Lecture 9: Percolation and stochastic topology

NSF/CBMS Conference

Sayan Mukherjee

Departments of Statistical Science, Computer Science, Mathematics
Duke University
www.stat.duke.edu/∼sayan

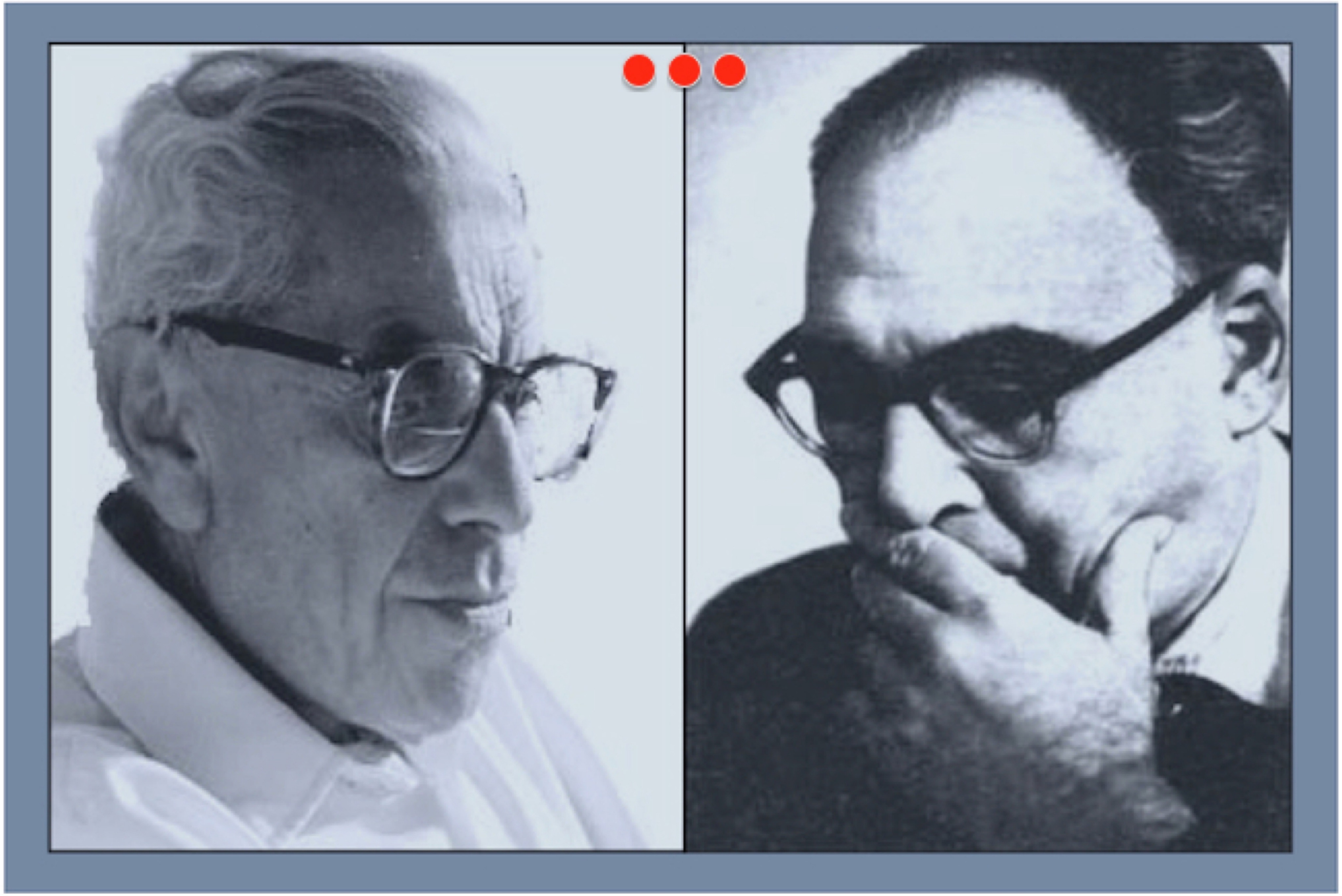
May 31, 2016

Stochastic topology

"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.

Random graphs

Erdős-Rényi



Erdős-Rényi random graph model

$G(n, p)$ is the probability space of graphs on vertex set $[n] = \{1, ...n\}$ with the probability of each edge p independently.

Erdős-Rényi random graph model

$G(n, p)$ is the probability space of graphs on vertex set $[n] = \{1, ...n\}$ with the probability of each edge p independently.

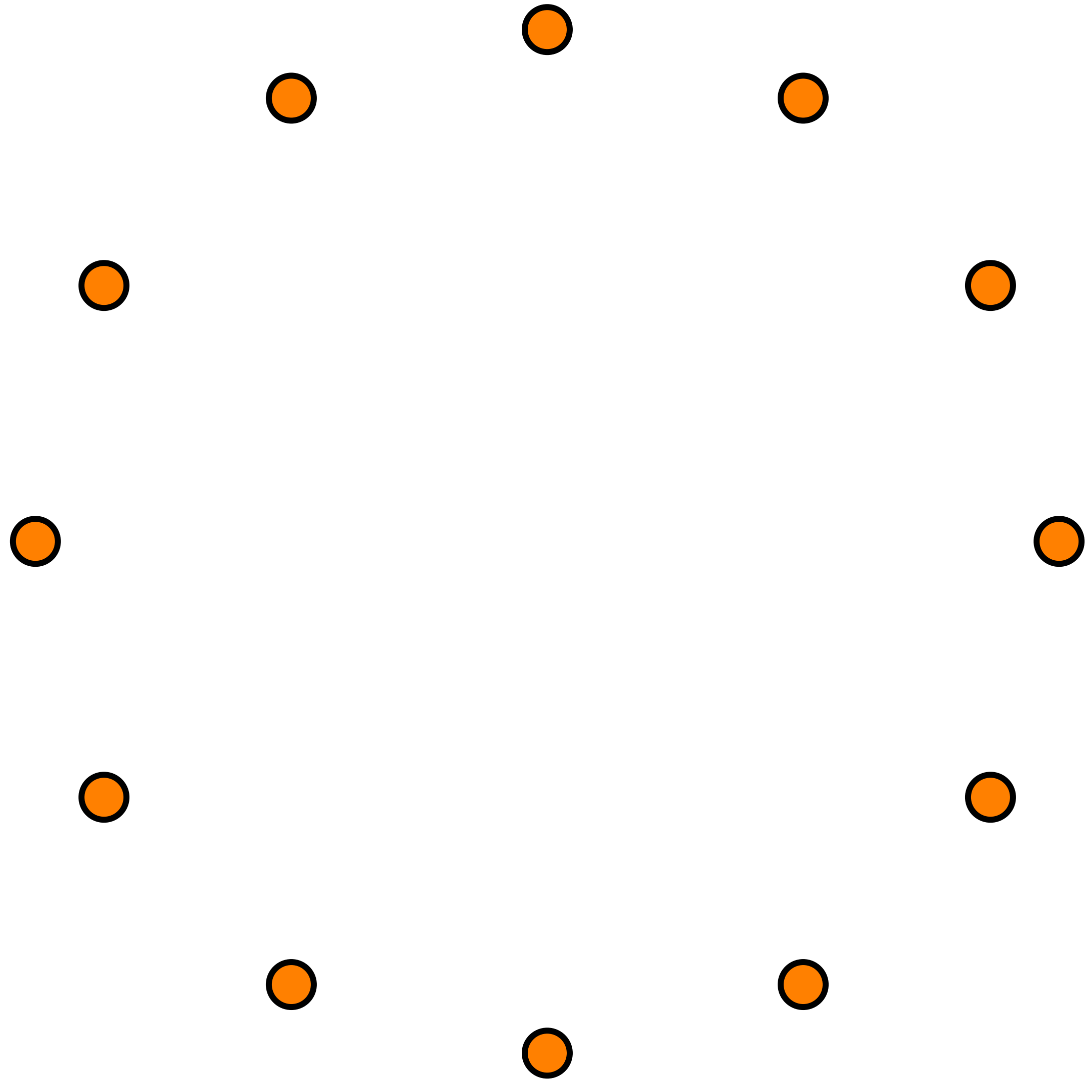
$G \sim G(n, p)$ is a draw of a graph G from the graph distribution.

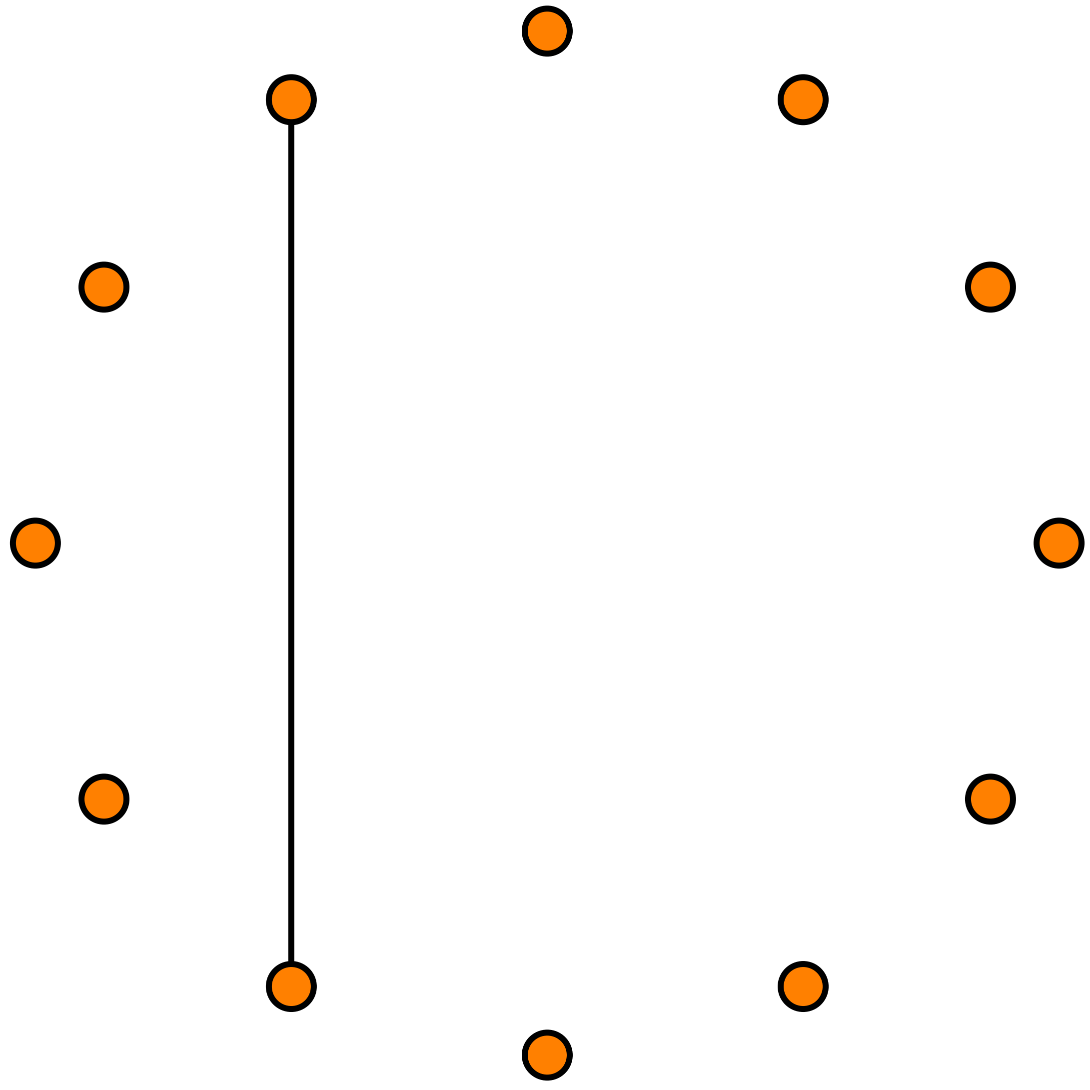
Erdős-Rényi random graph model

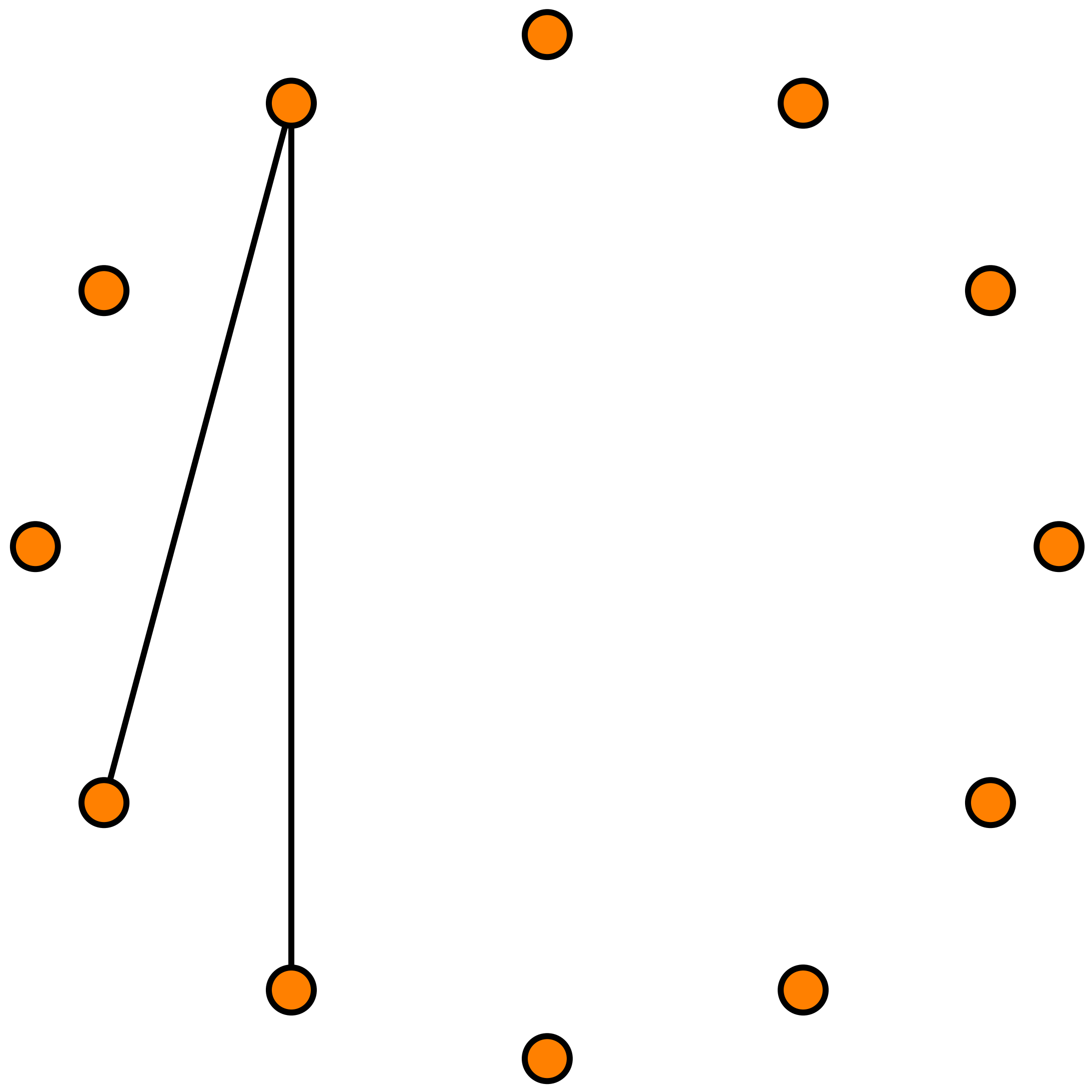
$G(n, p)$ is the probability space of graphs on vertex set $[n] = \{1, \ldots n\}$ with the probability of each edge p independently.

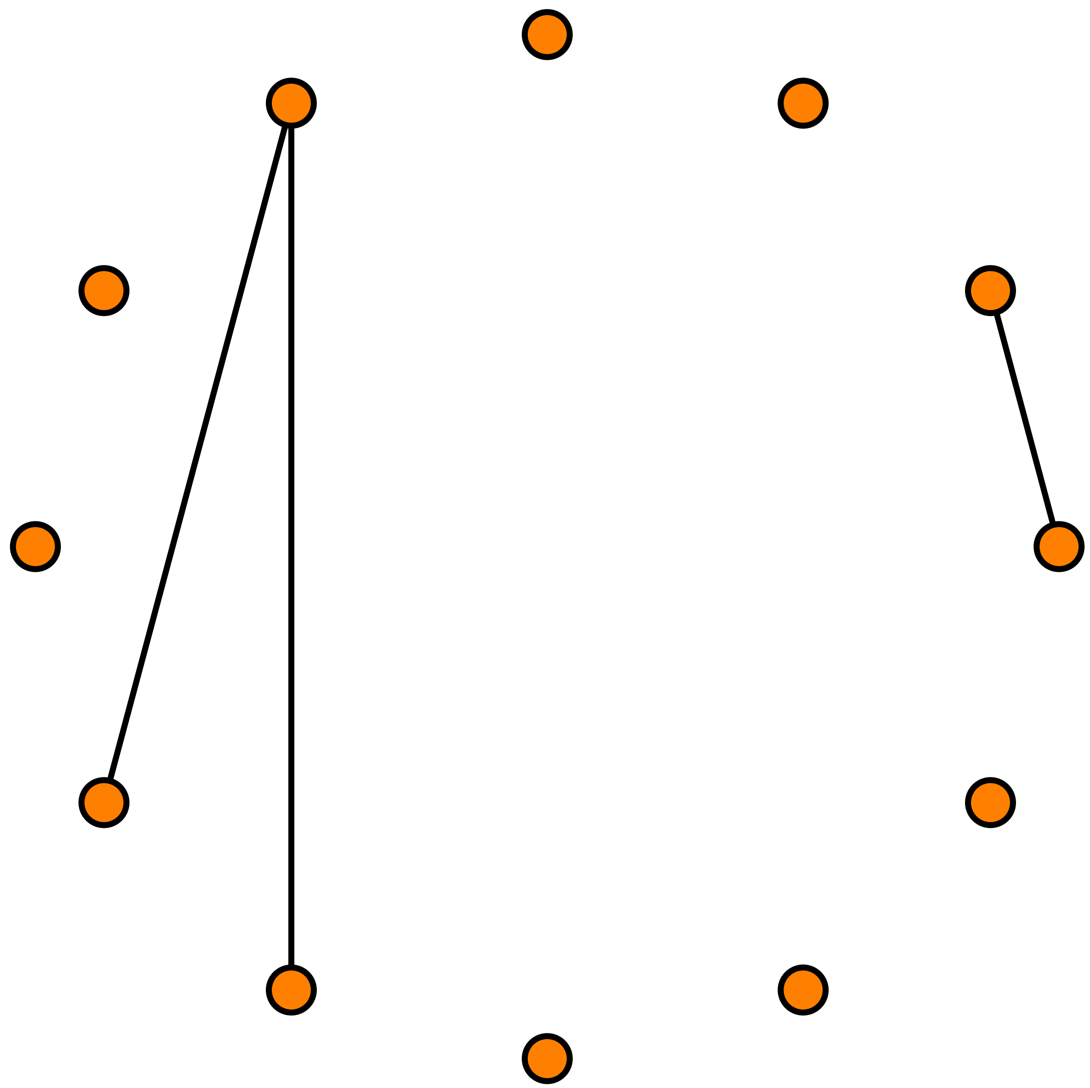
$G \sim G(n, p)$ is a draw of a graph G from the graph distribution.

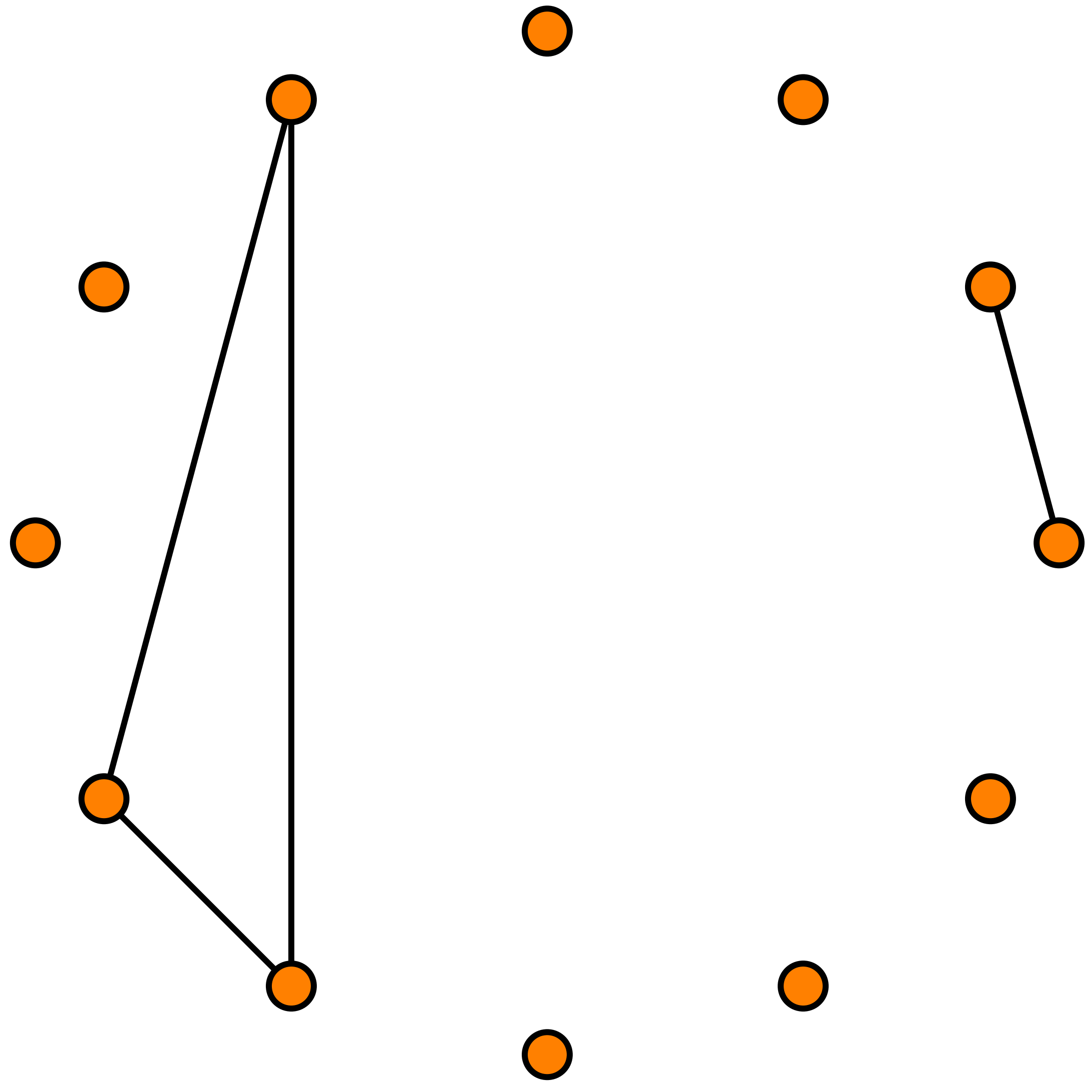
Consider $p(n)$ and $n \to \infty$ and ask questions about thresholds of graph properties.

