Perfect matchings in the random bipartite geometric graph

Xavier Pérez-Giménez
University of Nebraska, Lincoln

Joint with Abigail Raz

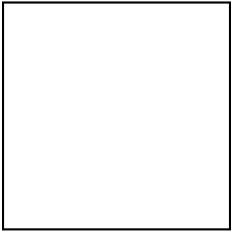
University of Groningen
April 16th, 2021

RANDOM BIPARTITE GEOMETRIC GRAPHS

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



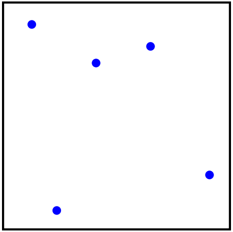
We consider the process where we add edges sequentially by their length, ending with those of length r.

RANDOM BIPARTITE GEOMETRIC GRAPHS

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



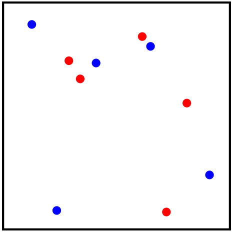
We consider the process where we add edges sequentially by their length, ending with those of length r.

RANDOM BIPARTITE GEOMETRIC GRAPHS

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



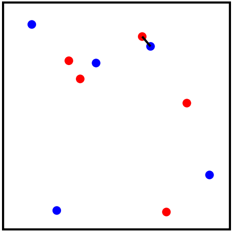
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



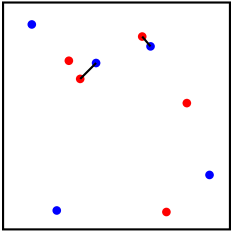
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



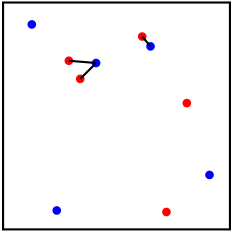
We consider the process where we add edges sequentially by their length, ending with those of length r.

RANDOM BIPARTITE GEOMETRIC GRAPHS

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



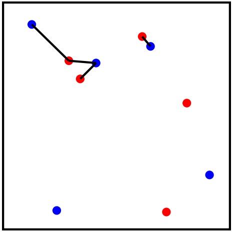
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



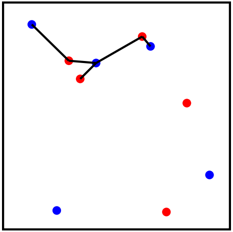
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



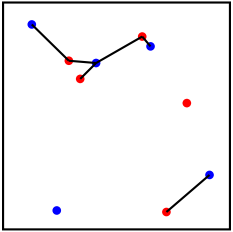
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



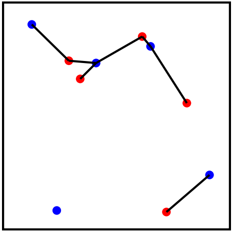
We consider the process where we add edges sequentially by their length, ending with those of length r.

Random bipartite geometric graphs

Definition

A random bipartite geometric graph $G(n, n, r)$ is defined by randomly placing n red vertices and n blue vertices on a given metric space and adding the edge (u, v), between vertices of different colors, if $d(u, v) \leq r$.

We focus our attention on the d-dimensional unit cube and torus with the euclidean metric.



We consider the process where we add edges sequentially by their length, ending with those of length r.

Natural question

Question

Does the first edge in the process that results in the minimum degree being at least one coincide, with high probability, with the first edge that creates a perfect matching?

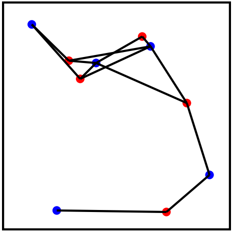
With high probability ("w.h.p.") means with probability tending to 1 as $n \to \infty$.

Natural question

Question

Does the first edge in the process that results in the minimum degree being at least one coincide, with high probability, with the first edge that creates a perfect matching?

