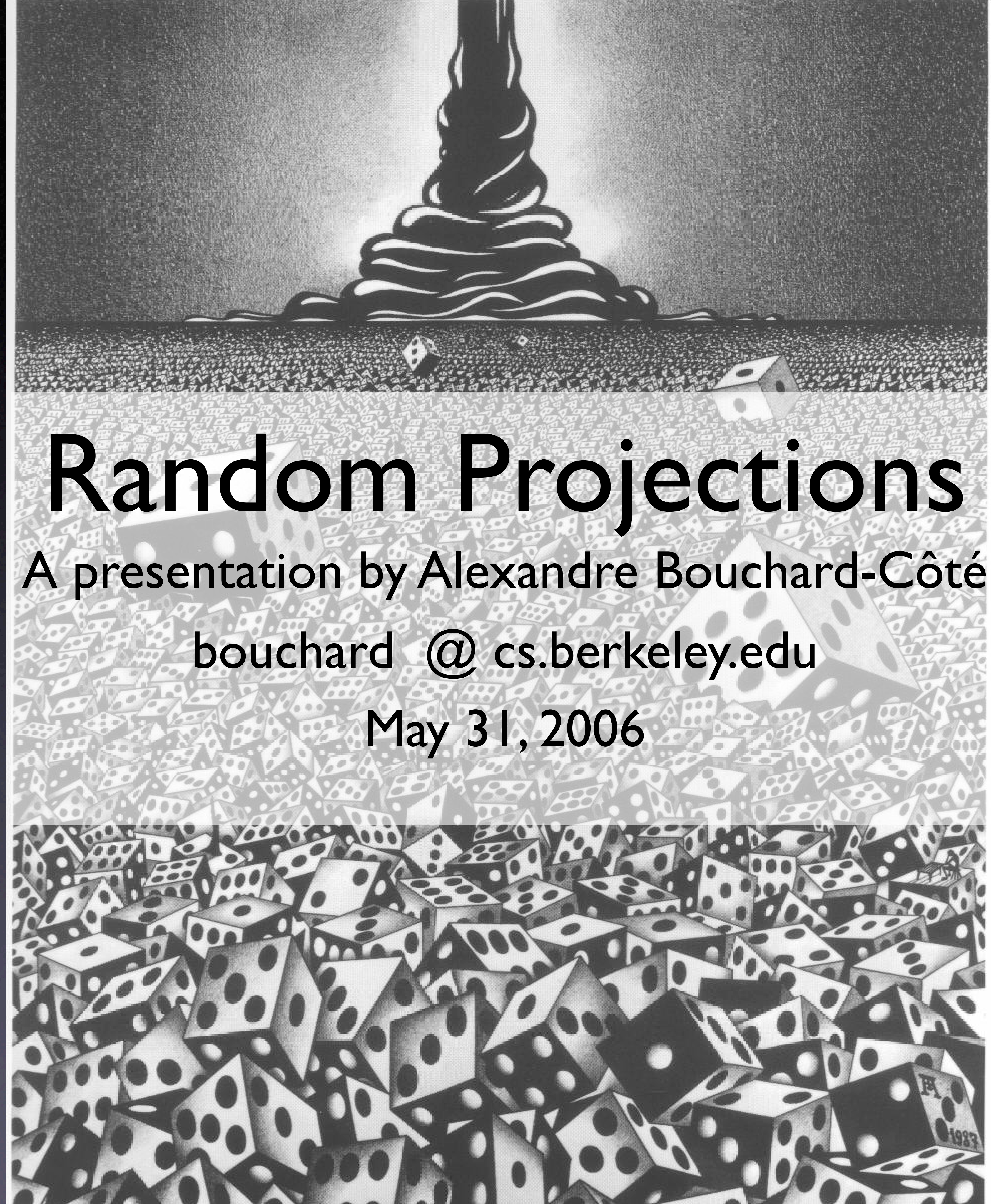


Random Projections

A presentation by Alexandre Bouchard-Côté

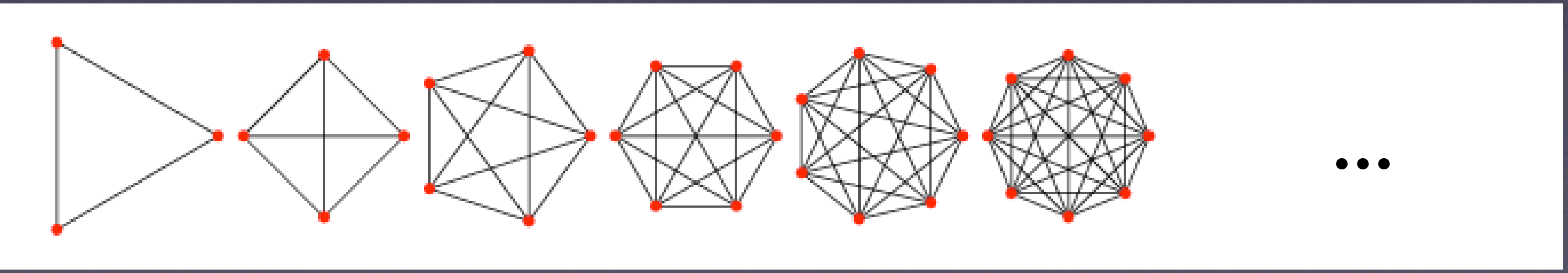
bouchard @ cs.berkeley.edu

May 31, 2006

A surprising result [1]

Any n points in Euclidean space can be embedded in $\Omega(\log n / \epsilon^2)$ dimensions without distorting the distances between any pair of points by more than a factor $(1 \pm \epsilon)$, for any $0 < \epsilon < 1$.