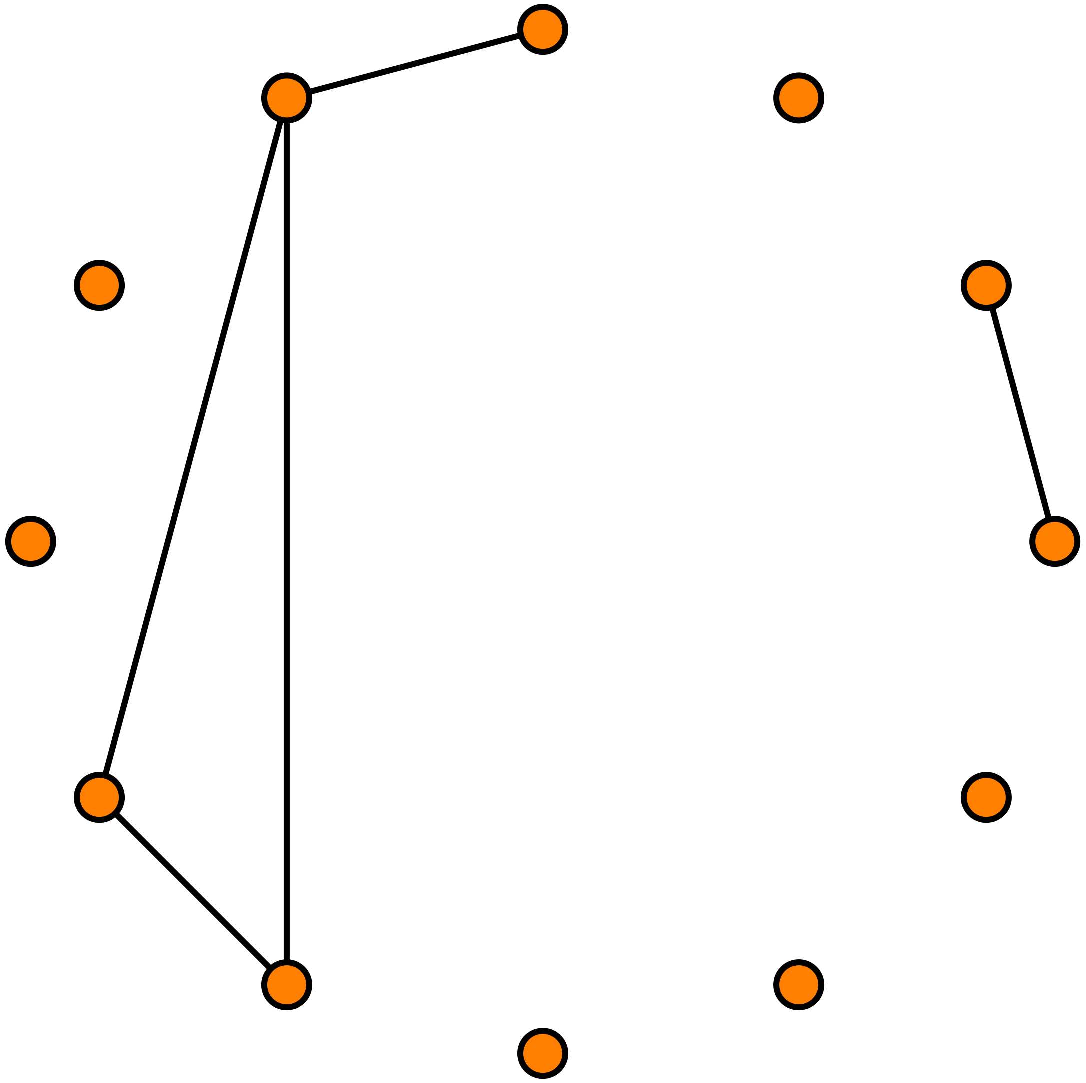


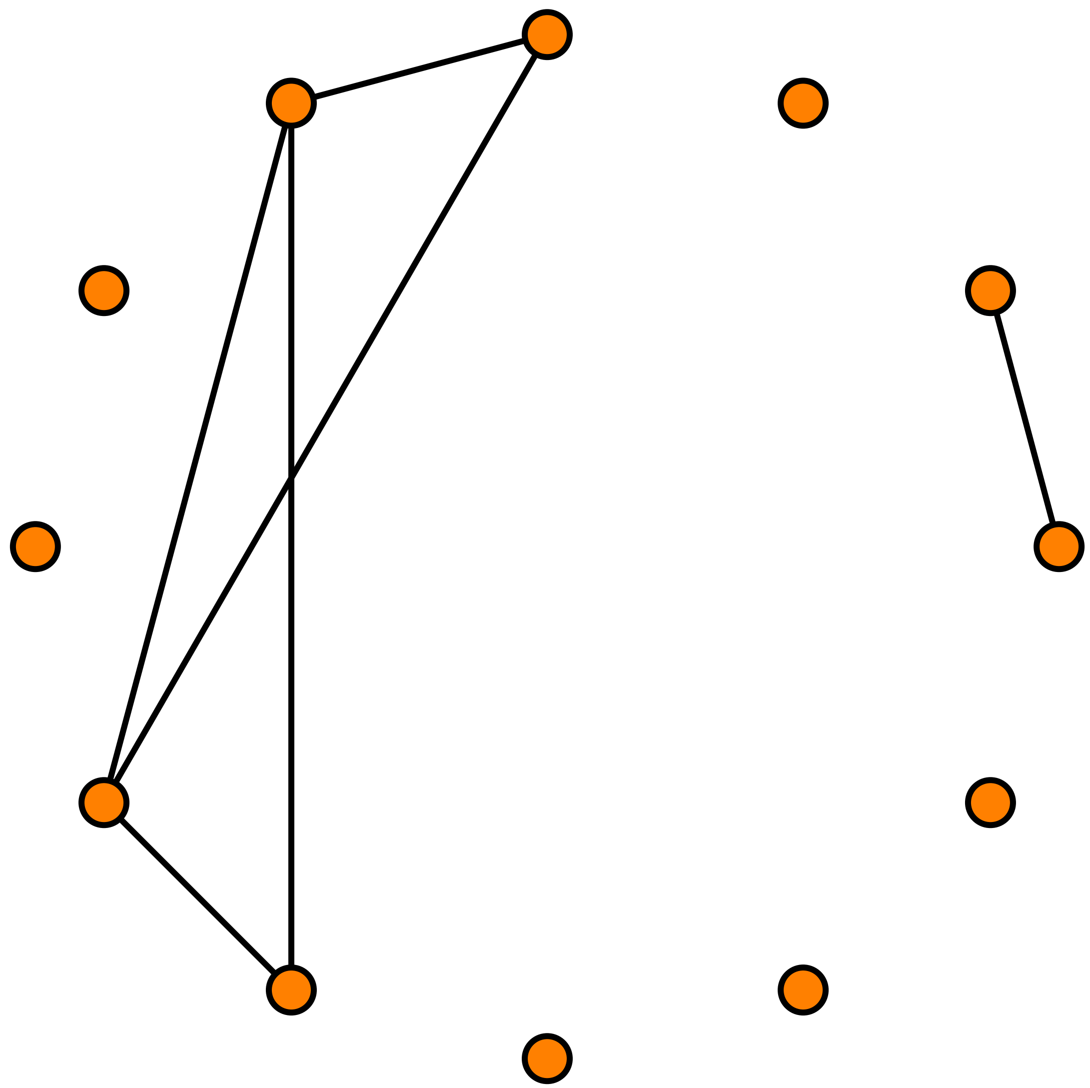


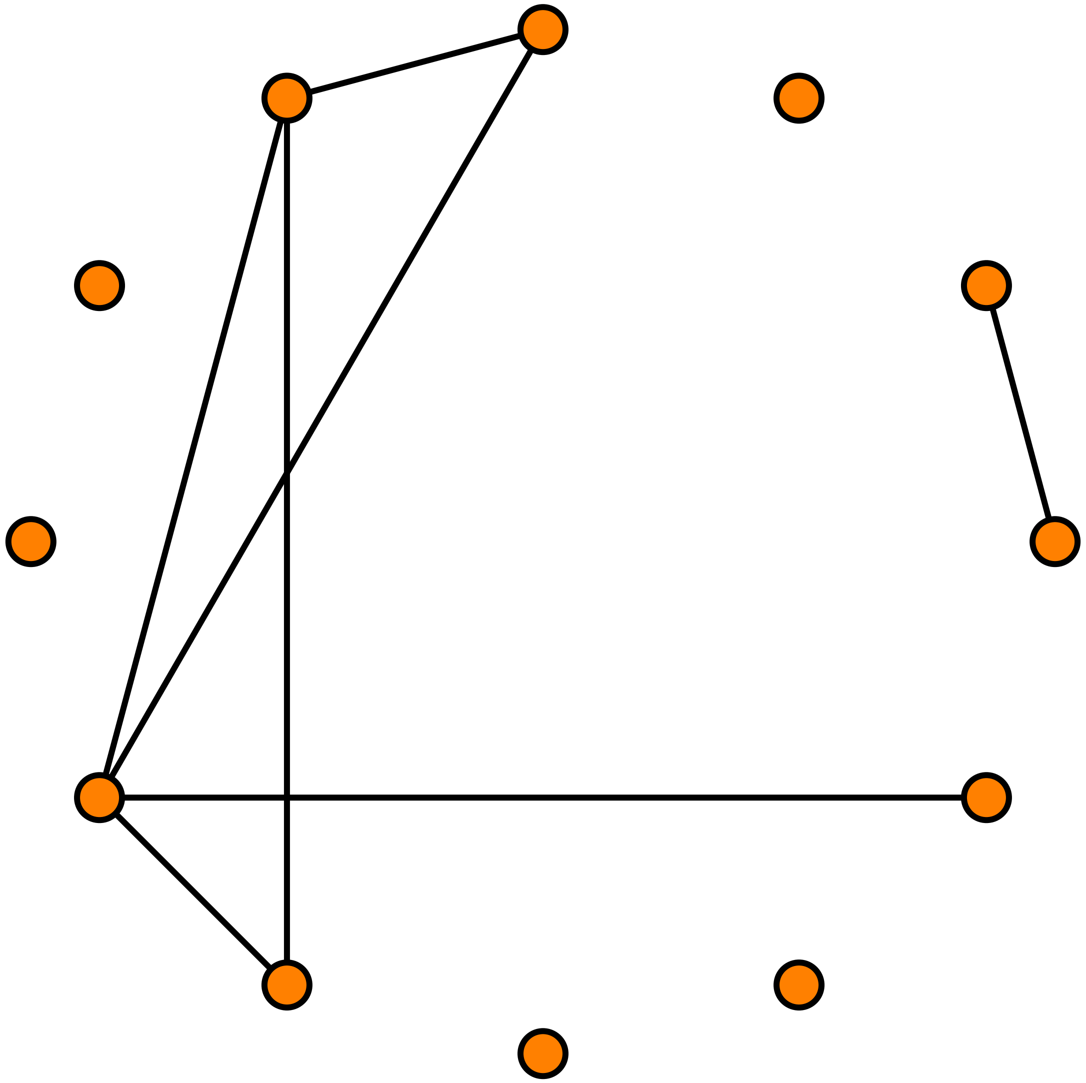


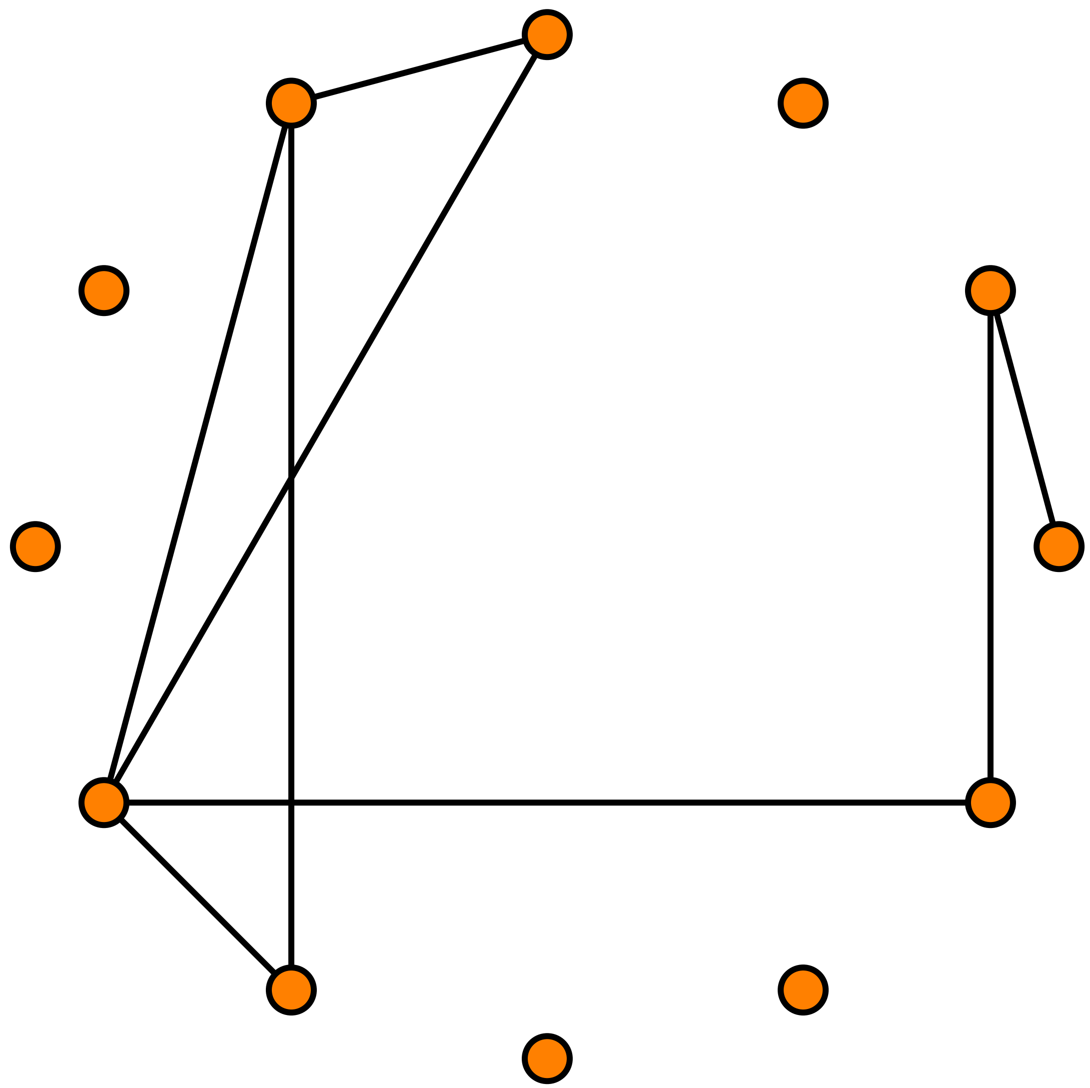


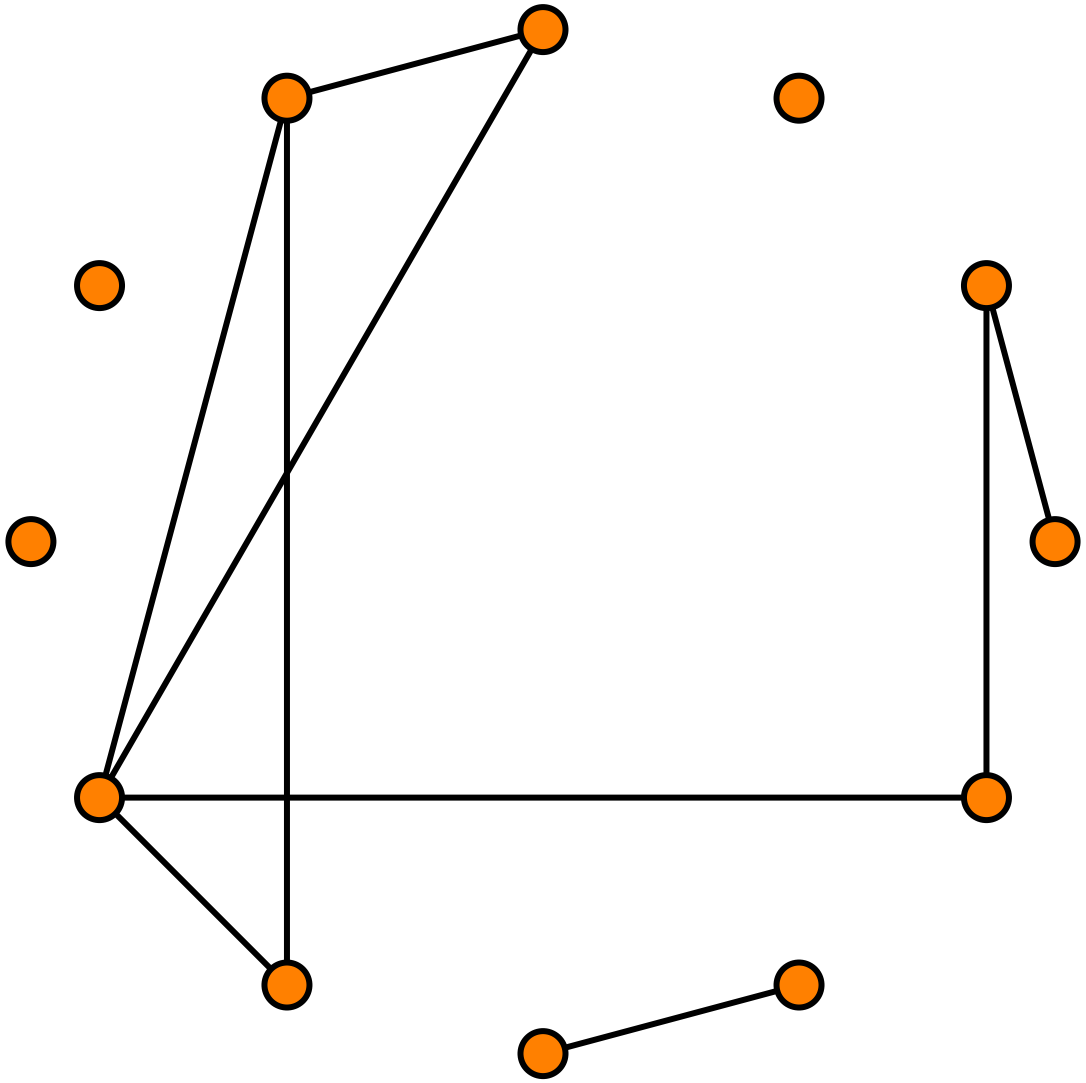


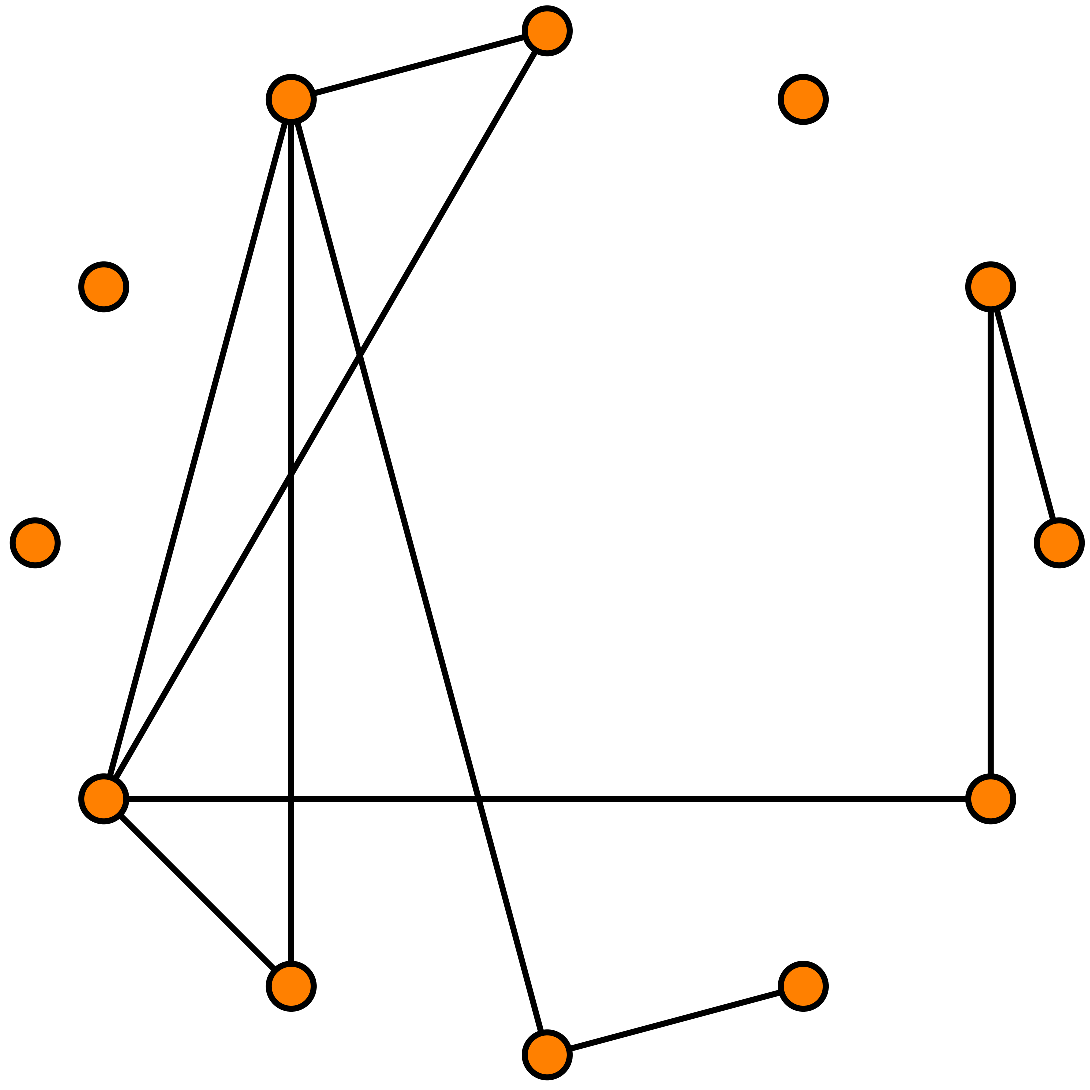


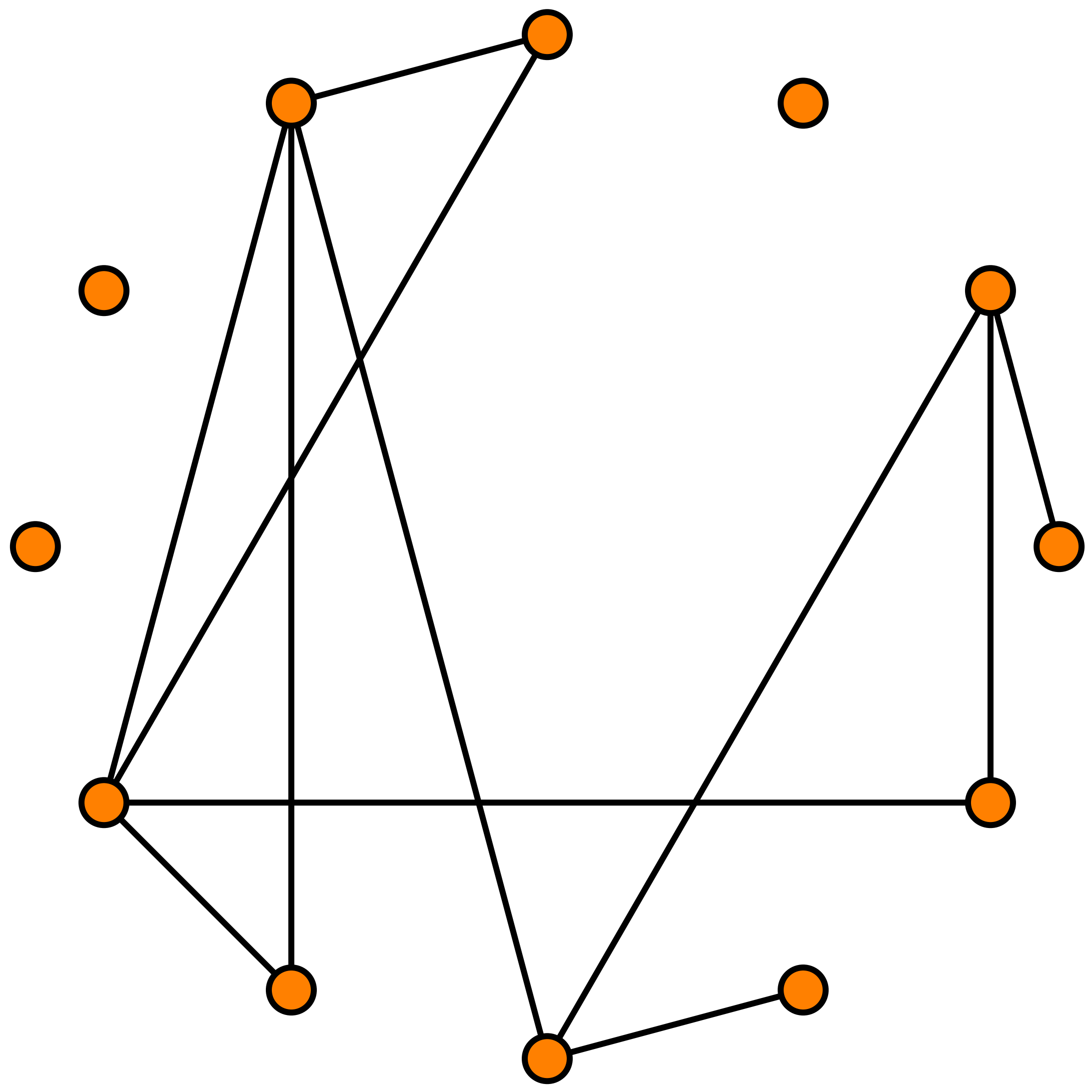


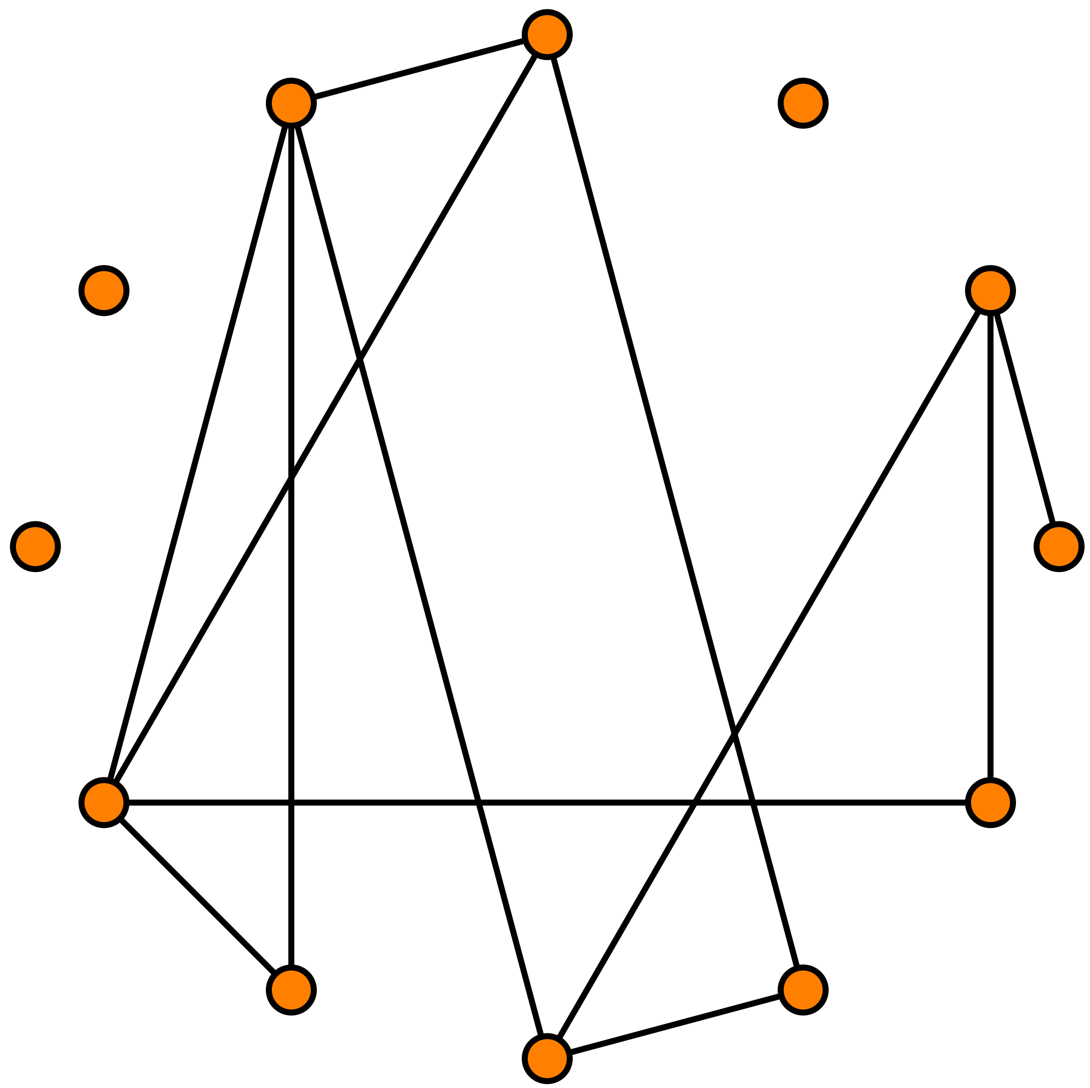


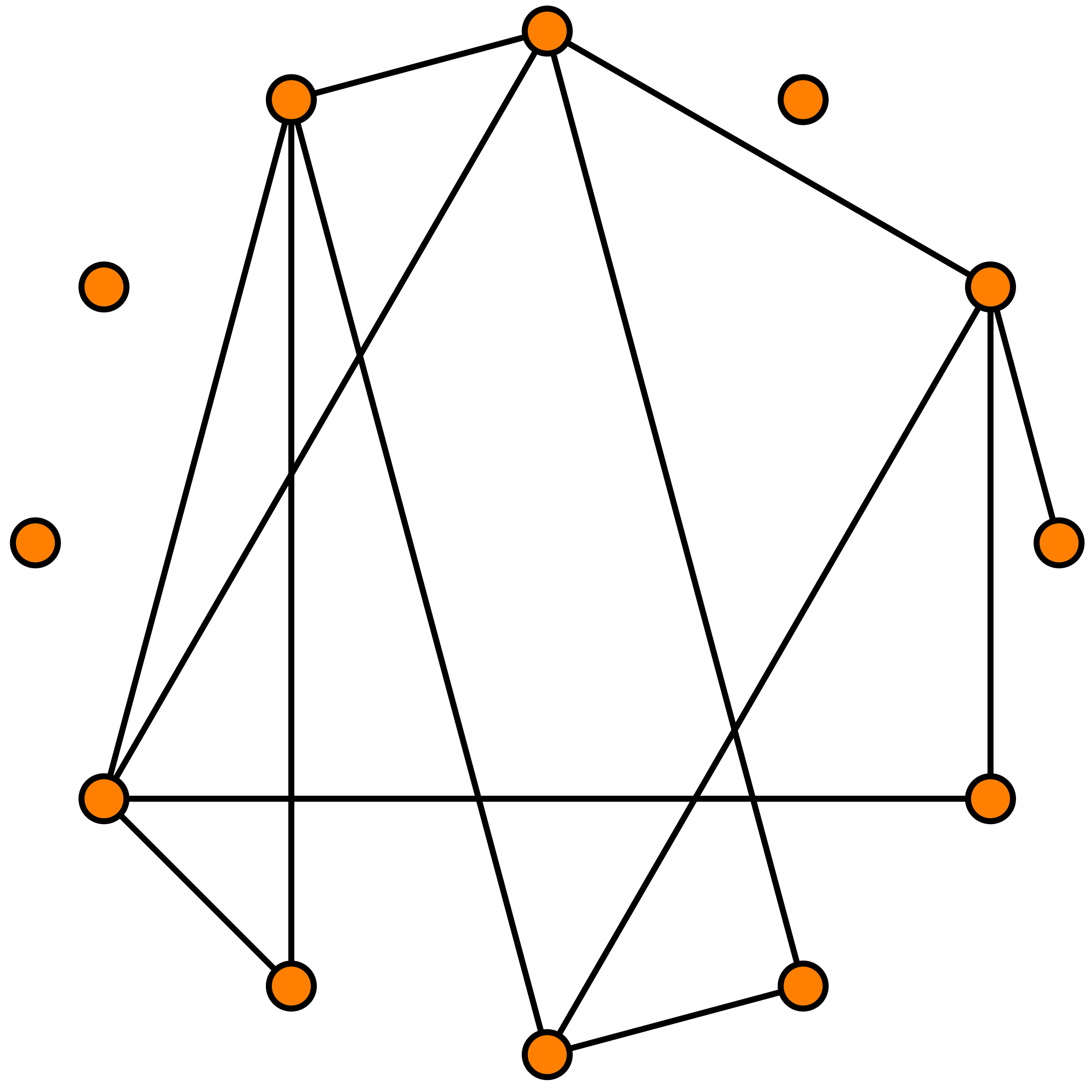


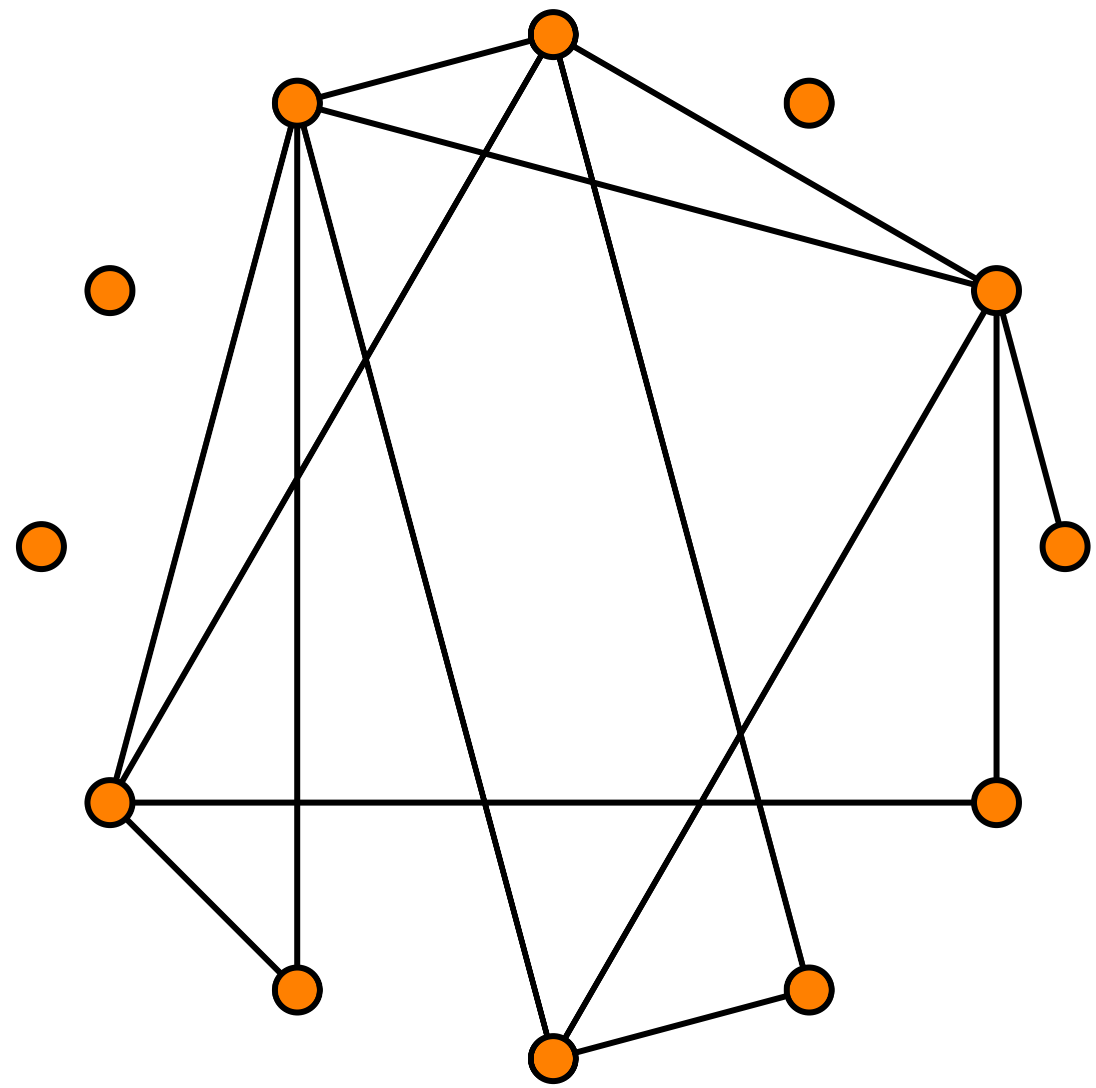


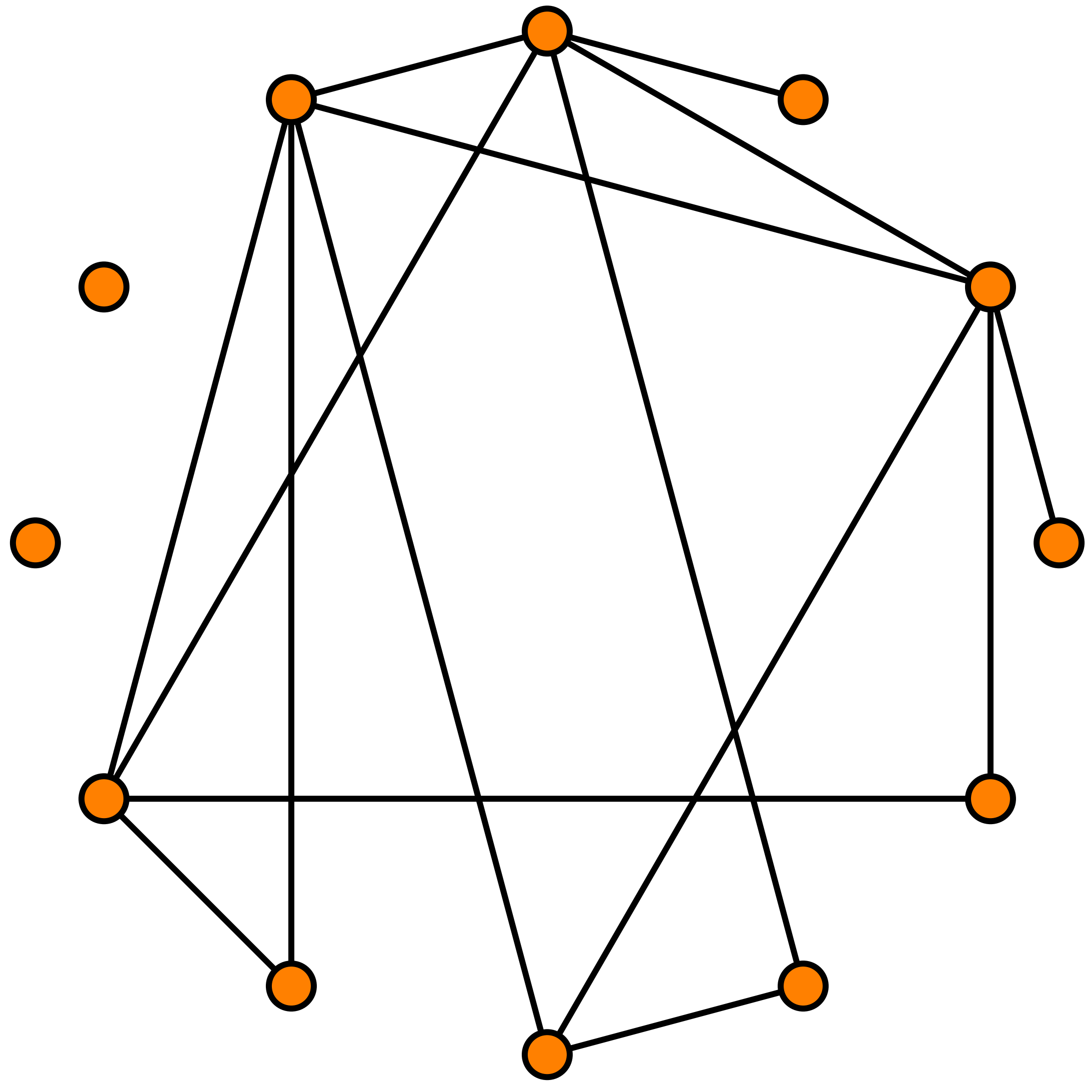


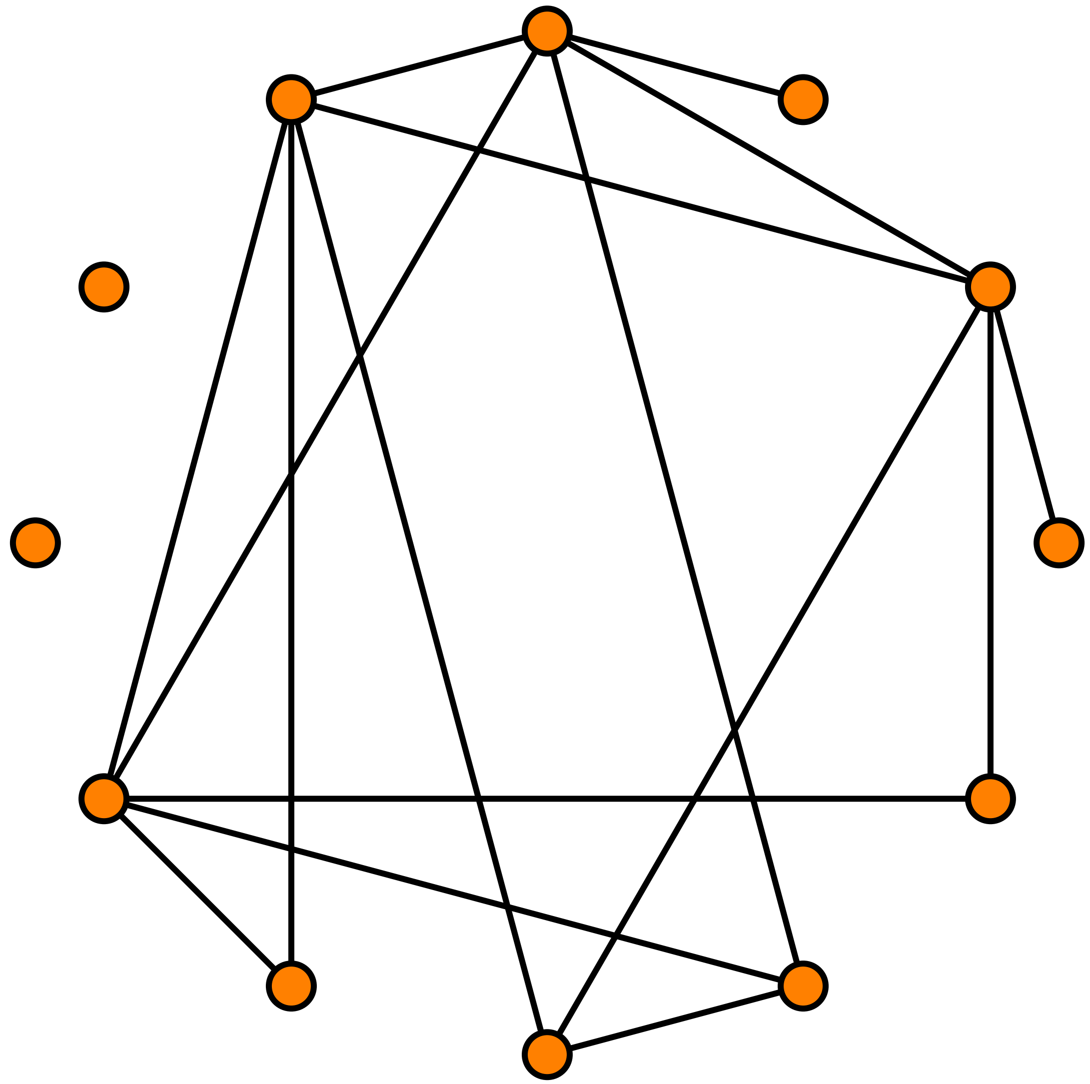


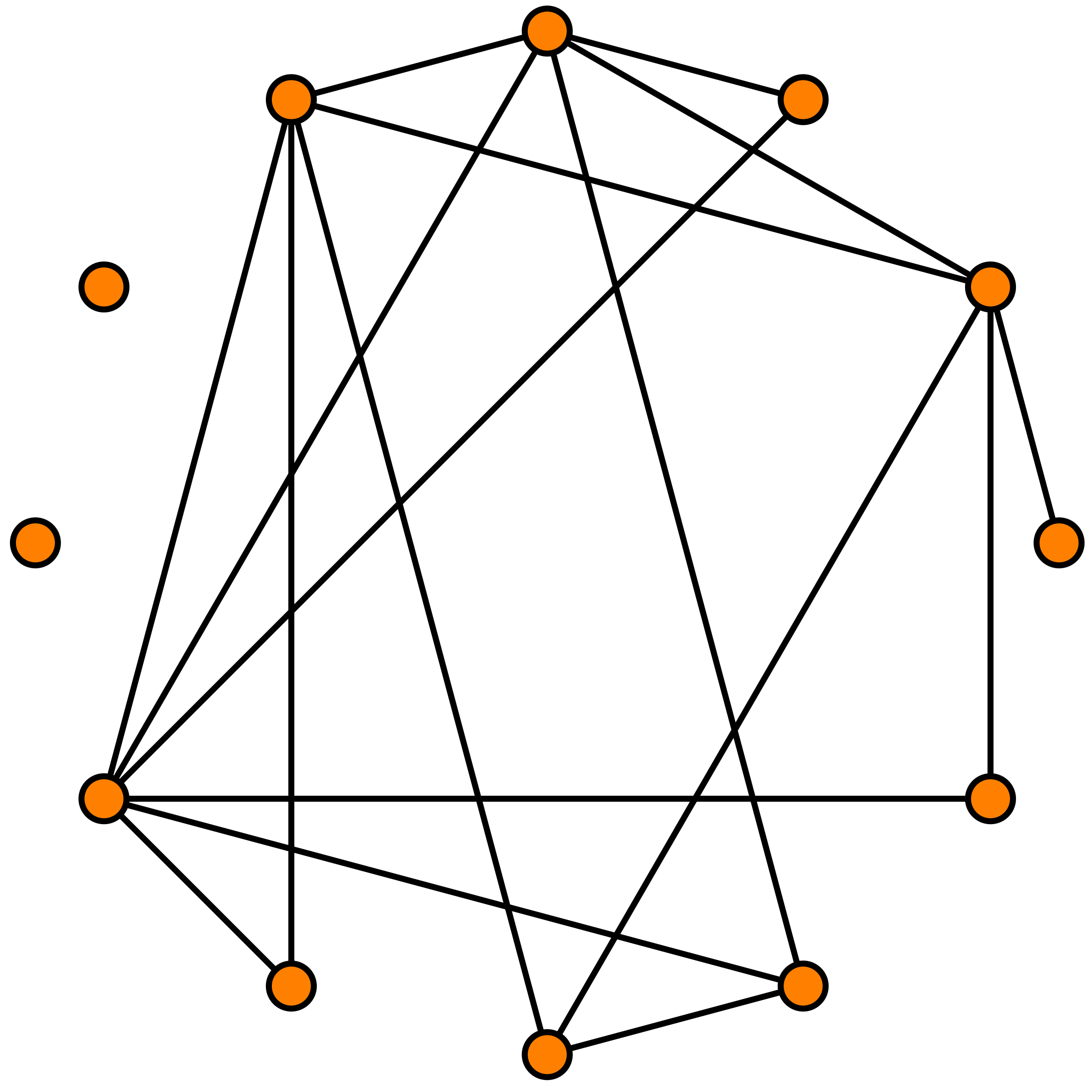


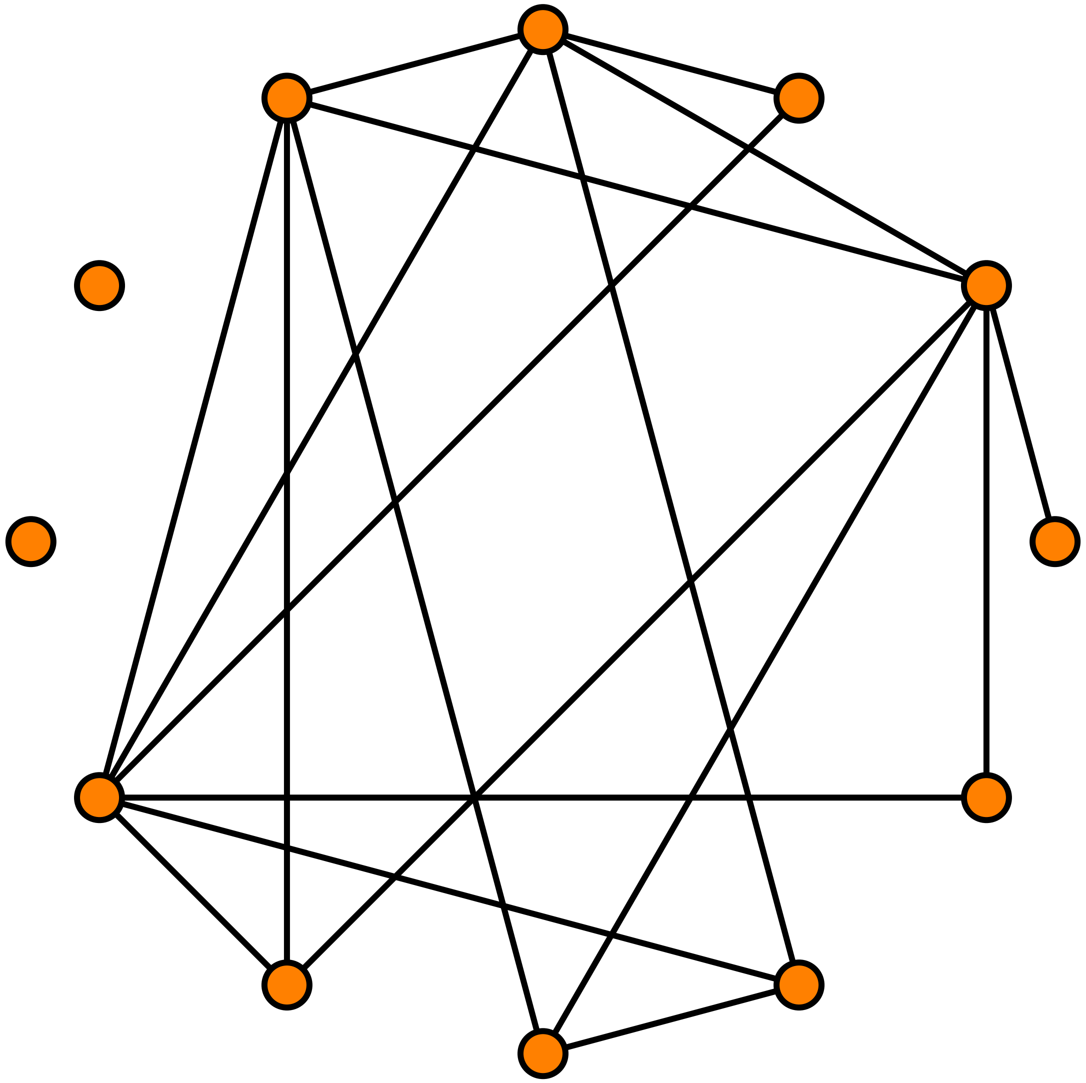


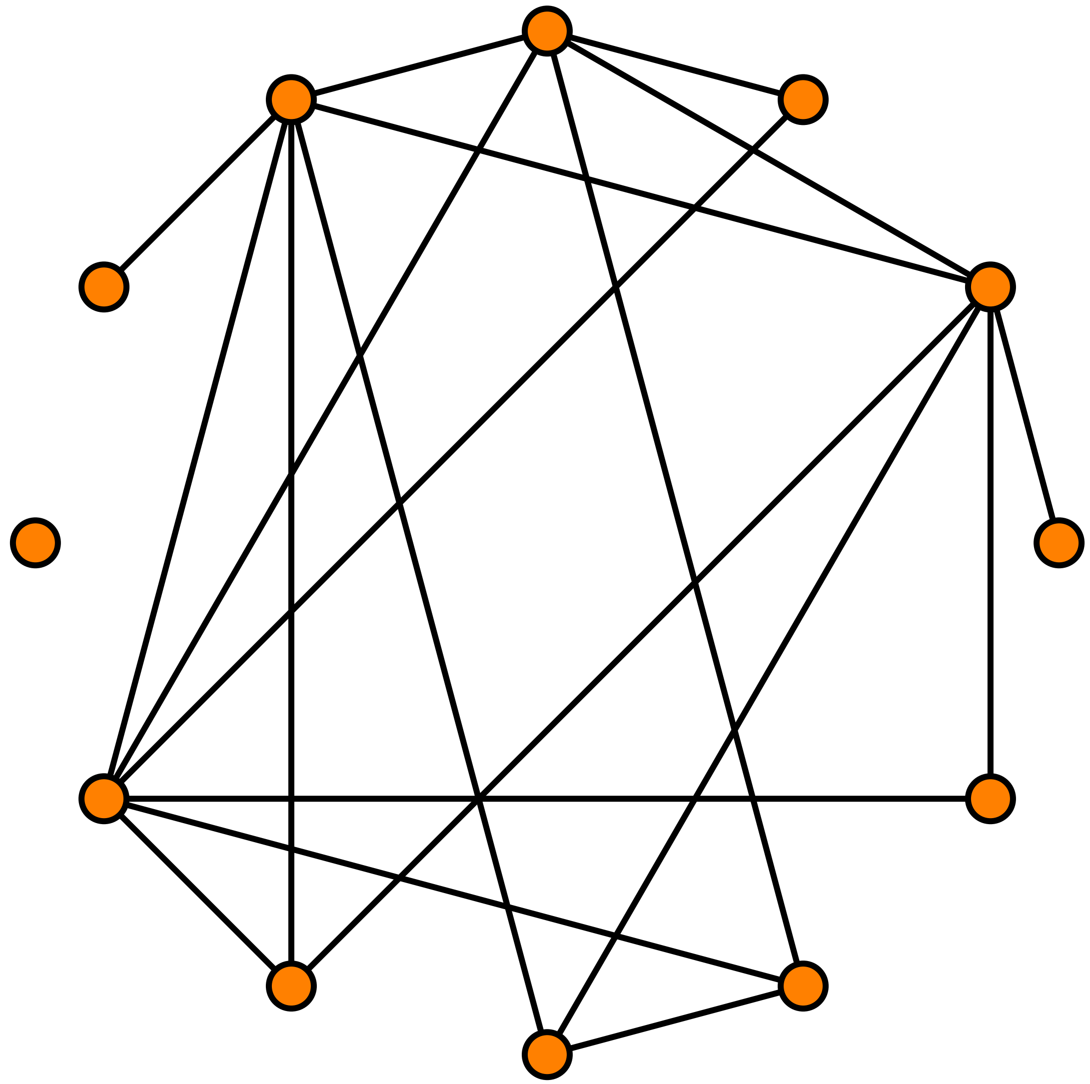