With high probability ("w.h.p.") means with probability tending to 1 as $n \to \infty$.

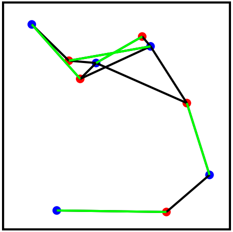


Natural question

Question

Does the first edge in the process that results in the minimum degree being at least one coincide, with high probability, with the first edge that creates a perfect matching?

With high probability ("w.h.p.") means with probability tending to 1 as $n \to \infty$.



Result for $d = 2$

For any $d \geq 2$ the threshold for having minimum degree at least 1 is $r = \Theta\left(\sqrt[d]{\frac{\log n}{n}}\right)$. However, in 1989, Leighton and Shor proved that when $d = 2$, with high probability, a perfect matching first appears when $r = \Theta\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)$.

Result for $d = 2$

For any $d \geq 2$ the threshold for having minimum degree at least 1 is $r = \Theta\left(\sqrt[d]{\frac{\log n}{n}}\right)$. However, in 1989, Leighton and Shor proved that when $d = 2$, with high probability, a perfect matching first appears when $r = \Theta\left(\frac{(\log n)^{3/2}}{\sqrt{n}}\right)$.

Shor and Yukich, 1991

When $d \geq 3$ a perfect matching first appears, w.h.p., when $r = \Theta\left(\sqrt[d]{\frac{\log n}{n}}\right)$.

Let $A^r = \{x : d(A, x) \leq r\}$. When $d = 2$ the ratio $\mathrm{Vol}(A^r)/\mathrm{Vol}(A)$ is not large enough as $\mathrm{Vol}(A)$ grows. This results in Hall's condition failing for "large" subsets of vertices.

Result

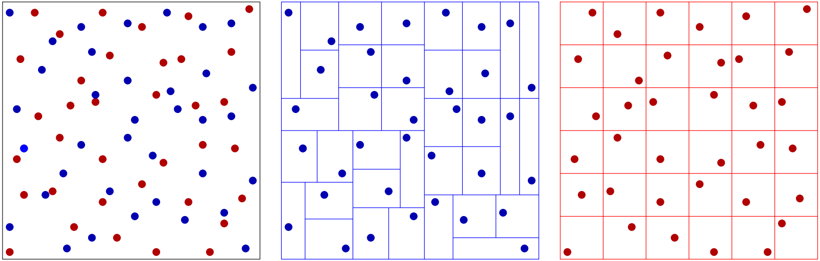
Theorem

For the d-dimensional cube and d-dimensional torus, with $d \geq 3$, the first edge in the process that results in the minimum degree being at least one coincides, with high probability, with the first edge that creates a perfect matching.

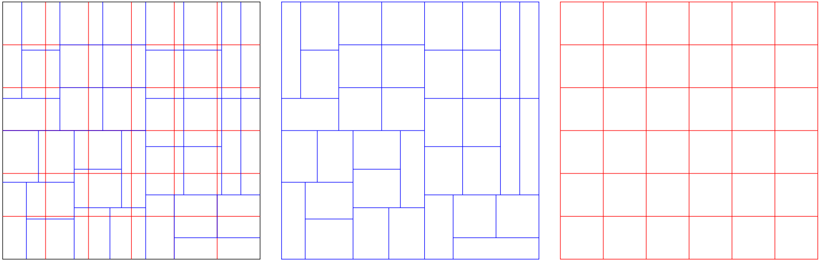
Note: The result also holds if you replace the random red points by lattice points (or any fixed set of points "evenly distributed" throughout the cube or torus.)

