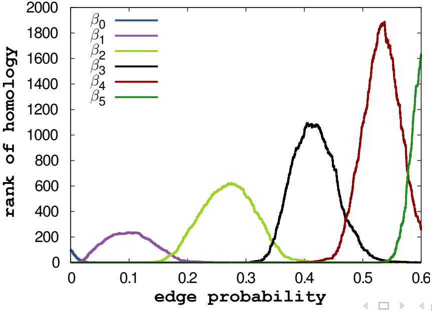
Topology of random flag complexes

Matthew Kahle

Ohio State University

2013–07–18



"I predict a new subject of statistical topology. Rather than count the number of holes, Betti numbers, etc., one will be more interested in the distribution of such objects on noncompact manifolds as one goes out to infinity." — Isadore Singer, 2004.

Some earlier work:

Some earlier work:

- Gaussian fields on manifolds — Adler and Taylor, 2003.

Some earlier work:

- Gaussian fields on manifolds — Adler and Taylor, 2003.
- Random simplicial complexes — Linial and Meshulam, 2006.

Some earlier work:

- Gaussian fields on manifolds — Adler and Taylor, 2003.
- Random simplicial complexes — Linial and Meshulam, 2006.
- Random triangulated surfaces — Pippenger and Schleich, 2006.

Some earlier work:

- Gaussian fields on manifolds — Adler and Taylor, 2003.
- Random simplicial complexes — Linial and Meshulam, 2006.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.

Some earlier work:

- Gaussian fields on manifolds — Adler and Taylor, 2003.
- Random simplicial complexes — Linial and Meshulam, 2006.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages — Farber and Kappeller, 2007.

Randomness models the natural world.

Randomness models the natural world.

- We need to put topological data analysis on firmer probabilistic foundations in order to quantify inference, etc.

Randomness models the natural world.

- We need to put topological data analysis on firmer probabilistic foundations in order to quantify inference, etc.
- Certain situations in physics seem to be well modeled by probabilistic topology.

Randomness models the natural world.

- We need to put topological data analysis on firmer probabilistic foundations in order to quantify inference, etc.
- Certain situations in physics seem to be well modeled by probabilistic topology.
- Why do so many groups / manifolds / simplicial complexes / etc. seem to have a certain topological property?

Randomness models the natural world.

- We need to put topological data analysis on firmer probabilistic foundations in order to quantify inference, etc.
- Certain situations in physics seem to be well modeled by probabilistic topology.
- Why do so many groups / manifolds / simplicial complexes / etc. seem to have a certain topological property?

  E.g. many simplicial complexes and posets arising in combinatorics are homotopy equivalent to bouquets of spheres. But why does this happen so often?

The probabilistic method provides existence proofs.

The probabilistic method provides existence proofs.

- Ramsey theory and extremal graph theory e.g. Erdős, ...

The probabilistic method provides existence proofs.

- Ramsey theory and extremal graph theory e.g. Erdős, ...
- Geometric group theory — e.g. Gromov, Żuk

The probabilistic method provides existence proofs.

- Ramsey theory and extremal graph theory e.g. Erdős, ...
- Geometric group theory — e.g. Gromov, Żuk
- Expander graphs — e.g. Pinsker, Barzdin & Kolmogorov, etc.

Random graphs

Define $G(n, p)$ to be the probability space of graphs on vertex set $[n] = \{1, 2, \ldots, n\}$, where each edge has probability p, independently.

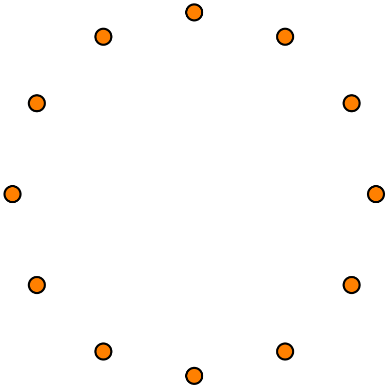
Define $G(n, p)$ to be the probability space of graphs on vertex set $[n] = \{1, 2, \ldots, n\}$, where each edge has probability p, independently.

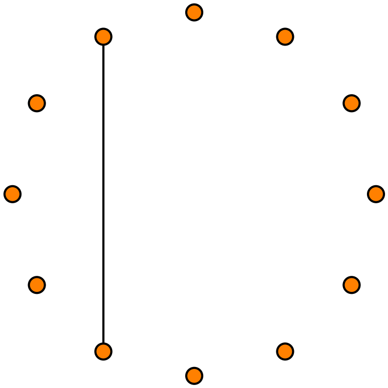
We use the notation $G \sim G(n, p)$ to indicate a graph chosen according to this distribution.

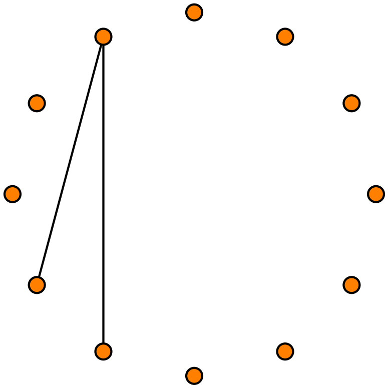
Define $G(n, p)$ to be the probability space of graphs on vertex set $[n] = \{1, 2, \ldots, n\}$, where each edge has probability p, independently.