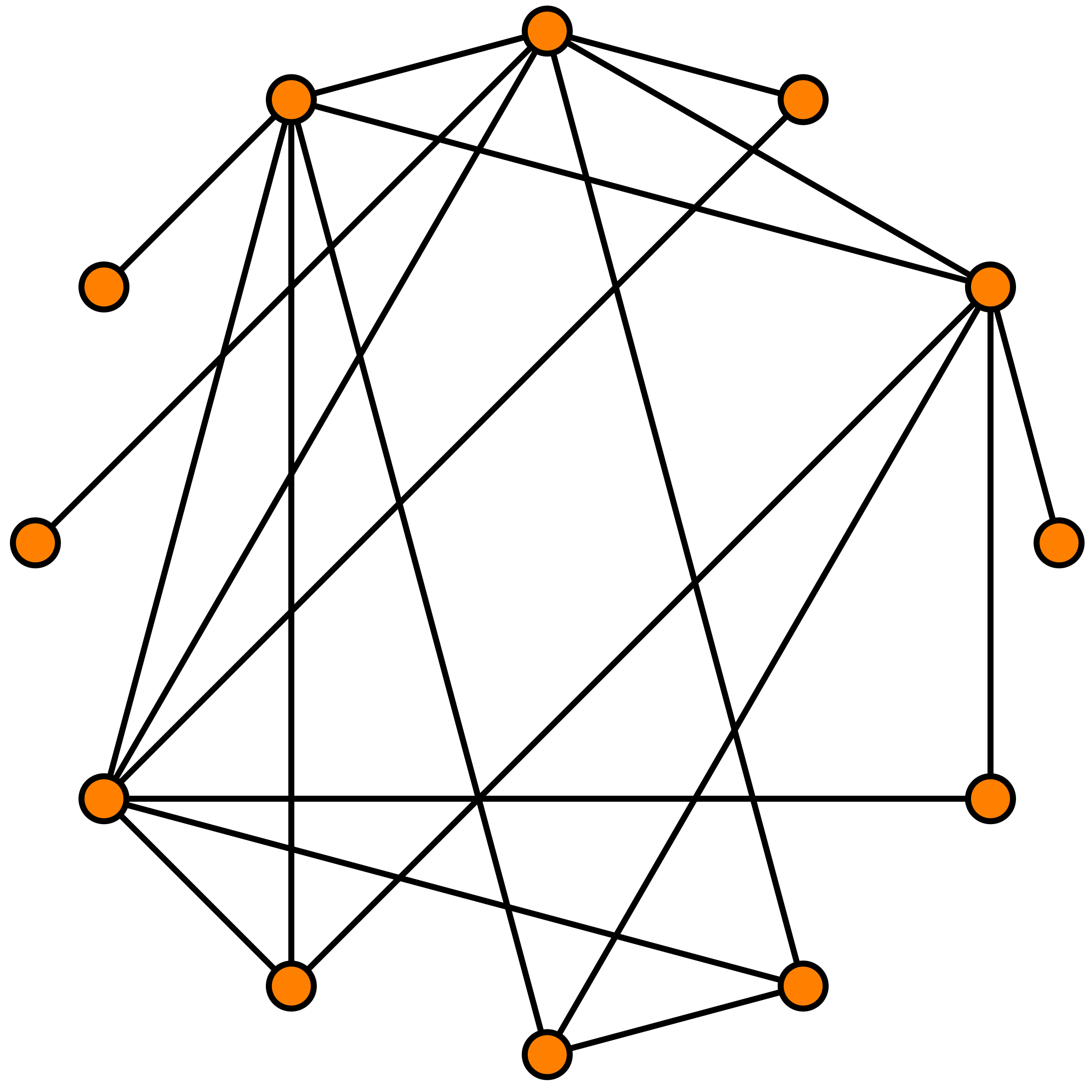












Erdős-Rényi theorem

Theorem (Erdős-Rényi)

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

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

Erdős-Rényi theorem

Theorem (Erdős-Rényi)

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

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

Theorem (Erdős-Rényi)

Let $c \in \mathbb{R}$ *be fixed and* $G \sim G(n,p)$. *If* $p = \frac{\log n + c}{n}$, *then* $\beta_0(G)$ *is asymptotically Poisson distributed with mean* e^{-c} *and*

$$\lim_{n\to\infty} \mathbb{P}[G \textit{ is connected }] = e^{-e^{-c}}.$$

Random simplicial complexes