Sharp threshold at $r \sim \sqrt[d]{\frac{\log n}{\alpha_d n}}$
(expected degree $D = \alpha_d r^d n \sim \log n$)

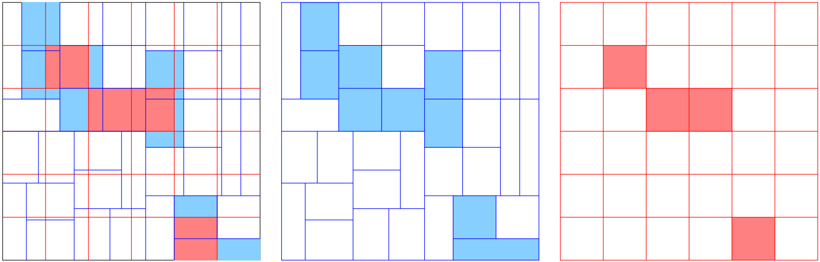
A SIMPLE POWERFUL IDEA



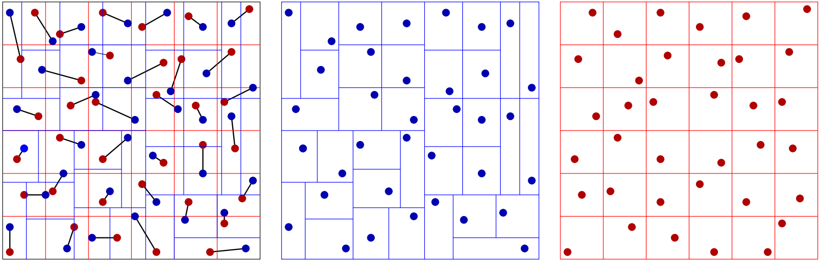
A simple powerful idea



A SIMPLE POWERFUL IDEA

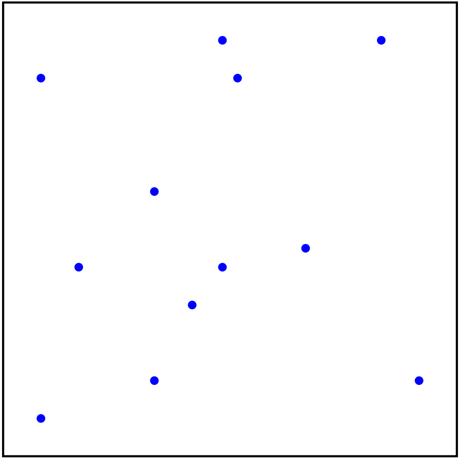


A SIMPLE POWERFUL IDEA



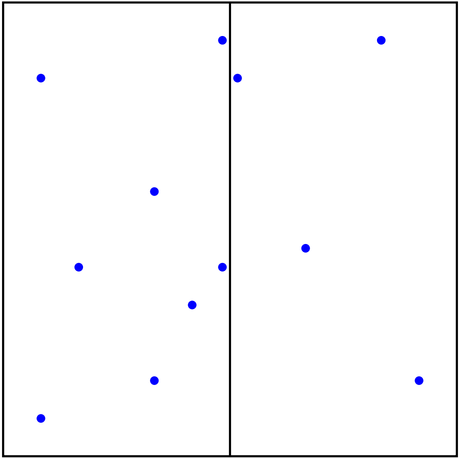
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



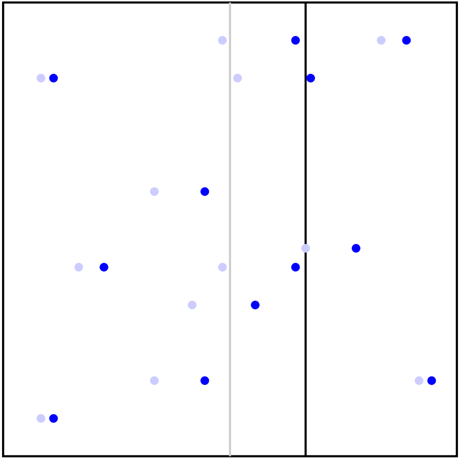
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



