Imaginary replica analysis of loopy regular random graphs

ACC Coolen

King's College London and Saddle Point Science

https://nms.kcl.ac.uk/ton.coolen

5th International Workshop on
Innovative Algorithms for Big Data

Kyoto, Nov 1st 2019

based on joint work with

Fabian Aguirre-Lopez, Paola Barucca and Mathilde Fekom

Tailoring random graphs

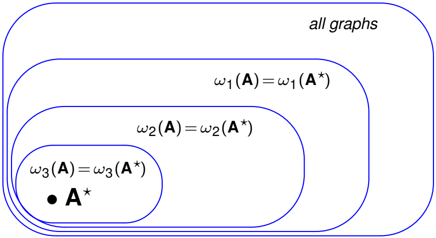
stat mech of process on complex network $\mathbf{A}^\star$,
use *random graph* $\mathbf{A}$ as proxy

- max entropy ensembles,
  constrained by values
  of observables $\omega_1(\mathbf{A}) \ldots \omega_L(\mathbf{A})$

$$\textit{hard constraints}: \quad p(\mathbf{A}) \propto \prod_{\ell \leq L} \delta_{\omega_\ell(\mathbf{A}), \omega_\ell(\mathbf{A}^\star)}$$

$$\textit{soft constraints}: \quad p(\mathbf{A}) \propto \mathrm{e}^{\sum_{\ell=1}^{L} \hat{\omega}_\ell \omega_\ell(\mathbf{A})}, \quad \langle \omega_\ell(\mathbf{A}) \rangle = \omega_\ell(\mathbf{A}^\star) \;\; \forall \ell$$

- larger $L \to$
  better approximation



The problem

- nearly all real-world networks: (physics, biology, economics, computer science, ...)

  $\rightarrow$ *sparse graphs,*
  $\rightarrow$ *many short loops*

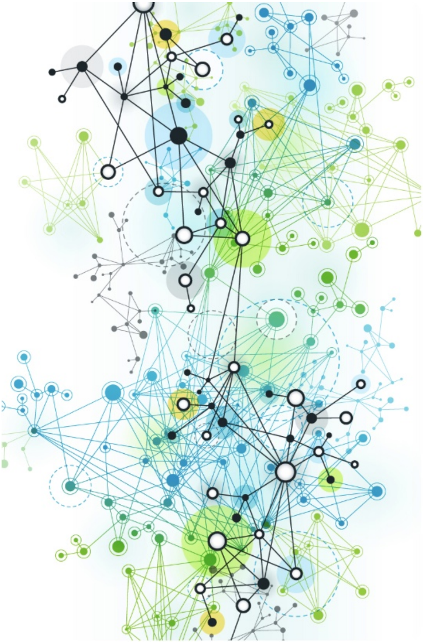
- max entropy graph ensembles with prescribed degree stats:

  $\rightarrow$ *sparse graphs,*
  $\rightarrow$ *locally tree-like*

- realistic tailoring of graphs requires enforcing short loops

  but

  most available analysis methods, e.g. replicas, GFA, cavity, BP, ...
  assume graphs are locally tree-like



Loopy random graph ensembles

- *simplest model*
  (Strauss '86, Jonasson '99)

  nondirected graphs

  maximum entropy ensemble,
  constrain average degree $\langle k \rangle$
  and density of triangles $\langle m \rangle$

$$p(\mathbf{A}) \propto \mathrm{e}^{\alpha \sum_{ij} A_{ij} + \beta \sum_{ijk} A_{ij} A_{jk} A_{ki}}$$

  generating
  function:

$$\phi(\alpha, \beta) = \frac{1}{N} \log \sum_{\mathbf{A}} \mathrm{e}^{\alpha \sum_{ij} A_{ij} + \beta \sum_{ijk} A_{ij} A_{jk} A_{ki}}$$

$$\langle k \rangle = \partial \phi / \partial \alpha$$

$$\langle m \rangle = \partial \phi / \partial \beta$$

  – tricky in finite connectivity regime (sum over graphs?)

  – pathological behaviour ...

better loopy graph models?

– maximum entropy
– hard constrained degrees
– soft constrained nrs of short loops

- $\mathrm{Tr}(\mathbf{A}^\ell) = N\int\!\mathrm{d}\mu\;\mu^\ell\varrho(\mu|\mathbf{A})$,
  $\varrho(\mu|\mathbf{A})$: spectrum of **A**

$$p(\mathbf{A}) \propto \mathrm{e}^{N\int\mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}\prod_i \delta_{k_i,\sum_j A_{ij}}$$

  generating function:

$$\phi[\hat{\varrho}] = \frac{1}{N}\log\sum_{\mathbf{A}} \mathrm{e}^{N\int\mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}\prod_i \delta_{k_i,\sum_j A_{ij}} \qquad \varrho(\mu) = \delta\phi/\delta\hat{\varrho}(\mu)$$

- *simplest example*

  $k_i = q$ (regular random graphs)
  $\hat{\varrho}(\mu) = \alpha\mu^3$ (constrain nr of triangles)

$$p(\mathbf{A}) \propto \mathrm{e}^{\alpha\mathrm{Tr}(\mathbf{A}^3)}\prod_i \delta_{q,\sum_j A_{ij}}, \qquad \phi(\alpha) = \frac{1}{N}\log\sum_{\mathbf{A}} \mathrm{e}^{\alpha\mathrm{Tr}(\mathbf{A}^3)}\prod_i \delta_{q,\sum_j A_{ij}}$$

Making calculations feasible

graph ensemble :
$$p(\mathbf{A}) = \frac{1}{Z[\hat{\varrho}]}\, \mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_i \delta_{k_i,\sum_j A_{ij}}$$

generating function :
$$\Phi[\hat{\varrho}] = \frac{1}{N}\log \sum_{\mathbf{A}} \mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} \prod_i \delta_{k_i,\sum_j A_{ij}}$$

- derive via Edwards-Jones spectrum formula:

$$\mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} = \lim_{\varepsilon,\Delta\downarrow 0}\; \lim_{n_\mu \to \frac{\mathrm{i}\Delta}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\hat{\varrho}(\mu)}\; \lim_{m_\mu \to -n_\mu} \prod_\mu \Big[Z(\mu+\mathrm{i}\varepsilon|\mathbf{A})^{n_\mu}\; \overline{Z(\mu+\mathrm{i}\varepsilon|\mathbf{A})}^{\,m_\mu} \Big]$$

$$Z(\mu|\mathbf{A}) = \int_{\mathbb{R}^N} \mathrm{d}\boldsymbol{\phi}\; \mathrm{e}^{-\frac{1}{2}\mathrm{i}\boldsymbol{\phi}\cdot[\mathbf{A}-\mu\mathbb{1}]\boldsymbol{\phi}}$$

