi} \angle(u, v) = \frac{\theta_{uv}}{\pi} \]
What’s known? **Binary Quantization**

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

- Let’s define

$$\psi(u) := \text{sign} (\Phi u) \iff \psi_j(u) = \text{sign} (\varphi_j \cdot u) \in \{\pm 1\}$$

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

$$\Rightarrow X_j = \frac{1}{2} |\psi_j(u) - \psi_j(v)| \sim \text{Bernoulli} (\frac{\theta_{uv}}{\pi}) \in \{0, 1\}$$
What’s known? Binary Quantization

Let’s define

\[ \psi(u) := \text{sign}(\Phi u) \iff \psi_j(u) = \text{sign}(\varphi_j \cdot u) \in \{\pm 1\} \]

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

\[ X_j = \frac{1}{2} |\psi_j(u) - \psi_j(v)| \sim \text{Bernoulli}(\frac{\theta_{uv}}{\pi}) \in \{0, 1\} \]

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

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

\[ \theta_{uv} - \epsilon \leq \frac{1}{2M} \|\psi(u) - \psi(v)\|_1 \leq \theta_{uv} + \epsilon, \]

for all \( u, v \in S \).
What’s known? Binary Quantization

Let’s define

\[ \psi(u) := \text{sign}(\Phi u) \iff \psi_j(u) = \text{sign}(\varphi_j \cdot u) \in \{\pm 1\} \]

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

\[ X_j = \frac{1}{2} |\psi_j(u) - \psi_j(v)| \sim \text{Bernoulli}(\frac{\theta_{uv}}{\pi}) \in \{0, 1\} \]

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

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

\[ \theta_{uv} - \epsilon \leq \frac{1}{2M} ||\psi(u) - \psi(v)||_1 \leq \theta_{uv} + \epsilon, \]

for all \( u, v \in S \).

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

[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.”
Buffon’s needle problem

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

with \( u \sim \mathcal{U}([0, \delta]) \) and \( \theta \sim \mathcal{U}([0, 2\pi]) \)
Fact 1: if $L < \delta$, $\mathbb{P} = \frac{2}{\pi\delta}L$

(small integral to solve)
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 $\pi$! (first "Monte Carlo" method)

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

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

Fact 2: if \( L \geq \delta \), \( \mathbb{P} \neq \frac{2}{\pi \delta} L \) but
\[
\mathbb{E}X = \frac{2}{\pi \delta} L,
\]
with \( X = \#\{ N(u, \theta) \cap \mathcal{G} \} \).

Proof: cut \( N \) in parts smaller than \( \delta \) and sum expectations!

with \( u \sim \mathcal{U}([0, \delta]) \) and \( \theta \sim \mathcal{U}([0, 2\pi]) \)
Buffon’s needle problem

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

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

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

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

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]
Buffon in $N$-D? [LJ, 2013]

Let $X = \#\{ N(u, \Omega) \cap \mathcal{G} \}$, with $\Omega \sim \mathcal{U}(SO(N))$, $u \sim \mathcal{U}([0, \delta])$.

random rotation in $\mathbb{R}^N$
Buffon in $N$-D?  [LJ, 2013]

Let $X = \#\{N(u, \Omega) \cap G\}$, 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 $\tau_N = \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi} \Gamma(\frac{N+1}{2})} \simeq N 1/\sqrt{N}$
Buffon in N-D? [LJ, 2013]

Let $X = \#\{ N(u, \Omega) \cap \mathcal{G} \}$,
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 $\tau_N = \frac{\Gamma\left(\frac{N}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{N+1}{2}\right)} \simeq_N 1/\sqrt{N}$

Moreover,

$p_k := \mathbb{P}[X = k]$ is computable!
(and all its moments: $\mathbb{E}X^q \leq c_q (L/\delta)^q$, $c_q > 0$)

We write:

$X \sim \text{Buffon}(L/\delta, N)$
with $0 \leq X \leq \lfloor L/\delta \rfloor$. 
4. Quantizing the J-L Lemma
-- epilogue --
Where’s the equivalence? (what’s the point?)

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

random Gaussian projection
\( \varphi \sim \mathcal{N}^{N \times 1}(0, 1) \)

random “dithering”
\( u \sim \mathcal{U}([0, \delta]) \)

Scalar Quantization
resolution \( \delta > 0 \)
Where's the equivalence?

