Super-logarithmic clique numbers in dense inhomogeneous random graphs

Gwen McKinley

May 10, 2019

Massachusetts Institute of Technology

Outline

- Background & definitions
- Previous work
- My results
- Proof methods
- Future directions

Background & Definitions

Background & Definitions

- The **Erdős-Rényi random graph,** $\mathbb{G}(\boldsymbol{n}, \boldsymbol{p})$, has n vertices, and each pair of vertices forms an edge independently with probability p.
- The **clique number**, $\boldsymbol{\omega}(\boldsymbol{G})$, of a graph G is the number of vertices in the largest complete subgraph of G.

Theorem

$$\omega(\mathbb{G}(n,p)) = (1+o(1))\frac{2}{\log(1/p)}\log(n)$$

with probability $1 - o(1)$, as $n \to \infty$.

Proof idea:

- **Upper bound** on $\omega(\mathbb{G}(n,p))$: find k for which the expected number of k-cliques in $\mathbb{G}(n,p)$ is asymptotically zero. (first moment method)
- **Lower bound** on $\omega(\mathbb{G}(n,p))$: for slightly smaller k, show that the number of k-cliques is highly concentrated around its expectation. (second moment method)

Background & Definitions

- More recently, interest in **inhomogeneous random graphs**.
- Edge probabilities not equal to a constant p and/or not independent.

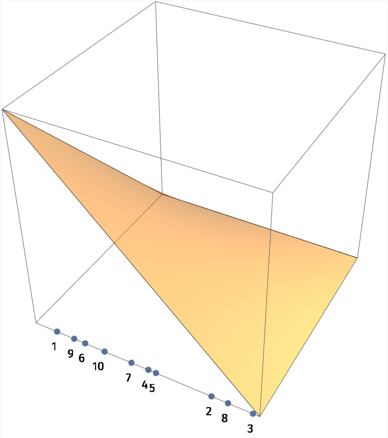
Background & Definitions

- A (dense) **graphon** is a symmetric, measurable function $W : [0,1]^2 \to [0,1]$.
- Can generate an inhomogeneous random graph from W by uniformly sampling n numbers $x_1, \ldots, x_n \in [0,1]$ and making each (i,j) an edge with probability $W(x_i, x_j)$.
- Call this a **W-random graph**, and write $\mathbb{G}(\boldsymbol{n}, \boldsymbol{W})$.

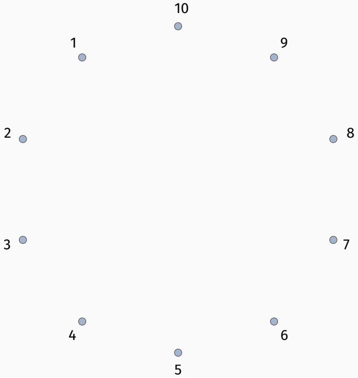
Background & Definitions

- Natural extension of the Erdős-Rényi random graph:
- Notice that $\mathbb{G}(n, p) = \mathbb{G}(n, W)$ for the constant graphon $W = p$.