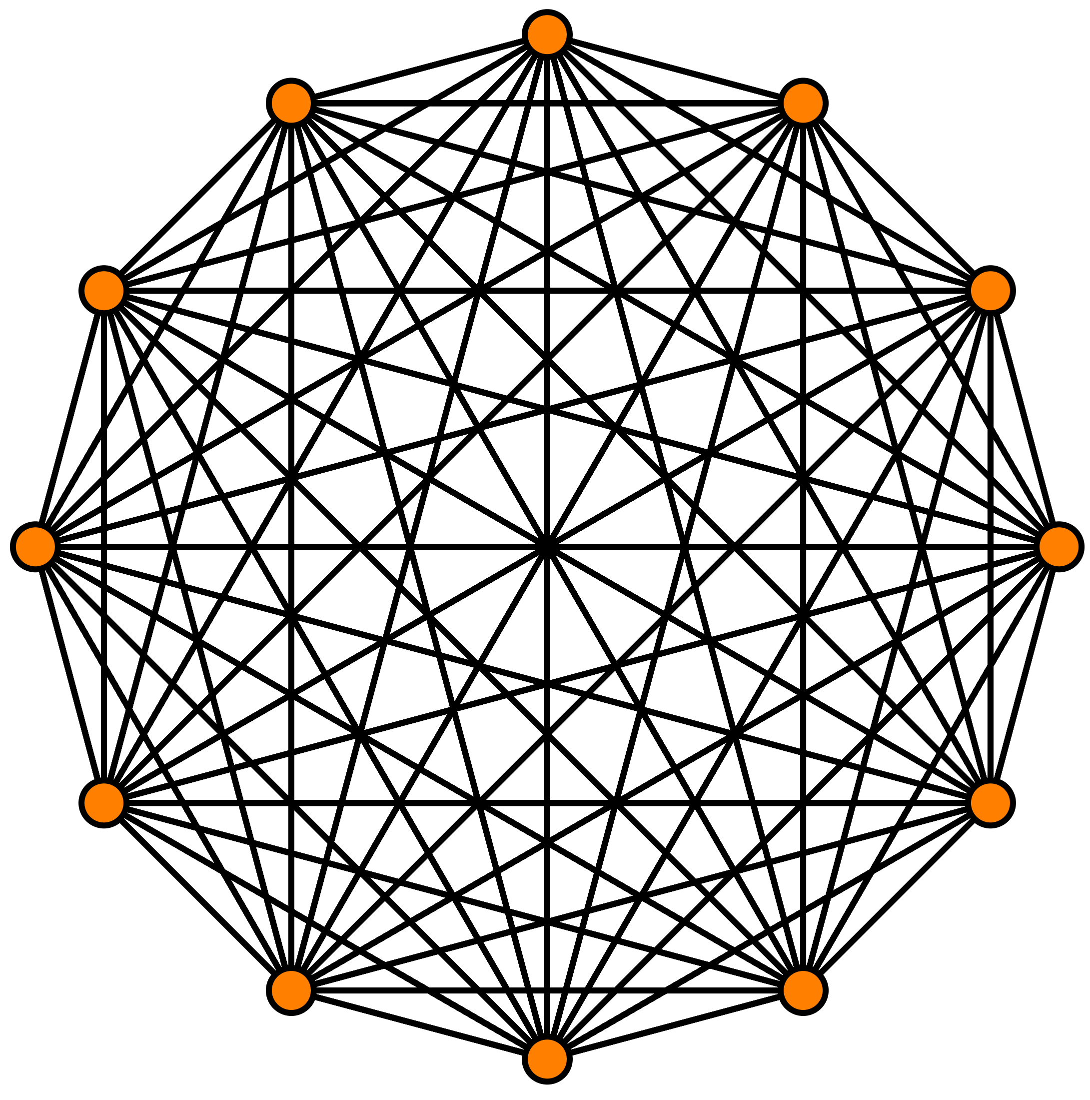
Lineal Meshulam 2-face model

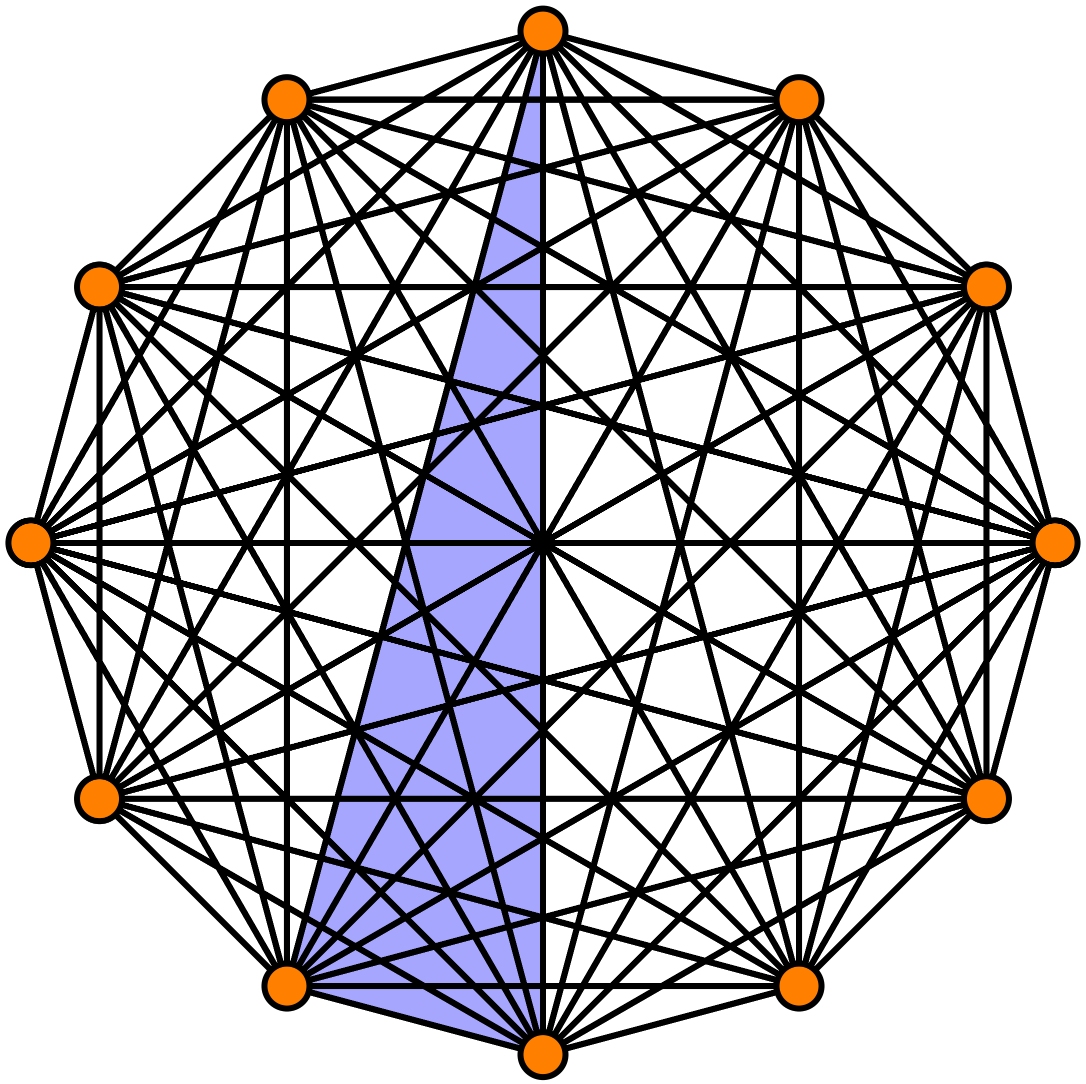
$Y_2(n, p)$ is the probability space of 2-dimensional simplicial complexes with vertex set $[n]$ and edge set $\binom{[n]}{2}$ and every 2-dimensional face is included with probability p, independently.

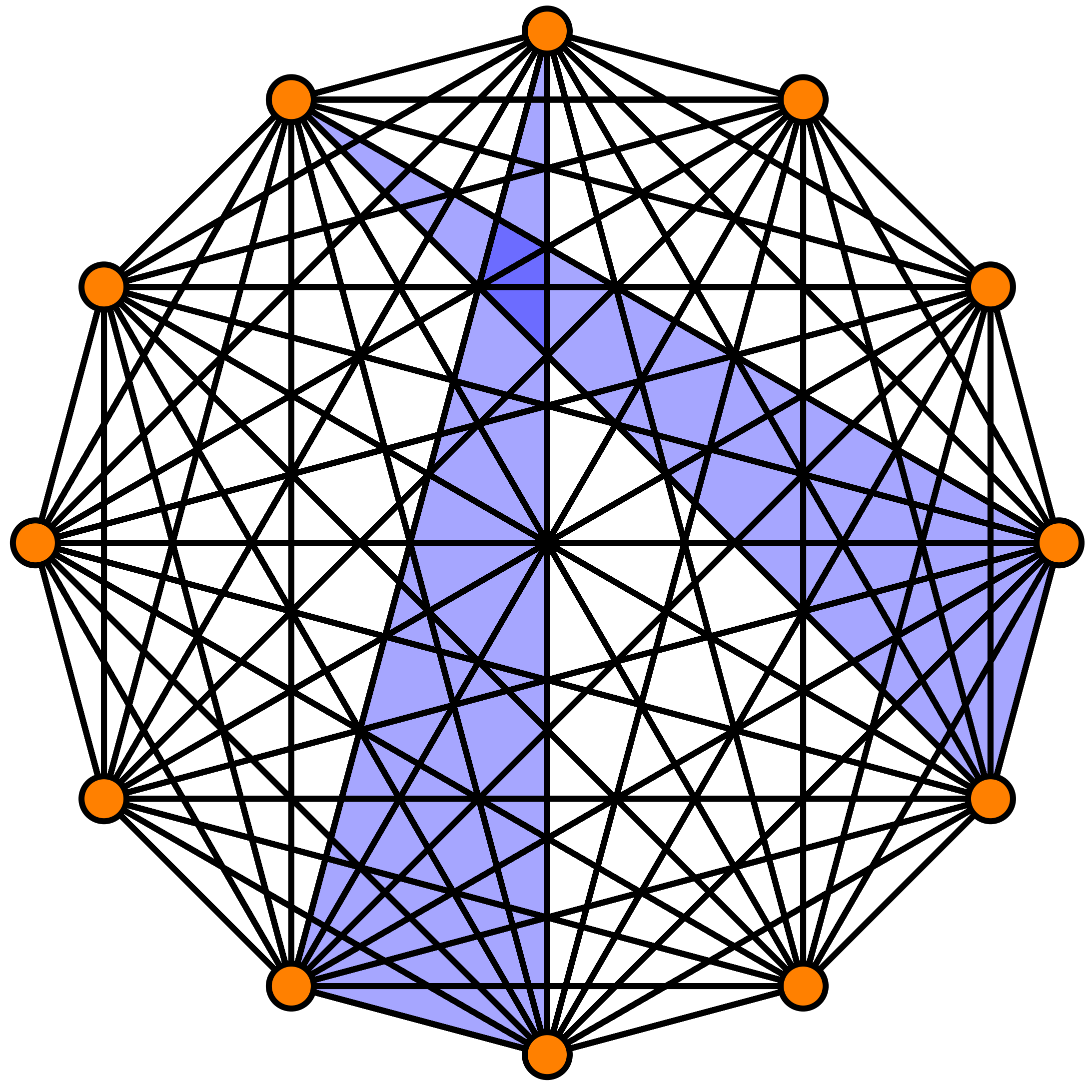
Lineal Meshulam 2-face model

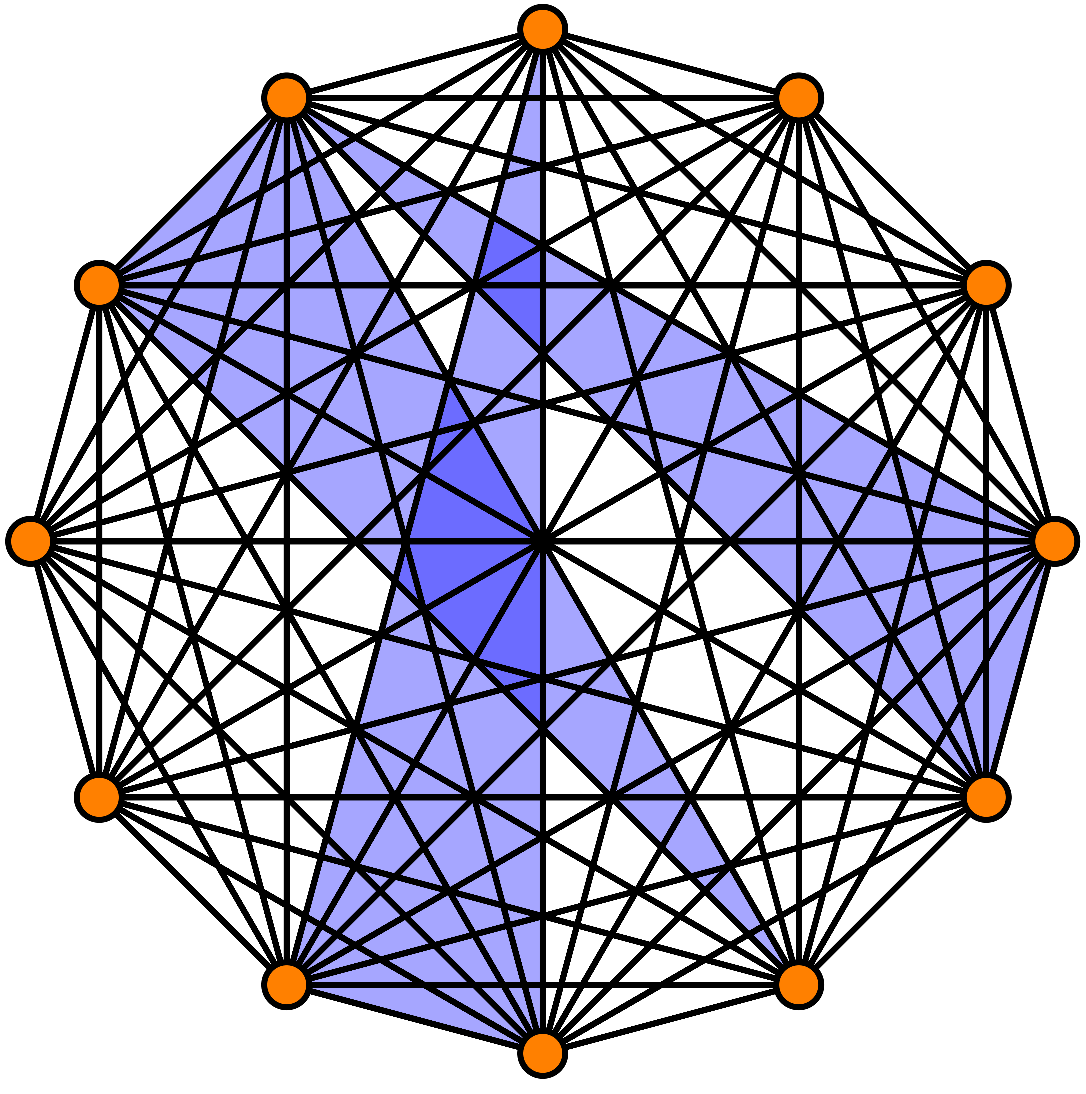
$Y_2(n, p)$ is the probability space of 2-dimensional simplicial complexes with vertex set $[n]$ and edge set $\binom{[n]}{2}$ and every 2-dimensional face is included with probability p, independently.