Example: 1 billion points on a billion dimensional simplex can be mapped to a million dimensional subspace with no more than 1% distortion on the lengths!

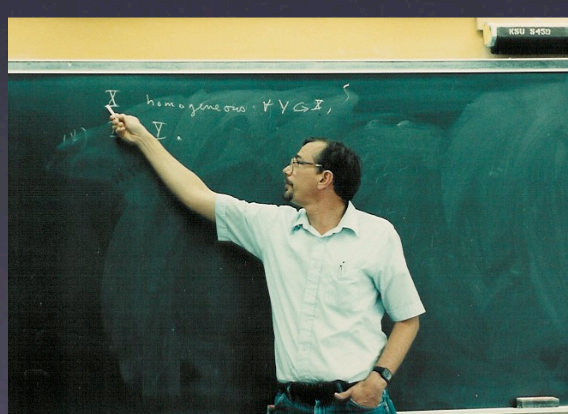


Outline

- Quick history of random projections
- Main result for today: Johnson-Lindenstrauss
- An elementary proof
- Survey of some applications

Genealogy


J. Lindenstrauss


W. Johnson


P. Frankl


H. Maehara


S. Dasgupta


A. Gupta


P. Indyk


R. Motwani

1984 [1] 1988 [2] 1998 [3] 2003 [4]

Johnson-Lindenstrauss [4]

- Fix an arbitrary $0 < \epsilon < 1$
- Consider a set V of n points in R^d
- We want to project them into a k dimensional subspace
- Assume k is large enough: $k \geq \frac{4\log(n)}{\frac{\epsilon^2}{2} - \frac{\epsilon^3}{3}}$

Johnson-Lindenstrauss

- Consequence: there is a map $f : R^d \to R^k$ such that for all pair of points $u, v \in V$:

$$(1-\epsilon)||u-v||^2 \leq ||f(u)-f(v)||^2 \leq (1+\epsilon)||u-v||^2$$

- Moreover:
  - f can be found in randomized poly-time
  - the one given by the proof is actually an orthogonal projection

Main steps of the proof [4]

1. Show that the length of a randomly projected vector is sharply concentrated about its mean
2. Apply this to one fixed pair of element in V
3. Extend to the n points in V and repeat to amplify the probability

First step

- Let $u \in R^d$, $||u|| = 1$ (fixed)
- $W \cong R^k$ (random)
- We want a concentration bound on:

$$||\pi_W(u)||^2$$

First step

- Let $u \in R^d$, $||u|| = 1$ (random)
- $W \cong R^k$ (fixed)
- We want a concentration bound on:

$$||\pi_W(u)||^2$$

First step

- Let $u \in R^d$, $||u|| = 1$ (random)
- $W = R^k$ (fixed)
- We want a concentration bound on:

$$||\pi_W(u)||^2$$

First step (con't)

- Note: if X_1 , ..., X_d are iid N(0, 1), then

$$Y = \frac{(X_1, \ldots, X_d)}{||X||} \sim Uni(S^{d-1})$$

- Let

$$L = ||\pi_{R^k}(Y)||^2$$

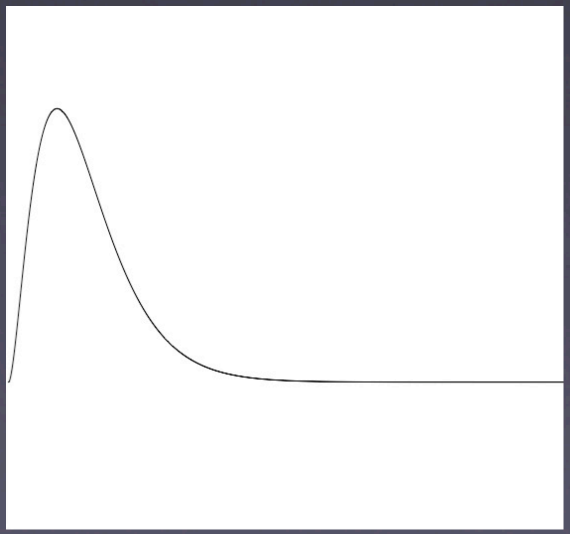
- Note: $E[L] = \frac{k}{d}$

First step (con't)