W-random graph example

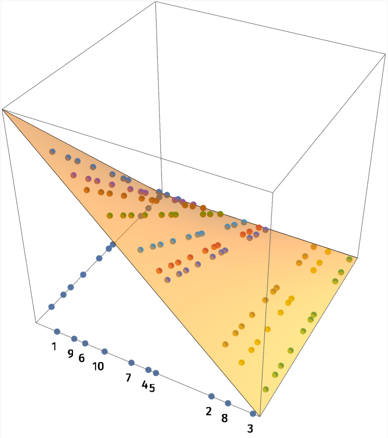


$$W(x, y) = (1 - x)(1 - y)$$

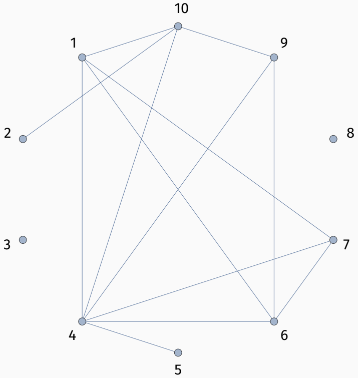


$$\mathbb{G}(n, W)$$

W-random graph example

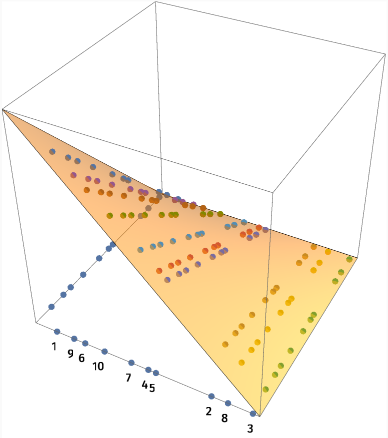


$W(x, y) = (1 - x)(1 - y)$

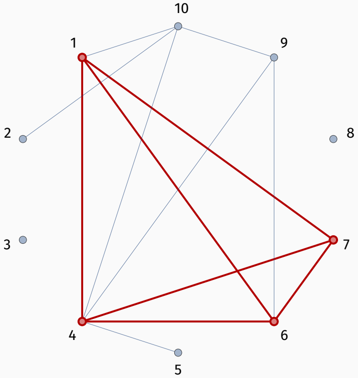


$\mathbb{G}(n, W)$

W-random graph example



$W(x, y) = (1 - x)(1 - y)$



$\mathbb{G}(n, W)$

Question: what can we say about clique numbers of W-random graphs?

Previous Work

Previous Results

Theorem (Doležal, Hladký, and Máthé, 2017)

For a graphon $W\colon \Omega^2 \to [0,1]$ that is essentially bounded away from 0 and 1,

$$\omega(\mathbb{G}(n,W)) = \boldsymbol{(1+o(1))\kappa(W)\log n},$$

with probability $1 - o(1)$ as $n \to \infty$, where

$$\kappa(W) = \sup\left\{\frac{2\|h\|_1^2}{\int_{(x,y)\in\Omega^2} h(x)h(y)\log(1/W(x,y))\,d(\nu^2)} \;:\; \text{h} \geq 0 \text{ an } L^1\text{-function on } \Omega\right\}.$$

"**essentially bounded**" = bound holds everywhere except perhaps on some set of measure zero.

Previous Results

Theorem (Doležal, Hladký, and Máthé, 2017)

For a graphon $W\colon \Omega^2 \to [0,1]$ that is essentially bounded away from 0 and 1,

$$\omega(\mathbb{G}(n,W)) = \boldsymbol{(1+o(1))\kappa(W)\log n}.$$

- Most important part is the characterization of $\boldsymbol{\kappa(W)}$:
- If W is bounded between $p_1 > 0$ and $p_2 < 1$, then can couple $\mathbb{G}(n,W)$ with $\mathbb{G}(n,p_1)$ and $\mathbb{G}(n,p_2)$ so that

$$\omega(\mathbb{G}(n,p_1)) \le \omega(\mathbb{G}(n,W)) \le \omega(\mathbb{G}(n,p_2)).$$

- Both these bounds are $\Theta(\log n)$, so $\omega(\mathbb{G}(n,W)) = \boldsymbol{\Theta(\log n)}$.

Previous Results

Clique numbers for other inhomogeneous random graphs:

- Graphs with a **power-law** degree distribution (Janson, Łuczak, and Norros, 2010)
- **Hyperbolic** random graphs (Bläsius, Friedrich, and Krohmer, 2018)
- **Rank-1 inhomogeneous** random graphs (Bogerd, Castro, and van der Hofstad, 2018); explicitly computed clique number when vertex weights bounded away from 1, showed 2-point concentration.

Graphons that approach 1

Returning to graphons,

- **Question:** what happens if W is not bounded away from 1?
- **Easy case:** if $W = 1$ on $S \times S$ for some set S of positive measure, then $\omega(\mathbb{G}(n, W))$ is linear a.a.s.
- **General case:** too hard! (or too weird...)