- 1. replica method
  2. path integral
  3. steepest descent for $N\to\infty$
  4. replica symmetric soln
  5. analytical continuation to *imaginary* (n_μ, m_μ)
  6. limits $\epsilon\downarrow 0$, $\Delta\downarrow 0$

Replica symmetric theory

$$\mathcal{W}[\{x\}] \;=\; \frac{1}{\mathcal{C}^2}\sum_{k>0} p(k)\frac{k}{\langle k\rangle}\frac{\mathcal{A}[\{x\}]\mathcal{F}_{k-1}[\{x\}]}{\int\{\mathrm{d}x'\}\mathcal{A}[\{x'\}]\mathcal{F}_k[\{x\}]}$$

$$\mathcal{C}^2 \;=\; \sum_{k>0} p(k)\frac{k}{\langle k\rangle}\frac{\int\{\mathrm{d}x\}\mathcal{A}[\{x\}]\mathcal{F}_{k-1}[\{x\}]}{\int\{\mathrm{d}x\}\mathcal{A}[\{x\}]\mathcal{F}_k[\{x\}]}$$

with

$$\mathcal{F}_k[\{x\}] \;=\; \Big[\prod_{\ell\leq k}\int\{\mathrm{d}x_\ell\}\;\mathcal{W}[\{x_\ell\}]\Big]\delta_{\mathrm{F}}\big[x - F[x_1,\ldots,x_k]\big]$$

$$\mathcal{A}[\{x\}] \;=\; \mathrm{e}^{-\frac{1}{2}\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\mathrm{sgn}[x(\mu)]}$$

spectrum:

$$\varrho(\mu) \;=\; -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\mu}\Big\{\sum_k p(k)\frac{\int\{\mathrm{d}x\}\;\mathcal{A}[\{x\}]\mathcal{F}_k[\{x\}]\mathrm{sgn}\big[x(\mu)\big]}{\int\{\mathrm{d}x\}\;\mathcal{A}[\{x\}]\mathcal{F}_k[\{x\}]}$$
$$+\langle k\rangle\mathcal{C}^2\int\{\mathrm{d}x\mathrm{d}x'\}\mathcal{W}[\{x\}]\mathcal{W}[\{x'\}]\mathcal{B}[\{x\},\{x'\}]$$
$$\times\theta[x(\mu)x'(\mu)]\theta[1-x(\mu)x'(\mu)]\mathrm{sgn}[x(\mu)+x'(\mu)]\Big\}$$

$$\mathcal{B}[\{x\},\{x'\}] = \mathrm{e}^{\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\frac{\mathrm{d}}{\mathrm{d}\mu}\left\{\theta[x(\mu)x'(\mu)]\theta[1-x(\mu)x'(\mu)]\mathrm{sgn}[x(\mu)+x'(\mu)]\right\}}$$

Limit of locally tree-like graphs

$\hat{\varrho}(\mu) \to 0$ for all μ: $\quad p(\mathbf{A}) \propto \prod_i \delta_{k_i, \sum_j A_{ij}}$

- entropy per node:

$$S = \frac{1}{2}\langle k \rangle \left[\log\left(\frac{N}{\langle k \rangle}\right) + 1 \right] + \sum_k p(k) \log \tilde{p}(k) + \epsilon_N$$ ✓

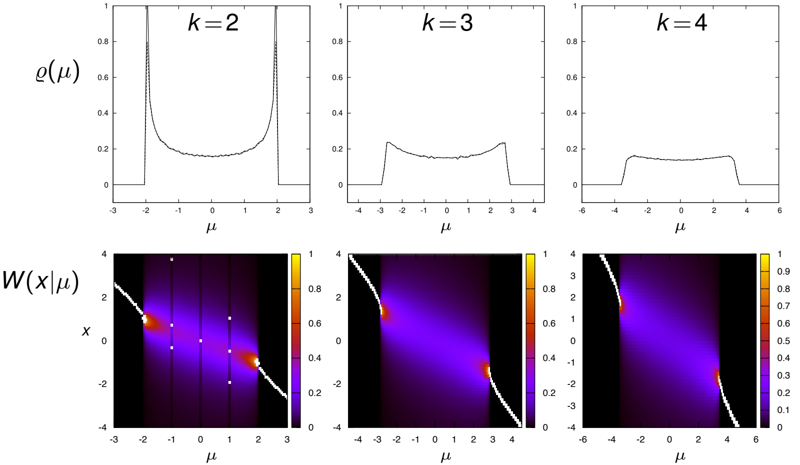
- spectrum for regular graphs, $k > 1$:
  McKay's '81 formula

$$\varrho(\mu) = \theta\left[2\sqrt{k-1} - |\mu|\right] \frac{k\sqrt{4(k-1)-\mu^2}}{2\pi(k^2-\mu^2)}$$ ✓

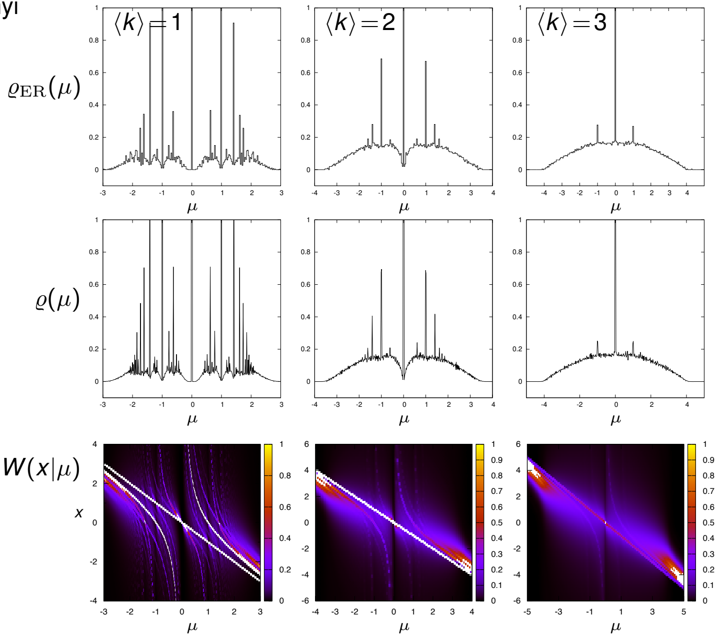
- spectrum for arbitrary $p(k)$:
  Dorogovtsev et al '03 formula

$$G(z|\mu) = 1 - \sqrt{z}\int_0^\infty \frac{\mathrm{d}y}{\sqrt{y}}\, \mathrm{e}^{\mathrm{i}y\mu}\, \Phi\big(G(y|\mu)\big) J_1(2\sqrt{yz})$$ ✓

Regular treelike graphs:



Erdös-Rényi graphs:

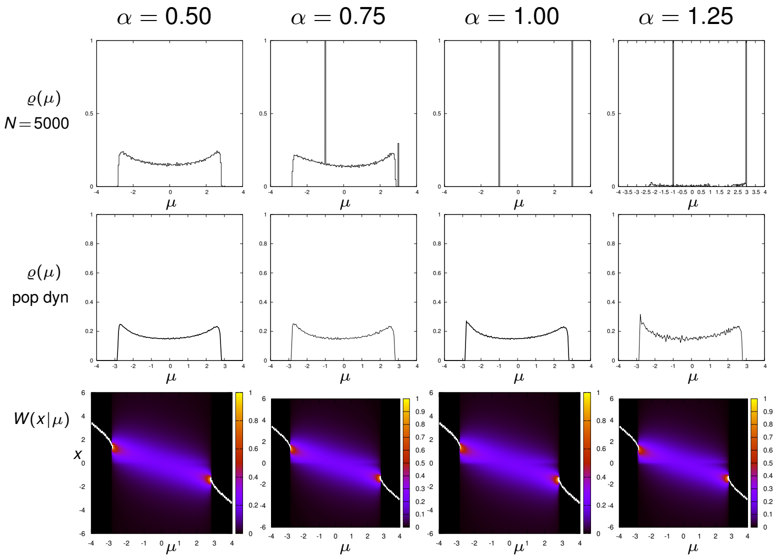


Spectra of loopy graphs

regular graphs, $k = 3$

$\hat{\varrho}(\mu) = \alpha \mathrm{Tr}(\mathbf{A}^3)$

2016 ...
what is going on?



Possible causes of the deviation

- Numerical precision?

  extend MCMC simulation times in graph generation,
  relation between μ-resolution in $x(\mu)$ and pop dynamics convergence

- Analytical technicalities?

  issues related to complex logarithm,
  other saddle-point types,
  order or limits $\epsilon\downarrow 0$, $\Delta\downarrow 0$, $n_\mu \to \mathrm{i}(\Delta/\pi)\frac{\mathrm{d}}{\mathrm{d}\mu}\hat{\varrho}(\mu)$, ...

- Some fundamental problem in the theory?

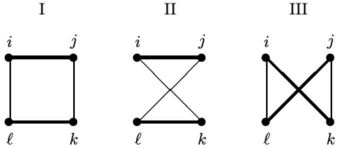
Numerical generation of tailored loopy graphs

- desired measure: $$p(\mathbf{A}) = \frac{\mathrm{e}^{-H(\mathbf{A})}}{Z} \prod_{i=1}^{N} \delta_{k_i,\sum_j A_{ij}}$$

- initialize **A** to *any* graph with correct degrees

- MCMC process:

  elementary moves:
  edge swaps
  (preserve degrees)



  move acceptance prob:

$$w(\mathbf{A} \to F\mathbf{A}) = \frac{1}{1 + \exp[\Delta H(\mathbf{A}) + \log(1+\Delta n(\mathbf{A})/n(\mathbf{A}))]}$$

- entropic term $\Delta n(\mathbf{A})/n(\mathbf{A})$ reflects *change* in nr of available moves

$$n(\mathbf{A}) = \underbrace{\frac{1}{4}N^2\langle k\rangle^2 + \frac{1}{4}N\langle k\rangle - \frac{1}{2}N\langle k^2\rangle}_{invariant} + \underbrace{\frac{1}{4}\mathrm{Tr}(\mathbf{A}^4) + \frac{1}{2}\mathrm{Tr}(\mathbf{A}^3) - \frac{1}{2}\sum_{ij} k_i A_{ij} k_j}_{state\ dependent}$$

Analytical technicalities

are we maybe too cavalier with complex logarithms?

- basic building block:
  Edwards-Jones formula

$$\varrho(\mu) = -\lim_{\epsilon\downarrow 0} \frac{2}{\pi N} \mathrm{Im} \frac{\partial}{\partial\mu} \log Z_N(\mu+\mathrm{i}\epsilon|\mathbf{A}), \qquad Z_N(\mu|\mathbf{A}) = \int_{\mathbb{R}^N} \mathrm{d}x\, \mathrm{e}^{-\frac{1}{2}\mathrm{i}\mathbf{x}\cdot(\mathbf{A}-\mu\mathbb{1})\mathbf{x}}$$

- naive derivation uses $\log\prod_i z_i = \sum_i \log z_i$
  (true only for $z_i \in \mathbb{R}$)

  any conventional defn of
  complex log (wherever the cut):

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \mathrm{Im} \log Z_N &= \lim_{N\to\infty} \frac{1}{N} \mathrm{Im}\Big(\log|Z_N| + \mathrm{i}\mathrm{Arg}(Z_N)\Big) \\ &= \lim_{N\to\infty} \frac{1}{N} \mathrm{Arg}(Z_N) = 0 \quad \ldots \end{aligned}$$

- E-J formula has been very successful,
  subtleties become problematic only for imaginary replicas?

Possible causes of the deviation

- Numerical precision?

  extend MCMC simulation times in graph generation,
  relation between μ-resolution in $x(\mu)$ and pop dynamics convergence

- Analytical technicalities?

  issues related to complex logarithm,
  other saddle-point types,
  order or limits $\epsilon \downarrow 0$, $\Delta \downarrow 0$, $n_\mu \to \mathrm{i}(\Delta/\pi)\frac{\mathrm{d}}{\mathrm{d}\mu}\hat{\varrho}(\mu)$, ...

- Some fundamental problem in the theory?

find exactly solvable loopy graph models
that don't require imaginary replicas

Definition of the toy model

- ensemble of degree-2 regular random graphs

$$p(\mathbf{A}) = \frac{1}{Z_N(\boldsymbol{\alpha})} \times \overbrace{e^{\sum_{\ell=3}^{K} \ell\alpha_\ell n_\ell(\mathbf{A})}}^{\textit{control nrs of cycles}} \times \overbrace{\prod_{i=1}^{N} \delta_{2,\sum_j A_{ij}}}^{\textit{graphs: sets of rings}}$$

$n_\ell(\mathbf{A})$: nr of length-ℓ cycles, K finite

- generating function

$$\phi_N(\boldsymbol{\alpha}) = \frac{1}{N}\log[Z_N(\boldsymbol{\alpha})/N!], \qquad Z_N(\boldsymbol{\alpha}) = \sum_{\mathbf{A}} e^{\sum_{\ell=3}^{K} \ell\alpha_\ell n_\ell(\mathbf{A})} \prod_{i=1}^{N} \delta_{2,\sum_j A_{ij}}$$

fraction of nodes in length-ℓ cycles

$$m_\ell = \frac{\ell}{N}\langle n_\ell(\mathbf{A})\rangle = \frac{\partial}{\partial \alpha_\ell}\phi_N(\boldsymbol{\alpha})$$