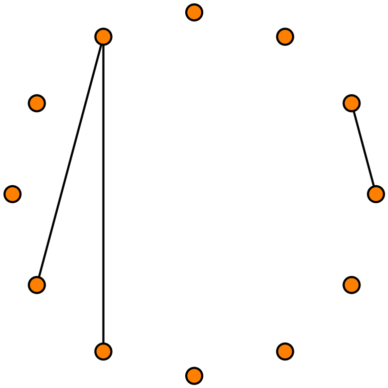
We use the notation $G \sim G(n, p)$ to indicate a graph chosen according to this distribution.

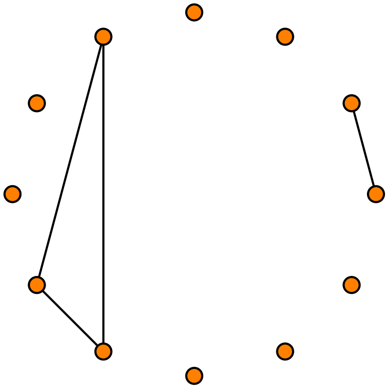
It is often useful to think of a growth process associated with $G(n, p)$ where edges are added one at a time.

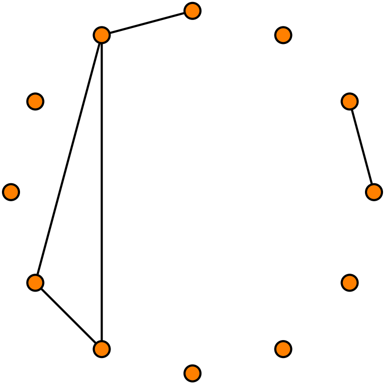


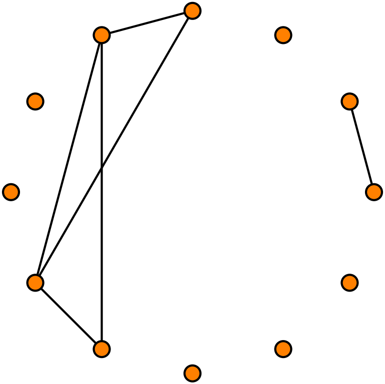


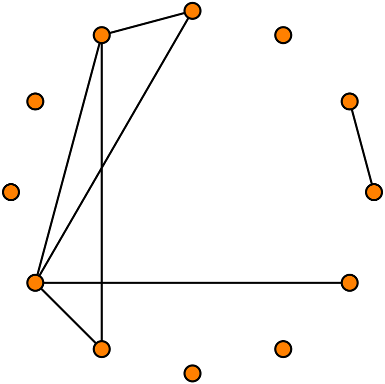


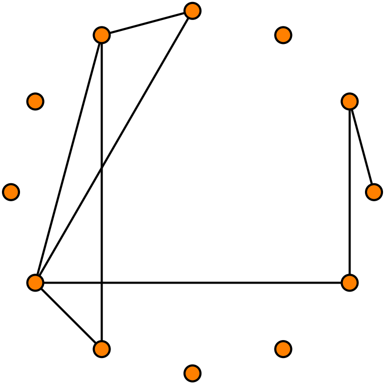


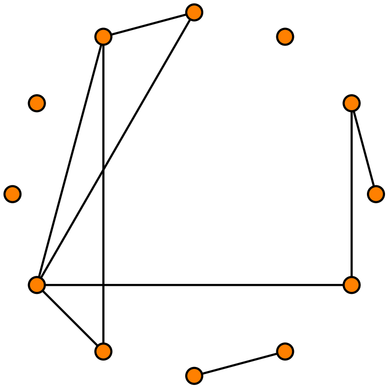


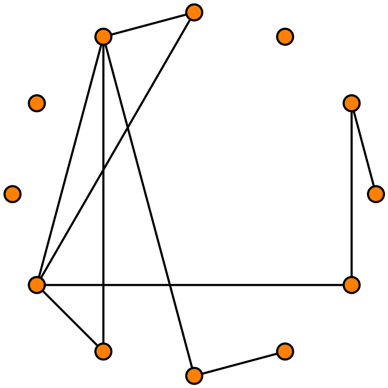


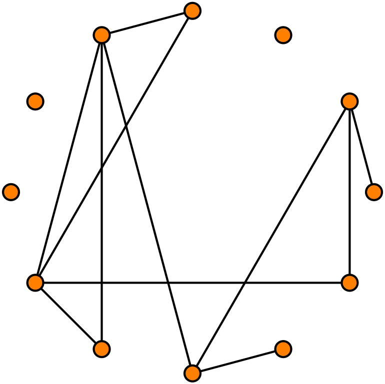


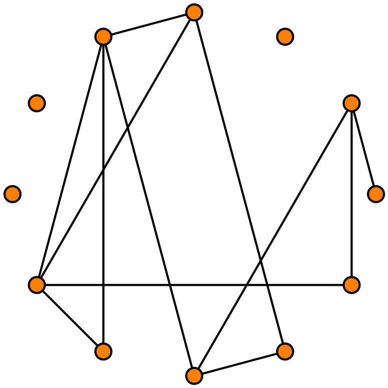


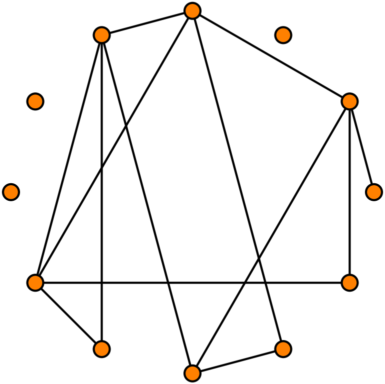


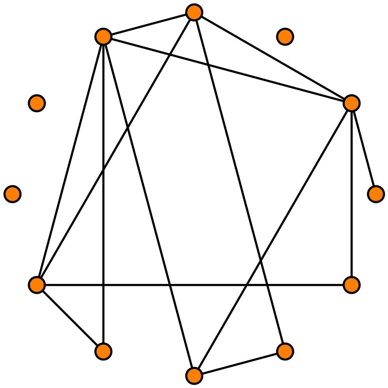


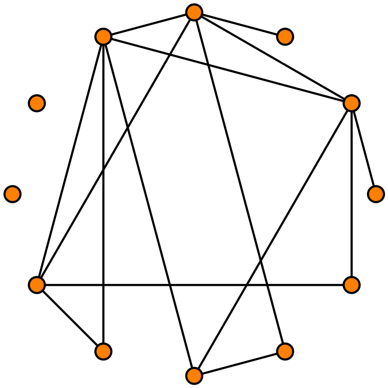


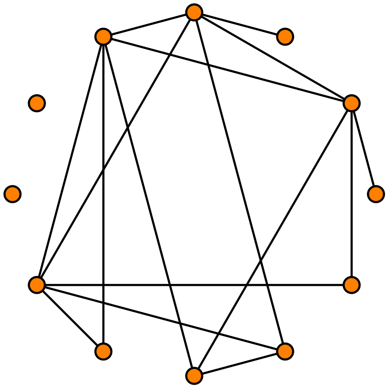


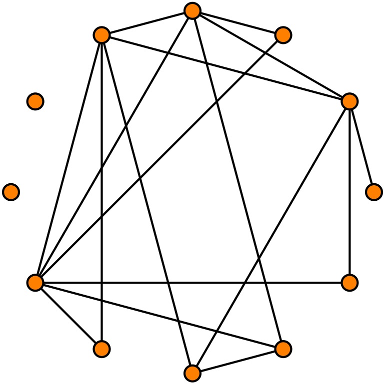


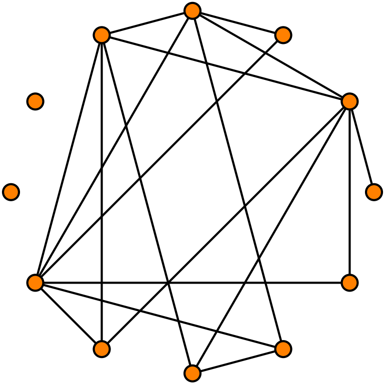


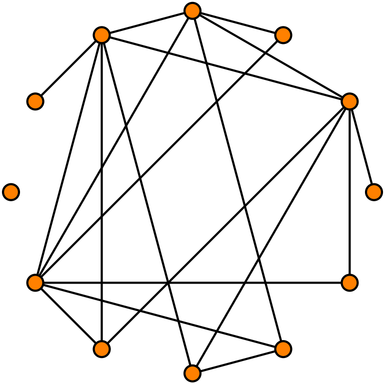


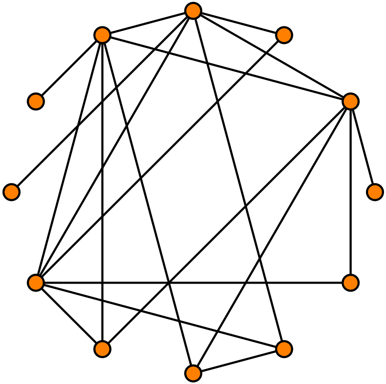