Graphons that approach 1

Example (Doležal, Hladký, and Máthé, 2017)

There exists a graphon W and a sequence of integers $n_1 < n_2 < \cdots$ such that $\omega(\mathbb{G}(n_i, W))$ alternates between at most $\log\log n_i$ and at least $\frac{n_i}{\log\log n_i}$ on elements of the sequence asymptotically almost surely.

- In fact, can replace $\log\log n$ with any $\omega(1)$ function.
- Shown for a highly discontinuous graphon $W : [0,1]^2 \to [0,1]$.
- **Question:** Even if W is not bounded away from 1, can we find a good characterization of $\omega(\mathbb{G}(n, W))$ as long as W is reasonably "well-behaved"?

Graphons that approach 1

- **Question:** Even if W is not bounded away from 1, can we find a good characterization of $\omega(\mathbb{G}(n, W))$ as long as W is reasonably "well-behaved"?
- **Note:** we only care about points *on the diagonal* (i.e. with $x = y$) where W approaches 1.

Lemma (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon whose essential supremum is strictly less than 1 in some neighborhood of each point (x, x) for $x \in [0,1]$. Then $\omega(\mathbb{G}(n, W)) = \boldsymbol{O(\log n)}$ asymptotically almost surely.

Graphons that approach 1

- **Updated question:** If W approaches 1 near at least one point (a, a), and W is sufficiently "well-behaved", can we find a good characterization of $\omega(\mathbb{G}(n, W))$?
- **Depends...**
  - How many points (a, a)?
  - How fast does W approach 1?
  - How well-behaved?

New Results

Moderate rate of approach

- Know that if W never approaches 1, then $\omega(\mathbb{G}(n, W)) = \boldsymbol{O(\log n)}$.
- **New fact:** If W approaches 1 "moderately fast" at a finite number of points, then $\omega(\mathbb{G}(n, W)) = \boldsymbol{\Theta(\sqrt{n})}$.

Theorem (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some collection of points $(a_1, a_1), \ldots, (a_k, a_k)$, and essentially bounded away from 1 in some neighborhood of (x, x) for each other $x \in [0,1]$. If all directional derivatives of W exist at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and are uniformly bounded away from 0 and $-\infty$, then $\boldsymbol{\omega(\mathbb{G}(n, W)) = \Theta(\sqrt{n})}$ asymptotically almost surely.

Moderate rate of approach

Even if W doesn't have directional derivatives, can sometimes obtain a similar bound:

Lemma (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some point (a,a). If W is locally Lipschitz continuous at (a,a), then $\omega(\mathbb{G}(n,W)) = \boldsymbol{\Omega(\sqrt{n})}$ asymptotically almost surely.

Immoderate rate of approach

So what *doesn't* have a clique number $\Theta(\sqrt{n})$?

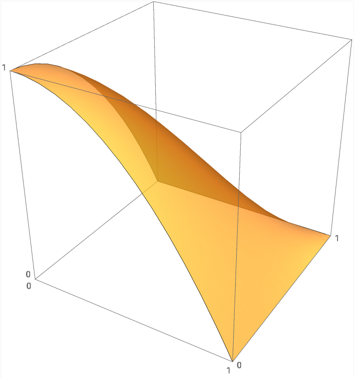
Example (M., 2019)

For any constant $r > 0$, define the graphon

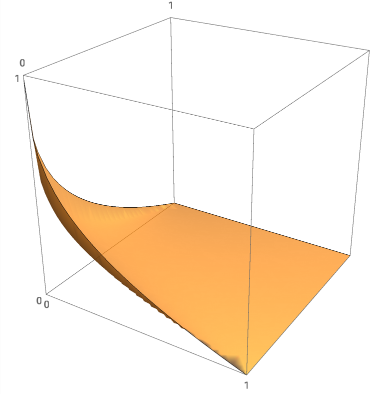
$$U_r(x, y) := (1 - x^r)(1 - y^r).$$

The random graph $\mathbb{G}(n, U_r)$ asymptotically almost surely has clique number $\boldsymbol{\Theta(n^{\frac{r}{r+1}})}$.

