High-dimensional permutations and discrepancy

Nati Linial
Joint (mostly) with Zur Luria

Colloquia in Combinatorics, London, May 2016

What are high dimensional permutations?

What are high dimensional permutations?

A permutation can be encoded by means of a permutation matrix.

What are high dimensional permutations?

A permutation can be encoded by means of a permutation matrix. As we all know, this is an $n \times n$ array of zeros and ones in which every line contains exactly one 1-entry.

What are high dimensional permutations?

A permutation can be encoded by means of a permutation matrix. As we all know, this is an $n \times n$ array of zeros and ones in which every line contains exactly one 1-entry.
A line here means either a row or a column.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d+1$ factors) of zeros and ones in which every line contains exactly one 1-entry.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d+1$ factors) of zeros and ones in which every line contains exactly one 1-entry.
Whereas a matrix has two kinds of lines, namely rows and columns, now there are $d+1$ kinds of lines.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d+1$ factors) of zeros and ones in which every line contains exactly one 1-entry.
Whereas a matrix has two kinds of lines, namely rows and columns, now there are $d+1$ kinds of lines.
A line is a set of n entries in the array that are obtained by fixing d out of the $d+1$ coordinates and the letting the remaining coordinate take all values from 1 to n.

The case $d = 2$. A familiar face?

According to our definition, a 2-dimensional permutation on $[n]$ is an $[n] \times [n] \times [n]$ array of zeros and ones in which every row, every column, and every shaft contains exactly one 1-entry.

The case $d = 2$. A familiar face?

According to our definition, a 2-dimensional permutation on $[n]$ is an $[n] \times [n] \times [n]$ array of zeros and ones in which every row, every column, and every shaft contains exactly one 1-entry.
An equivalent description can be achieved by using a topographical map of this terrain.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows:

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i,j) entry in this array is k where k is the "height above the ground" of the unique 1-entry in the shaft $(i,j,*)$.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i, j) entry in this array is k where k is the "height above the ground" of the unique 1-entry in the shaft $(i, j, *)$.
It is easily verified that the defining condition is that in this array every row and every column contains every entry $n \geq i \geq 1$ exactly once.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i, j) entry in this array is k where k is the "height above the ground" of the unique 1-entry in the shaft $(i, j, *)$.
It is easily verified that the defining condition is that in this array every row and every column contains every entry $n \geq i \geq 1$ exactly once.
In other words: Two-dimensional permutations are synonymous with Latin Squares.

Where do we go from here?

We seek high-dimensional counterparts of known phenomena in ("classical" = "one-dimensional") permutations.

Where do we go from here?

We seek high-dimensional counterparts of known phenomena in ("classical" = "one-dimensional") permutations. Specifically, we wish to:

Where do we go from here?

We seek high-dimensional counterparts of known phenomena in ("classical" = "one-dimensional") permutations. Specifically, we wish to:

- Enummerate d-dimensional permutations.

Where do we go from here?

We seek high-dimensional counterparts of known phenomena in ("classical" = "one-dimensional") permutations. Specifically, we wish to:

- Enummerate d-dimensional permutations.
- Find how to generate them randomly and efficiently and describe their typical behavior.

Where do we go from here?

We seek high-dimensional counterparts of known phenomena in ("classical" = "one-dimensional") permutations. Specifically, we wish to:

- Enummerate d-dimensional permutations.
- Find how to generate them randomly and efficiently and describe their typical behavior.
- Investigate analogs of the Birkhoff von-Neumann Theorem on doubly stochastic matrices.

... and more and more and more....

- Of Erdős-Szekeres.

… and more and more and more….

- Of Erdős-Szekeres. Of the solution to Ulam's Problem.

... and more and more and more....

- Of Erdős-Szekeres. Of the solution to Ulam's Problem.
- Find out how small their discrepancy can be.

... and more and more and more....

- Of Erdős-Szekeres. Of the solution to Ulam's Problem.
- Find out how small their discrepancy can be.
- Use low-discrepancy permutations to construct high-dimensional expanders.

… and more and more and more….

- Of Erdős-Szekeres. Of the solution to Ulam's Problem.
- Find out how small their discrepancy can be.
- Use low-discrepancy permutations to construct high-dimensional expanders.
- Apply them to study multiparty communication complexity.