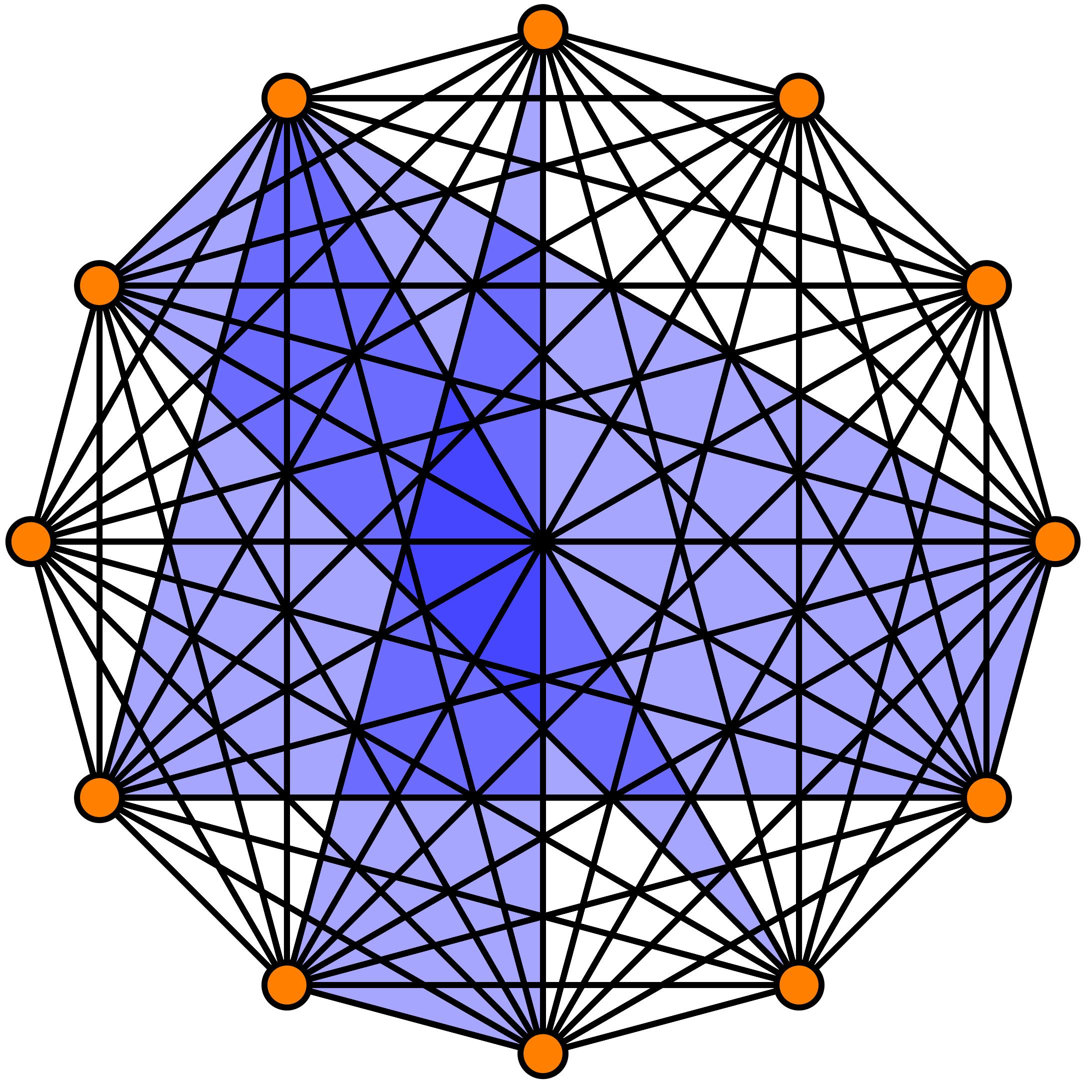
$Y \sim Y_2(n, p)$ is a draw of a random 2-simplicial complex Y from the distribution $Y_2(n, p)$

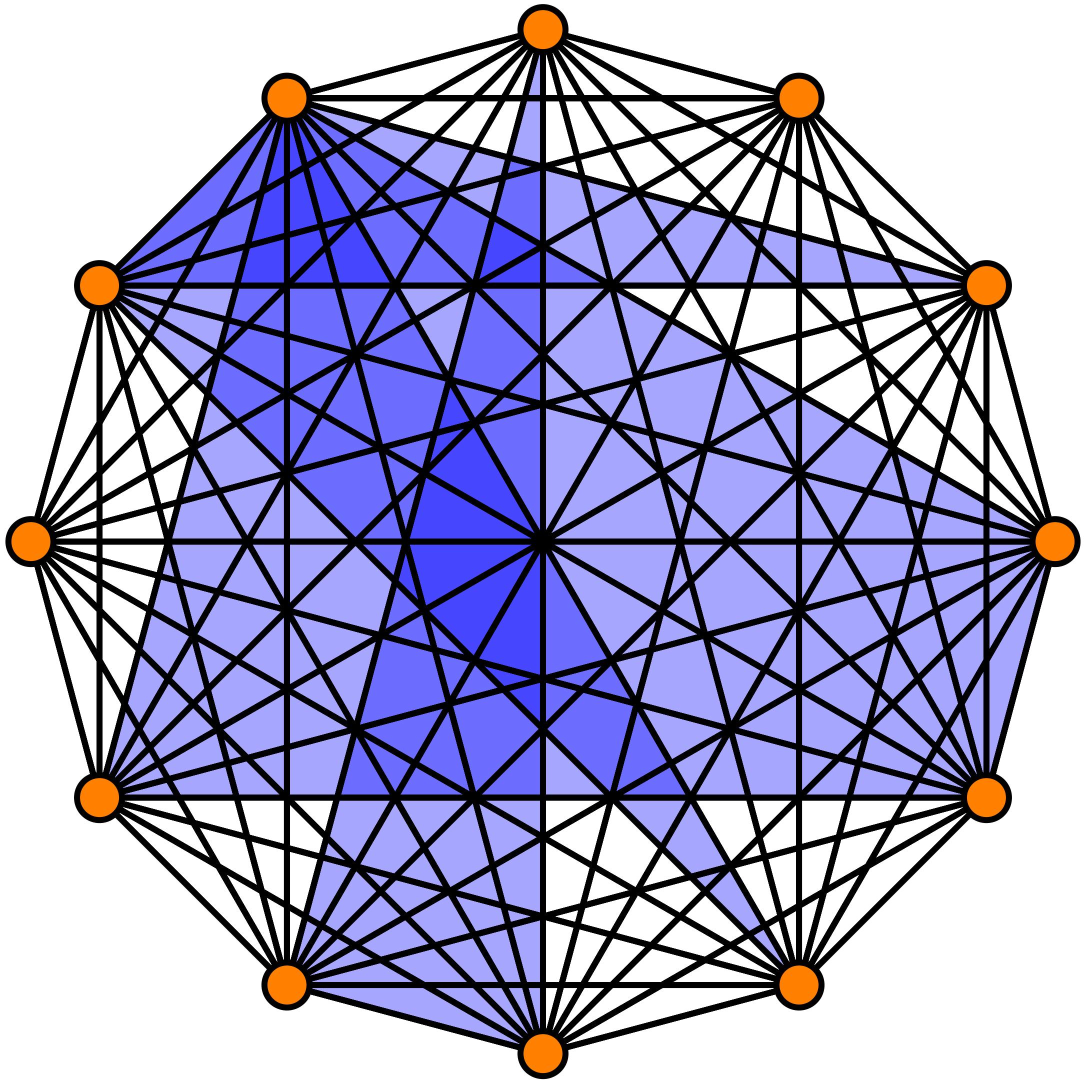


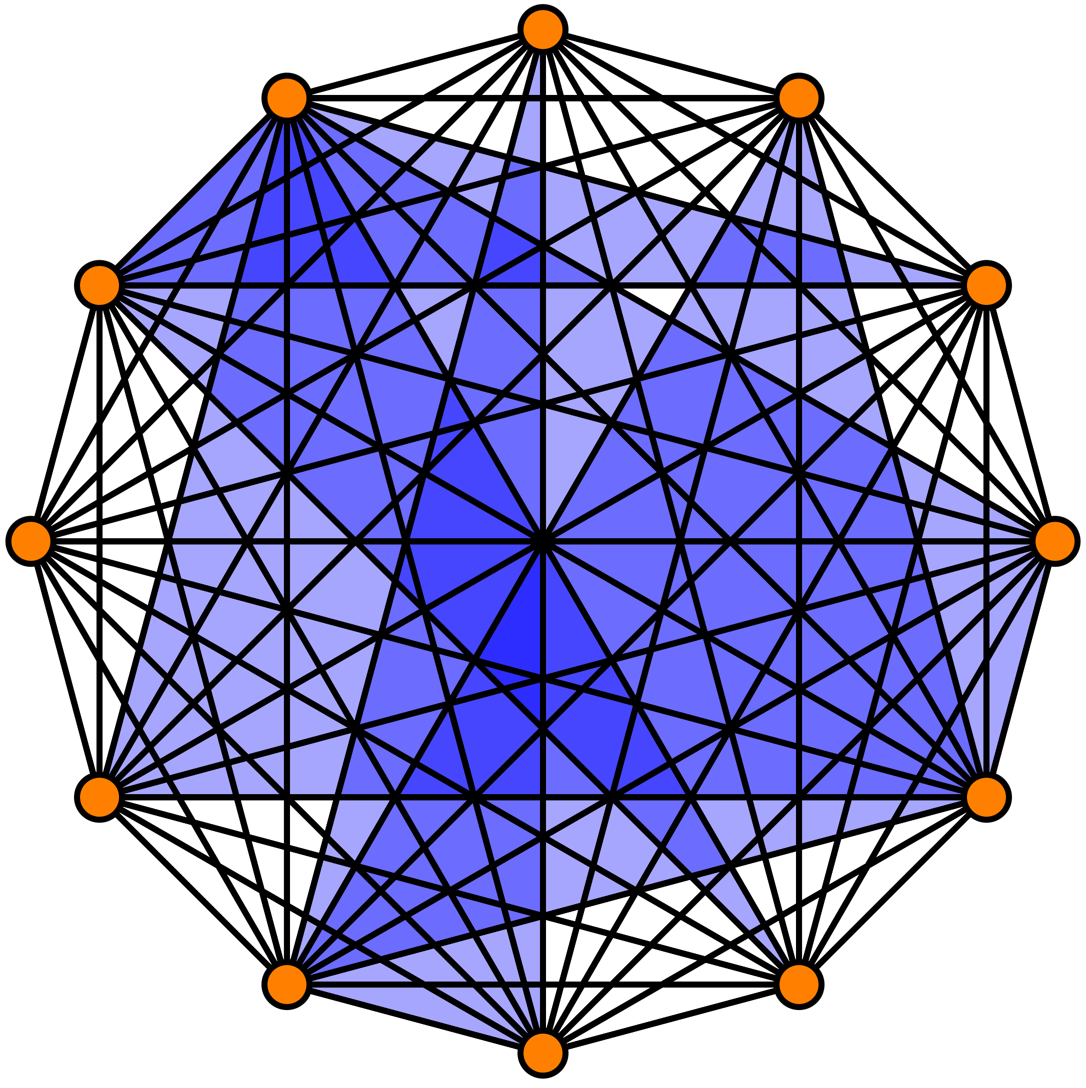


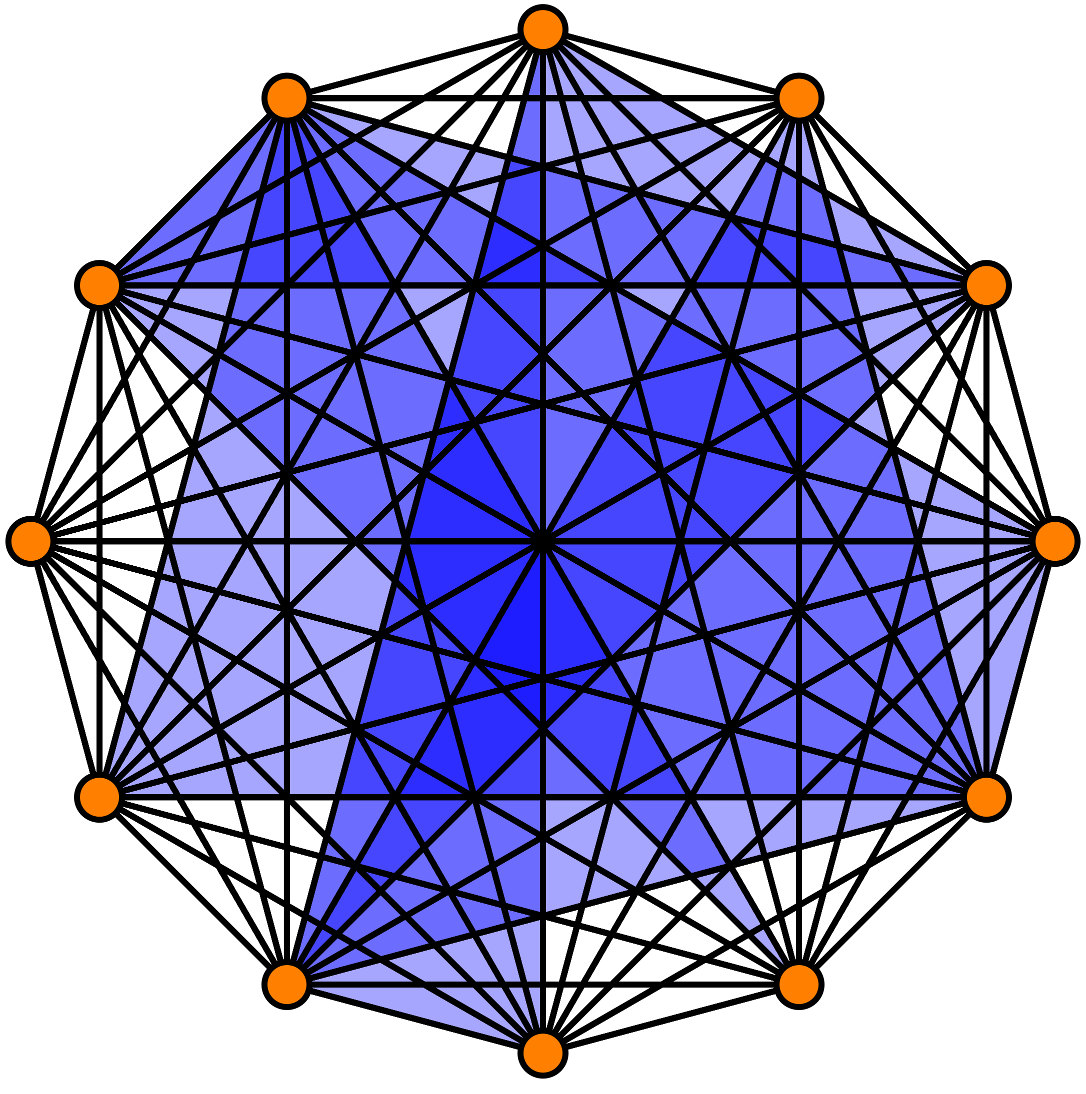


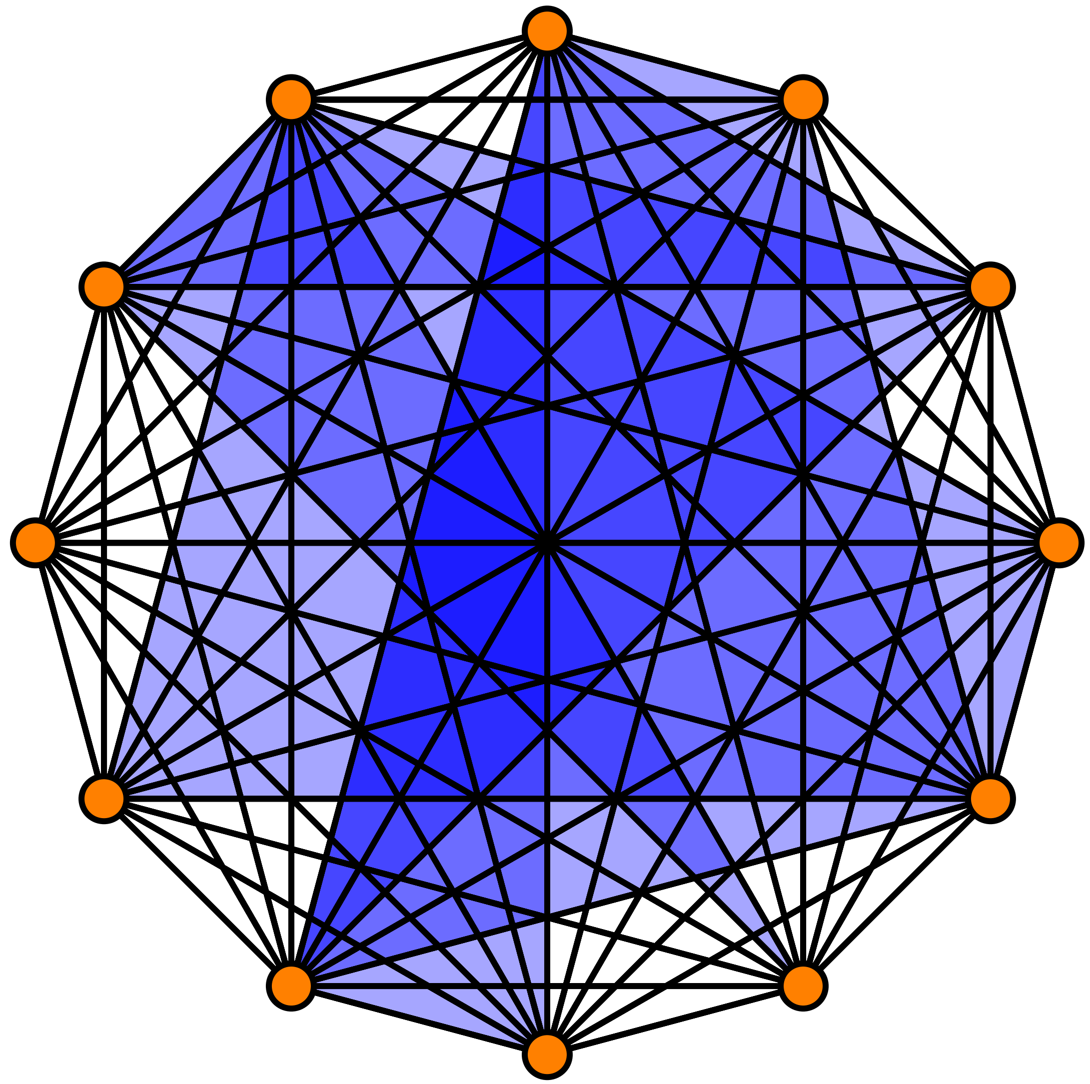


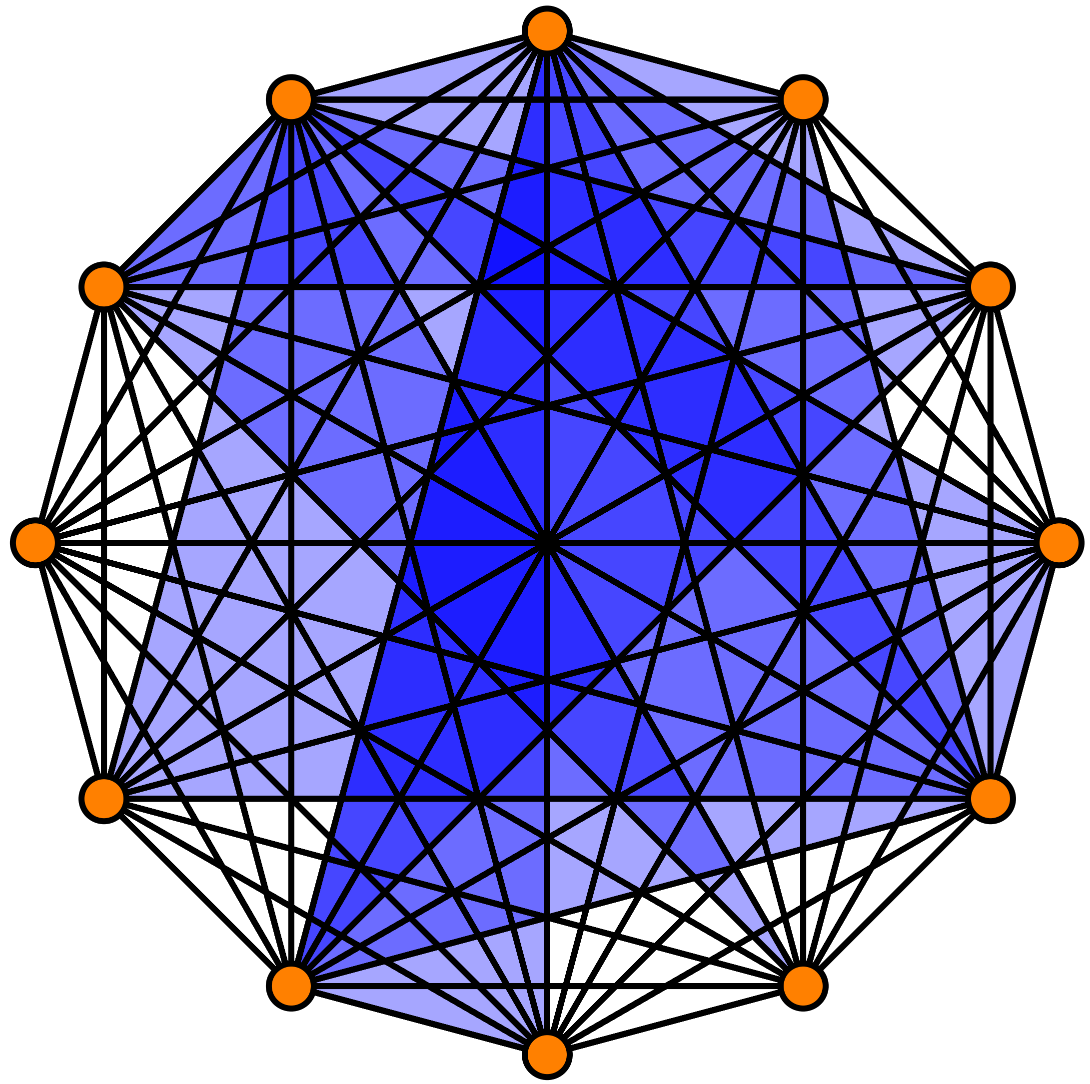












Lineal Meshulam model

$Y_d(n, p)$ is the probability space of d-dimensional simplicial complexes over all simplicial complexes on n vertices with complete $d - 1$ skeleton and every d-dimensional face is included with probability p, independently.

Lineal Meshulam model

$Y_d(n, p)$ is the probability space of d-dimensional simplicial complexes over all simplicial complexes on n vertices with complete $d - 1$ skeleton and every d-dimensional face is included with probability p, independently.

$Y \sim Y_d(n, p)$ is a draw of a random d-simplicial complex Y from the distribution $Y_d(n, p)$

Percolation on complexes

Theorem (Linial,Meshulam)

Let $\epsilon > 0$ *be fixed and* $Y \sim Y_2(n, p)$*. Then*

$$\mathbb{P}[H_1(Y, \mathbb{Z}/2) = 0] \longrightarrow \begin{cases} 1 & : p \geq (2+\epsilon)\log n/n \\ 0 & : p \leq (2-\epsilon)\log n/n. \end{cases}$$

Percolation on complexes

Theorem (Linial,Meshulam)

Let $\epsilon > 0$ *be fixed and* $Y \sim Y_2(n, p)$. *Then*

$$\mathbb{P}[H_1(Y, \mathbb{Z}/2) = 0] \longrightarrow \begin{cases} 1 & : p \geq (2 + \epsilon) \log n/n \\ 0 & : p \leq (2 - \epsilon) \log n/n. \end{cases}$$

Theorem (Meshulam, Wallach)

Let $d \geq 2$, $\ell \geq 2$, *and* $\epsilon > 0$ *be fixed and* $Y \sim Y_d(n, p)$. *Then*

$$\mathbb{P}[H_{d-1}(Y, \mathbb{Z}/\ell) = 0] \longrightarrow \begin{cases} 1 & : p \geq (d + \epsilon) \log n/n \\ 0 & : p \leq (d - \epsilon) \log n/n. \end{cases}$$

Simple connectivity

Theorem (Babson, Hoffman, Kahle)

Let $\epsilon > 0$ *be fixed and* $Y \sim Y_2(n, p)$. *Then*

$$\mathbb{P}[\pi_1(Y) = 0] \longrightarrow \begin{cases} 1 & : p \geq \frac{n^{\epsilon}}{\sqrt{n}} \\ 0 & : p \leq \frac{n^{-\epsilon}}{\sqrt{n}} \end{cases}$$

Embedding

Every d-dimensional simplicial complex is embeddable in $\mathbb{R}^{2d+1}$ but not necessarily in $\mathbb{R}^{2d}$.

Embedding

Every d-dimensional simplicial complex is embeddable in $\mathbb{R}^{2d+1}$ but not necessarily in $\mathbb{R}^{2d}$.

Theorem (Wagner)

There exist constants $c_1, c_2 > 0$ such that for $Y \sim Y_d(n, p)$

- *if $p < c_1/n$ then with high probability Y is embeddable $\mathbb{R}^{2d}$,*
- *if $p > c_2/n$ then with high probability Y is not embeddable $\mathbb{R}^{2d}$.*

Random clique complex

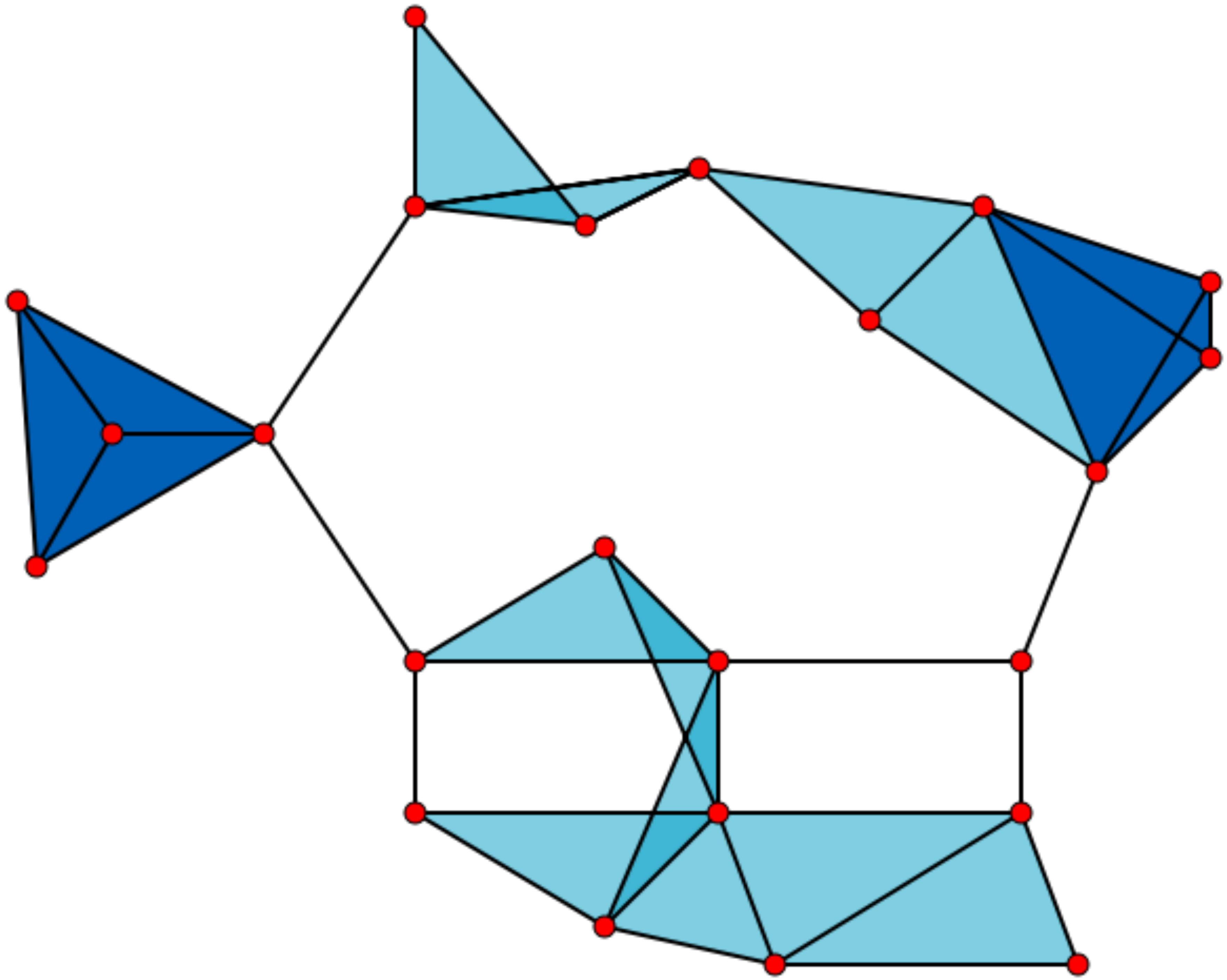
$X_d(n, p)$ is the probability space of d-dimensional random clique complexes. Given a random graph $G \sim G(n, p)$, the clique complex X is generated by including as a faces of the clique complex $X(G)$ complete subgraphs of G.