- solution:

combinatorics of dividing N nodes into closed rings of different lengths

Solution

- result of combinatorics

$$\phi_N(\boldsymbol{\alpha}) \;=\; \frac{1}{N}\log\int_{-\pi}^{\pi}\frac{\mathrm{d}\omega}{2\pi}\,\mathrm{e}^{Nf_N(\omega,\boldsymbol{\alpha})}$$

$$f_N(\omega,\boldsymbol{\alpha}) \;=\; \mathrm{i}\omega+\sum_{\ell=3}^{K}\frac{\mathrm{e}^{(\alpha_\ell-\mathrm{i}\omega)\ell}}{2\ell N}+\sum_{\ell=K+1}^{N}\frac{\mathrm{e}^{-\mathrm{i}\omega\ell}}{2\ell N}$$

- Large N

$$\phi(\boldsymbol{\alpha})=\lim_{N\to\infty}\mathrm{extr}_\omega\; f_N(\omega,\boldsymbol{\alpha})$$

  - finite $\{\alpha_\ell\}$: $m_\ell=0$ for all ℓ,

    $\{\alpha_\ell\}$ must scale differently with N
    to induce finite cycle densities!

  - $\alpha_\ell=\tilde{\alpha}_\ell+\ell^{-1}\log N$

$$\tilde{f}_N(\omega,\boldsymbol{\alpha}) \;=\; \mathrm{i}\omega+\sum_{\ell=3}^{K}\frac{\mathrm{e}^{(\tilde{\alpha}_\ell-\mathrm{i}\omega)\ell}}{2\ell}+\sum_{\ell=K+1}^{N}\frac{\mathrm{e}^{-\mathrm{i}\omega\ell}}{2\ell N}$$

Phenomenology of the toy model

fraction of nodes in large rings

$$m_\infty = 1 - \sum_{\ell=3}^{K} m_\ell$$

graph spectrum:

$$\varrho(\mu) = \sum_{\ell=3}^{K} m_\ell \, \varrho_\ell(\mu) + m_\infty \frac{\theta(2-|\mu|)}{\pi\sqrt{4-\mu}}$$

$$\varrho_\ell(\mu) = \frac{1}{\ell} \sum_{r=0}^{\ell-1} \delta[\mu - 2\cos(2\pi r/\ell)]$$

- Two phases

connected phase: $m_\infty > 0$

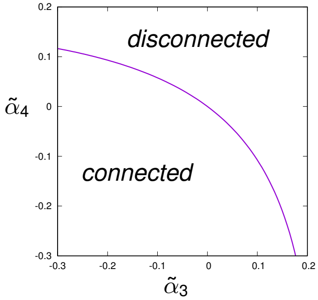
presence of extensively large rings,
found for small $\{\alpha_\ell\}$

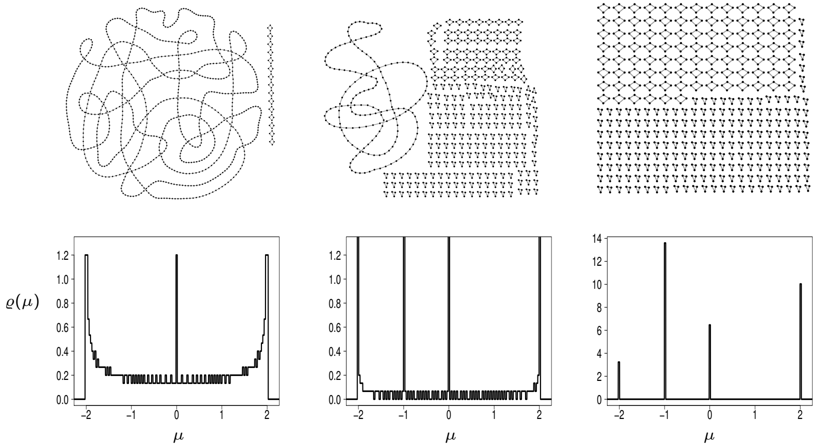
disconnected phase: $m_\infty = 0$

only short rings,
found for large $\{\alpha_\ell\}$

percolation transition: $\sum_{\ell=3}^{K} \mathrm{e}^{\ell\tilde{\alpha}_\ell} = 2$

$K = 4$





$$\textit{disconnected phase}: \qquad m_\ell = \frac{1}{2} x^\ell \mathrm{e}^{\ell \tilde{\alpha}_\ell}, \qquad 1 = \frac{1}{2} \sum_{\ell=3}^{K} x^\ell \mathrm{e}^{\ell \tilde{\alpha}_\ell}$$

$$\textit{connected phase}: \qquad m_\ell = \frac{1}{2} \mathrm{e}^{\ell \tilde{\alpha}_\ell}, \qquad m_\infty = 1 - \frac{1}{2} \sum_{\ell=3}^{K} \mathrm{e}^{\ell \tilde{\alpha}_\ell}$$

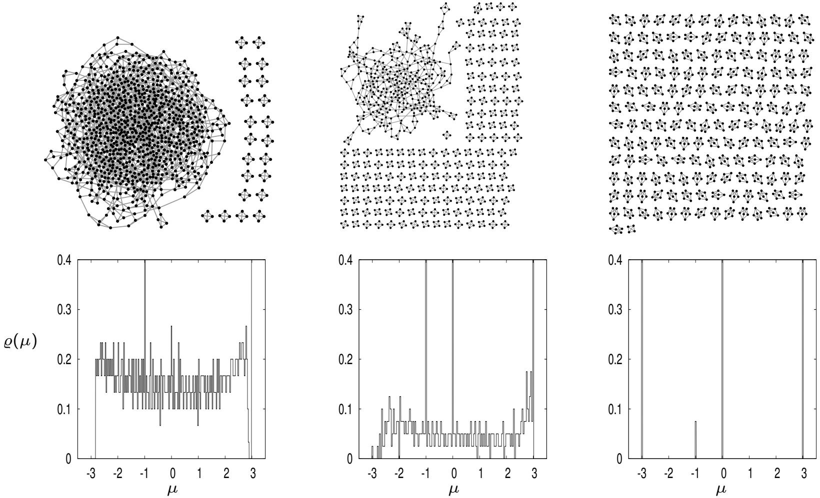
Graphs with degree $k \geq 3$

$$p(\mathbf{A}) = \frac{1}{Z_N(\boldsymbol{\alpha})} \mathrm{e}^{\alpha_3 n_3(\mathbf{A}) + \alpha_4 n_4(\mathbf{A})} \prod_{i=1}^{N} \delta_{3,\sum_j A_{ij}}$$

$$\alpha_\ell = \tilde{\alpha}_\ell + \ell^{-1} \log N$$

no longer solvable directly,
but similar phenomenology ...