Theorem (Erdős–Rényi, 1959)

Let $\epsilon > 0$ *be fixed and* $G \sim G(n,p)$*. Then*

$$\mathbb{P}[G \textit{ is connected}] \to \begin{cases} 1 & : p \geq (1+\epsilon)\log n/n \\ 0 & : p \leq (1-\epsilon)\log n/n \end{cases}$$

Theorem (Erdős–Rényi, 1959)

Let $\epsilon > 0$ *be fixed and* $G \sim G(n,p)$. *Then*

$$\mathbb{P}[G \textit{ is connected}] \to \begin{cases} 1 & : p \geq (1+\epsilon)\log n/n \\ 0 & : p \leq (1-\epsilon)\log n/n \end{cases}$$

They actually proved a slightly sharper result.

Theorem (Erdős–Rényi, 1959)

Let $c \in \mathbb{R}$ be fixed and $G \sim G(n,p)$. If

$$p = \frac{\log n + c}{n},$$

then $\widetilde{\beta}_0(G)$ is asymptotically Poisson distributed with mean e^{-c} and in particular

$$\mathbb{P}[G \text{ is connected}] \to e^{-e^{-c}}$$

as $n \to \infty$.

The first step is to show that if $p \approx \log n/n$, then the probability that there are any components of order i, with $2 \leq i \leq n/2$ tends to 0 as $n \to \infty$.