The count - An interesting numerology

As we all know (Stirling's formula)

$$n! = \left((1 + o(1))\frac{n}{e}\right)^n$$

The count - An interesting numerology

As we all know (Stirling's formula)

$$n! = \left((1 + o(1))\frac{n}{e}\right)^n$$

As van Lint and Wilson showed, the number of order-n Latin squares is

$$|\mathcal{L}_n| = \left((1 + o(1))\frac{n}{e^2}\right)^{n^2}$$

So, let us conecture

Conjecture

The number of d-dimensional permutations on [n] is

$$|S_n^d| = \left((1 + o(1))\frac{n}{e^d}\right)^{n^d}$$

and what we actually know

At present we can only prove the upper bound

Theorem (NL, Zur Luria '14)

The number of d-dimensional permutations on $[n]$ is

$$|S_n^d| \le \left((1+o(1))\frac{n}{e^d}\right)^{n^d}$$

How van Lint and Wilson enumerated Latin Squares

Recall that the permanent of a square matrix is a "determinant without signs".

$$per(A) = \sum_{\sigma \in S_n} \prod a_{i,\sigma(i)}$$

This is a curious and fascinating object. E.g.

This is a curious and fascinating object. E.g.

- It counts perfect matchings in bipartite graphs.

This is a curious and fascinating object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.

This is a curious and fascinating object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.
- It is $\#$-P-hard to calculate the permanent exactly, even for a $0/1$ matrix.

This is a curious and fascinating object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.
- It is $\#$-P-hard to calculate the permanent exactly, even for a $0/1$ matrix.
- On the other hand, there is an efficient approximation scheme for permanents of nonnegative matrices.

This is a curious and fascinating object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.
- It is $\#$-P-hard to calculate the permanent exactly, even for a $0/1$ matrix.
- On the other hand, there is an efficient approximation scheme for permanents of nonnegative matrices.
- The most important open problem in algebraic computational complexity is to separate permanents from determinants.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, can we, at least, bound it?

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, can we, at least, bound it? We say that A is a doubly stochastic matrix provided that

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, can we, at least, bound it? We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, can we, at least, bound it? We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.
- The sum of entries in every row is 1.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, can we, at least, bound it? We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.
- The sum of entries in every row is 1.
- The sum of entries in every column is 1.

By the marriage theorem, a doubly stochastic matrix has a positive permanent.

By the marriage theorem, a doubly stochastic matrix has a positive permanent. The set of doubly stochastic matrices is a convex polytope. The permanent is a continuous function, so: What is min $per\ A$ over $n \times n$ doubly-stochastic matrices?

By the marriage theorem, a doubly stochastic matrix has a positive permanent. The set of doubly stochastic matrices is a convex polytope. The permanent is a continuous function, so: What is min *per* A over $n \times n$ doubly-stochastic matrices? As conjectured by van der Waerden in the 20's and proved over 50 years later, in the minimizing matrix all entries are $\frac{1}{n}$.

Theorem (Falikman; Egorichev '80-81)

The permanent of every $n \times n$ doubly stochastic matrix is $\geq \frac{n!}{n^n}$.

An upper bound on permanents

The following was conjectured by Minc

Theorem (Brégman '73)

Let A be an $n \times n$ $0/1$ matrix with r_i ones in the i-th row $i = 1, \ldots, n$. Then per $A \leq \prod_i (r_i!)^{1/r_i}$. The bound is tight.

How we proved the upper bound on the number of d-dimensional permutations

Our proof can be viewed as an extension of the Minc-Brégman theorem. Specifically, we use ideas from papers of Schrijver and Radhakrishnan elaborating on Brégman's proof.

How we proved the upper bound on the number of d-dimensional permutations

Our proof can be viewed as an extension of the Minc-Brégman theorem. Specifically, we use ideas from papers of Schrijver and Radhakrishnan elaborating on Brégman's proof.
This gave us an upper bound on the number of d-dimensional permutations.

How we proved the upper bound on the number of d-dimensional permutations

Our proof can be viewed as an extension of the Minc-Brégman theorem. Specifically, we use ideas from papers of Schrijver and Radhakrishnan elaborating on Brégman's proof.
This gave us an upper bound on the number of d-dimensional permutations.
What about a matching lower bound?

How we proved the upper bound on the number of d-dimensional permutations

Our proof can be viewed as an extension of the Minc-Brégman theorem. Specifically, we use ideas from papers of Schrijver and Radhakrishnan elaborating on Brégman's proof.
This gave us an upper bound on the number of d-dimensional permutations.
What about a matching lower bound?
Here the plot thickens...

