Random graphs, the multiplicative coalescent and percolation of tree-averages

First half of talk: Review of first-order and second-order analysis of component sizes in the random graph process. The latter based on Aldous (1997).

Brief mention of subsequent related work

Second half of talk: Some work-in-progress.

Recall the **random graph** $\mathcal{G}(n, p(n))$ has n vertices, and each potential edge is present with probability $p(n)$.

For first-order analysis, set $p(n) = t/n$ for "time" $0 \leq t$. For our purposes, key observation is that the graph looks locally like a Galton-Watson branching process with Poisson(t) offspring distribution [call this PGW(t)]. Easy to guess that

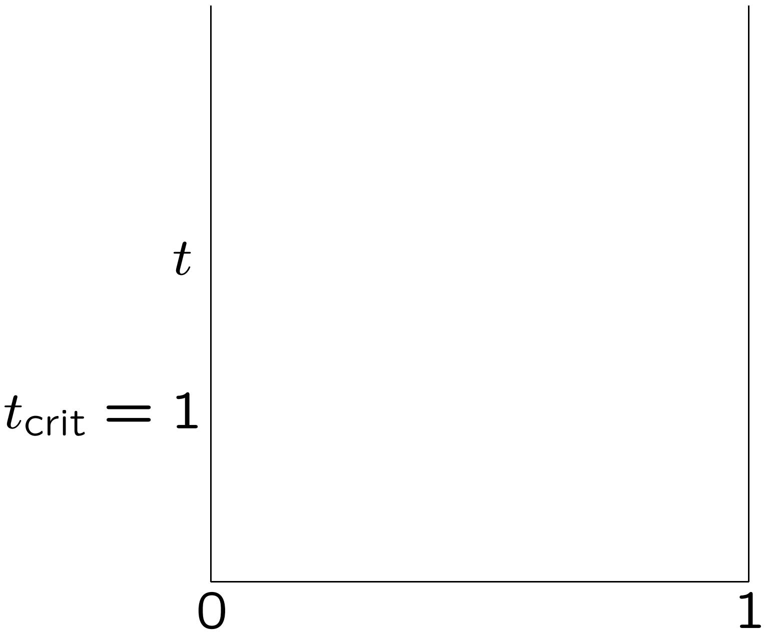
$C_1^n(t) :=$ size of largest component of $\mathcal{G}(n, t/n)$

has limit behavior

$$\lim_n n^{-1} C_1^n(t) = \Phi(t)$$

$$\Phi(t) := P(\text{non-extinction of PGW}(t)).$$

The elementary fixed-point equation for Φ leads to an explicit expression and to the qualitative behavior



$\Phi(t) = $ (size giant component$/n$)

Emphasize two points.

Scaling exponent is 1.

$$\Phi(t) \sim (t - t_{\text{crit}})^1 \text{ as } t \downarrow t_{\text{crit}}.$$

No autonomous limit process in this scaling.

Second order analysis. At critical point, size of largest component grows as $\Theta(n^{2/3})$. Easy to deduce (next slide) that correct time scaling is $p(t) = \frac{1}{n} + \frac{t}{n^{4/3}}$.

Theorem. [Aldous 1997]
Write $C_1^n(t) \geq C_2^n(t) \geq \ldots$ for component sizes of $\mathcal{G}(n, \frac{1}{n} + \frac{t}{n^{4/3}})$. Then

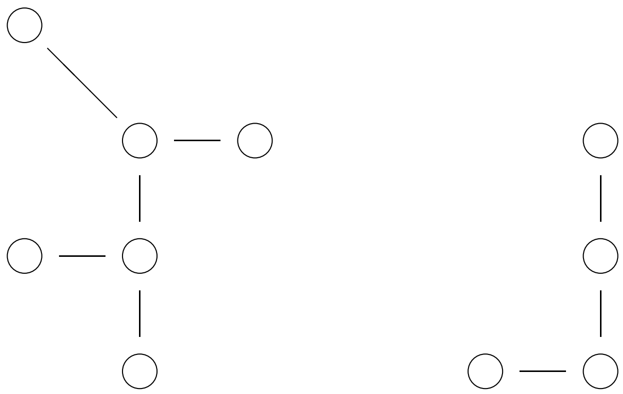
$$n^{-2/3}(C_1^n(t), C_2^n(t), \ldots) \ \xrightarrow{d}$$

$$\mathbf{X}(t) = (X_1(t), X_2(t), \ldots)$$

where the limit process, indexed by $-\infty < t < \infty$, evolves as the **multiplicative coalescent**:

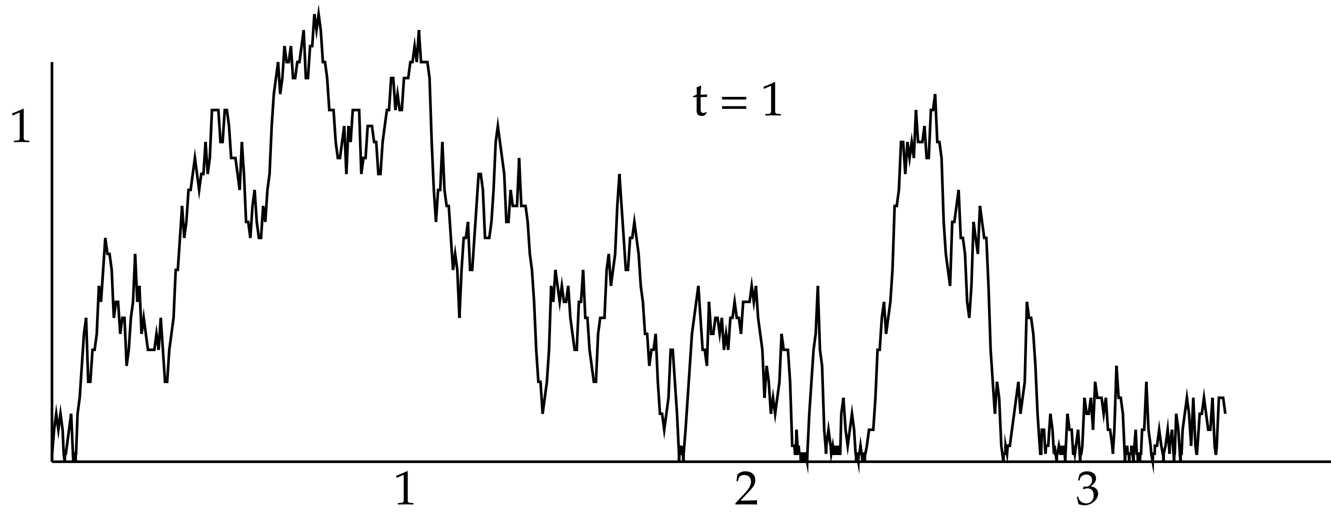
For each pair of clusters, masses $\{x_i, x_j\}$ say, they coalesce into one cluster of mass $x_i + x_j$ at rate $x_i x_j$.

$$\mathcal{G}(n, \tfrac{1}{n} + \tfrac{t}{n^{4/3}})$$



size $xn^{2/3}$ size $yn^{2/3}$

When we increment t to $t + dt$, each of the $xyn^{4/3}$ possible edges appears with probability $dt/n^{4/3}$ and so there is rate $xy\ dt$ of merging into a component of size $(x + y)n^{2/3}$.



Inhomogeneous reflecting Brownian motion
$(B^t(s), 0 \leq s < \infty)$ with drift $= t - s$ at time s.

.

$\mathbf{X}(t) = (X_1(t), X_2(t), \ldots)$ is the ranked list of lengths of excursions from 0 of $B^t(\cdot)$.

.

As time t increases from $-\infty$ to $+\infty$ the size of the largest cluster grows from 0 to ∞.

.

Step i of walk:

$X(i) - X(i-1) = -1 +$ # children of vertex i.