The first step is to show that if $p \approx \log n/n$, then the probability that there are any components of order i, with $2 \leq i \leq n/2$ tends to 0 as $n \to \infty$.

A “union bound” argument shows that it is sufficient to show that

$$\sum_{k=2}^{\lfloor n/2 \rfloor} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} \to 0,$$

as $n \to \infty$.

The first step is to show that if $p \approx \log n/n$, then the probability that there are any components of order i, with $2 \leq i \leq n/2$ tends to 0 as $n \to \infty$.

A "union bound" argument shows that it is sufficient to show that

$$\sum_{k=2}^{\lfloor n/2 \rfloor} \binom{n}{k} k^{k-2} p^{k-1} (1-p)^{k(n-k)} \to 0,$$

as $n \to \infty$.

So w.h.p. $G(n, p)$ consists of a giant component and isolated vertices.

Now set $p = (\log n + c)/n$, where $c \in \mathbb{R}$ is fixed. By linearity of expectation, the expected number of isolated vertices V is

$$\mathbb{E}[V] = n(1-p)^{n-1},$$

and since $1 - p \approx e^{-p}$ for $p \approx 0$, we have

$$\mathbb{E}[V] \to e^{-c},$$

as $n \to \infty$.

Now set $p = (\log n + c)/n$, where $c \in \mathbb{R}$ is fixed. By linearity of expectation, the expected number of isolated vertices V is

$$\mathbb{E}[V] = n(1-p)^{n-1},$$

and since $1 - p \approx e^{-p}$ for $p \approx 0$, we have

$$\mathbb{E}[V] \to e^{-c},$$

as $n \to \infty$.

By computing the higher moments, one can show that V approaches a Poisson distribution with mean e^{-c}.

Comments:

Comments:

- The proof is phrased in terms of cohomology rather than in terms of homology.

Comments:

- The proof is phrased in terms of cohomology rather than in terms of homology.
- The ultimate obstruction to connectivity is isolated vertices.

Comments:

- The proof is phrased in terms of cohomology rather than in terms of homology.
- The ultimate obstruction to connectivity is isolated vertices.
- Once p is a little bit past the connectivity threshold, $G \sim G(n, p)$ is an *expander*.

Theorem (Hoffman, K., Paquette, 2012)

Fix $k \geq 0$ *and* $\epsilon > 0$ *be fixed. Let* $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n \leq 2$ *be the eigenvalues of the normalized Laplacian of the random graph* $G(n, p)$*. There is a constant* C_k *so that when*

$$p \geq \frac{(k+1)\log n + C_k \sqrt{\log n} \log \log n}{n}$$

is satisfied, then

$$\lambda_2 > 1 - \epsilon,$$

with probability at least $1 - o(n^{-k})$*.*

It is worth noting that there is another topological phase transition for $G(n, p)$, namely when cycles first appear.

Theorem (Pittel)

Let $c > 0$ be fixed, $p = c/n$, and $G \sim G(n,p)$. Then

$$\mathbb{P}[H_1(G) = 0] \to \begin{cases} 0 & : c \geq 1 \\ \frac{\sqrt{1-c}}{\exp(c/2+c^2/4)} & : c < 1 \end{cases}$$

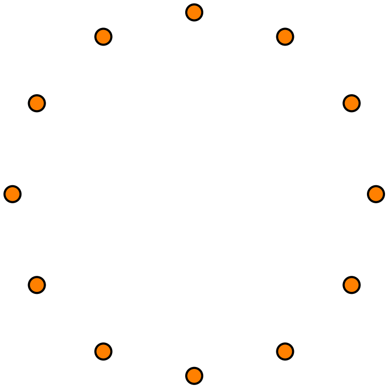
as $n \to \infty$.

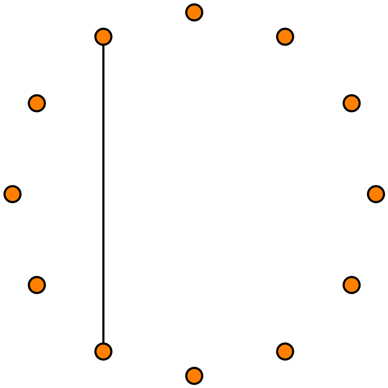
Random flag complexes

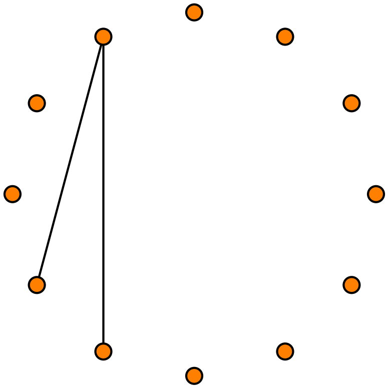
Let $X \sim X(n, p)$ be the clique complex (or flag complex) of $G \sim G(n, p)$, i.e. the maximal simplicial complex compatible with G.

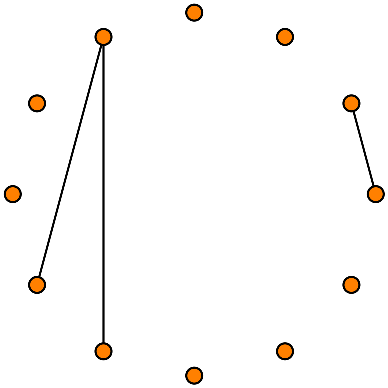
Let $X \sim X(n, p)$ be the clique complex (or flag complex) of $G \sim G(n, p)$, i.e. the maximal simplicial complex compatible with G.