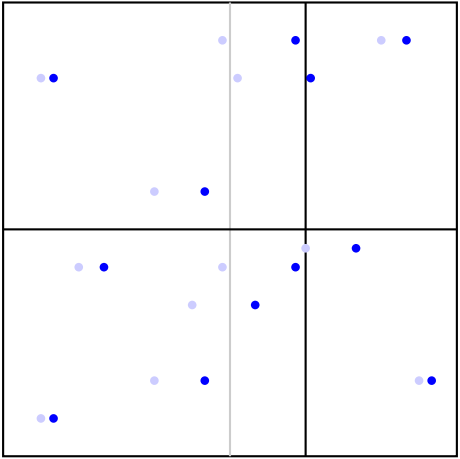
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



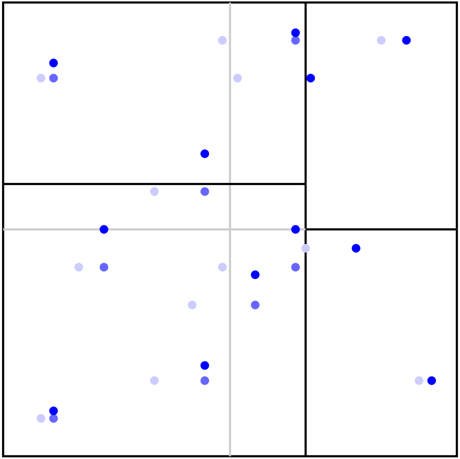
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



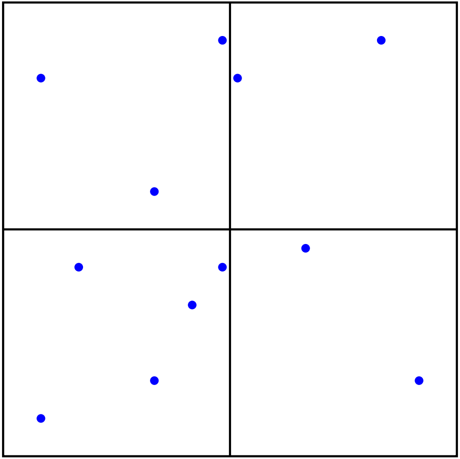
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



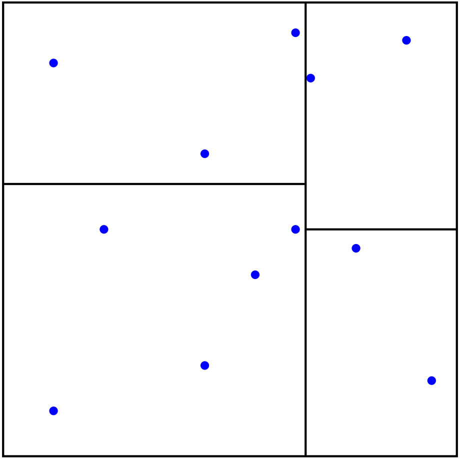
Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



Initial transformation of the blue vertices

To begin we recursively subdivide our cube following an algorithm first used by Ajtai, Komlós, and Tusnády in 1984 and then by Shor and Yukich in 1991.



Let r be slightly below the threshold for isolated vertices, $r \sim \sqrt[d]{\frac{\log n}{\alpha_d n}}$, the expected degree is $\log n - \omega(1)$.

Proposition

Let $L = L(\epsilon)$ be large constant. After carrying out this process for $k = \log_2(L^{-1}n/\log n)$ steps the cube is partitioned into 2^k boxes, and the following hold with high probability:

1. the number of blue vertices in each box is $(1 \pm \epsilon)2^{-k}n$, and thus the volume of each box is $(1 \pm \epsilon)2^{-k}$;
2. every blue vertex has been shifted by, for example, at most $4^{d+3}\epsilon r$;
3. the aspect ratio of every box is, for example, at most $1 + 10\epsilon$.