- Concentration of normal r.v. is very well studied
- Using independence and properties of the moment generating function of normal distributions, we obtain:

$$(\forall \beta > 1) \;\; P[L \geq \beta\mu] \leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right),$$

$$(\forall \beta < 1) \;\; P[L \leq \beta\mu] \leq \exp\left(\frac{k}{2}(1-\beta+\log\beta)\right).$$



Second step

- Let f be a projection to a random k-subspace
- Fix one pair u, v $\in$ V and let

$$L = ||f(u) - f(v)||^2, \mu = (k/d)||u - v||^2$$

- Applying the first step to L, we obtain

$$\begin{aligned} P[L \leq (1-\epsilon)\mu] &\leq \tfrac{1}{n^2} \\ P[L \geq (1+\epsilon)\mu] &\leq \tfrac{1}{n^2} \end{aligned}$$

Final step

- Choose the map $f' = \left(\frac{n}{k}\right)^{1/2} f$
- By the last step, for any fixed pair of points u and v,

$$P\left\{ \frac{||f'(v) - f'(u)||^2}{||u - v||^2} \notin [1-\epsilon, 1+\epsilon] \right\} \leq \frac{2}{n^2}$$

- By the union bound: the probability that the event in the LHS occurs is

$$\leq \binom{n}{2} \frac{2}{n^2} = \left(1 - \frac{1}{n}\right).$$

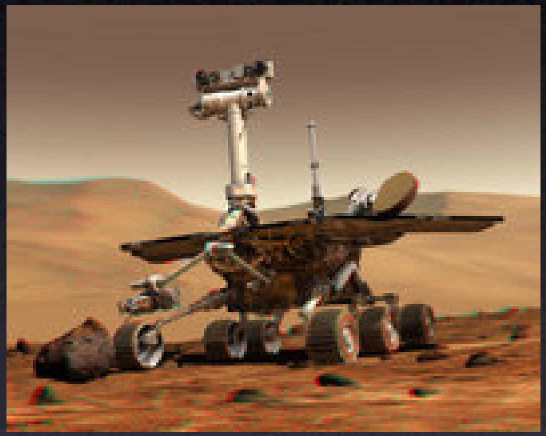
∎

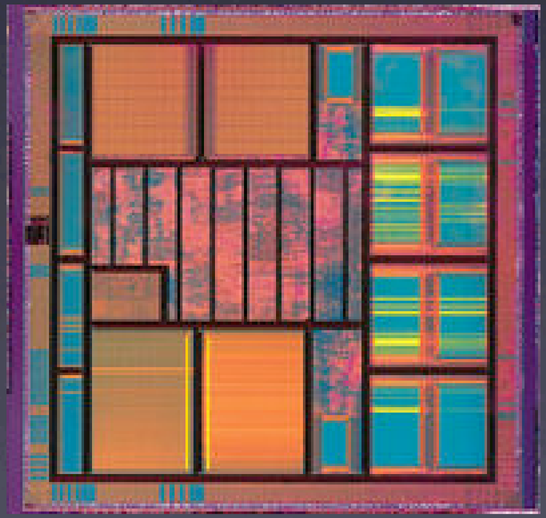
Some applications

- Fast dimensionality reduction[5]
  - Given a m by n matrix (for instance, m documents and n word types), the goal is to quickly compute a rank k approximation
  - Classically: take the top k terms in a SVD. This takes time $O(m\ n^2)$
  - Alternative: apply the random projection first to decrease the dimensionality to $l \geq k$, then, apply SVD. Slight loss of accuracy but takes time $O(m\ n \log n)$.

More applications

- Asymmetric image compression [6]
  - space probes, bio probes, ...
- Approximation algorithms for NP-hard optimization problems [5]
  - VLSI design
- many more...





References

- [1] W. Johnson and J. Lindenstrauss, Extensions of Lipschitz maps into a Hilbert space. Contemp. Math. 26, 1984, pp. 189-206
- [2] P. Frankl and H. Maehara, The Johnson-Lindenstrauss lemma and the sphericity of some graphs. J. Combin. Theory Ser. B 44(3), 1988, pp. 355-362
- [3] P. Indyk and R. Motwani, Approximate Nearest Neighbors: Towards Removing the curse of dimensionality. Proc. 30th Symposium on Theory of Computing, 1998, pp. 604-613
- [4] S. Dasgupta and A. Gupta, An elementary proof of a theorem of Johnson and Lindenstrauss. Random Structures and Algorithms, 22(1): 60-65, 2003.
- [5] S. Vempala, The Random Projection Method. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 2004.
- [6] E. Candès and J. Romberg, Practical signal recovery from random projections. 2005 http://www.acm.caltech.edu/~emmanuel/papers/PracticalRecovery.pdf, accessed May 7, 2006.

- The slides are available online: http://www.cs.berkeley.edu/~bouchard/pub/random_projection_presentation.pdf

- Acknowledgments: Guillaume Obozinski for introducing me to the random projection method and for conversations related to the subject, le Fond Québécois de la Recherche sur la Nature et les Technologies for financial support.

- The illustration on the first slide is from A.T. Fomenko, "Geometry and probability" (1987).