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then *per* $A > 0$.

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then *per* $A > 0$.
The analogous statement is no longer true in higher dimensions.

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then *per* $A > 0$.
The analogous statement is no longer true in higher dimensions.
Here is an example of a $4 \times 4 \times 4$ array with two zeros and two ones in every line which contains no 2-permutation.

An example

$$\begin{array}{cccc} * & 0 & * & 0 \\ 0 & 1 & 0 & 1 \\ * & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array} \qquad \begin{array}{cccc} 0 & * & 0 & * \\ 0 & 0 & 1 & 1 \\ * & * & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}$$

$$\begin{array}{cccc} 0 & 0 & * & * \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{array} \qquad \begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}$$

Don't panic

In a recent breakthrough P. Keevash solved a 160-years old problem and showed the existence of combinatorial designs.

Don't panic

In a recent breakthrough P. Keevash solved a 160-years old problem and showed the existence of combinatorial designs.
His work yields as well the tight lower bound on $|\mathcal{L}_\mathrm{n}|$.

Don't panic

In a recent breakthrough P. Keevash solved a 160-years old problem and showed the existence of combinatorial designs.

His work yields as well the tight lower bound on $|\mathcal{L}_{\rm n}|$.

It is conceivable that an appropriate adaptation of his method will prove the tight lower bound in all dimensions.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dimensional permutation) A, layer by layer.

Approximately counting Latin squares

The general scheme: We consider a Latin square ($=$ a 2-dimensional permutation) A, layer by layer. Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dimensional permutation) A, layer by layer.
Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.
Note that every layer in A is a permutation matrix.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dimensional permutation) A, layer by layer.
Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.
Note that every layer in A is a permutation matrix.
Given several layers in A, how many permutation matrices can play the role of the next layer?

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B.

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B. Therefore, there are exactly $\text{per}B$ possibilities for the next layer.

How many choices for the next layer?

To estimate the number of Latin squares we bound at each step the number of possibilities for the next layer ($=\mathrm{per}B$) from above and from below using Minc-Brégman and van der Waerden, respectively.

How many choices for the next layer?

To estimate the number of Latin squares we bound at each step the number of possibilities for the next layer ($=\text{per}B$) from above and from below using Minc-Brégman and van der Waerden, respectively.

Repeat....

Our high-dim Minc-Brégman Theorem

Definition

Let A be an $[n]^{d+1}$ array of $0/1$. Define $per_d(A)$ to be the number of distinct d-dimensional permutations that are contained in A.

Our high-dim Minc-Brégman Theorem

Definition

Let A be an $[n]^{d+1}$ array of $0/1$. Define $per_d(A)$ to be the number of distinct d-dimensional permutations that are contained in A.

Note that for $d = 1$ we recover the usual definition of the permanent.

Our high-dim Minc-Brégman Theorem

Definition

Let A be an $[n]^{d+1}$ array of $0/1$. Define $per_d(A)$ to be the number of distinct d-dimensional permutations that are contained in A.

Note that for $d = 1$ we recover the usual definition of the permanent.

Theorem

$$per_d(A) \leq \prod_{\mathbf{i}} \exp(f(d, r_{\mathbf{i}})),$$

where $r_{\mathbf{i}}$ is the number of 1's in the line $l_{\mathbf{i}}$.

Our high-dimensional Minc-Brégman Theorem (contd.)

We define $f(d,r)$ via $f(0,r) = \log r$, and

$$f(d,r) = \frac{1}{r} \sum_{k=1,\dots,r} f(d-1,k).$$

Note that $f(1,r) = \frac{\log(r!)}{r}$ and we recover Brégman's inequality.

Our high-dimensional Minc-Brégman Theorem (contd.)

We define $f(d,r)$ via $f(0,r) = \log r$, and

$$f(d,r) = \frac{1}{r} \sum_{k=1,\ldots,r} f(d-1,k).$$

Note that $f(1,r) = \frac{\log(r!)}{r}$ and we recover Brégman's inequality. In general

$$f(d,r) = \log r - d + O_d(\frac{\log^d r}{r})$$

Discrepancy and high-dimensional expansion

Quite a few brave argonauts are on a mission in quest of high-dimensional expanders.

Discrepancy and high-dimensional expansion

Quite a few brave argonauts are on a mission in quest of high-dimensional expanders. This has already generated some very impressive by body of work, but my belief is that we still do not have the right definition which should be based on the structure of random objects and the ideas of small discrepancy and sparsity.