W-random graph example



U_2, clique number $= \Theta(n^{2/3})$



$U_{\frac{1}{2}}$, clique number $= \Theta(n^{1/3})$

Immoderate rate of approach

Example (M., 2019)

For any constant $r > 0$, define the graphon

$$U_r(x, y) := (1 - x^r)(1 - y^r).$$

The random graph $\mathbb{G}(n, U_r)$ asymptotically almost surely has clique number $\boldsymbol{\Theta(n^{\frac{r}{r+1}})}$.

- If $r > 1$, the directional derivatives are 0, and $\frac{r}{r+1} > \frac{1}{2}$.
- If $r < 1$, the directional derivatives are $-\infty$, and $\frac{r}{r+1} < \frac{1}{2}$.

Immoderate rate of approach

A bit more general:

Lemma (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at some point (a,a). If W is locally α-Hölder continuous at (a,a) for some constant α, then $\omega(\mathbb{G}(n,W)) = \boldsymbol{\Omega(n^{\frac{\alpha}{\alpha+1}})}$ asymptotically almost surely.

Immoderate rate of approach

Can also show

- If W is equal to 1 at some point (a, a), and all directional derivatives of W at (a, a) exist and are equal to **zero**, then $\omega(\mathbb{G}(n, W)) = \boldsymbol{\omega(\sqrt{n})}$ a.a.s.
- If W is equal to 1 at the points $(a_1, a_1), \ldots, (a_k, a_k)$, and bounded away from 1 near all other (x, x), and all directional derivatives of W at $(a_1, a_1), \ldots, (a_k, a_k)$ are equal to $-\infty$, then $\omega(\mathbb{G}(n, W)) = \boldsymbol{o(\sqrt{n})}$ a.a.s.

Methods

Methods

Before we get to some proofs, one thing to notice...

Wait...what happened to the constants?

- In all the results above, clique number is given up to a constant factor.
- But for Erdős-Rényi random graphs, and the result by Doležal, Hladký, and Máthé, a correct constant is given.
- Does a correct constant exist for the examples considered here? Yes!

Theorem (Doležal, Hladký, and Máthé, 2017)

For any graphon W, with probability $1 - o(1)$,

$$\omega(\mathbb{G}(n, W)) = (1 + o(1)) \cdot \mathbb{E}[\omega(\mathbb{G}(n, W))].$$

Wait...what happened to the constants?

- Correct constant exists - why haven't we found it?
- Recall method for finding clique number for Erdős-Rényi random graphs:
  - **Upper bound:** find k for which the expected number of k-cliques in $\mathbb{G}(n, p)$ is asymptotically zero. (first moment method)
  - **Lower bound:** show that the number of k-cliques has low variance $\Rightarrow$ close to its expectation with high probability. (second moment calculation)
- **Problem here:** the number of k-cliques has *high* variance in most of the examples above (for k in relevant range).

Wait...what happened to the constants?

- **Problem here:** the number of k-cliques has *high* variance in most of the examples above (for k in relevant range).
- A few ways possible around this:
  - Doležal, Hladký, and Máthé applied the second moment method to a carefully chosen restriction of the original graphon.
  - Could try using tools from large deviations theory (as in Achlioptas, Peres, 2003).
- *However*, to get a lower bound tight up to a constant, there is a simpler way!

Proof sketch

Keep that in mind - we'll come back to it in a moment! For now, look at the overall proof strategy for some of the results above.

Proof sketch

Overall approach to prove results above:

- Compute the clique numbers associated to some family of *specific* graphons with the desired local behavior.
- **Fact:** the clique number of $\mathbb{G}(n, W)$ is primarily determined by local behavior of W near points with $W(x,x) = 1$.
- Use elements of specific family, together with fact above, to determine the clique number associated to *any* graphon with comparable local behavior.