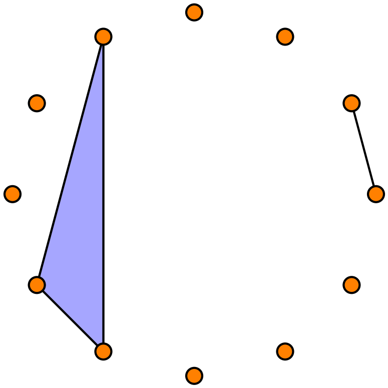
Note: every simplicial complex is homeomorphic to a flag complex, e.g. by barycentric subdivision, so $X(n, p)$ puts a measure on a wide range of topologies.

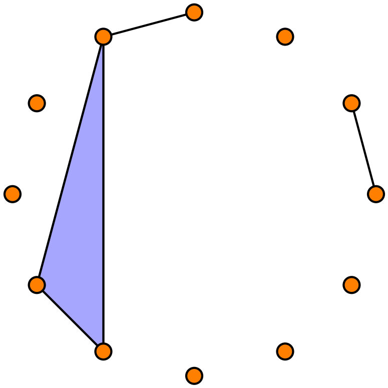


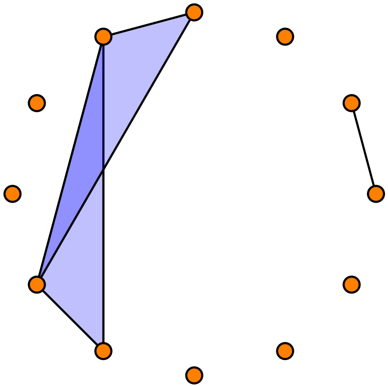


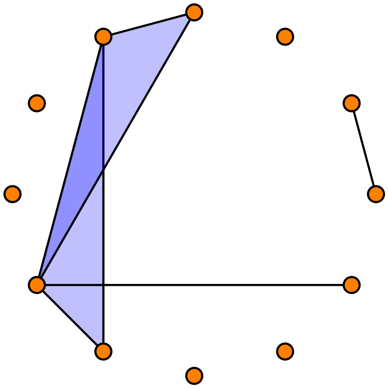


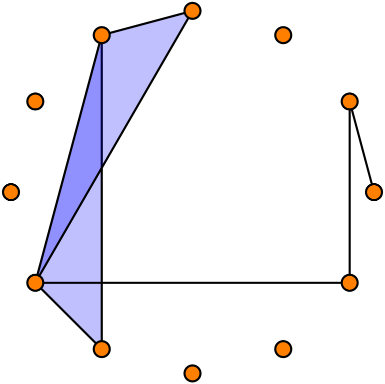


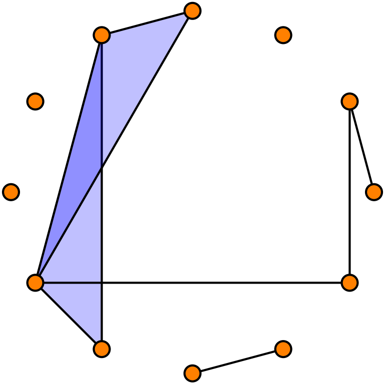


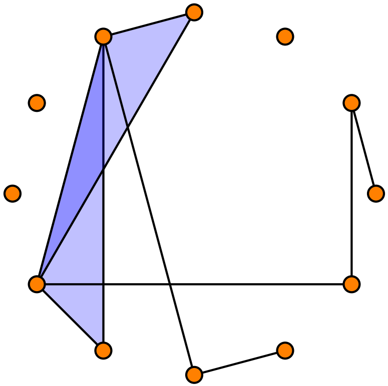


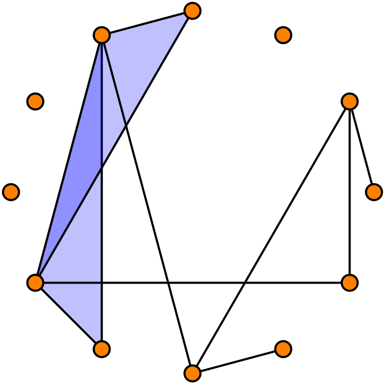


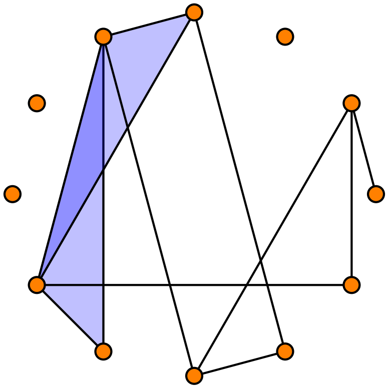


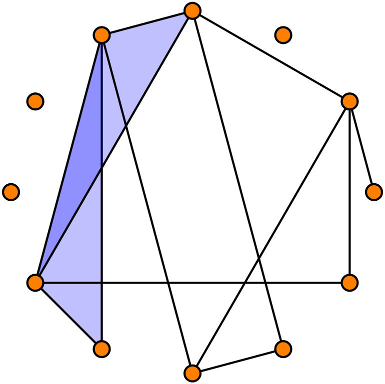


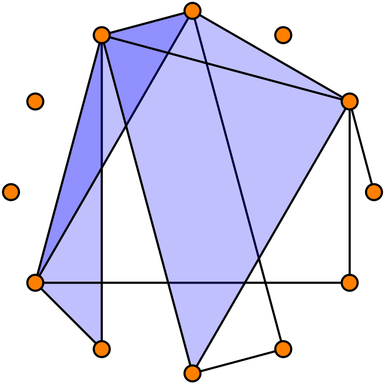


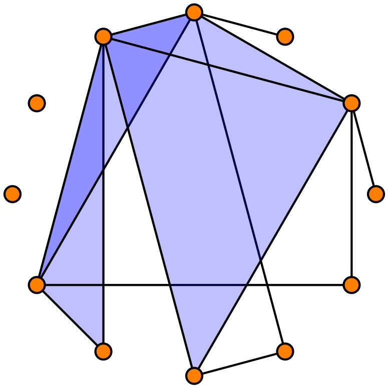