Discrepancy and high-dimensional expansion

Quite a few brave argonauts are on a mission in quest of high-dimensional expanders. This has already generated some very impressive by body of work, but my belief is that we still do not have the right definition which should be based on the structure of random objects and the ideas of small discrepancy and sparsity.

Let's see what the study of high-dimensional permutations has to offer in this direction.

A little background - An example of discrepancy in geometry

Theorem (van Aardenne-Ehrenfest '45, Schmidt '75)

- *There is a set of N points $X \subset [0,1]^2$, s.t.* $||X \cap R| - N \cdot area(R)| \leq O(\log N)$ *for every axis-parallel rectangle $R \subseteq [0,1]^2$.*

A little background - An example of discrepancy in geometry

Theorem (van Aardenne-Ehrenfest '45, Schmidt '75)

- *There is a set of N points $X \subset [0,1]^2$, s.t.* $||X \cap R| - N \cdot area(R)| \leq O(\log N)$ *for every axis-parallel rectangle $R \subseteq [0,1]^2$.*
- *On the other hand, for every set of N points $X \subset [0,1]^2$ there is an axis-parallel rectangle R for which* $||X \cap R| - N \cdot area(R)| \geq \Omega(\log N)$.

Discrepancy in graph theory

Theorem (Alon Chung, '88 "The expander mixing lemma")

Discrepancy in graph theory

Theorem (Alon Chung, '88 "The expander mixing lemma")

Let $G = (V, E)$ be an n-vertex d-regular graph, and let λ be the largest absolute value of a nontrivial eigenvalue of G's adjacency matrix. Then for every $A, B \subset V$,

$$\left| e(A, B) - \frac{d}{n}|A||B| \right| \leq \lambda\sqrt{|A||B|}.$$

Discrepancy in high-dimensional permutations

Conjecture (NL and Zur Luria '15)

Discrepancy in high-dimensional permutations

Conjecture (NL and Zur Luria '15)

There exist order-N Latin squares such that for every $A, B, C \subseteq [N]$ there holds

$$\left| |L \cap (A \times B \times C)| - \frac{|A||B||C|}{N} \right| \leq O(\sqrt{|A||B||C|}).$$

Discrepancy in high-dimensional permutations

Conjecture (NL and Zur Luria '15)

There exist order-N Latin squares such that for every $A, B, C \subseteq [N]$ there holds

$$\left| |L \cap (A \times B \times C)| - \frac{|A||B||C|}{N} \right| \leq O(\sqrt{|A||B||C|}).$$

Moreover, this holds for almost every Latin square.

Discrepancy in high-dimensional permutations

It is interesting to restrict this conjecture to the case of empty boxes, i.e., deal with the case where $L \cap (A \times B \times C) = \emptyset$. The conjecture reads

Discrepancy in high-dimensional permutations

It is interesting to restrict this conjecture to the case of empty boxes, i.e., deal with the case where $L \cap (A \times B \times C) = \emptyset$. The conjecture reads

Conjecture

There exist order-N Latin squares in which every empty box has volume $O(N^2)$.

Discrepancy in high-dimensional permutations

It is interesting to restrict this conjecture to the case of empty boxes, i.e., deal with the case where $L \cap (A \times B \times C) = \emptyset$. The conjecture reads

Conjecture

There exist order-N Latin squares in which every empty box has volume $O(N^2)$. Moreover, this holds for almost every Latin square.

Discrepancy in high-dimensional permutations

It is interesting to restrict this conjecture to the case of empty boxes, i.e., deal with the case where $L \cap (A \times B \times C) = \emptyset$. The conjecture reads

Conjecture

There exist order-N Latin squares in which every empty box has volume $O(N^2)$. Moreover, this holds for almost every Latin square.

Note

Every Latin square has an empty box of volume $\Omega(N^2)$.

Discrepancy in high-dimensional permutations

Theorem (NL and Zur Luria)

- *There exist order-N Latin squares in which every empty box has volume $O(N^2)$.*

Discrepancy in high-dimensional permutations

Theorem (NL and Zur Luria)

- *There exist order-N Latin squares in which every empty box has volume $O(N^2)$.*
- *In almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$.*

Discrepancy in high-dimensional permutations

Can we explicitly construct such Latin squares?

Discrepancy in high-dimensional permutations

Can we explicitly construct such Latin squares? The multiplication table of a finite group is a Latin square.