Size of j'th component $= T_j - T_{j-1}$

$$T_j := \min\{i : X(i) = -j\}.$$

If original graph is from the probability model $\mathcal{G}(n, \frac{1}{n} + \frac{t}{n^{4/3}})$ then we can write exact expressions for conditional mean and variance step size.

$$\begin{array}{rcl} \text{cond. mean} & = & -1 + p(n) \times (\#\text{ available vertices}) \\ & \approx & -1 + (\frac{1}{n} + \frac{t}{n^{4/3}})(n-i) \\ & \approx & \frac{t - in^{-2/3}}{n^{1/3}}. \\ \text{cond. var.} & \approx & 1. \end{array}$$

Rescale:

$$\begin{aligned} &(\text{"time"}) \quad && i \leftarrow sn^{2/3} \\ &(\text{size}) \quad && x \leftarrow yn^{1/3} \end{aligned}$$

For rescaled process, cond. mean drift rate

$$= \frac{t-s}{n^{1/3}} \times n^{2/3} \times n^{-1/3} = t - s$$

time change   size change

cond. variance rate

$$= 1 \times n^{2/3} \times (n^{-1/3})^2 = 1$$

time change   size change

Then standard results on martingales $\to_{dist}$ Brownian motion imply limit process is the inhomogeneous Brownian motion $(B^t(s), 0 \leq s < \infty)$ with drift $= t - s$ at time s.

Related work

- prior detailled calculations with near-critical random graphs (Janson-Knuth-Luczak-Pittel 1993)

- Large deviation limit (Puhalskii 2003)

- critical random hypergraphs (Goldschmidt 2003)

- entrance boundary ($-\infty \uparrow t$) behavior of multiplicative coalescent (Aldous - Limic 1998)

- Immigrating vertices (Aldous - Pittel 2000)

- Brownian excursion representations of additive coalescent (Bertoin 2000-2003; Aldous - Pitman 1998, Miermont 2001, Chassaing - Louchard 2002)

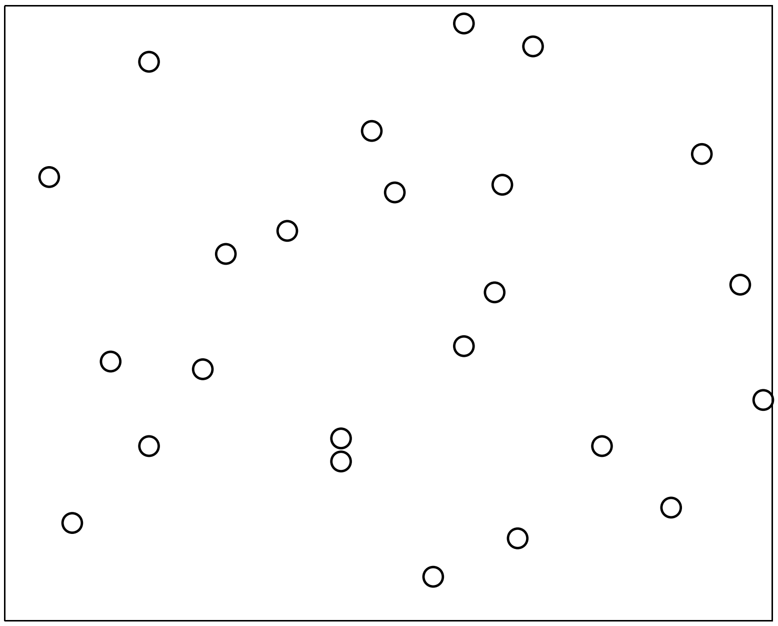
- but no progress (?) on other gelling kernels such as

$$K(x,y) = (xy)^{\gamma},\ 1/2 < \gamma < 1.$$

And now for something completely* different.

(* at first sight)

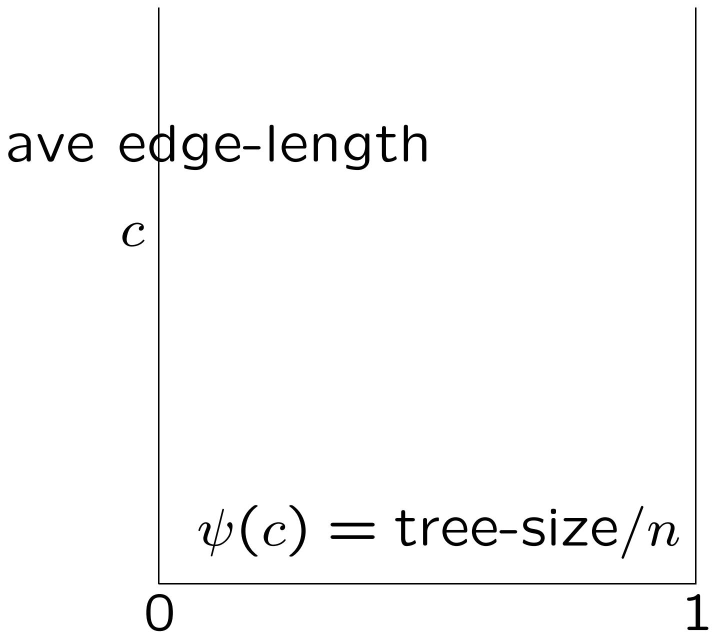
Here are some problems that seem too difficult to solve. Setting: n random points in square of area n in two dimensions. Note the scaling: distance from point to nearest neighbor is order 1. We'll be looking at $n \to \infty$ limits.