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

\[ \begin{align*}
\text{Idea: in } N\text{-D} & \\
\bullet & \text{ grid } G \leftrightarrow \text{ quant. } Q \text{ with resol. } \delta \\
\bullet & \text{ needle } N \leftrightarrow \text{ segment } \overline{xx}' \\
\bullet & \text{ fixed grid } G / \text{random } N \leftrightarrow \text{ random grid } G / \text{fixed } \overline{xx}'
\end{align*} \]
For $M$ quantized projections?

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

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

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

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

**Proposition** For each $1 \leq j \leq M$ and conditionally to the knowledge of $r_j = \|\varphi_j\|$, we have

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

*Proof:* 1 page but intuition was given before.
For $M$ quantized projections?

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

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

**Proposition** For each $1 \leq j \leq M$ and conditionally to the knowledge of $r_j = \|\varphi_j\|$, we have

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

*Proof:* 1 page but intuition was given before.

Moreover,

$$\mathbb{E} X_j = \mathbb{E}_{r_j} \left( \mathbb{E}(X_j | r_j) \right) = \tau_N \mathbb{E}_{r_j} \left( \frac{r_j}{\delta} \|x - x'\| \right) = \frac{\sqrt{2}}{\sqrt{\pi}} \frac{1}{\delta} \|x - x'\|,$$

since $\mathbb{E} \|\varphi_j\| = \frac{\sqrt{2}}{\sqrt{\pi} \tau_N}!$ (Chi($N$) distr.) coincidences happen!
and finally ...

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

$$\frac{1}{M} \sum_j X_j = \frac{1}{\delta M} \|\psi(x) - \psi(x')\|_1$$
and finally ...

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

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

**Quasi-isometry!**

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

$$M \geq M_0 = O(\varepsilon^{-2} \log |S|),$$

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

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

for all $x, x' \in S$. 

(reminder: $\psi(x) = Q(\Phi x + u)$)
and finally ...

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

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

Quasi-isometry!

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

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

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

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

for all $x, x' \in S$.

decaying multiplicative and additive errors!
5. A few numerical tests
Simulations

Idea: testing $V_{ψ} = \frac{\sqrt{π}}{M \sqrt{2}} \left\| Q(Φx + u) - Q(Φx' + u) \right\|_1$

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

Expectation of $V_{ψ}$?

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

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

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

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

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

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

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

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

(demo available on http://tinyurl.com/quantJL)
6. Conclusions
Conclusions and perspectives

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

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

How to recover $\mathbf{x}$?
Guarantees if $\mathbf{x}^*$ both sparse and consistent with $\mathbf{q}$?

Lower bound $\|\mathbf{x} - \mathbf{x}^*\| = \Omega(K/M)$
Thank you!

Your next dinner?
A few references ...


Appendix
Linear Dimensionality Reduction


proof sketch:
- Randomness helps! (Achlioptas 2003)
- and “measure concentration” (Ledoux, Talagrand, ...)

Gaussian vector in $\mathbb{R}^N$
Linear Dimensionality Reduction


proof sketch:

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

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

$$
P\left[ \left| \| \Phi (u - v) \|_2^2 - \| u - v \|_2^2 \right| \geq \varepsilon \| u - v \|_2^2 \right] \leq 2e^{-M\varepsilon^2/3},$$

- Union bound on $(\binom{|S|}{2}) = O(|S|^2)$ pairs in $S$: $P[\bigcup_j E_j] \leq \sum_j P[E_j]$
Linear Dimensionality Reduction


**Proof sketch:**

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

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

\[
\mathbb{P} \left[ \left| \| \Phi(u - v) \| - \| u - v \| \right| \geq \epsilon \| u - v \| \right] \leq 2e^{-Me^2/3},
\]

- Union bound on \( \binom{|S|}{2} = O(|S|^2) \) pairs in \( S \):

\[
\mathbb{P}(\cup_j \mathcal{E}_j) \leq \sum_j \mathbb{P} \left[ \mathcal{E}_j \right]
\]

\[
\Rightarrow \mathbb{P}(\exists \text{ failure for one pair in } S) \leq 2e^{2 \log |S| - Me^2/3} < 2/3
\]

\( e.g. \) if \( M \geq M_0 = O(\epsilon^{-2} \log |S|) \)
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