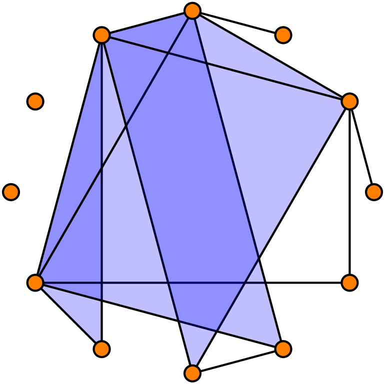


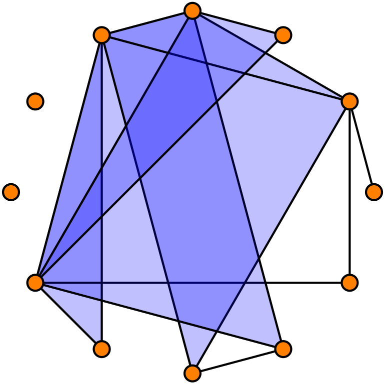


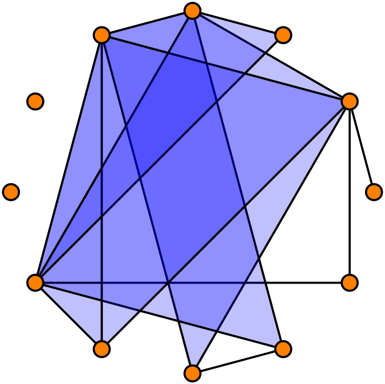


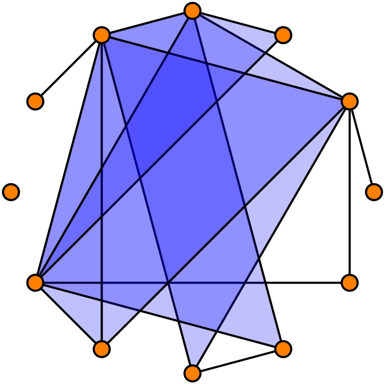


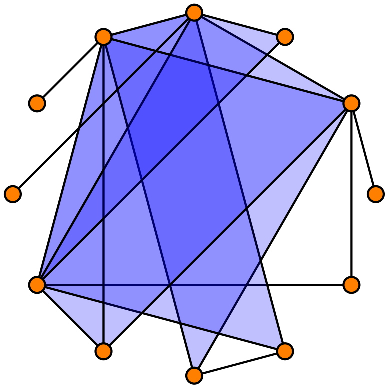


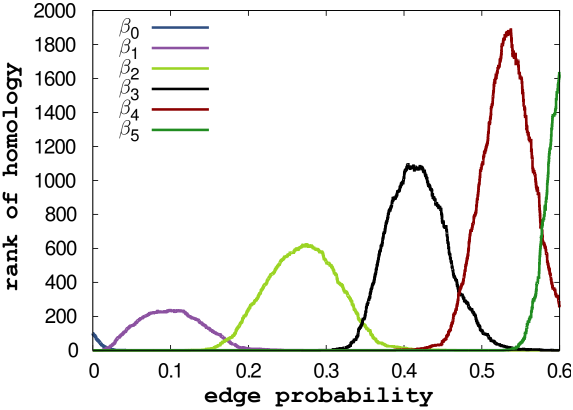












Betti numbers of random flag complexes on $n = 100$ vertices, with $0 \leq p \leq 0.6$.

Theorem (K., 2007)

Fix $k \geq 1$, and let $X \sim X(n,p)$. If

$$\frac{1}{n^{1/k}} \ll p \ll \frac{1}{n^{1/(k+1)}}$$

then

$$\mathbb{P}[H_k(X) = 0] \to 0$$

as $n \to \infty$.

Theorem (K., 2007)

Fix $k \geq 1$, and let $X \sim X(n, p)$.

- *If*

$$p \ll \frac{1}{n^{1/k}}$$

  then

$$\mathbb{P}[H_k(X) = 0] \to 1$$

  as $n \to \infty$.