Proof sketch - moderate rate of approach

Use this framework to sketch a proof of the following:

Theorem (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at the point $(0,0)$, and essentially bounded away from 1 in some neighborhood of (x,x) for each other $x \in [0,1]$. If all directional derivatives of W exist at $(0,0)$, and are uniformly bounded away from 0 and $-\infty$, then $\boldsymbol{\omega(\mathbb{G}(n, W)) = \Theta(\sqrt{n})}$ a.a.s.

Proof sketch - moderate rate of approach

First, a family of graphons with the desired local behavior:

Definition

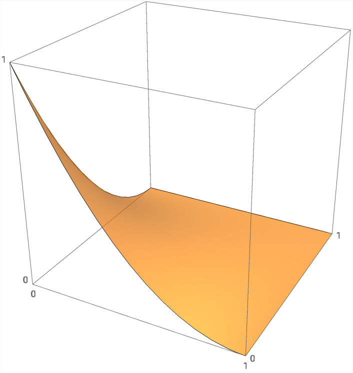
For any $r > 0$, define the graphon

$$W_r(x, y) = (1 - x)^r (1 - y)^r.$$

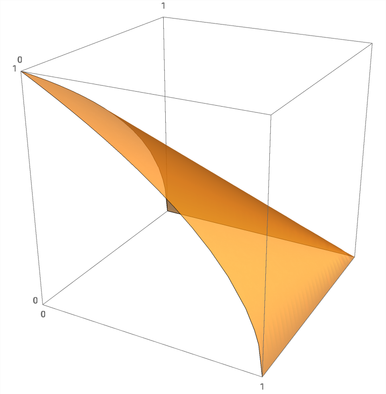
W_r has directional derivatives between $-r$ and $-\frac{r}{\sqrt{2}}$ at $(0, 0)$.

Proof sketch - moderate rate of approach

Family of graphons: $W_r(x, y) = (1 - x)^r(1 - y)^r$



W_2



$W_{1/2}$

Proof sketch - moderate rate of approach

Now, find the clique number of $\mathbb{G}(n, W_r)$. For upper bound, use first moment method:

$$
\begin{aligned}
\mathbb{E}[\#\text{ k-cliques in } \mathbb{G}(n, W_r)] &= \binom{n}{k} \cdot \Pr([k] \text{ is a clique}) \\
&= \binom{n}{k} \int_{[0,1]^k} \prod_{\ell \neq m \in [k]} W_r(x_\ell, x_m)\ d\vec{x} \\
&= \binom{n}{k} \left(\int_0^1 (1-x)^{r(k-1)} dx \right)^k \\
&= \binom{n}{k} \left(\frac{1}{r(k-1)+1} \right)^k .
\end{aligned}
$$

Proof sketch - moderate rate of approach

$$\mathbb{E}[\#\text{ k-cliques in } \mathbb{G}(n, W_r)] \;=\; \binom{n}{k}\left(\frac{1}{r(k-1)+1}\right)^k.$$

- If k is slightly more than $\left(\frac{e}{r}\right)^{1/2} \cdot \sqrt{n}$, then this is $o(1)$.
- $\Rightarrow$ by Markov, with probability $(1 - o(1))$, have

$$\omega(\mathbb{G}(n, W_r)) \leq (1 + o(1)) \left(\frac{e}{r}\right)^{1/2} \cdot \sqrt{n}.$$

Proof sketch - moderate rate of approach

Now for a **lower bound** on $\omega(\mathbb{G}(n, W_r))$.
We want a clique of size $\boldsymbol{\Theta(\sqrt{n})}$. Recall that

$$W_r(x, y) = (1 - x)^r (1 - y)^r.$$

- **Question:** Which vertices are likely to form a large clique?
- **Answer:** Vertices $\boldsymbol{i}$ with $\boldsymbol{x_i}$ close to 0.