$k = 3$:



Lessons from the toy model

- to achieve $\mathcal{O}(1)$ short loop densities and modified spectra, we need (for regular random graphs):

$$\hat{\varrho}(\mu) \quad \rightarrow \quad \log N\, \varphi(\mu) + \hat{\varrho}(\mu)$$

  but this gives integrals in the replica theory
  that may no longer be amenable to steepest descent ...

- options

  1. keep $\hat{\varrho}(\mu) = \mathcal{O}(1)$:

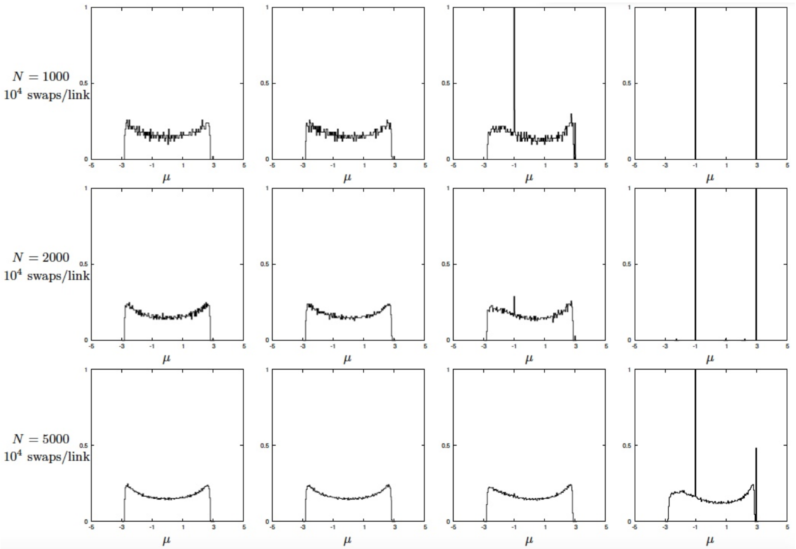
     study *finite size corrections* to loop numbers and spectra,
     extend theory to include also sub-leading order in N

  2. choose $\hat{\varrho}(\mu) = \mathcal{O}(\log N)$:

     requires extending steepest descent arguments
     to expressions of the form

$$f = \lim_{M\to\infty} \frac{1}{M} \log \int_{\mathbb{R}^M} \mathrm{d}x\, \mathrm{e}^{MF(\mathbf{x})}$$

simulations indeed confirm that with $\hat{\varrho}(\mu) = \alpha\mu^3$, the deviations from McKay's spectrum, unexplained by the replica theory, were *finite size effects ...*



steepest descent type integrals in high dimensions

- define

$$I(d,N) = \frac{1}{N}\log\int_{\mathbb{R}^d} \mathrm{d}\mathbf{x}\ \mathrm{e}^{-N\Phi(\mathbf{x})}, \qquad d,N\to\infty, \quad d/N=\alpha \ \textit{finite}$$

- standard steepest descent:
  $d/N \to 0$, sufficiently fast

- For $\alpha > 0$:

$$I(d,N) = \min_{\mathbf{x}}\Phi(\mathbf{x}) + \frac{1}{2}\alpha\log(\frac{d}{2\pi\alpha}) + \frac{\alpha}{2d}\sum_{i=1}^{d}\log a_i + o(1)$$

  (as if approx by Gaussian integral)
  provided:

$$\lim_{d\to\infty}\frac{1}{d}\sum_{i=1}^{d}\frac{1}{a_i} = 0 \quad \textit{and} \quad \lim_{d\to\infty}\frac{1}{d^2}\sum_{i=1}^{d}\frac{1}{a_i^2} = 0$$

  $\{a_i\}$: eigenvals of curvature matrix of Φ at minimum

Imaginary replica analysis in two leading orders

ensemble:

$$p(\mathbf{A}) = \frac{\mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}}{Z[\hat{\varrho}]}\prod_{i=1}^{N}\delta_{q,\sum_j A_{ij}} \qquad Z[\hat{\varrho}] = \sum_{\mathbf{A}} \mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})}\prod_{i=1}^{N}\delta_{q,\sum_j A_{ij}}$$

- find $\phi[\hat{\varrho}]$ in leading two orders in N

$$\phi[\hat{\varrho}] = \frac{1}{N}\log Z[\hat{\varrho}], \qquad \varrho(\mu) = \frac{\delta\phi[\hat{\varrho}]}{\delta\hat{\varrho}(\mu)}$$

$$\mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} = \lim_{\Delta\to 0}\;\lim_{n_\mu\to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)}\;\lim_{m_\mu\to -n_\mu}\prod_\mu Z(\mu_\epsilon|\mathbf{A})^{n_\mu}\,\overline{Z(\mu_\epsilon|\mathbf{A})}^{m_\mu}$$

$$Z(\mu_\epsilon|\mathbf{A}) = \int\prod_{i=1}^{N}\mathrm{d}\phi^i\,\exp\Big[-\frac{\mathrm{i}}{2}\sum_{ij}\phi^i(A_{ij}-\mu_\epsilon\delta_{ij})\phi^j\Big], \qquad \mu_\epsilon \equiv \mu+\mathrm{i}\varepsilon$$

replicas:

$$\prod_\mu Z(\mu_\epsilon|\mathbf{A})^{n_\mu}\,\overline{Z(\mu_\epsilon|\mathbf{A})}^{m_\mu} = \int \mathrm{d}\mathbf{\Phi}\,\mathrm{d}\mathbf{\Psi}\,\mathrm{e}^{-\frac{\mathrm{i}}{2}\sum_{\mu,\alpha_\mu}\sum_{ij}\phi^i_{\mu,\alpha_\mu}\phi^j_{\mu,\alpha_\mu}(A_{ij}-\mu_\epsilon\delta_{ij})+\frac{\mathrm{i}}{2}\sum_{\mu,\beta_\mu}\sum_{ij}\psi^i_{\mu,\beta_\mu}\psi^j_{\mu,\beta_\mu}(A_{ij}-\overline{\mu_\epsilon}\delta_{ij})}$$

connect notation to Metz et al (2014) and Lucibello et al (2014)
(finite size calculations for *treelike* graphs)

- $\boldsymbol{\varphi} \in \mathbb{R}^d$, where $d = \sum_\mu n_\mu + \sum_\mu m_\mu$:

$$\boldsymbol{\varphi} = \begin{pmatrix} \{\phi_{\mu,\alpha_\mu}\} \\ \{\psi_{\mu,\beta_\mu}\} \end{pmatrix}, \qquad \boldsymbol{\varphi}\cdot\boldsymbol{\varphi}' = \sum_{\mu,\alpha_\mu} \phi_{\mu,\alpha_\mu}\phi'_{\mu,\alpha_\mu} - \sum_{\mu,\beta_\mu} \psi_{\mu,\beta_\mu}\psi'_{\mu,\beta_\mu}$$