Discrepancy in high-dimensional permutations

Can we explicitly construct such Latin squares? The multiplication table of a finite group is a Latin square. However,

Theorem (Kedlaya '95)

The Latin square of every order-N group contains an empty box of volume $\geq \Omega(N^{2.357\ldots})$ (this exponent is $\frac{33}{14}$).

Gowers has examples of groups where all empty boxes have volume $N^{8/3}$.

Discrepancy in high-dimensional permutations

Theorem (NL and Zur Luria)

- *There exist order-N Latin squares in which every empty box has volume $O(N^2)$.*
- *In almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$.*

A word about the proof

We construct Latin squares with no large empty boxes using Keevash's construction of Steiner systems.

A word about the proof

We construct Latin squares with no large empty boxes using Keevash's construction of Steiner systems.

- To every Steiner triple system X we associate a Latin square L where $\{i, j, k\} \in X$ implies $L(i, j, k) = \ldots = L(k, j, i) = 1$ (six terms). Also, for all i, let $L(i, i, i) = 1$.

A word about the proof

We construct Latin squares with no large empty boxes using Keevash's construction of Steiner systems.

- To every Steiner triple system X we associate a Latin square L where $\{i, j, k\} \in X$ implies $L(i, j, k) = \ldots = L(k, j, i) = 1$ (six terms). Also, for all i, let $L(i, i, i) = 1$.
- Keevash's method starts with a random greedy choice of triples. His main argument shows that whp this triple system can be completed to an STS. We show that this initial phase suffices to hit all boxes of volume above Cn^2.

A word about the proof

Why is it that in almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$?

A word about the proof

Why is it that in almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$?

Fix a box $A \times B \times C$ and note that the probability that it is empty in a random Latin square is:

A word about the proof

Why is it that in almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$?
Fix a box $A \times B \times C$ and note that the probability that it is empty in a random Latin square is:

$$\frac{\mathrm{per}_d X}{|\mathcal{L}_\mathrm{n}|}$$

where X is the $n \times n \times n$ array whose entries are zero in $A \times B \times C$ and one otherwise.

A word about the proof

Why is it that in almost every order-N Latin squares all empty boxes have volume $O(N^2 \log^2 N)$?
Fix a box $A \times B \times C$ and note that the probability that it is empty in a random Latin square is:

$$\frac{\mathrm{per}_d X}{|\mathcal{L}_\mathrm{n}|}$$

where X is the $n \times n \times n$ array whose entries are zero in $A \times B \times C$ and one otherwise. Our Brégman-type upper bound on $\mathrm{per}_d X$ yields the conclusion fairly straightforwardly.

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations:

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$.

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$. Consider the bipartite graph $G_{L,Z} = ([n],[n];E)$, where $xy \in E$ iff $L(x,y,z) = 1$ for some $z \in Z$.

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$. Consider the bipartite graph $G_{L,Z} = ([n],[n];E)$, where $xy \in E$ iff $L(x,y,z) = 1$ for some $z \in Z$.

Question

Do there exist Latin squares L such that for every $Z \subseteq [n]$ the bipartite graph $G_{L,Z}$ is as good an expander as the best $|Z|$-regular expanders?

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$. Consider the bipartite graph $G_{L,Z} = ([n],[n];E)$, where $xy \in E$ iff $L(x,y,z) = 1$ for some $z \in Z$.

Question

Do there exist Latin squares L such that for every $Z \subseteq [n]$ the bipartite graph $G_{L,Z}$ is as good an expander as the best $|Z|$-regular expanders?

... and as usual: Is this true for most LS's?

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$. Consider the bipartite graph $G_{L,Z} = ([n],[n];E)$, where $xy \in E$ iff $L(x,y,z) = 1$ for some $z \in Z$.

Question

Do there exist Latin squares L such that for every $Z \subseteq [n]$ the bipartite graph $G_{L,Z}$ is as good an expander as the best $|Z|$-regular expanders?

... and as usual: Is this true for most LS's? Can we explicitly construct such LS's?

...and one more word about sparsity

There is an even stronger conjecture on sparse HD permutations: Let L be an order-n Latin square and let $Z \subseteq [n]$. Consider the bipartite graph $G_{L,Z} = ([n],[n];E)$, where $xy \in E$ iff $L(x,y,z) = 1$ for some $z \in Z$.

Question

Do there exist Latin squares L such that for every $Z \subseteq [n]$ the bipartite graph $G_{L,Z}$ is as good an expander as the best $|Z|$-regular expanders?