Proof sketch - moderate rate of approach

$$W_r(x, y) = (1 - x)^r (1 - y)^r.$$

- **Question:** Which vertices are likely to form a large clique?
- **Answer:** Vertices $\boldsymbol{i}$ with $\boldsymbol{x_i}$ close to 0.
- **Fact:** For any constant c, there are at least $\boldsymbol{c\sqrt{n}}$ vertices with $x_i \leq (1 + o(1))\frac{\boldsymbol{c}}{\sqrt{\boldsymbol{n}}}$, a.a.s. (Chebyshev)

Proof sketch - moderate rate of approach

- Have $c\sqrt{n}$ vertices with $x_i \leq (1 + o(1))\frac{c}{\sqrt{n}}$.
- In general, they *won't* form a clique, but...
- With probability $1 - o(1)$, the subgraph they induce will be missing at most $\frac{c\sqrt{n}}{2}$ edges if $c \leq \left(\frac{1}{3er}\right)^{1/2}$.
  definition of W_r + union bound + some algebra
- Delete one vertex from each non-edge $\Rightarrow$ remaining vertices form a clique of size $\frac{c\sqrt{n}}{2} = \Theta(\sqrt{n})$.

Proof sketch - moderate rate of approach

Quick recap:

- For any $r > 0$, the graphon $W_r(x, y) = (1-x)^r(1-y)^r$ has clique number $\Theta(\sqrt{n})$.
- **Upper bound:** $(1+o(1))\left(\frac{e}{r}\right)^{1/2} \cdot \sqrt{n}$, by first moment method.
- **Lower bound:** $\frac{1}{2}\left(\frac{1}{3er}\right)^{1/2}\sqrt{n}$, by guessing which vertices are likely to form a large clique, and showing this indeed happens with high probability.

Proof sketch - moderate rate of approach

Recall proof strategy for general W:

- Compute the clique numbers associated to some *specific* family of graphons with the desired local behavior. **Done.**
- **Fact:** the clique number of $\mathbb{G}(n, W)$ is primarily determined by local behavior of W near points with $W(x, x) = 1$.
- Use elements of specific family, together with fact above, to determine the clique number associated to *any* graphon with comparable local behavior.

Proof sketch - moderate rate of approach

Recall proof strategy for general W:

- Compute the clique numbers associated to some *specific* family of graphons with the desired local behavior.
- **Fact:** the clique number of $\mathbb{G}(n, W)$ is primarily determined by local behavior of W near points with $W(x, x) = 1$.
- Use elements of specific family, together with fact above, to determine the clique number associated to *any* graphon with comparable local behavior.

Proof sketch - moderate rate of approach

Fact: the clique number of $\mathbb{G}(n, W)$ is primarily determined by local behavior of W near points with $W(x,x) = 1$. Explicitly...

Lemma (M., 2019)

Let $W, U : [0,1] \to [0,1]^2$ be graphons equal to 1 at some point (a,a), and essentially bounded away from 1 in some neighborhood of (x,x) for all other $x \in [0,1]^2$. If there exists a neighborhood N of (a,a) on which $W(x,y) \leq U(x,y)$, then a.a.s.,

$$\omega(\mathbb{G}(n, W)) \leq (1 + o(1)) \cdot \omega(\mathbb{G}(n, U)) + O(\log n).$$

Proof sketch - moderate rate of approach

Now, putting it all together...

- Let W be *any* graphon equal to 1 at $(0,0)$, bounded away from 1 near all other points (x,x), and with directional derivatives bounded away from 0 and $-\infty$ at $(0,0)$.
- **Recall:** W_r has directional derivatives between $-r$ and $-\frac{r}{\sqrt{2}}$.
- So if we take $\boldsymbol{r_1}$ sufficiently large and $\boldsymbol{r_2}$ sufficiently small, we can bound W between $\boldsymbol{W_{r_1}}$ and $\boldsymbol{W_{r_2}}$ in some neighborhood of $(0,0)$.