$d \times d$ matrix $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_{1\,\varepsilon}\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \dots & & & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & & & & \\ & \mathbf{0} & \mu_{M\,\varepsilon}\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & & \vdots \\ \vdots & & \mathbf{0} & \overline{\mu_{1\,\varepsilon}}\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & & \\ & & & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & & \dots & & \mathbf{0} & \overline{\mu_{M\,\varepsilon}}\mathbf{I}_{m_{\mu_M}} \end{pmatrix}$$

and

$$\nu(\boldsymbol{\varphi}) = \mathrm{e}^{\frac{1}{2}\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\mu}\boldsymbol{\varphi}}, \qquad \lim = \lim_{\Delta\to 0}\ \lim_{n_\mu \to \mathrm{i}\frac{\Delta}{\pi}\hat{\varrho}'(\mu)}\ \lim_{m_\mu\to -n_\mu}$$

Result over sum over **A**
and expansion:

$$\phi[\hat{\varrho}] = \lim \frac{1}{N} \log \int \mathcal{D}P\mathcal{D}\hat{P}\, e^{N\, S[P,\hat{P}]} + \mathcal{O}(\frac{1}{N^2}),$$

$$\begin{aligned} S[P,\hat{P}] &= -\frac{q}{2}(1+\frac{q-2}{2N}) + \mathrm{i}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\hat{P}(\boldsymbol{\varphi},\omega) \\ &\quad + \frac{q}{2}(1+\frac{q}{N})\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\ P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-\mathrm{i}\omega-\mathrm{i}\omega'} \\ &\quad - \frac{q^2}{4N}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\mathrm{d}\omega'\ P(\boldsymbol{\varphi},\omega)P(\boldsymbol{\varphi}',\omega')\mathrm{e}^{-2\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'-2\mathrm{i}\omega-2\mathrm{i}\omega'} \\ &\quad - \frac{q}{2N}\int \mathrm{d}\boldsymbol{\varphi}\mathrm{d}\omega\ P(\boldsymbol{\varphi},\omega)\mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}-2\mathrm{i}\omega} + \log\int \frac{\mathrm{d}\omega\mathrm{d}\boldsymbol{\varphi}}{2\pi}\nu(\boldsymbol{\varphi})\mathrm{e}^{\mathrm{i}\omega q-\mathrm{i}\hat{P}(\boldsymbol{\varphi},\omega)} \end{aligned}$$

same form as Metz et al (2014),
but different defs of $\boldsymbol{\varphi}$ and $\nu(\boldsymbol{\varphi})$

- retrace calculations of Metz et al ...

$$\phi[\hat{\varrho}] = \lim \left\{ \log Z_q + \sum_{\ell=3}^{\infty} \frac{(q-1)^\ell}{2N\ell} \mathrm{Tr}(\mathbf{M}^\ell) \right\}$$

with

$$Z_q = \int \mathrm{d}\boldsymbol{\varphi}\ \nu(\boldsymbol{\varphi}) \Big[\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') \Big]^q, \qquad U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') = \mathrm{e}^{-\mathrm{i}\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}'}$$

$$M(\boldsymbol{\varphi}, \boldsymbol{\varphi}') = \frac{\nu(\boldsymbol{\varphi})}{Z_q} \Big[\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') \Big]^{q-2} U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}'),$$

order par eqn:

$$W_0(\boldsymbol{\varphi}) = \frac{\nu(\boldsymbol{\varphi})}{Z_q} \Big[\int \mathrm{d}\boldsymbol{\varphi}'\ U_1(\boldsymbol{\varphi}, \boldsymbol{\varphi}') W_0(\boldsymbol{\varphi}') \Big]^{q-1}$$

same form as Metz et al (2014),
but different defs of $\boldsymbol{\varphi}$ and $\nu(\boldsymbol{\varphi})$

- RS ansatz,

  now differences with Metz et al
  become apparent

$$W_0(\varphi) = \mathcal{C} \int \mathrm{d}\mathbf{X}\; W(\mathbf{X}) \big[\mathrm{e}^{-\frac{1}{2}\mathrm{i}\varphi\cdot\mathbf{X}\varphi} / Z(\mathbf{X}) \big]$$

$$\int \mathrm{d}\mathbf{X}\; W(\mathbf{X}) = 1, \qquad Z(\mathbf{X}) = \prod_{\mu} \Big(\frac{2\pi}{\mathrm{i}x(\mu)}\Big)^{\frac{n_\mu}{2}} \overline{\Big(\frac{2\pi}{\mathrm{i}x(\mu)}\Big)}^{\frac{m_\mu}{2}}$$

$$\mathbf{X} = \begin{pmatrix} x(\mu_1)\,\mathbf{I}_{n_{\mu_1}} & \mathbf{0} & & \dots & & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & & & \\ & \mathbf{0} & x(\mu_M)\,\mathbf{I}_{n_{\mu_M}} & \mathbf{0} & & \vdots \\ \vdots & & \mathbf{0} & \overline{x(\mu_1)}\,\mathbf{I}_{m_{\mu_1}} & \mathbf{0} & \\ & & & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & & \dots & & \mathbf{0} & \overline{x(\mu_M)}\,\mathbf{I}_{m_{\mu_M}} \end{pmatrix}$$

- RS order par eqn:

$$W(\mathbf{X}) = \frac{Z(\mathbf{X})}{\mathcal{Z}_{q-1}} \int \Big(\prod_{k=1}^{q-1} \mathrm{d}\mathbf{X}_k W(\mathbf{X}_k)\Big)\, \delta\Big(\mathbf{X}+\boldsymbol{\mu}+\sum_{k=1}^{q-1}\mathbf{X}_k^{-1}\Big)$$

$$\mathcal{Z}_q = \int \mathrm{d}\mathbf{X}\Big(\prod_{k=1}^{q} \mathrm{d}\mathbf{X}_k W(\mathbf{X}_k)\Big)\, Z(\mathbf{X})\, \delta\Big(\mathbf{X}+\boldsymbol{\mu}+\sum_{k=1}^{q}\mathbf{X}_k^{-1}\Big)$$

- special soln:

$$W(\mathbf{X}) = \delta(\mathbf{X}-\mathbf{X}^\star), \qquad x^\star(\mu) = -\frac{1}{2}\mu_\epsilon - \frac{1}{2}\mathrm{i}\sqrt{4(q-1)-\mu_\epsilon^2}$$

$$\phi[\hat{\varrho}] = \lim \Big\{\frac{1}{2}q\log Z(\mathbf{X}^\star) - \frac{1}{2}(q-2)\log Z(\mathbf{X}^\star - (\mathbf{X}^\star)^{-1}) + \sum_{\ell=3}^{\infty}\frac{(q-1)^\ell}{2N\ell}\frac{1}{Z^\ell(\mathbf{X}^\star)}\prod_{\mu}\Big[Z(\mu_\epsilon|\mathbf{A}^\star_{\ell,\mu})^{n_\mu}\,\overline{Z(\mu_\epsilon|\mathbf{A}^\star_{\ell,\mu})}^{m_\mu}\Big]\Big\}$$