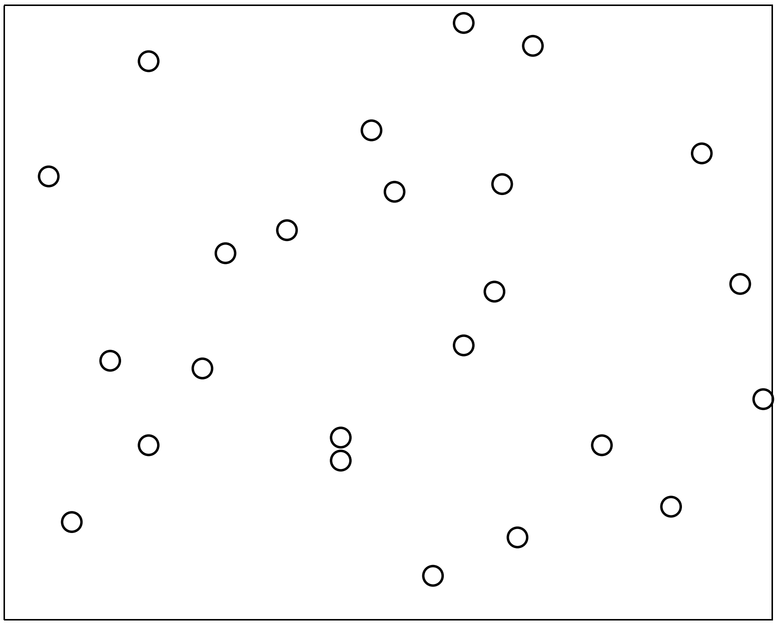


To any tree $\mathbf{t}$ on a subset of these points, associate two numbers

$$\left(\frac{\#\text{ edges of } \mathbf{t}}{n-1}\ ,\ \text{ave edge-length in } \mathbf{t} \right).$$

As $\mathbf{t}$ varies we get a region, and we study the boundary curve.

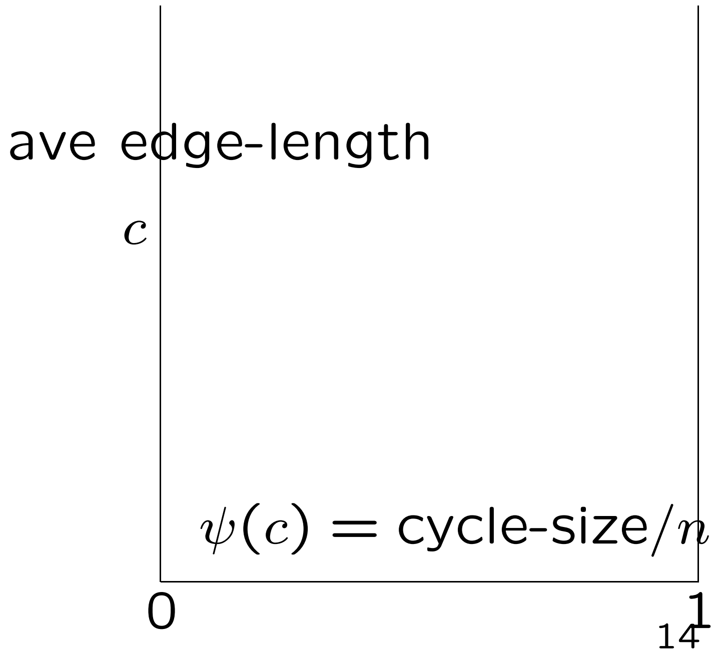




To any cycle $\mathbf{c}$ on a subset of these points, associate two numbers

$$\left(\frac{\#\text{ edges of }\mathbf{c}}{n}\,,\ \text{ave edge-length in }\mathbf{c}\right).$$