Random clique complex

$X_d(n,p)$ is the probability space of d-dimensional random clique complexes. Given a random graph $G \sim G(n,p)$, the clique complex X is generated by including as a faces of the clique complex $X(G)$ complete subgraphs of G.

$X \sim X_d(n,p)$ is a draw of a random d-clique complex X from the distribution $X_d(n,p)$.

Random clique complex



Homology vanishing results

Theorem (Kahle)

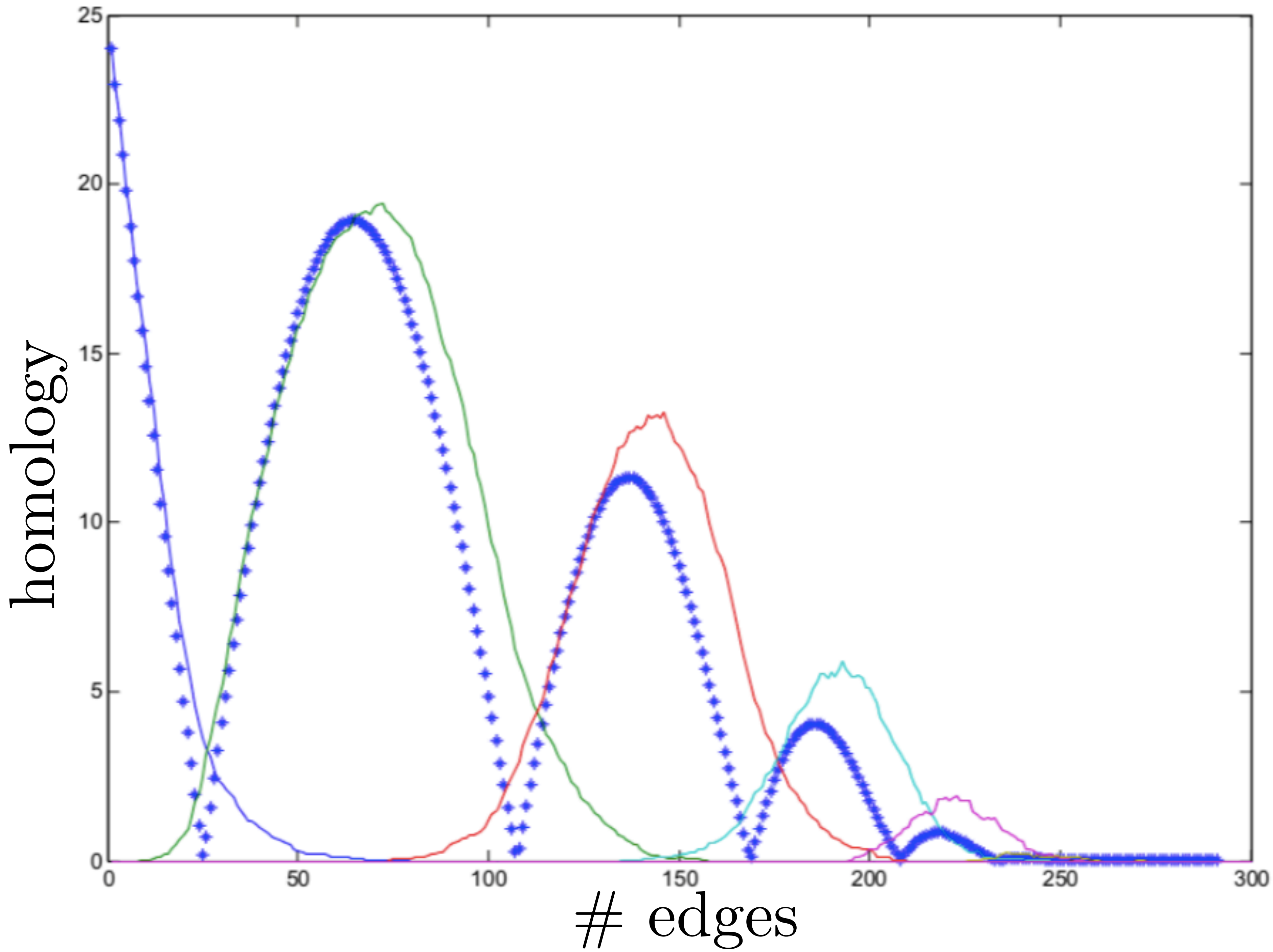
Fix $k \geq 1$ and $X \sim X_d(n, p)$ and $\omega(1)$ a function that tends to infinity arbitrarily slowly. Then with high probability $H_k(X, \mathbb{R}) = 0$ if

$$p \geq \left(\frac{(k/2+1)\log n + \frac{k}{2}\log\log n + \omega(1)}{n} \right)^{1/(k+1)}$$

and $H_k(X, \mathbb{R}) \neq 0$ if

$$p \in \left[1/n^k, \left(\frac{(k/2+1)\log n + \frac{k}{2}\log\log n - \omega(1)}{n} \right)^{1/(k+1)} \right].$$

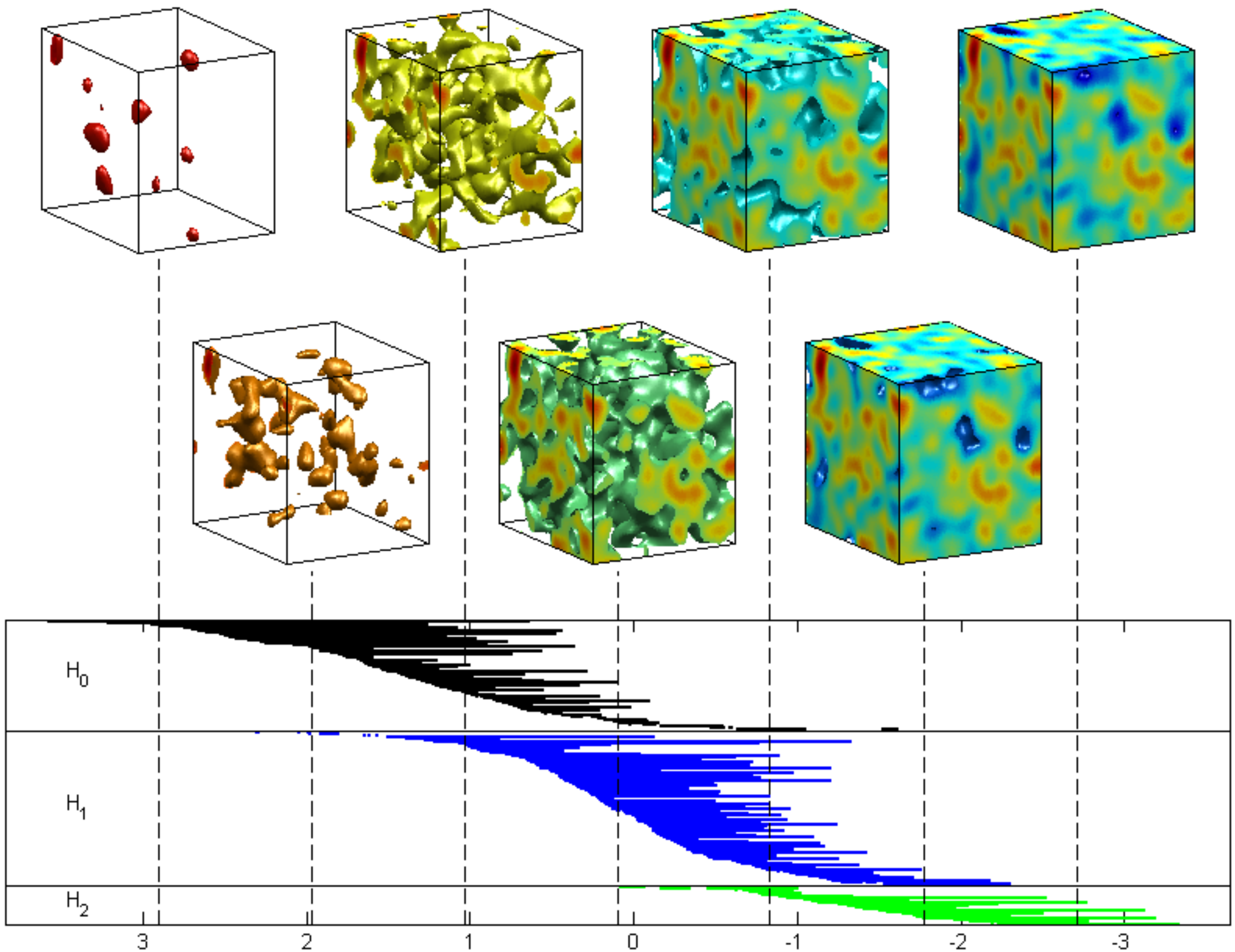
Random flag complex



Kahle and Nanda: $\mathbb{E}[\chi(X)]$ in blue for $X(25, p)$ with β_1 in green, β_2 in red, β_3 in cyan, and β_5 in purple.

Percolation and manifolds

Stochastic topology



Robert Adler

Problem statement

Given points $\mathcal{P} = \{X_i, ..., X_n\} \overset{iid}{\sim} \rho$ where the support of ρ is a manifold $\mathcal{M}$.

Problem statement

Given points $\mathcal{P} = \{X_i, ..., X_n\} \overset{iid}{\sim} \rho$ where the support of ρ is a manifold $\mathcal{M}$.

The empirical object is

$$\mathcal{U}(\mathcal{P}, r) = \bigcup_{p \in \mathcal{P}} B_{r(n)}(p).$$

Topology of noise

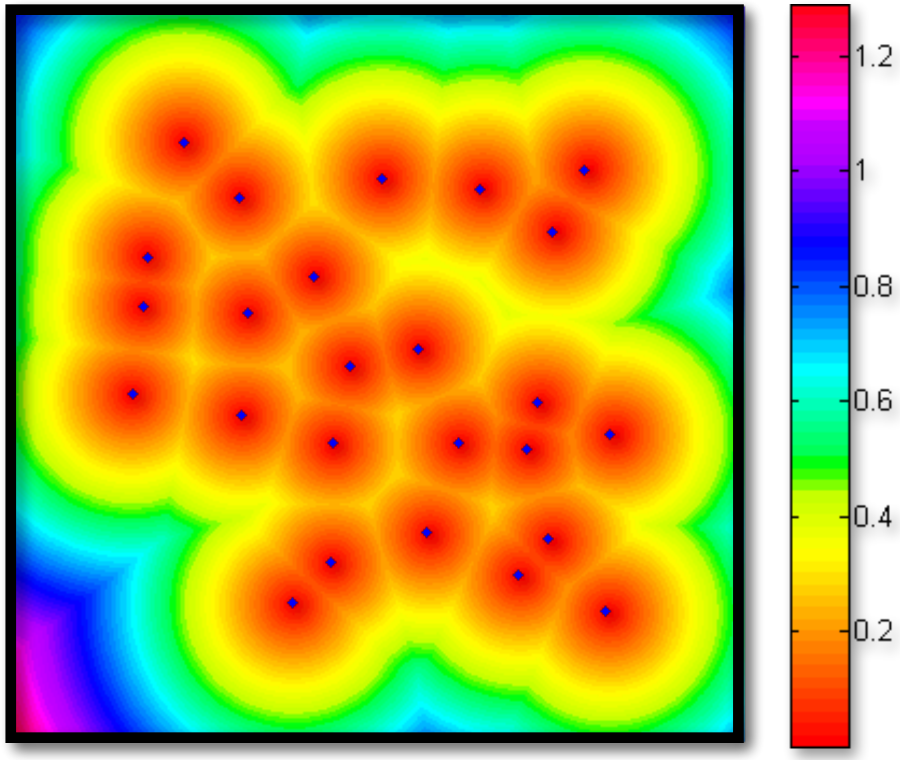
Random point cloud

$$\mathcal{P} \subset \mathbb{R}^d$$

Objects of study:

1. Union of Balls - $\mathcal{U}(\mathcal{P}, r)$

2. Distance Function - $d_{\mathcal{P}}(x)$



Homology (Betti numbers)

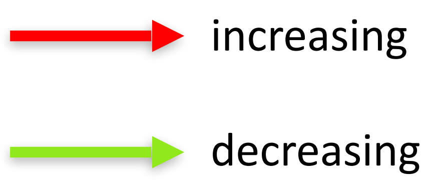
Critical points

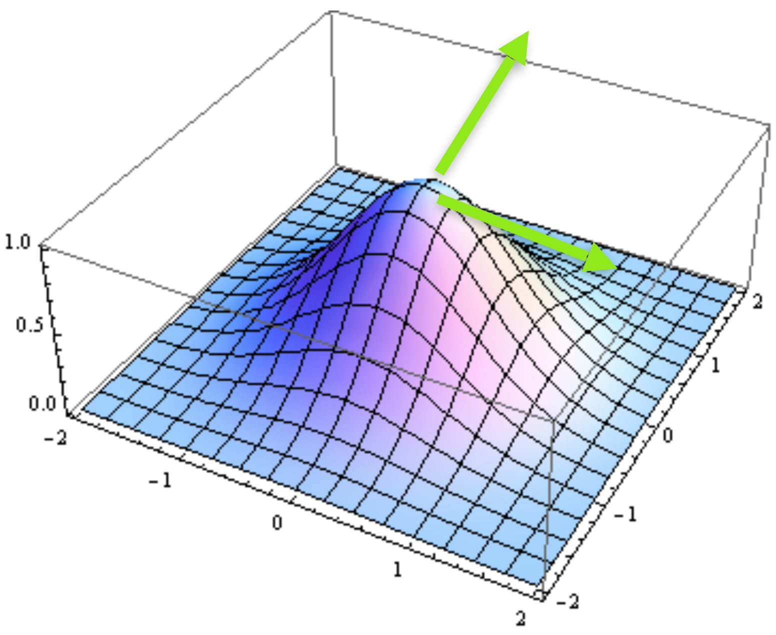
Critical points

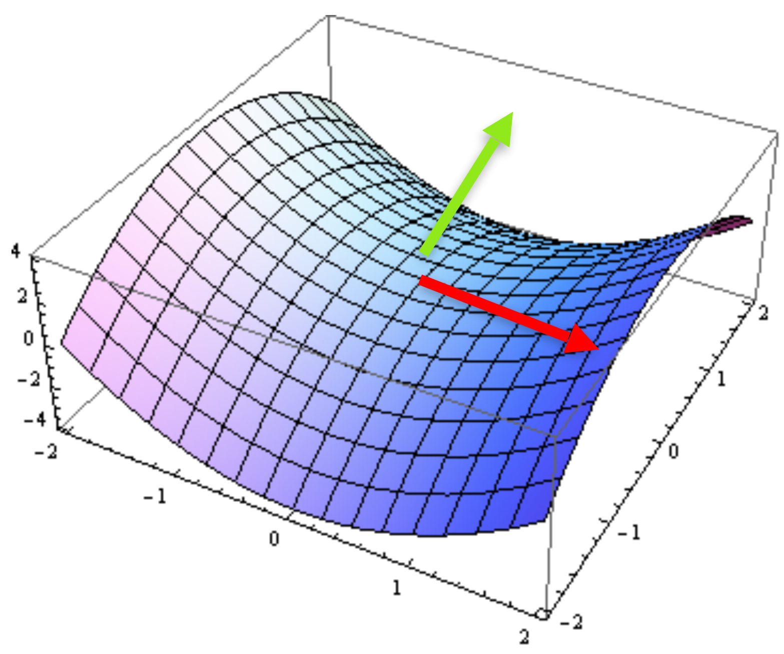
$$f : \mathbb{R}^d \to \mathbb{R}$$

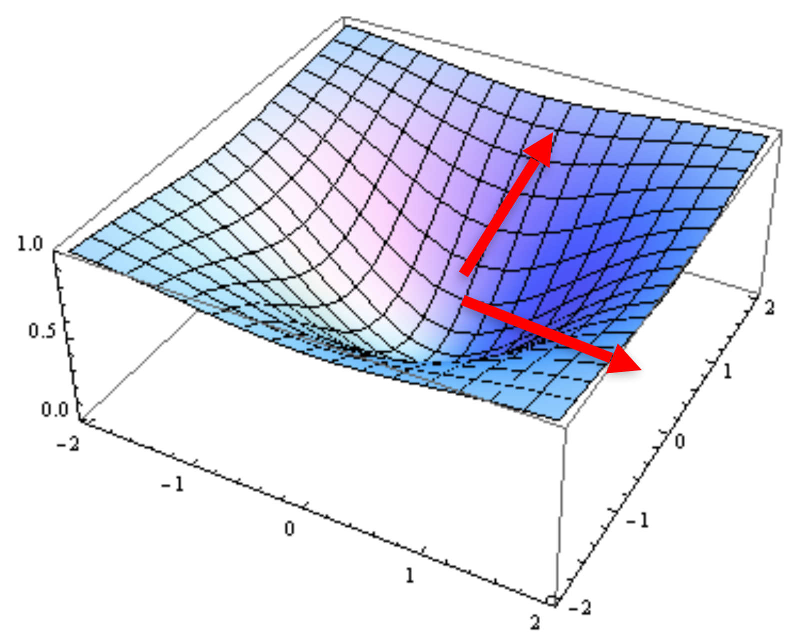
$$c \in \mathbb{R}^d \qquad \nabla f(c) = 0$$

$$\mu(c)$$









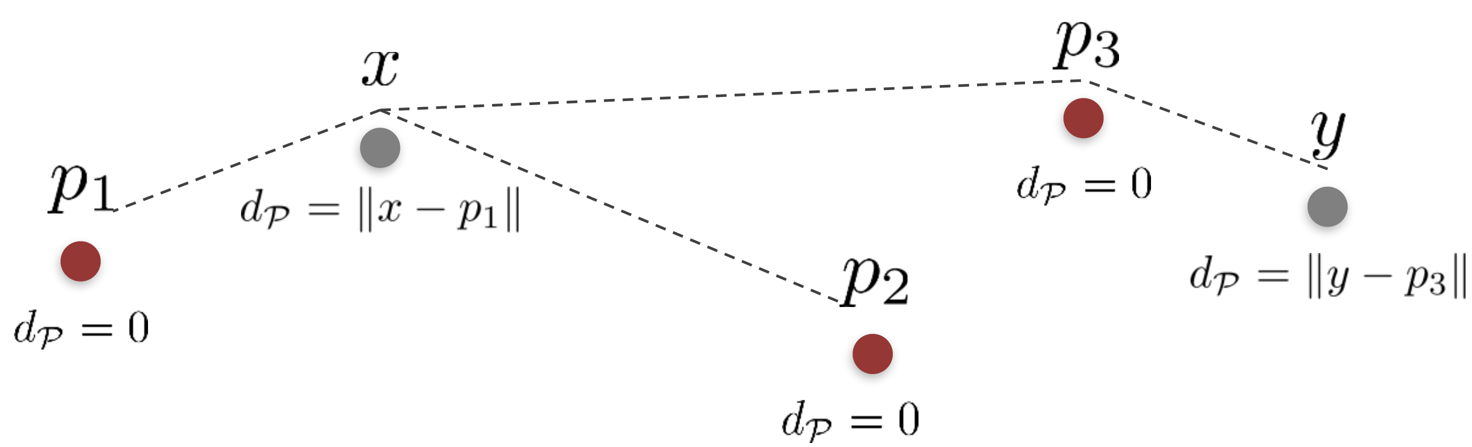
Critical points

The Distance Function:

For a finite set

$$d_{\mathcal{P}}(x) = \min_{p \in \mathcal{P}} \|x - p\|, \quad x \in \mathbb{R}^d$$

Example: $\mathcal{P} = \{p_1, p_2, p_3\}$



Goal: Study critical points of $d_{\mathcal{P}}$ for a random $\mathcal{P}$

Critical points

$$\mathcal{P} = \{p_1, p_2, p_3\} \subset \mathbb{R}^2$$

$$d_{\mathcal{P}} : \mathbb{R}^2 \to \mathbb{R}^+$$

Index 0 (minimum)

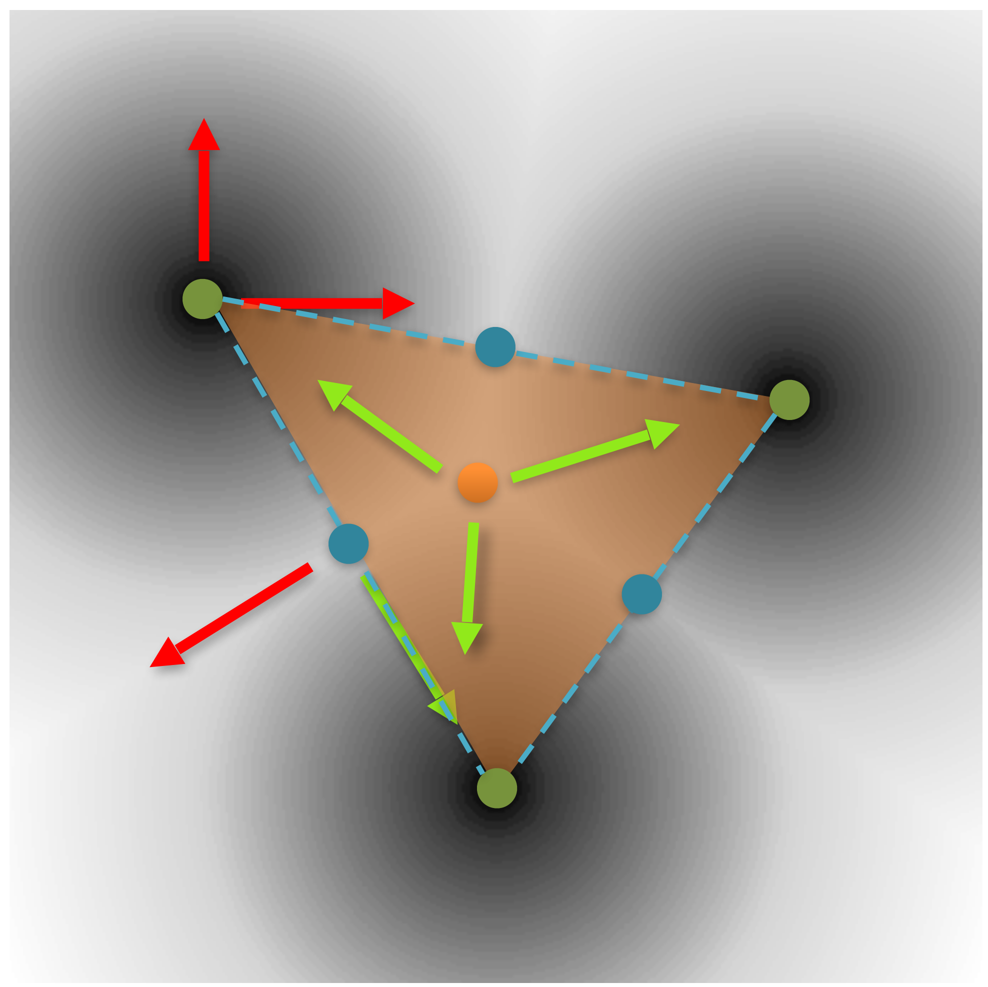
Generated by 1 point

Index 1 (saddle)

Generated by 2 points

Index 2 (maximum)