Theorem (K., 2007)

Fix $k \geq 1$, and let $X \sim X(n,p)$.

- *If*
$$p \ll \frac{1}{n^{1/k}}$$
then
$$\mathbb{P}[H_k(X) = 0] \to 1$$
as $n \to \infty$.
- *Also, if*
$$p \gg \frac{1}{n^{1/(2k+1)}}$$
then
$$\mathbb{P}[H_k(X) = 0] \to 1$$
as $n \to \infty$.

In fact, much more can be said about the size of homology in this regime. The limiting expectation has a nice formula.

In fact, much more can be said about the size of homology in this regime. The limiting expectation has a nice formula.

Theorem (K., 2007)

If

$$\frac{1}{n^{1/k}} \ll p \ll \frac{1}{n^{1/(k+1)}}$$

then

$$\frac{\mathbb{E}[\beta_k]}{\binom{n}{k+1} p^{\binom{k+1}{2}}} \to 1$$

as $n \to \infty$.

Moreover, the kth Betti number satisfies a central limit theorem.

Moreover, the kth Betti number satisfies a central limit theorem.

Theorem (K.–Meckes, 2010))

If

$$\frac{1}{n^{1/k}} \ll p \ll \frac{1}{n^{1/(k+1)}}$$

then

$$\frac{\beta_k - \mathbb{E}[\beta_k]}{\sqrt{Var\left[\beta_k\right]}} \to \mathcal{N}(0,1)$$

as $n \to \infty$.

The following new result gives a sharp vanishing threshold for cohomology.

The following new result gives a sharp vanishing threshold for cohomology.

Theorem (K., 2012)

Let $0 < \epsilon < 1$ and k be fixed and $X \sim X(n,p)$. Then

$$\mathbb{P}[H^k(Y,\mathbb{Q}) = 0] \to \begin{cases} 1 & : p \geq \left(\frac{(k/2+1+\epsilon)\log n}{n} \right)^{1/(k+1)} \\ 0 & : \frac{1}{n^{1/k}} \ll p \leq \left(\frac{(k/2+1-\epsilon)\log n}{n} \right)^{1/(k+1)} \end{cases}$$

The following new result gives a sharp vanishing threshold for cohomology.

Theorem (K., 2012)

Let $0 < \epsilon < 1$ *and* k *be fixed and* $X \sim X(n,p)$. *Then*

$$\mathbb{P}[H^k(Y,\mathbb{Q}) = 0] \to \begin{cases} 1 & : p \geq \left(\frac{(k/2+1+\epsilon)\log n}{n}\right)^{1/(k+1)} \\ 0 & : \frac{1}{n^{1/k}} \ll p \leq \left(\frac{(k/2+1-\epsilon)\log n}{n}\right)^{1/(k+1)} \end{cases}$$

This provides a generalization of the Erdős–Rényi Theorem.

Besides the earlier spectral gap theorem, the main tool is the following.

Besides the earlier spectral gap theorem, the main tool is the following.

Theorem

(Garland, 1973, Ballman–Świątkowski, 1997) If Δ is a pure k-dimensional simplicial complex, such that the link $lk_\Delta(\sigma)$ of every $(k-2)$-face σ is connected and has spectral gap satisfying

$$\lambda_2[lk_\Delta(\sigma)] > 1 - 1/k,$$

then $H^{k-1}(\Delta, \mathbb{Q}) = 0$.

We have the following corollary.

Corollary

Fix $d \geq 0$, and let $X \sim X(n, p)$ be a random flag complex, where

$$\frac{1}{n^{2/d}} \ll p \ll \frac{1}{n^{2/(d+1)}}.$$

Then w.h.p. X is d-dimensional and $\widetilde{H}_i(X, \mathbb{Q}) = 0$ unless $i = \lfloor d/2 \rfloor$.

We have the following corollary.

Corollary

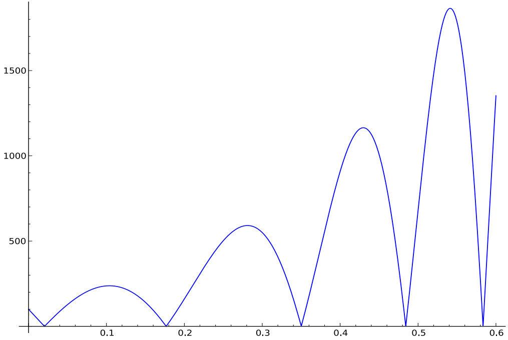
Fix $d \geq 0$, and let $X \sim X(n,p)$ be a random flag complex, where

$$\frac{1}{n^{2/d}} \ll p \ll \frac{1}{n^{2/(d+1)}}.$$

Then w.h.p. X is d-dimensional and $\widetilde{H}_i(X,\mathbb{Q}) = 0$ unless $i = \lfloor d/2 \rfloor$.

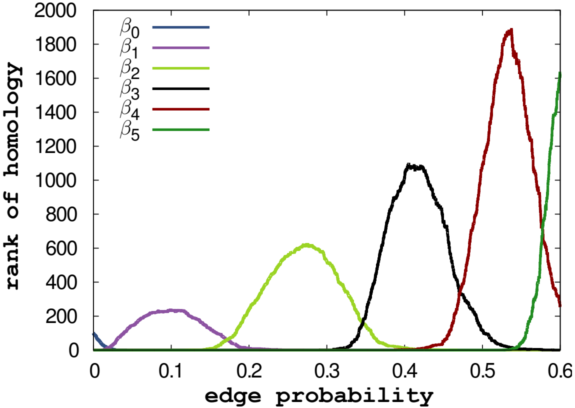
Moreover, if $d \geq 6$ then w.h.p. X is rationally homotopy equivalent to a wedge of $\lfloor d/2 \rfloor$-spheres.