... and as usual: Is this true for most LS's? Can we explicitly construct such LS's? Can we do this in all dimensions?

Erdős-Szekeres, Ulam & Co.

As every(?) undergraduate knows:

Erdős-Szekeres, Ulam & Co.

As every(?) undergraduate knows:

Theorem (Erdős-Szekeres '35)

Every permutation in S_n has a monotone subsequence of length $\geq \sqrt{n}$. The bound is tight.

Erdős-Szekeres, Ulam & Co.

As every(?) undergraduate knows:

Theorem (Erdős-Szekeres '35)

Every permutation in S_n has a monotone subsequence of length $\geq \sqrt{n}$. The bound is tight.

Several years later Ulam asked:

Problem (Ulam)

Let $L_n(\pi)$ be the length of the longest increasing subsequence in the permutation $\pi \in S_n$.

Erdős-Szekeres, Ulam & Co.

As every(?) undergraduate knows:

Theorem (Erdős-Szekeres '35)

Every permutation in S_n has a monotone subsequence of length $\geq \sqrt{n}$. The bound is tight.

Several years later Ulam asked:

Problem (Ulam)

Let $L_n(\pi)$ be the length of the longest increasing subsequence in the permutation $\pi \in S_n$. How is L_n distributed when π is drawn uniformly from S_n?

Erdős-Szekeres, Ulam & Co.

It is not hard to see that with high probability

$$c_1\sqrt{n} > L_n(\pi) > c_2\sqrt{n}$$

for some absolute $c_1 > c_2 > 0$.

Erdős-Szekeres, Ulam & Co.

It is not hard to see that with high probability

$$c_1\sqrt{n} > L_n(\pi) > c_2\sqrt{n}$$

for some absolute $c_1 > c_2 > 0$.
A lot of beautiful work was done over the years.
Here are two highlights.

Erdős-Szekeres, Ulam & Co.

It is not hard to see that with high probability

$$c_1\sqrt{n} > L_n(\pi) > c_2\sqrt{n}$$

for some absolute $c_1 > c_2 > 0$.
A lot of beautiful work was done over the years.
Here are two highlights.

Theorem (Logan and Shepp '77, Vershik and Kerov '77)

$$\lim_{n\to\infty} \frac{\mathbb{E}L_n}{\sqrt{n}} = 2.$$

Erdős-Szekeres, Ulam & Co.

Theorem (Baik, Deift, Johansson '99)

$$\frac{L_n - 2\sqrt{n}}{n^{1/6}} \xrightarrow{d} \textit{Tracy-Widom distribution.}$$

Erdős-Szekeres, Ulam & Co.

Theorem (NL and Michael Simkin)

- *Every d-dimensional length-n permutation has a monotone subsequence of length $\Omega_d(\sqrt{n})$. The bound is tight (but we still do not know the d-dependent factor).*

Erdős-Szekeres, Ulam & Co.

Theorem (NL and Michael Simkin)

- *Every d-dimensional length-n permutation has a monotone subsequence of length $\Omega_d(\sqrt{n})$. The bound is tight (but we still do not know the d-dependent factor).*
- *The longest monotone subsequence in almost every permutation has length $\Theta(n^{\frac{d}{d+1}})$.*

Erdős-Szekeres, Ulam & Co.

Theorem (NL and Michael Simkin)

- *Every d-dimensional length-n permutation has a monotone subsequence of length $\Omega_d(\sqrt{n})$. The bound is tight (but we still do not know the d-dependent factor).*
- *The longest monotone subsequence in almost every permutation has length $\Theta(n^{\frac{d}{d+1}})$. We have no further details about this distribution.*

Doubly stochastic matrices and Birkhoff-von Neumann

Recall that a real nonnegative matrix A is called doubly stochastic if all row and column sums in A equal 1.

Doubly stochastic matrices and Birkhoff-von Neumann

Recall that a real nonnegative matrix A is called doubly stochastic if all row and column sums in A equal 1. The set Ω_n of $n \times n$ doubly-stochastic matrices is clearly a polytope. The vertex set of Ω_n has a very pleasing description:

Doubly stochastic matrices and Birkhoff-von Neumann

Recall that a real nonnegative matrix A is called doubly stochastic if all row and column sums in A equal 1. The set Ω_n of $n \times n$ doubly-stochastic matrices is clearly a polytope. The vertex set of Ω_n has a very pleasing description:

Theorem (Birkhoff von-Neumann '46)

The vertex set of Ω_n is the set of all $n \times n$ permutation matrices.