$\mathbf{A}^\star_{\ell,\mu}$: $\ell\times\ell$ adjacency matrix of length-ℓ loop, with complex field on diagonal:

$$\big(\mathbf{A}^\star_{\ell,\mu}\big)_{kk'} = \delta_{k,k'+1}+\delta_{k,k'-1}+\frac{2-q}{x^\star(\mu)}\delta_{kk'} \qquad (\textit{with } k \bmod \ell)$$

- imaginary replica limits

$$\phi[\hat{\varrho}] = \int \mathrm{d}\mu\ \hat{\varrho}(\mu)\varrho_0(\mu) + \frac{1}{N}\sum_{\ell=3}^{\infty}\frac{(q-1)^{\ell}}{2\ell}\ \mathrm{e}^{\ell\int\mathrm{d}\mu\ \hat{\varrho}(\mu)G_\ell(\mu)} + o(\frac{1}{N})$$

$$\varrho(\mu) = \varrho_0(\mu) + \overbrace{\frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell}\mathrm{e}^{\ell\int\mathrm{d}\mu\ \hat{\varrho}(\mu)G_\ell(\mu)}G_\ell(\mu) + o(\frac{1}{N})}^{\textit{finite size corrections}}$$

with

$$\varrho_0(\mu) = \frac{q}{2\pi}\frac{\sqrt{4(q-1)-\mu^2}}{q^2-\mu^2}\,\theta\left[2\sqrt{q-1}-|\mu|\right] \qquad (\textit{McKay's formula})$$

$$G_\ell(\mu) = \frac{1}{\ell\pi}\mathrm{Im}\left\{\sum_{k=1}^{\ell}\frac{1+\mathrm{i}\frac{q-2}{x^\star(\mu)}[4(q-1)-\mu^2]^{-\frac{1}{2}}}{2\cos(2\pi k/\ell)-\frac{q-2}{x^\star(\mu)}-\mu}\right\} - \frac{1}{\pi}\frac{\theta\left(2\sqrt{q-1}-|\mu|\right)}{\sqrt{4(q-1)-\mu^2}}$$

- interpretation

$(q-1)^{\ell}/2\ell$: expected nr of length-ℓ loops in random regular graph

$\ell G_\ell(\mu)$: correction to McKay formula due to single length-ℓ loop

$\rightarrow$ finite size corrections from well separated loops

- remaining integrals
  that involve $\hat{\varrho}(\mu)$:

$$\varrho(\mu) \;=\; \varrho_0(\mu) + \frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell}\mathrm{e}^{\ell\mathcal{J}_\ell[\hat{\varrho}]}\; G_\ell(\mu) + o(\frac{1}{N})$$

$$\mathcal{J}_\ell[\hat{\varrho}] \;=\; \sum_{p=1}^{\infty}\frac{1}{(q-1)^{p\ell/2}}\;\frac{2}{\pi}\int_{-1}^{1}\frac{\mathrm{d}t}{\sqrt{1-t^2}}\;\hat{\varrho}(2t\sqrt{q-1})\,T_{p\ell}(t).$$

$T_n(t)$: Chebyshev polynomials, $t\in[-1,1]$:

$$T_n(t) = \mathrm{Re}\big(t+\mathrm{i}\sqrt{1-t^2}\big)^n : \qquad T_0(t)=1,\quad T_1(t)=t,\quad T_2(t)=2t^2-1,\ldots$$

$$\frac{2}{\pi}\int_{-1}^{1}\frac{\mathrm{d}t}{\sqrt{1-t^2}}\;T_n(t)\,T_m(t) = \delta_{nm}(1+\delta_{m0})$$

Tests of the theory

- *flat measures:*

  $\hat{\varrho}(\mu) = 0$

  $$\varrho(\mu) = \varrho_0(\mu) + \frac{1}{2N}\sum_{\ell=3}^{\infty}(q-1)^{\ell} G_{\ell}(\mu) + o(N^{-1})$$

  reproduces formula of Metz et al (2014) ✓

- *toy model:*

  $q = 2, \ \hat{\varrho}(\mu) = \alpha\mu^3$

  $$\varrho(\mu) = \varrho_0(\mu) + \frac{1}{2N}\sum_{\ell=3}^{\infty} \mathrm{e}^{6\alpha\delta_{\ell 3}}\Big\{\frac{1}{\ell}\sum_{k=1}^{\ell}\delta\Big[\mu - 2\cos(\frac{2\pi k}{\ell})\Big] - \frac{1}{\pi}\frac{\theta(2-|\mu|)}{\sqrt{4-\mu^2}}\Big\} + o(\frac{1}{N}) \quad ✓$$

  triangle density:

  $$m(\alpha) \equiv \int \mathrm{d}\mu\ \varrho(\mu)\mu^3 = N^{-1}\mathrm{e}^{6\alpha} \quad ✓$$

Degree-q triangularly constrained regular graphs

$\hat{\varrho}(\mu) = \alpha\mu^3$

- finite size corrections to spectrum:

$$\varrho(\mu) = \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) + \frac{1}{N}\tilde{\varrho}_1(\mu) + o(\frac{1}{N}),$$

$\varrho_0(\mu)$: McKay formula

$\varrho_1(\mu)$: corrections for non-deformed measure

$\tilde{\varrho}_1(\mu)$: impact of promoting triangles

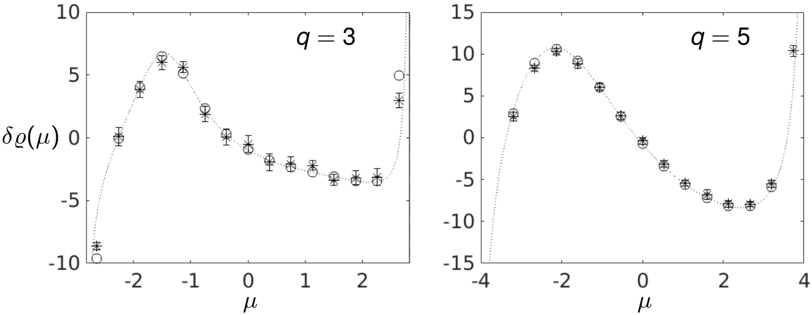
$$\tilde{\varrho}_1(\mu) = \frac{(q-1)^3}{2\pi}(\mathrm{e}^{6\alpha}-1)\frac{\theta[2\sqrt{q-1}-|\mu|]}{\sqrt{4(q-1)-\mu^2}}\left\{\frac{q-2}{3}\left[\frac{2q+\mu}{q^2-3(q-1)+\mu q+\mu^2}+\frac{1}{q-\mu}\right]-1\right\}$$

- triangle density:

$$m(\alpha) = \frac{1}{N}(q-1)^3\mathrm{e}^{6\alpha} + o(\frac{1}{N}).$$

- accurate for $\alpha < \alpha_1(N)$,
  theory implicitly assumes triangles are well separated
  $\alpha > \alpha_1(N)$: alternative saddle point?