As $\mathbf{c}$ varies we get a region, and we study the boundary curve.



xxx hand write as overlays

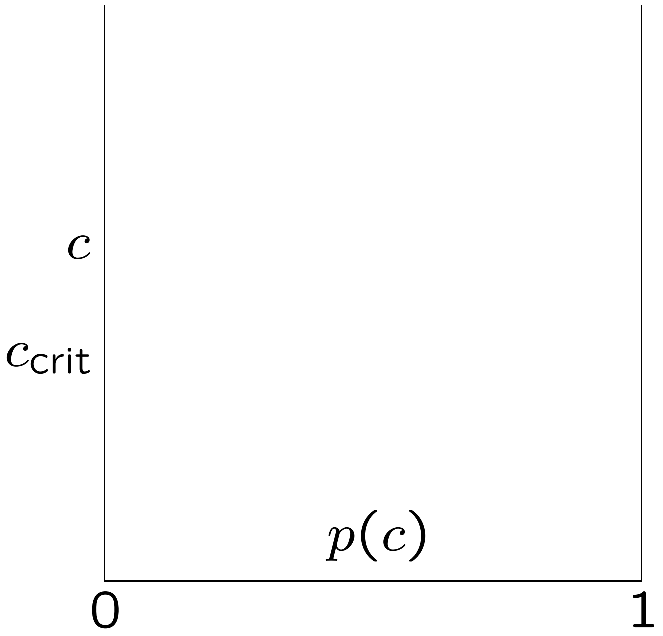
The Beardwood-Halton-Hammersley (1959) constant for length of TSP tour through random points.

Left-end behavior analogous to critical points and scaling exponents in percolation.

Consider **continuum percolation**: put edge between any two points at distance $\leq c$. Write

$p(c) =$ proportion of vertices in giant component.

Expect critical point and scaling exponent.



$$p(c) \asymp (c - c_{\text{crit}})^{\alpha}$$

Can write down definitions of $p(c)$ and $\psi(c)$ which agree except for one word.

$\psi(c) = \lim_n \max\{n^{-1}\#\mathbf{t} \, : \, \mathbf{t} \text{ tree, } \text{ave}_{e \in \mathbf{t}} L(e) \leq c\}$

where $L(e) =$ length of edge e.

$p(c) = \lim_n \max\{n^{-1}\#\mathbf{t} \, : \, \mathbf{t} \text{ tree, } \max_{e \in \mathbf{t}} L(e) \leq c\}$.

Bad news: Can't do any of these new problems in 2 dimensions

Good news: but can do (semi-rigorous) in the mean-field model of distance (just an augmentation of the random graph model). Get interesting results for scaling exponents:
1 : percolation
2 : trees
3 : cycles.

Bad news: method throws no light on why these particular numbers.

Interesting thought: maybe instead of looking at these "first-order" effects we should go straight to "second-order" and look for a process analogous to the multiplicative coalescent but where clusters have both mass and another hidden variable . . . .

Stochastic mean-field model of distance: $D(n)$. Take complete graph on n points. Give each edge e a random length $L(e)$ with uniform $U[0, n]$ distribution, independent for all $\binom{n}{2}$ edges.

So subgraph on edges e with $L(e) \leq t$ is just the random graph $\mathcal{G}(n, t/n)$.

So continuum percolation in this model gives connected components of the random graph.

We have a methodology which (semi-rigorously!) determines the limit function ψ in problems like this. Central idea: there is a $n = \infty$ problem which, because of tree-like structure of geometry $\mathcal{D}(\infty)$, has solutions expressible in terms of solutions of fixed point equations for probability distributions. See Aldous - Percus (2003) for illustration of methodology.

This is all a $\{$ variation? specialization? formalization? $\}$ of the **cavity method** from statistical physics, for which see 2003 survey by Mézard-Parisi.