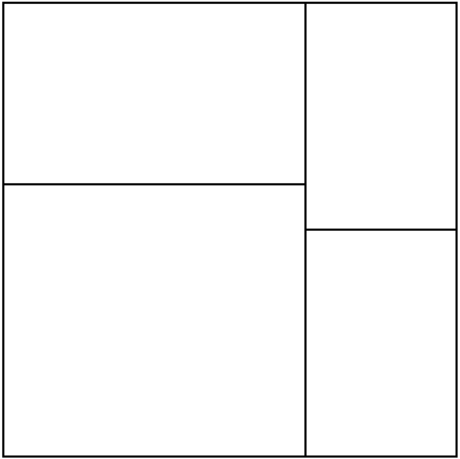
At this point every box has side lengths $(1 \pm O(\epsilon))\sqrt[d]{L \log n/n}$ and $(1 \pm \epsilon)L \log n$ blue vertices.

Deviation from the proof of Shor and Yukich

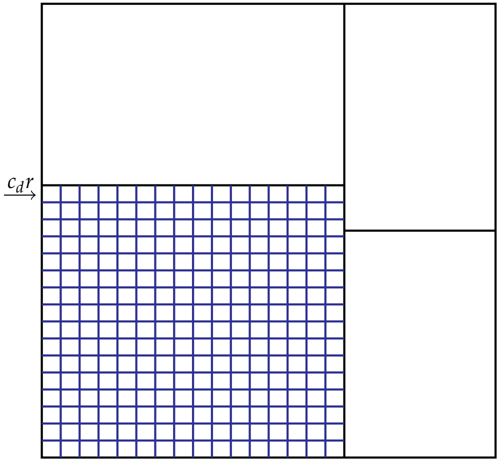
If you continue this process until each box contains exactly one blue vertex (w.h.p. $O(\log n)$ more steps suffice) then you can easily verify Hall's condition using edges of length $O\left(\sqrt[d]{\frac{\log n}{n}}\right)$.

However, as we now require additional control over the edge lengths we must use more care in verifying Hall's condition.