Higher-dimensional Birkhoff-von Neumann?

An $n \times n \times n$ array A of nonnegative reals is called a tri-stochastic array if the sum of entries in every line in A equals 1.

Higher-dimensional Birkhoff-von Neumann?

An $n \times n \times n$ array A of nonnegative reals is called a tri-stochastic array if the sum of entries in every line in A equals 1. It is not hard to see that

Higher-dimensional Birkhoff-von Neumann?

An $n \times n \times n$ array A of nonnegative reals is called a tri-stochastic array if the sum of entries in every line in A equals 1. It is not hard to see that

- The set of all order-n tri-stochastic arrays is a polytope Π_n.

Higher-dimensional Birkhoff-von Neumann?

An $n \times n \times n$ array A of nonnegative reals is called a tri-stochastic array if the sum of entries in every line in A equals 1. It is not hard to see that

- The set of all order-n tri-stochastic arrays is a polytope Π_n.
- Every Latin square of order n is a vertex of Π_n.

Higher-dimensional Birkhoff-von Neumann?

An $n \times n \times n$ array A of nonnegative reals is called a tri-stochastic array if the sum of entries in every line in A equals 1. It is not hard to see that

- The set of all order-n tri-stochastic arrays is a polytope Π_n.
- Every Latin square of order n is a vertex of Π_n.

Is the converse also true?

No two-dimensional Birkhoff-von Neumann

Theorem (NL and Zur Luria '14)

The polytope Π_n has at least $|\mathcal{L}_{\rm n}|^{3/2}$ vertices.

No two-dimensional Birkhoff-von Neumann

Theorem (NL and Zur Luria '14)

The polytope Π_n *has at least* $|\mathcal{L}_\mathrm{n}|^{3/2}$ *vertices.*
(So Latin squares are just a negligible minority of all vertices).

No two-dimensional Birkhoff-von Neumann - some comments

- It is not hard to see that Π_n has at most $|\mathcal{L}_{\rm n}|^3$ vertices.

No two-dimensional Birkhoff-von Neumann - some comments

- It is not hard to see that Π_n has at most $|\mathcal{L}_\mathrm{n}|^3$ vertices.
- Presumably something similar holds in higher dimensions as well, but we do not know how to prove it.

No two-dimensional Birkhoff-von Neumann - some comments

- It is not hard to see that Π_n has at most $|\mathcal{L}_{\mathrm{n}}|^3$ vertices.
- Presumably something similar holds in higher dimensions as well, but we do not know how to prove it.
- The basic idea of the proof starts at the counterexample to high-dimensional van der Waerden that we saw before.

This is a vertex for tri-stochastic arrays

$$\begin{array}{cccc} * & 0 & * & 0 \\ 0 & 1 & 0 & 1 \\ * & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array} \qquad \begin{array}{cccc} 0 & * & 0 & * \\ 0 & 0 & 1 & 1 \\ * & * & 0 & 0 \\ 1 & 0 & 1 & 0 \end{array}$$

$$\begin{array}{cccc} 0 & 0 & * & * \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{array} \qquad \begin{array}{cccc} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{array}$$

Back to basics - Reproving Brégman's theorem

One of the insights gained about Brégman's theorem is that it is useful to interpret it using the notion of entropy.

Back to basics - Reproving Brégman's theorem

One of the insights gained about Brégman's theorem is that it is useful to interpret it using the notion of entropy.
So let us review the basics of this method.

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is defined as

$$H(X) := -\sum p_i \log p_i$$

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is defined as

$$H(X) := -\sum p_i \log p_i$$

In particular if the range of X has cardinality N, then $H(X) \leq \log N$ with equality iff X is distributed uniformly.

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is defined as

$$H(X) := -\sum p_i \log p_i$$

In particular if the range of X has cardinality N, then $H(X) \leq \log N$ with equality iff X is distributed uniformly.
All logarithms here are to base e. This is not the convention when it comes to entropy, but it will make things more convenient for us.

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_{y} Pr(Y = y)H(X|Y = y).$$

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y)H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy:

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y)H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy: If $X_1, \ldots, X_n$ are discrete random variables defined on the same probability space, then

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y)H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy: If $X_1, \ldots, X_n$ are discrete random variables defined on the same probability space, then

$$H(X_1, \ldots, X_n) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly per A.