One consequence is that now we can predict Betti numbers of random flag complexes fairly well, even for small n...

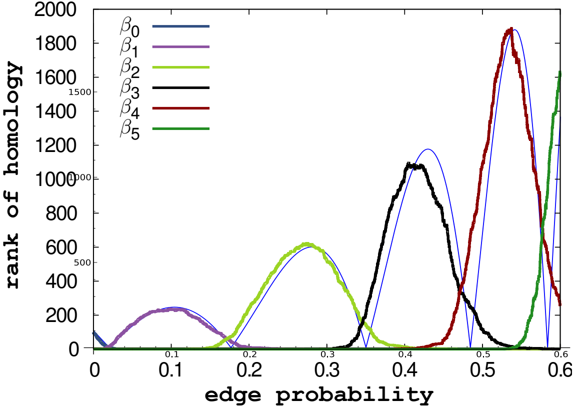


The predicted Betti numbers are given by the absolute Euler characteristic.

$$|\mathbb{E}[\chi]| = \left| \binom{100}{1} - \binom{100}{2} p + \binom{100}{3} p^3 - \dots \right| .$$



The actual Betti numbers.


2000
1800
1600
1500
1400
1200
1000
1000
800
600
500
400
200
0
rank of homology
β_0
β_1
β_2
β_3
β_4
β_5
0
0.1
0.2
0.3
0.4
0.5
0.6
edge probability

The new results also say something with respect to *persistent homology* of the random filtration associated with $X(n, p)$.

Corollary

Fix $\epsilon > 0$ and $k \geq 1$. Let b, d be the birth time and death time, respectively of the longest bar in persistent H_k. Then w.h.p.

$$\frac{\log d - \log b}{\log n} \approx \frac{1}{k(k+1)}.$$

Open problems

It might be possible to slightly sharpen the main result.

Conjecture

If

$$p = \left(\frac{(k/2+1)\log n + (k/2)\log\log n + c}{n} \right)^{1/(k+1)},$$

where $c \in \mathbb{R}$ is constant, then the dimension of kth cohomology β^k approaches a Poisson distribution with mean

$$\mu = \frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c}.$$

In particular,

$$\mathbb{P}[H^k(X, \mathbb{Q}) = 0] \to \exp\left[-\frac{(k/2+1)^{k/2}}{(k+1)!} e^{-c} \right],$$

as $n \to \infty$.

How to handle torsion in homology of random complexes?

How to handle torsion in homology of random complexes?

There exists a 2-dimensional $\mathbb{Q}$-acyclic simplicial complex S on 31 vertices with

$$|H_1(S, \mathbb{Z})| = 736712186612810774591.$$

How to handle torsion in homology of random complexes?

There exists a 2-dimensional $\mathbb{Q}$-acyclic simplicial complex S on 31 vertices with

$$|H_1(S, \mathbb{Z})| = 736712186612810774591.$$

Gil Kalai showed that, on average, $\mathbb{Q}$-acyclic complexes have enormous torsion in homology.

Still, I conjecture the following.

Still, I conjecture the following.

Conjecture

If $X \sim X(n,p)$ is a random flag complex with

$$\frac{1}{n^{2/d}} \ll p \ll \frac{1}{n^{2/(d+1)}},$$

where $d \geq 6$ is fixed, then w.h.p. X is homotopy equivalent to a wedge of $\lfloor d/2 \rfloor$-spheres.

Still, I conjecture the following.

Conjecture

If $X \sim X(n, p)$ is a random flag complex with

$$\frac{1}{n^{2/d}} \ll p \ll \frac{1}{n^{2/(d+1)}},$$

where $d \geq 6$ is fixed, then w.h.p. X is homotopy equivalent to a wedge of $\lfloor d/2 \rfloor$-spheres.

By uniqueness of Moore spaces, this is equivalent to showing that $H_*(X)$ is torsion-free.

Thanks for your time and attention!

Acknowledgements

Collaborators: Chris Hoffman, Elizabeth Meckes, Elliot Paquette

Support: Stanford University (2007–10), Institute for Advanced Study (2010–11), Ohio State University (2011–), Alfred P. Sloan Foundation (2012–14)

References

- M. Kahle, Topology of random simplicial complexes: a survey (submitted, arXiv:1301.7165)
- M. Kahle, Sharp vanishing thresholds for cohomology of random flag complexes (submitted, arXiv:1207.0149)
- C. Hoffman, M. Kahle, and E. Paquette, A sharp threshold for Kazhdan's Property (T) (submitted, arXiv:1201.0425)
- M. Kahle and E. Meckes, Limit theorems for Betti numbers of random simplicial complexes (Homology, Homotopy and Applications, 2013)
- M. Kahle, Topology of random clique complexes (Discrete Math. 309 (2009), no. 6, 1658–1671.)