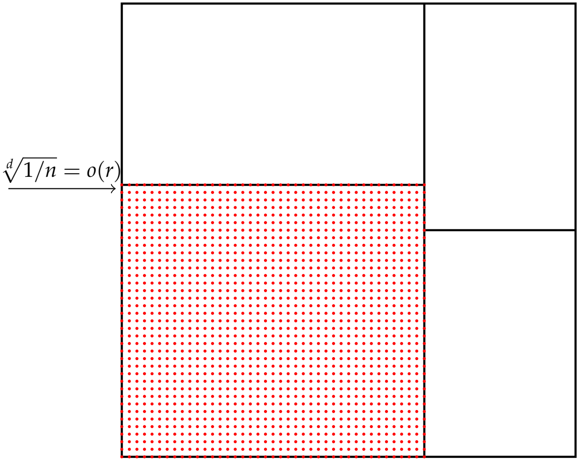
BOXES & CELLS



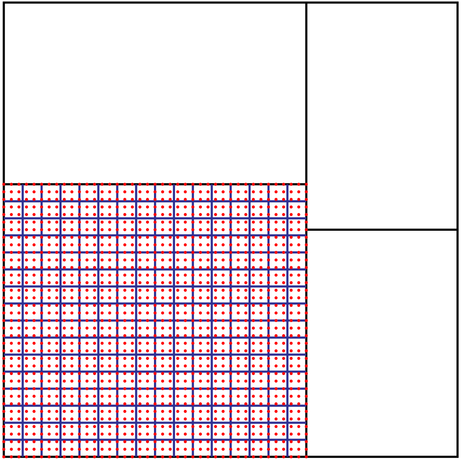
Boxes & cells



Boxes & cells

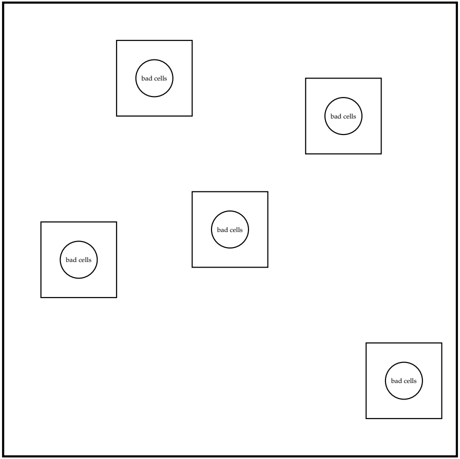


Boxes & cells



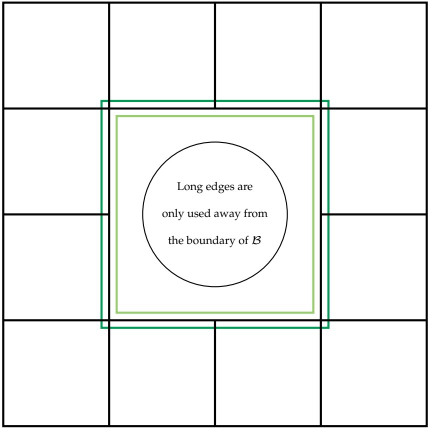
If the number of blue vertices in each cell is proportional to the volume of the cell then we can match the blue and red vertices using edges of length at most the diameter of the cell (plus $\sqrt[d]{1/n} = o(r)$).

We can isolate the “bad” cells (the cells whose number of blue vertices is not proportional to the volume of a cell).



We can form a matching for all of the blue vertices outside of these bad areas using “short” edges.

What happens to the bad areas?



We use the Brunn-Minkowski inequality to verify Hall's condition and form local matchings around the bad cells.

Proof Idea

A few notes on the proof:

- In the case where both the red and blue vertices are random we run the initial transformation separately for the red and blue vertices. We form the local matchings for boxes containing bad blue cells, bad red cells, or both.

- In the case of the d-dimensional cube additional care must be used close to the faces; however, the techniques remain the same. This is unsurprising as in the cube the last isolated vertices are close to faces (specifically 1 or 2 dimensional faces).

Application

Definition

The Prokhorov metric $\rho(\mu, \nu)$ is a metric on the space of probability measures on $[0, 1]^d$

$$\rho(\mu, \nu) := \inf\{\epsilon > 0 : \mu(A) \leq \nu(A^\epsilon) + \epsilon \text{ for all Borel sets } A\}.$$

Definition

Let λ be the Lebesgue measure on $[0, 1]^d$. Define λ_n, the nth empirical measure, by letting $X_1, \ldots X_n$ be i.i.d random variables drawn according to λ. Then if A is a measurable subset of $[0, 1]^d$ let

$$\lambda_n(A) = \frac{1}{n} \sum_{i=1}^{n} I_A(X_i),$$

where I_A is the indicator function for A.

Application

In 1969, Dudley began investigating the order of convergence for $\rho(\lambda_n, \lambda)$.

Theorem (Massart 1988)

There exists some c_d and C_d such that w.h.p.

$$c_d \leq \liminf_{n \to \infty} \sqrt[d]{\frac{n}{\log n}} \rho(\lambda_n, \lambda)$$

and

$$\limsup_{n \to \infty} \sqrt[d]{\frac{n}{\log^2 n}} \rho(\lambda_n, \lambda) \leq C_d.$$

Application

Shor and Yukich noted in their 1991 paper that any upper bound on the minimum r required for a perfect matching directly translates to an upper bound on $\rho(\lambda_n, \lambda)$.

Thus they proved

Corollary (Shor and Yukich 1991)

For $d \geq 3$ there exists some C_d such that w.h.p.

$$\limsup_{n\to\infty} \sqrt[d]{\frac{n}{\log n}}\rho(\lambda_n, \lambda) \leq C_d.$$

Our result provides an explicit constant in the upper bound. Unfortunately, it is not close to Massart's explicit lower bound; specifically, our upper bound grows with d while Massart's lower bound shrinks.

THANK YOU!

Questions?