Generated by 3 points



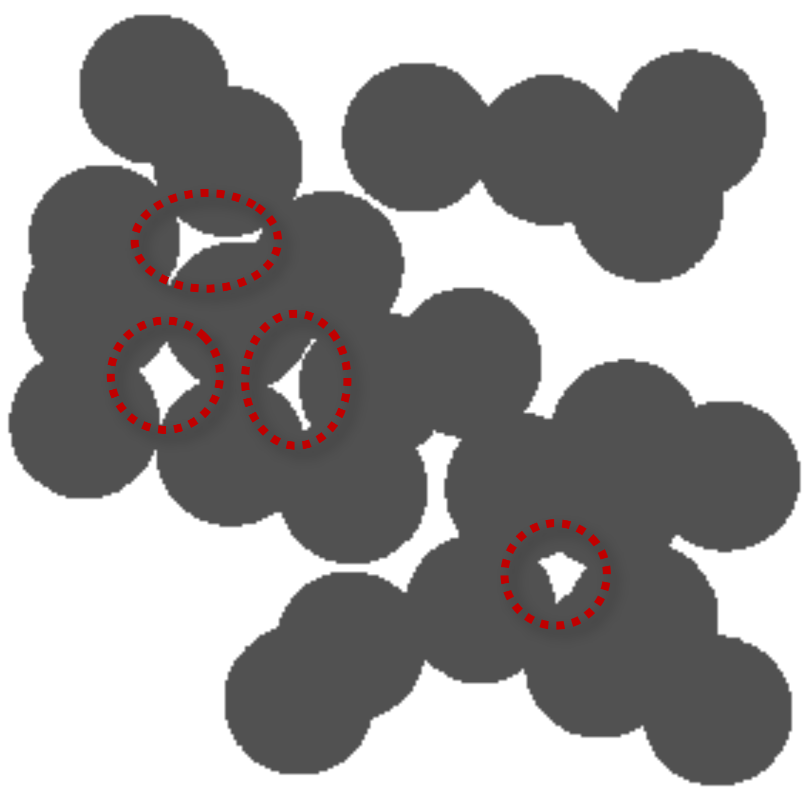
increasing

decreasing

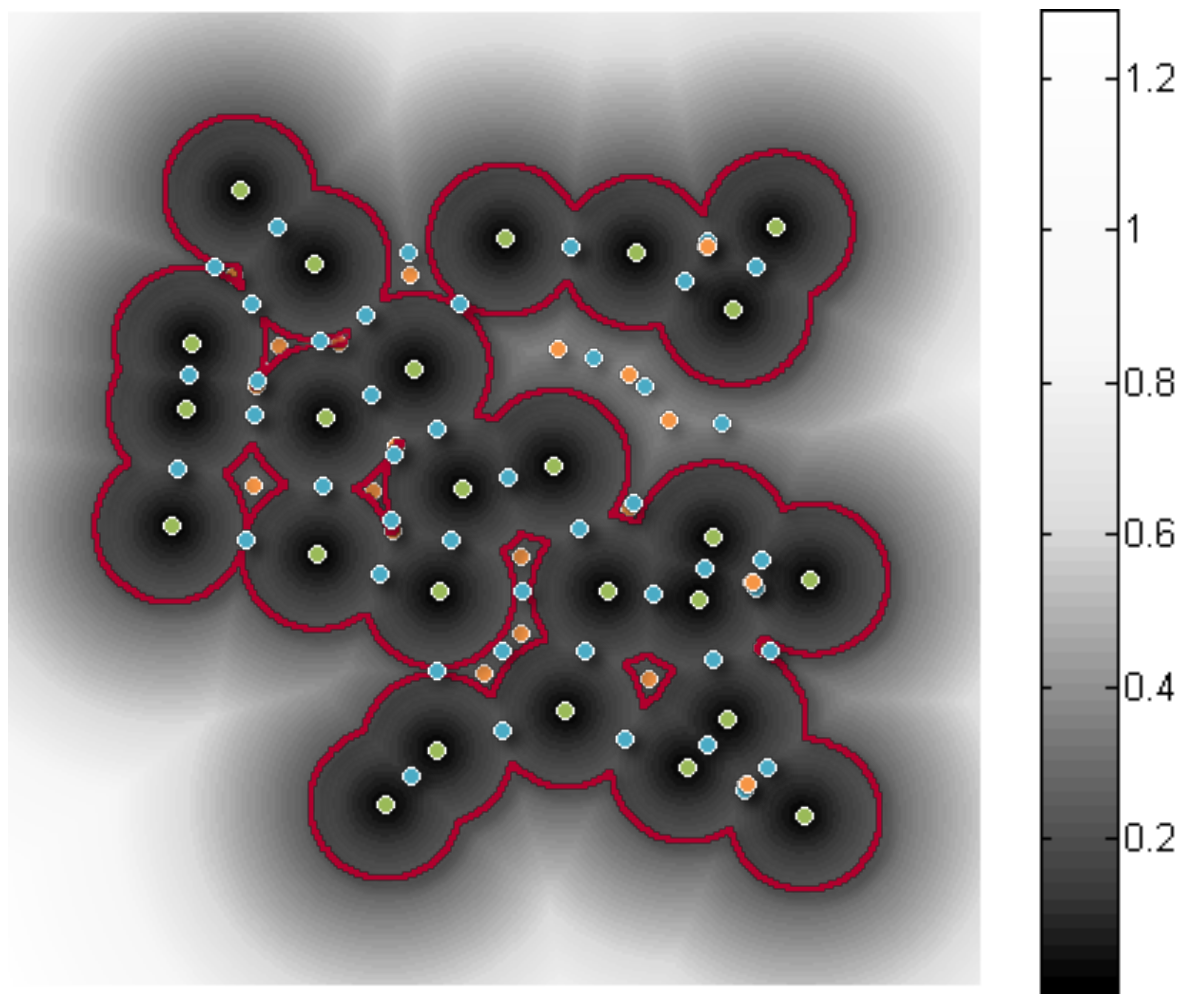
Index *k* critical points are "generated" by subsets of *k+1* points

Distributions on compact manifolds

Union of Balls - $\mathcal{U}(\mathcal{P}, r)$



Distance Function - $d_{\mathcal{P}}(x)$



$\beta_0 = 2$

$\beta_1 = 4$

Morse Theory

$\chi = \beta_0 - \beta_1 = -2 = N_0 - N_1 + N_2$

$N_0 = 25$

$N_1 = 30$

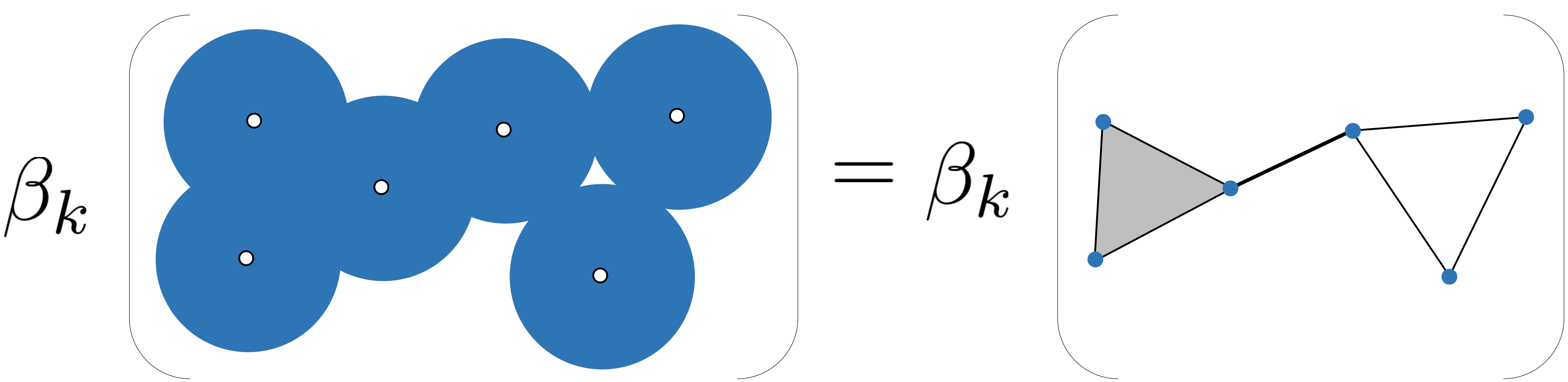
$N_2 = 3$

Nerve lemma

Theorem (Borsuk, 1948)

The Čech complex $\check{C}(\mathcal{P}, r)$ is homotopy equivalent to $\bigcup_{p \in \mathcal{P}} B_r(p)$. In particular,

$$\beta_k \left(\bigcup_{p \in \mathcal{P}} B_r(p) \right) = \beta_k(\check{C}(\mathcal{P}, r)).$$



Problem statement

As $n \to \infty$ and as $r \to 0$ where are limiting distributions of

(1) Betti numbers of $\mathcal{U}(\mathcal{P}, r)$.

(2) The number of critical points of

$$d_{\mathcal{P}}(x) := \min_{p \in \mathcal{P}} \|x - p\|, \quad x \in \mathbb{R}^d.$$

Earlier work

- Early mention by Milnor—1964

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages— Farber and Kappeller, 2007.

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages— Farber and Kappeller, 2007.
- Random simplicial complexes — Linial and Meshulam, 2006

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages— Farber and Kappeller, 2007.
- Random simplicial complexes — Linial and Meshulam, 2006
- Recovery of homology — Niyogi, Smale, and Weinberger, 2008

Earlier work

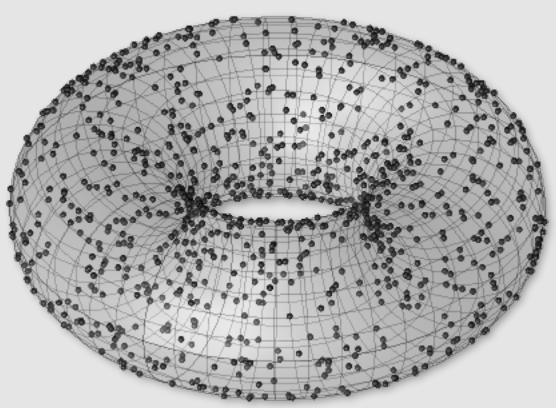
- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages— Farber and Kappeller, 2007.
- Random simplicial complexes — Linial and Meshulam, 2006
- Recovery of homology — Niyogi, Smale, and Weinberger, 2008
- Random geometric graphs — Penrose, 2003

Earlier work

- Early mention by Milnor—1964
- Gaussian fields on manifolds —-Adler and Taylor, 2003.
- Random triangulated surfaces — Pippenger and Schleich, 2006.
- Random 3-manifolds — Dunfield and Thurston, 2006.
- Random planar linkages— Farber and Kappeller, 2007.
- Random simplicial complexes — Linial and Meshulam, 2006
- Recovery of homology — Niyogi, Smale, and Weinberger, 2008
- Random geometric graphs — Penrose, 2003
- Thresholds for cohomology — Kahle, Meckes 2010, Kahle, 2014.

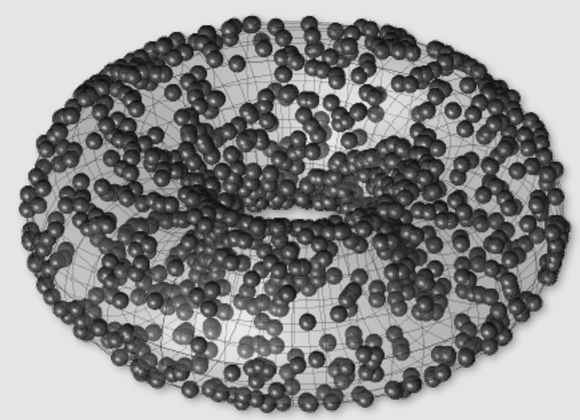
Three regimes

Subcritical



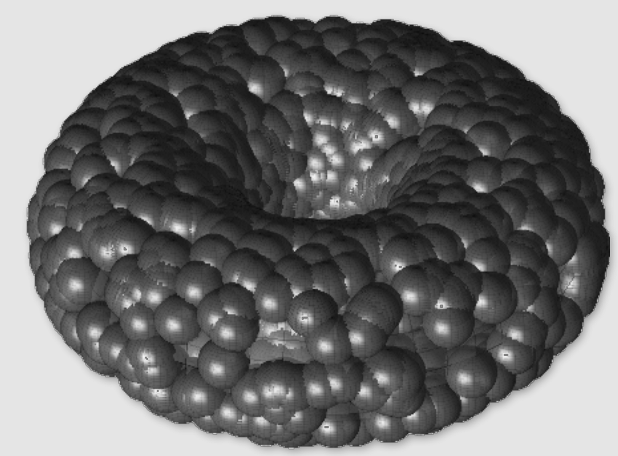
$nr^{d} \to 0$

Critical



$nr^{d} \to \lambda \in (0, \infty)$

Supercritical



$nr^{d} \to \infty$

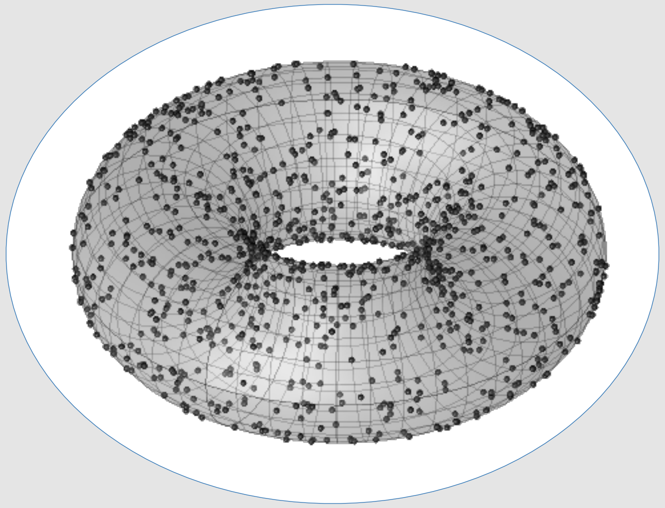
$\mathcal{X}_n \subset \mathbb{R}^{D}$

$d = \dim(\mathcal{M})$

$$nr^d \propto \mathbb{E}[\text{number of points in a ball of radius } r]$$

Subcritical regime

Subcritical



$$nr^d \to 0$$

Subcritical regime

Theorem: Bobrowski, M

If $nr^d \to 0$, then

- $\lim_{n\to\infty}(n^{k+2}r^{d(k+1)})^{-1}\mathbb{E}\{\beta_{k,n}(r)\} = \mu_k^b, \qquad 1 \le k \le d-1$
- $\lim_{n\to\infty}(n^{k+1}r^{dk})^{-1}\mathbb{E}\{N_{k,n}(r)\} = \mu_k^c, \qquad 1 \le k \le d$

Expected number of k-cycles:

$$\mathbb{E}\{\beta_k\} \sim n^{k+2} \times (r^d)^{k+1} = n\Lambda^{k+1} \quad (\Lambda = nr^d)$$

- subsets with k+2 vertices

all points are within a ball of radius r

$$\mu_k^b = \frac{1}{(k+2)!}\int_{\mathcal{M}} f^{k+2}(x)dx \int_{(\mathbb{R}^d)^{k+1}} h_1^b(0,\mathbf{y})d\mathbf{y} \qquad \mu_k^c = \frac{1}{(k+1)!}\int_{\mathcal{M}} f^{k+1}(x)dx \int_{(\mathbb{R}^d)^k} h_1^c(0,\mathbf{y})d\mathbf{y}$$

Subcritical regime

$$(n^{k+1}r^{dk})^{-1}\mathbb{E}\{N_{k,n}(r)\} \to \mu_k^c$$

Theorem: Bobrowski, M

If $nr^d \to 0$, and $k \geq 1$,

- $(n^{k+1}r^{dk})^{-1}\mathrm{Var}(N_{k,n}) \to \mu_k^c$
- $n^{k+1}r^{dk} \to 0 \quad \Rightarrow \quad N_{k,n} \xrightarrow{L^2} 0$
- $n^{k+1}r^{dk} \to a \in (0,\infty) \quad \Rightarrow \quad N_{k,n} \xrightarrow{D} \mathrm{Poisson}\,(a\mu_k^c)$
- $n^{k+1}r^{dk} \to \infty \quad \Rightarrow \quad \dfrac{N_{k,n} - \mathbb{E}\{N_{k,n}\}}{\sqrt{n^{k+1}r^{dk}}} \xrightarrow{D} \mathcal{N}(0,\mu_k^c)$

Similar results for $\beta_{k,n}$ [Kahle-Meckes].

Subcritical regime

Exact limit values are known, as well as limit distributions

$$\beta_{k,n} \sim n^{k+2} r^{d(k+1)}, \quad k = 1, \ldots, d-1$$

$$\Rightarrow \quad n \approx \beta_{0,n} \gg \beta_{1,n} \gg \beta_{2,n} \gg \cdots \gg \beta_{d-1,n}$$

Phase transitions:

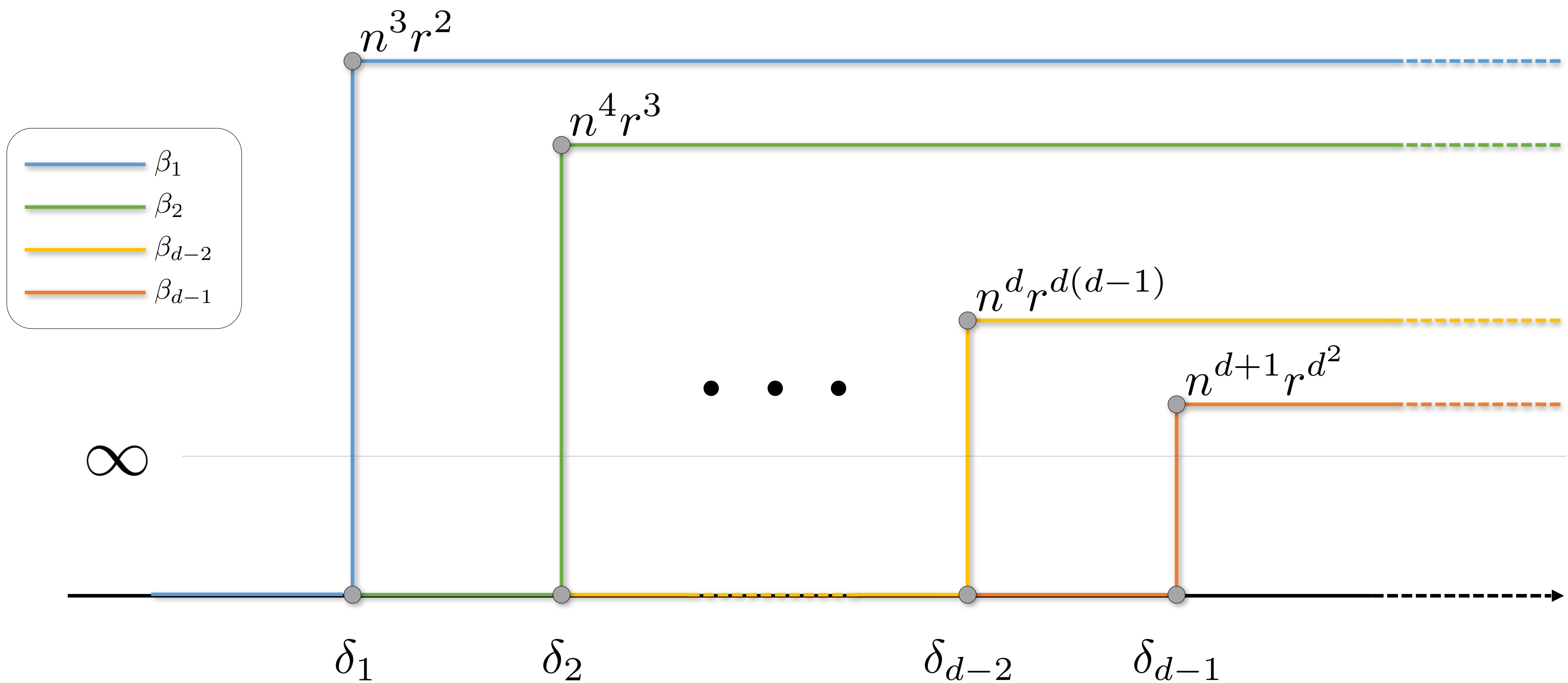
Critical radius:

$$n{\delta_k}^d = n^{-\frac{1}{k+1}}$$

$$\lim_{n\to\infty} \mathbb{E}\left\{\beta_{k,n}(r)\right\} = \begin{cases} 0 & r = o(\delta_k) \\ c & r = \Theta(\delta_k) \\ \infty & r = \omega(\delta_k) \end{cases} \Rightarrow \lim_{n\to\infty} \mathbb{P}\left(\beta_{k,n}(r) > 0\right) = \begin{cases} 0 & r = o(\delta_k) \\ e^{-c} & r = \Theta(\delta_k) \\ 1 & r = \omega(\delta_k) \end{cases}$$

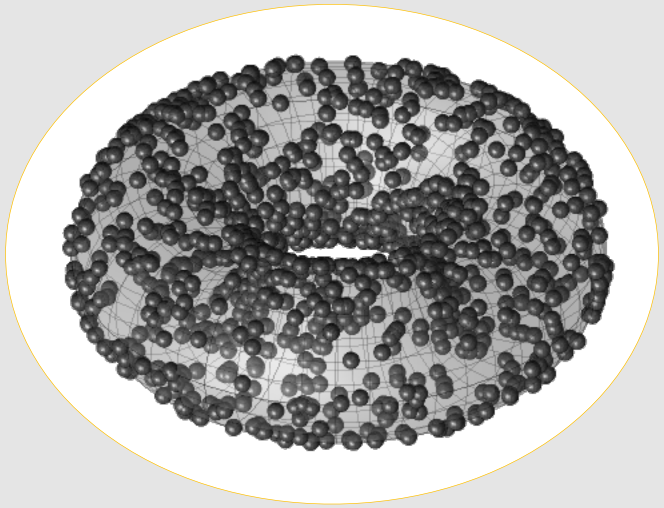
Subcritical regime

$$n{\delta_k}^d = n^{-\frac{1}{k+1}}$$



Critical regime

Critical



$$nr^d \to \lambda \in (0, \infty)$$

Critical regime

Theorem

If $nr^d \to \lambda$, then: $\mathbb{E}\{\beta_k\} = \Theta(n), \qquad 1 \le k \le d-1$

Theorem

If $nr^d \to \lambda$, and $1 \le k \le d$, then

$$\lim_{n\to\infty} n^{-1}\mathbb{E}\{N_{k,n}\} = \gamma_k(\lambda)$$

where

$$\gamma_k(\lambda) = \frac{\lambda^k}{(k+1)!}\int_{\mathcal{M}}\int_{(\mathbb{R}^d)^k} f^{k+1}(x) h_1^c(0,\mathbf{y}) e^{-\lambda\omega_d f(x) R^d(0,\mathbf{y})} d\mathbf{y}\, dx$$

Critical regime

Theorem: Bobrowski, M

If $nr^d \to \lambda$, then: $\mathbb{E}\{\beta_k\} = \Theta(n), \quad 1 \le k \le d-1$

Theorem: Bobrowski, M

If $nr^d \to \lambda$, and $1 \le k \le d-1$

- $\lim_{n\to\infty} \frac{\mathbb{E}[N_{k,n}]}{n} = \gamma_k(\lambda)$,
- $\lim_{n\to\infty} \frac{\mathrm{Var}[N_{k,n}]}{n} = \sigma_k^2(\lambda)$,
- $\lim_{n\to\infty} \frac{N_{k,n} - \mathbb{E}[N_{k,n}]}{\sqrt{n}} \xrightarrow{D} \mathcal{N}(0, \sigma_k^2(\lambda))$,

Critical regime

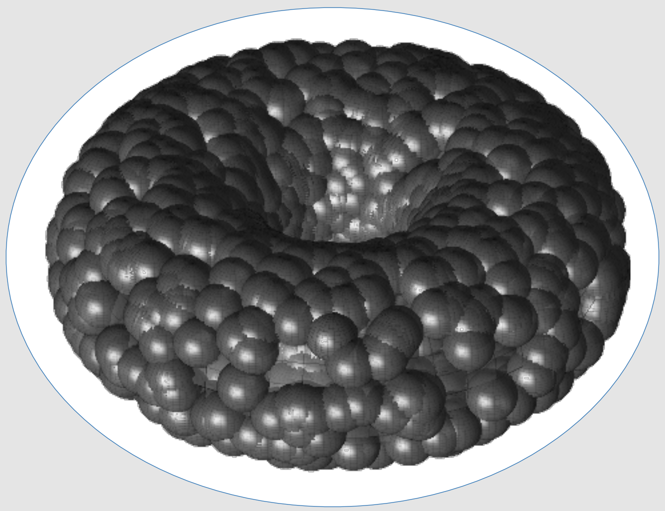
By Morse theory

$$\chi_n(r) = \sum_{k=0}^{d-1} (-1)^k \beta_{k,n}(r) = \sum_{k=0}^{d} (-1)^k N_{k,n}(r),$$

implies

$$n^{-1}\mathbb{E}[\chi_n(r)] \to 1 + \sum_{k=1}^{d} (-1)^k \gamma_k(\lambda).$$

Supercritical regime



$$nr^d \to \infty$$

Supercritical regime

Theorem: Bobrowski, M

Assume $f_{\min} := \inf_{x\in\mathcal{M}} f(x) > 0$ and $nr^d \to \infty$ then for $1 \le k \le d$

- $\lim_{n\to\infty} \frac{\mathbb{E}[N_{k,n}]}{n} = \gamma_k(\infty)$,
- $\lim_{n\to\infty} \frac{\mathrm{Var}[N_{k,n}]}{n} = \sigma_k^2(\infty)$,
- $\lim_{n\to\infty} \frac{N_{k,n}-\mathbb{E}[N_{k,n}]}{\sqrt{n}} \xrightarrow{D} \mathcal{N}(0, \sigma_k^2(\infty))$.