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly per A. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly per A. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

$$H(X) = \log(\text{per } A).$$

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly per A. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

$$H(X) = \log(\text{per } A).$$

Therefore, an upper bound on $H(X)$ yields an upper bound on per A, which is what we want.

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row.

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row. How should we interpret the relation

$$H(X) = H(X_1) + H(X_2 | X_1) + H(X_3 | X_1, X_2) + \ldots$$

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row. How should we interpret the relation

$$H(X) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$

In particular, what can we say about $H(X_i|X_1, \ldots, X_{i-1})$?

Let us fix a specific value for X, i.e., some generalized diagonal in A.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row. Since X takes on a generalized diagonal, we know that there are r_i indices j for which the index X_j coincides with one of these r_i positions.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row. Since X takes on a generalized diagonal, we know that there are r_i indices j for which the index X_j coincides with one of these r_i positions. If such a j is smaller than i, we say that j is shading the i-th row.

Clearly X_i can take on only unshaded values.

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i | X_1, \ldots, X_{i-1}) \leq \log N_i$$

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i|X_1,\ldots,X_{i-1}) \leq \log N_i$$

since the conditioned random variable $(X_i|X_1,\ldots,X_{i-1})$ can take at most N_i values.

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i | X_1, \ldots, X_{i-1}) \leq \log N_i$$

since the conditioned random variable $(X_i | X_1, \ldots, X_{i-1})$ can take at most N_i values.
Very nice. The trouble is that we know very little about the random variable N_i.

A good trick

The way around this difficulty is not to sum the terms $H(X_i|X_1, \ldots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows.

A good trick

The way around this difficulty is not to sum the terms $H(X_i|X_1, \dots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows. We add the corresponding terms in the order σ and finally we average over the random choice of σ.

A good trick

The way around this difficulty is not to sum the terms $H(X_i|X_1, \ldots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows. We add the corresponding terms in the order σ and finally we average over the random choice of σ.
What can we say about the expectation of $\log N_i^\sigma$?

This can be restated as follows: You are expecting r visitors who arrive at random, independently chosen times.

This can be restated as follows: You are expecting r visitors who arrive at random, independently chosen times.

One of the visitors is your guest of honor and you are interested in this guest's (random) arrival rank among the r visitors.

Clearly this rank N is uniformly distributed over $1, \ldots, r$.

Clearly this rank N is uniformly distributed over $1, \ldots, r$. In particular, the expectation

$$\mathbb{E}(\log N) = \frac{1}{r} \sum_{j=1,\ldots,r} \log j = \frac{\log r!}{r} = \log(r!)^{1/r}.$$

Clearly this rank N is uniformly distributed over $1, \ldots, r$. In particular, the expectation

$$\mathbb{E}(\log N) = \frac{1}{r} \sum_{j=1,\ldots,r} \log j = \frac{\log r!}{r} = \log(r!)^{1/r}.$$

By summing over all rows, the Brégman bound is established

$$H(X) = \log(\text{per } A) \leq \sum_i \log(r_i!)^{1/r_i}.$$

Doing the d-dimensional case

If A is an $n^{d+1} = n \times n \times \ldots \times n$ array of $0/1$ we define $\mathrm{per}_d(A)$ to be the number of d-permutations that are included in A. Let's consider all lines in A in the same direction, say lines of the form $l_{\mathbf{i}} = (i_1, \ldots, i_d, *)$. Let $r_{\mathbf{i}}$ be the number of 1's in the line $l_{\mathbf{i}}$.

Many questions remain open....

- Find the asymptotic number of high-dimensional permutations.

Many questions remain open....

- Find the asymptotic number of high-dimensional permutations.
- Do Birkhoff von-Neumann in higher dimensions.

Many questions remain open....

- Find the asymptotic number of high-dimensional permutations.
- Do Birkhoff von-Neumann in higher dimensions.
- Settle the discrepancy and sparsity problems.

Many questions remain open....

- Find the asymptotic number of high-dimensional permutations.
- Do Birkhoff von-Neumann in higher dimensions.
- Settle the discrepancy and sparsity problems.
- Find explicit constructions of low-discrepancy high-dimensional permutations.

Many questions remain open....

- Find the asymptotic number of high-dimensional permutations.
- Do Birkhoff von-Neumann in higher dimensions.
- Settle the discrepancy and sparsity problems.
- Find explicit constructions of low-discrepancy high-dimensional permutations.
- Find how to sample high-dimensional permutations and determine their typical behavior.