Proof sketch - moderate rate of approach

- W is locally bounded between $\boldsymbol{W_{r_1}}$ and $\boldsymbol{W_{r_2}}$ in some neighborhood of $(0,0)$ for some $\boldsymbol{r_1}, \boldsymbol{r_2}$.
- So by the Lemma, $\omega(\mathbb{G}(n, W))$ is bounded between $\omega(\mathbb{G}(n, W_{r_1})) = \boldsymbol{\Theta(\sqrt{n})}$ and $\omega(\mathbb{G}(n, W_{r_2})) = \boldsymbol{\Theta(\sqrt{n})}$.
- Thus $\omega(\mathbb{G}(n, W)) = \boldsymbol{\Theta(\sqrt{n})}$.

Proof sketch - moderate rate of approach

We've proved:

Theorem (M., 2019)

Let $W\colon [0,1]^2 \to [0,1]$ be a graphon equal to 1 at the point $(0,0)$, and essentially bounded away from 1 in some neighborhood of (x,x) for each other $x \in [0,1]$. If all directional derivatives of W exist at $(0,0)$, and are uniformly bounded away from 0 and $-\infty$, then $\boldsymbol{\omega(\mathbb{G}(n,W)) = \Theta(\sqrt{n})}$ a.a.s.

Bonus!

Can use the same ideas to prove:

Example (M.,, 2019)

Define the graphon $W : [0,1]^2 \to [0,1]$ by

$$W = (1 - f(x))(1 - f(y)), \text{ where } f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

The random graph $\mathbb{G}(n, W)$ has clique number $n^{1-o(1)}$ a.a.s.

Bonus!

Example (M., 2019)

$$W = (1 - f(x))(1 - f(y)), \text{ where } f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}.$$

The random graph $\mathbb{G}(n, W)$ has clique number $\boldsymbol{n^{1-o(1)}}$ a.a.s.

- Infinitely many zero derivatives at $(0, 0)$.
- $\Rightarrow$ locally bound below by $U_r = (1 - x^r)(1 - y^r)$ for any r.
- $\Rightarrow$ bound $\omega(\mathbb{G}(n, W))$ below by $\omega(\mathbb{G}(n, U_r)) = \boldsymbol{\Theta(n^{\frac{r}{r+1}})}$.

Future directions

Constants

Determining correct constants: Could perhaps use tools from large deviations theory, or a very technical second moment calculation.

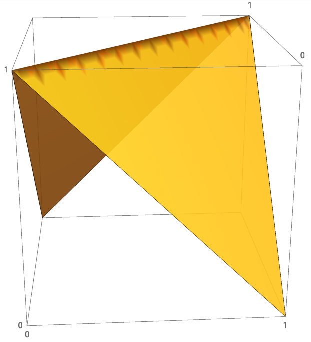
Finding large cliques

- **Finding a large clique:** In $G_{n,1/2}$, no known poly-time algorithm finds a clique more than half the size of the max clique (e.g. Krivelevich, Sudakov, 1998; problem due to Karp, 1976). Could we do better here?
- **Finding a planted clique:** Plant a clique of size k in a random graph, ask for poly-time algorithm to recover it. For $G_{n,p}$, algorithm known only if $k \geq \sqrt{n}$ (e.g. Alon et al, 1998), much larger than $\omega(G_{n,p})$. Could we do better here?

More points equal to 1

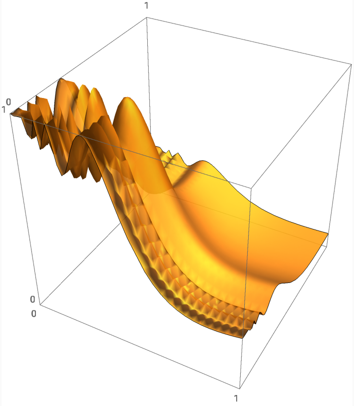
What about graphons with $W(x,x) = 1$ at infinitely many points?

For example,



$W(x,y) = 1 - |x - y|$

or



$$W(x,y) = \left(1 - x\sin^2\left(\tfrac{1}{x}\right)\right) \cdot \left(1 - y\sin^2\left(\tfrac{1}{y}\right)\right)$$

Thank you!