Supercritical regime

Coverage theorem: Bobrowski, M

If $nr_n^k > C \log n$

1. If $C > (\omega_k f_{\min})^{-1}$, then

$$\lim_{n\to\infty} \mathbb{P}\left(N_{k,n} = \hat{N}_{k,n}, \forall 1 \le k \le d\right) = 1.$$

2. If $C > 2(\omega_k f_{\min})^{-1}$, then there exists $M > 0$ such that for $n > M$

$$N_{k,n} \overset{a.s.}{=} \hat{N}_{k,n}, \forall 1 \le k \le d.$$

Supercritical regime

Convergance of Betti numbers: Bobrowski, M

If $r_n \to 0$ and $nr_n^k > C \log n$, then

1. If $C > (\omega_k f_{\text{min}})^{-1}$, then

$$\lim_{n\to\infty} \mathbb{P}\left(\beta_{k,n} = \beta_k(\mathcal{M}), \forall 0 \le k \le d\right) = 1.$$

2. If $C > 2(\omega_k f_{\text{min}})^{-1}$, then there exists $M > 0$ such that for $n > M$

$$\beta_{k,n} \overset{a.s.}{=} \beta_k(\mathcal{M}), \forall 0 \le k \le d.$$

Supercritical regime

Vacancy probability:

$$\mathbb{P}\left(B_r(x) \cap \mathcal{X}_n = \emptyset\right) \approx e^{-n\mathbb{P}(B_r(x))} \le e^{-n f_{\min}\omega_d r^d}$$

Coverage threshold:

$$n{\rho_c}^d = (\omega_d f_{\min})^{-1} \log n$$

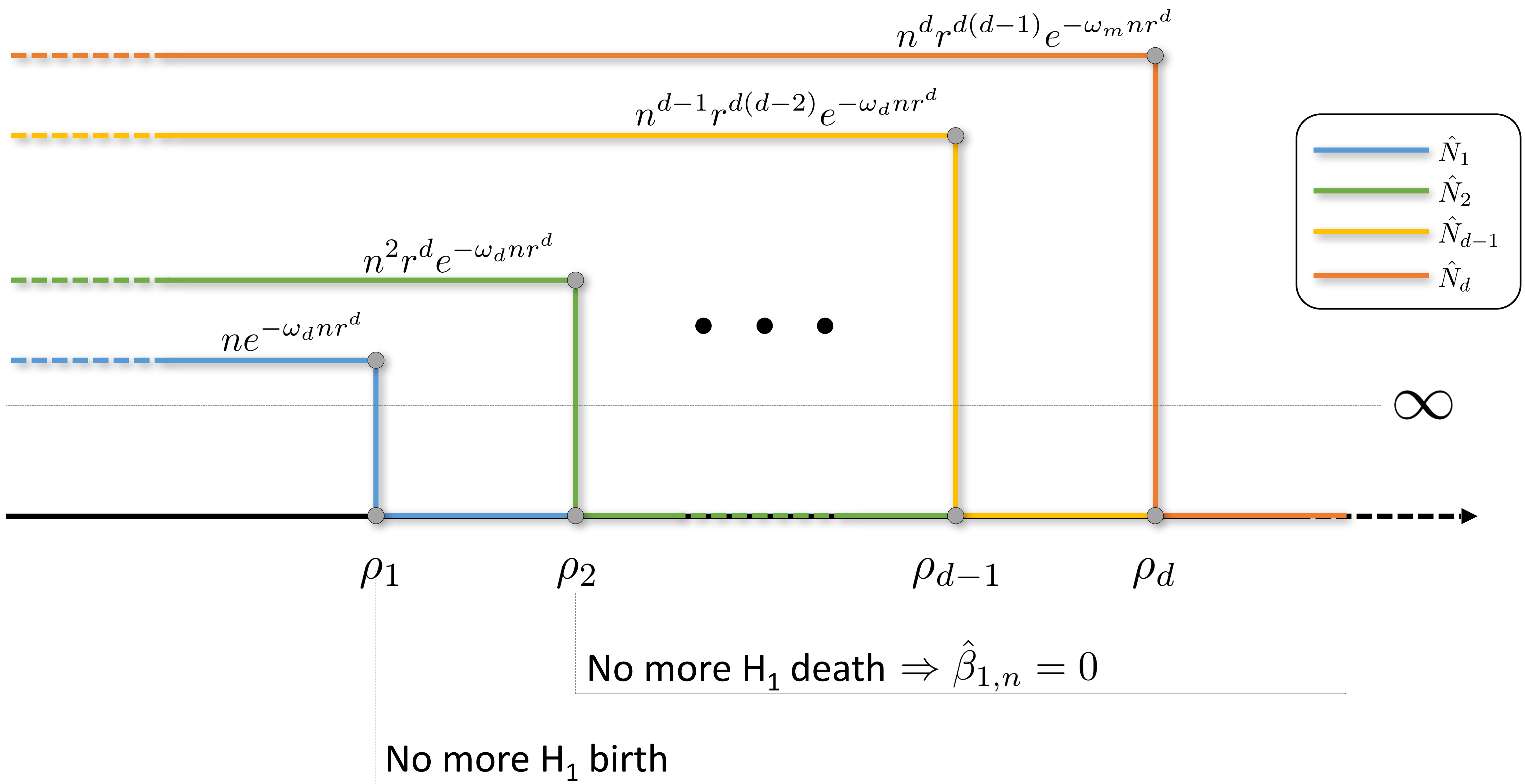
Theorem [Bobrowski,M]

If $r \ge (1+\epsilon)\rho_c$ for some $\epsilon > 0$, then

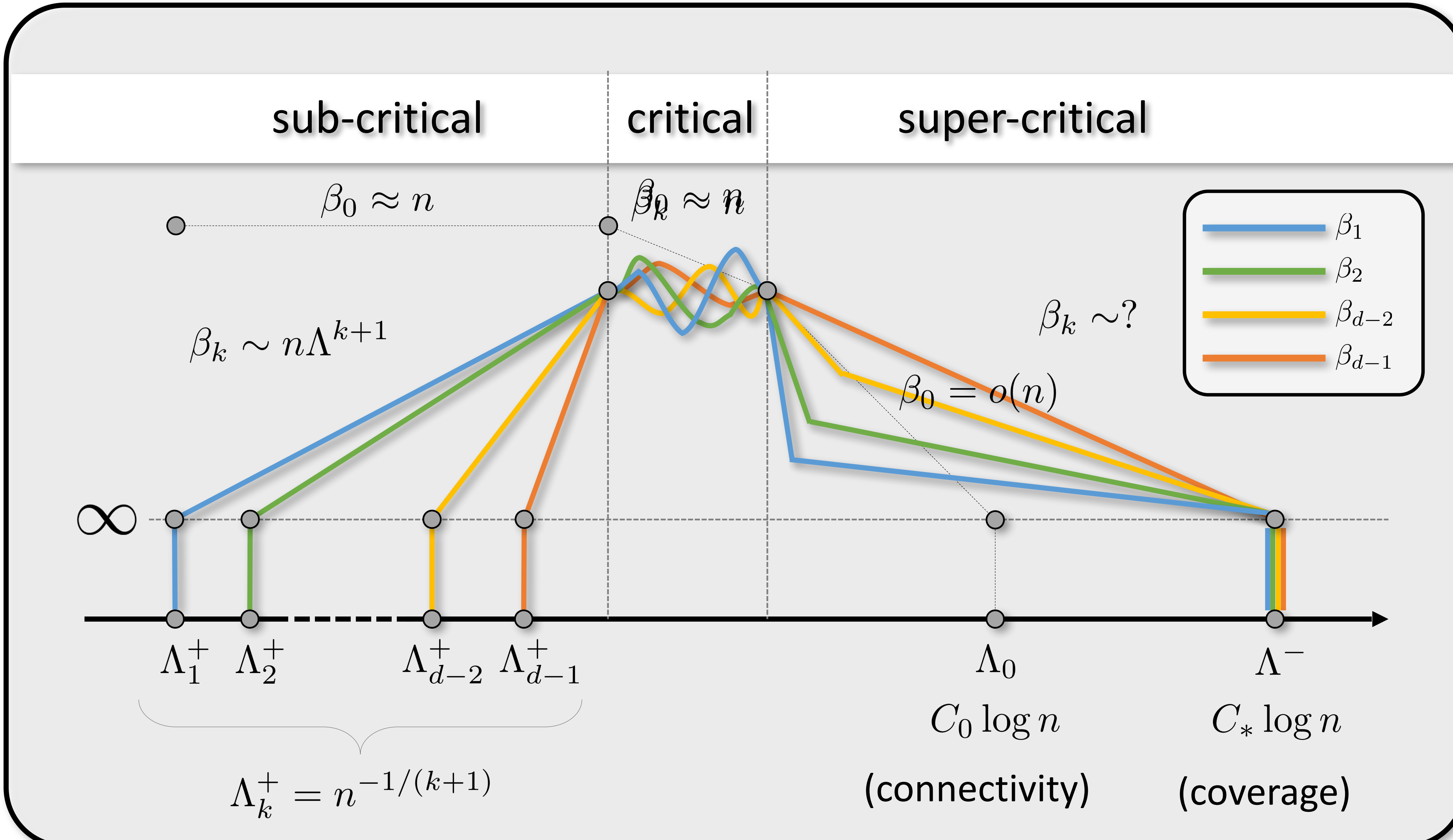
$$\lim_{n\to\infty} \mathbb{P}\left(\beta_{k,n}(r) = \beta_k(\mathcal{M}),\ \forall 0 \le k \le d\right) = 1$$

Supercritical regime

$$n{\rho_k}^d = \omega_d^{-1}(\log n + (k-1)\log\log n)$$



Review of scaling limits



Bobrowski & Kahle –*Topology of Random Geometric Complexes: a survey*

Open questions and problems

(1) Betti numbers outside the critical regime.

Open questions and problems

(1) Betti numbers outside the critical regime.

(2) Curved spaces.

Open questions and problems

(1) Betti numbers outside the critical regime.

(2) Curved spaces.

(3) Random walks and spectral simiplicial theory.

Open questions and problems

(1) Betti numbers outside the critical regime.
(2) Curved spaces.
(3) Random walks and spectral simiplicial theory.
(4) Higher-order expanders.

Open questions and problems

(1) Betti numbers outside the critical regime.
(2) Curved spaces.
(3) Random walks and spectral simiplicial theory.
(4) Higher-order expanders.
(5) Stochastic Hodge theory.

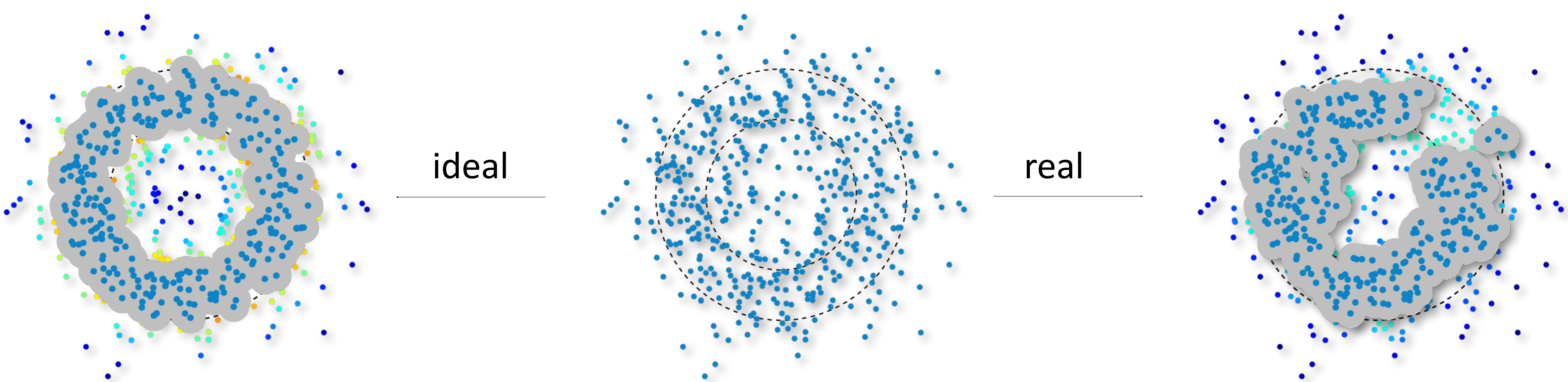
Homology of level sets

Homology of level sets

Density $f : \mathbb{R}^d \rightarrow \mathbb{R}^+$ we consider $D_L = \{x : f(x) > L\}$.

If we know $f(x_i)$ at sample points and pick $f(x_i) > L$ to construct homology.

We can use $\hat{f}(x) = n^{-1} \sum_{i=1}^{n} K(x, x_i)$ to pick points.



Homology of level sets

Recovering $H_k(D_L)$ is hard, noise and homology can be brittle.

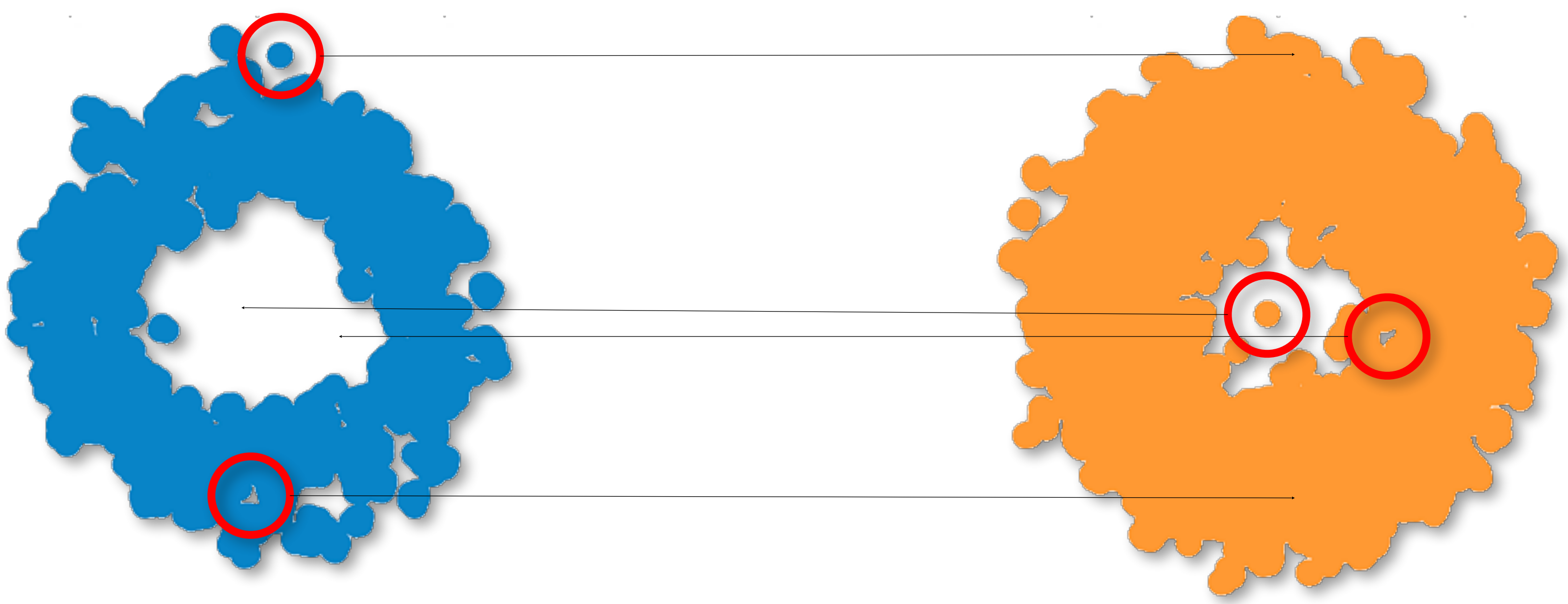
Homology of level sets

Recovering $H_k(D_L)$ is hard, noise and homology can be brittle.

Instead look at $D_{L+\epsilon} \hookrightarrow D_{L-\epsilon}$

$\beta_0(L+\epsilon) = 4$

$\beta_1(L+\epsilon) = 10$



$\beta_0(L-\epsilon) = 3$

$\beta_1(L-\epsilon) = 7$

$$\beta_0(L+\epsilon \to L-\epsilon) = 1, \qquad \beta_1(L+\epsilon \to L-\epsilon) = 1$$

Estimating homology of level sets

Define the following as procedure P1:

(1) Given $(X_1, ..., X_n)$ compute

$$
\begin{aligned}
\hat{D}_{L-\epsilon}(n, r) &:= \{X_i : \hat{f}_n(X_i) \geq L - \epsilon; 1 \leq i \leq n\} \\
\hat{D}_{L+\epsilon}(n, r) &:= \{X_i : \hat{f}_n(X_i) \geq L + \epsilon; 1 \leq i \leq n\}
\end{aligned}
$$

Estimating homology of level sets

Define the following as procedure P1:

(1) Given $(X_1, \ldots, X_n)$ compute

$$\begin{aligned} \hat{D}_{L-\epsilon}(n, r) &:= \{X_i : \hat{f}_n(X_i) \geq L - \epsilon; 1 \leq i \leq n\} \\ \hat{D}_{L+\epsilon}(n, r) &:= \{X_i : \hat{f}_n(X_i) \geq L + \epsilon; 1 \leq i \leq n\} \end{aligned}$$

(2) Consider the homology map

$$\iota_* : H_*(\hat{D}_{L+\epsilon}(n, r)) \longrightarrow \hat{D}_{L-\epsilon}(n, r).$$

Estimating homology of level sets

Define the following as procedure P1:

(1) Given $(X_1, ..., X_n)$ compute

$$\begin{aligned}\hat{D}_{L-\epsilon}(n,r) &:= \{X_i : \hat{f}_n(X_i) \geq L-\epsilon; 1 \leq i \leq n\}\\ \hat{D}_{L+\epsilon}(n,r) &:= \{X_i : \hat{f}_n(X_i) \geq L+\epsilon; 1 \leq i \leq n\}\end{aligned}$$

(2) Consider the homology map

$$\iota_* : H_*(\hat{D}_{L+\epsilon}(n,r)) \longrightarrow \hat{D}_{L-\epsilon}(n,r).$$

(3) Define $\hat{H}_*(L, \epsilon; n) := \mathrm{Im}(\iota_*)$.

Estimating homology of level sets

Theorem (Bobrowski,M,Taylor)

Let $L > 0$ *and* $\epsilon \in (0, L/2)$ *be such that the density* $f(x)$ *has no critical values in* $[L - 2\epsilon, L + 2\epsilon]$*. If* $r \to 0$ *and* $nr^d\infty$ *then for* n *large enough*

$$\mathbb{P}\left(\hat{H}_*(L, \epsilon; n) \cong H_*(D_L)\right) \geq 1 - 6ne^{-C^*_{\epsilon/2}nr^d}.$$

Estimating homology of level sets

Theorem (Bobrowski,M,Taylor)

Let $L > 0$ and $\epsilon \in (0, L/2)$ be such that the density $f(x)$ has no critical values in $[L - 2\epsilon, L + 2\epsilon]$. If $r \to 0$ and $nr^d \infty$ then for n large enough

$$\mathbb{P}\left(\hat{H}_*(L, \epsilon; n) \cong H_*(D_L)\right) \geq 1 - 6ne^{-C^*_{\epsilon/2} nr^d}.$$

In particular, if $nr^d \geq D \log n$ with $D > (C^*_{\epsilon/2})^{-1}$ then

$$\lim_{n\to\infty} \mathbb{P}\left(\hat{H}_*(L, \epsilon; n) \cong H_*(D_L)\right) = 1.$$

The longest bar

Random persistence

Define $\mathcal{M} = [0,1]^d$ and

$$\pi(\gamma) := \frac{\gamma_{\mathsf{death}}}{\gamma_{\mathsf{birth}}}.$$

Random persistence

Define $\mathcal{M} = [0,1]^d$ and

$$\pi(\gamma) := \frac{\gamma_{\text{death}}}{\gamma_{\text{birth}}}.$$

How does the maximal persistence over all k-cycles scale

$$\Pi_k(n) = \max_{\gamma \in PH_k(n)} \pi(\gamma).$$

Random persistence

Define $\mathcal{M} = [0,1]^d$ and

$$\pi(\gamma) := \frac{\gamma_{\text{death}}}{\gamma_{\text{birth}}}.$$

How does the maximal persistence over all k-cycles scale

$$\Pi_k(n) = \max_{\gamma \in PH_k(n)} \pi(\gamma).$$

Theorem (Bobrowski,Kahle,Skraba)

Let $\mathcal{P}_n$ be a Poisson process on $[0,1]^d$ and let $PH_k(n)$ be the k-th persistent homology. Then $\Pi_k(n)$ scales as

$$\Delta_k(n) := \left(\frac{\log n}{\log \log n} \right)^{1/k}.$$

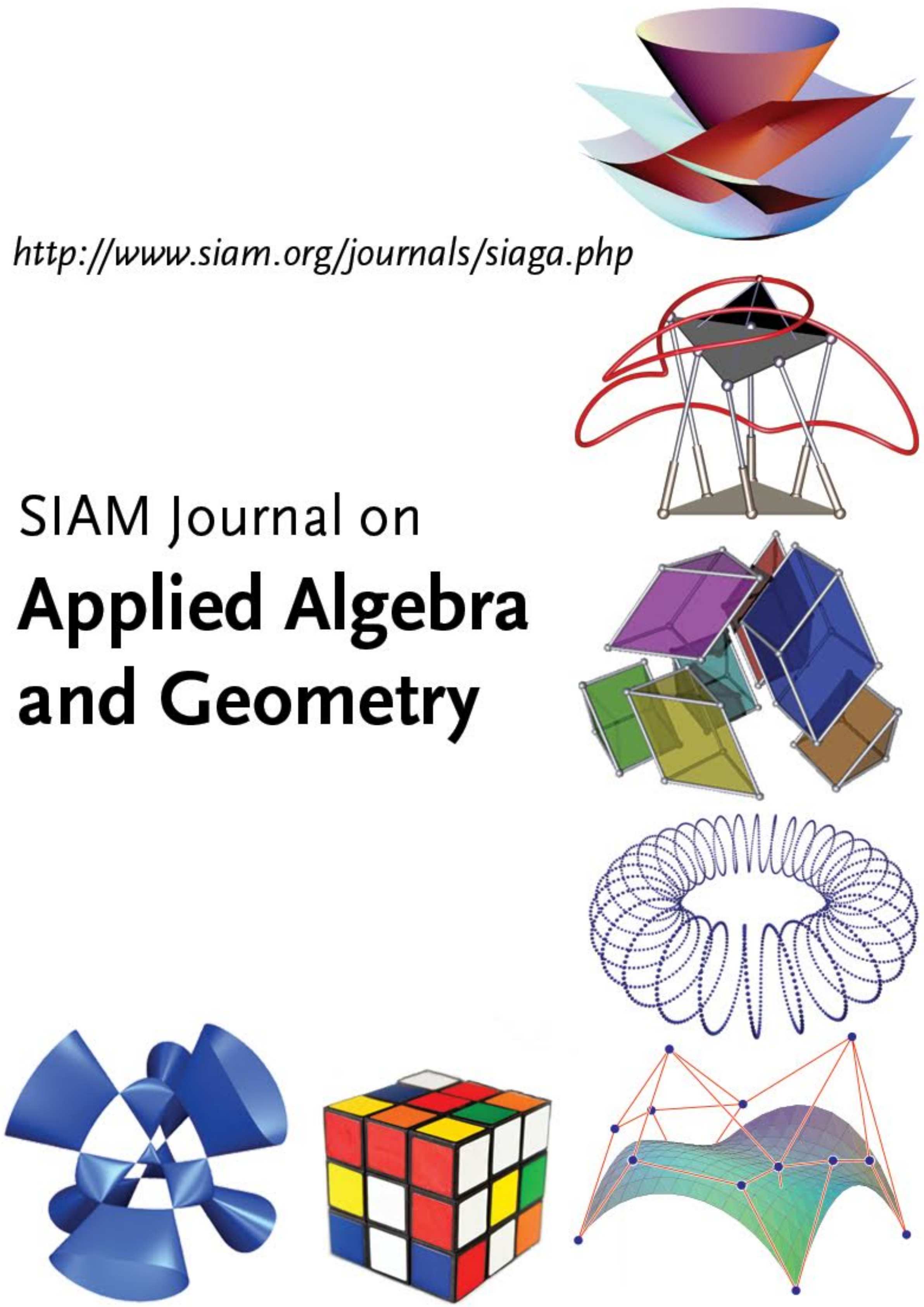
Acknowledgements

Many people.

Funding:

- Center for Systems Biology at Duke
- NSF DMS and CCF
- AFOSR, DARPA
- NIH

New journal



http://www.siam.org/journals/siaga.php
SIAM Journal on
Applied Algebra
and Geometry