- predictions versus simulations



$\delta\varrho(\mu) = N[\varrho(\mu) - \varrho_0(\mu)]$

left: $q = 3$, $N = 1000$, $\alpha = 0.416$
right: $q = 5$, $N = 2000$, $\alpha = 0.431$

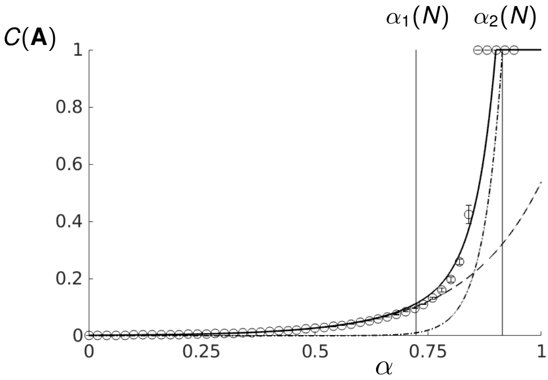
both examples:
clustering coeff $\langle C(\mathbf{A})\rangle \approx 0.02$

Beyond the region of validity

$\alpha > \alpha_1(N)$?

add contributions
from triangles
in cliques

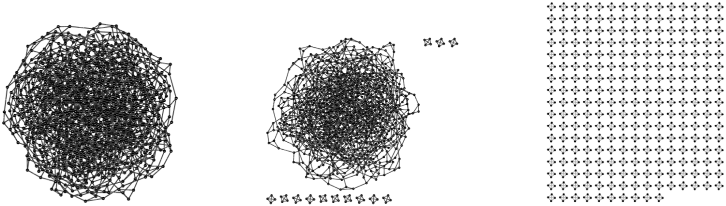
here: $q = 3$



dashed: *(i) present replica theory, isolated triangles*
dotted-dashed: *(ii) triangles in cliques*
solid: *(i)+(ii)*

$\alpha_1(N)$ and $\alpha_2(N)$ grow logarithmically with N

Qualitative picture for $q \geq 3$



connected $\alpha < \alpha_1(N)$ | *partially connected* $\alpha_1(N) < \alpha < \alpha_2(N)$ | *disconnected* $\alpha > \alpha_2(N)$

- *small* α:

  isolated triangles embedded in giant component

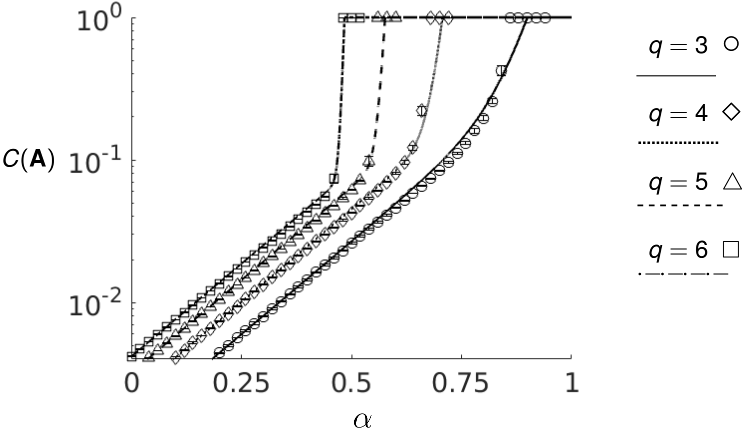
- *intermediate* α:

  triangles start touching, increasing nr of cliques of $q+1$ nodes

- *large* α:

  graphs shattered fully into cliques of $q+1$ nodes
  (graphs with max nr of triangles, ground state)

comparison to simulations with $N = 1000$
degree $q = 3 \ldots 6$



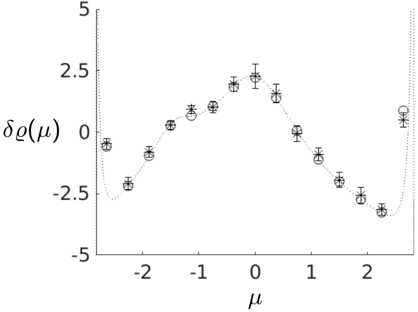
Other ensembles with $q = 3$

$$p(\mathbf{A}) = \frac{\mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)+\beta \mathrm{Tr}(\mathbf{A}^4)}}{Z(\alpha,\beta)} \prod_{i=1}^{N} \delta_{3,\sum_j A_{ij}}, \qquad Z(\alpha,\beta) = \sum_{\mathbf{A}\in G} \mathrm{e}^{\alpha \mathrm{Tr}(\mathbf{A}^3)+\beta \mathrm{Tr}(\mathbf{A}^4)} \prod_{i=1}^{N} \delta_{3,\sum_j A_{ij}}$$

prediction:

$$\begin{aligned} \varrho(\mu) &= \varrho_0(\mu) + \frac{1}{N}\varrho_1(\mu) + \frac{(q-1)^3}{2N}(\mathrm{e}^{6\alpha}-1)G_3(\mu) \\ &\quad + \frac{(q-1)^4}{2N}(\mathrm{e}^{8\beta}-1)G_4(\mu) + o(N^{-1}) \end{aligned}$$

$N = 2000$,
$\alpha = \beta = 0.2$

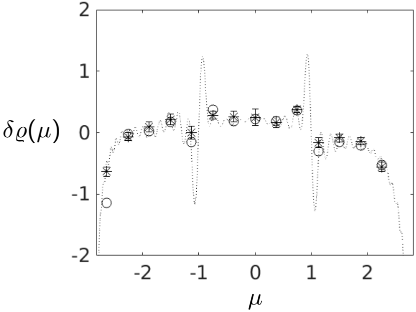


$$p(\mathbf{A}) = \frac{\mathrm{e}^{\alpha \int_{-1}^{1} \mathrm{d}\mu \;\; \varrho(\mu|\mathbf{A})}}{Z(\alpha)} \prod_{i=1}^{N} \delta_{3,\sum_j A_{ij}}, \qquad Z(\alpha,\beta) = \sum_{\mathbf{A}\in G} \mathrm{e}^{\alpha \int_{-1}^{1} \mathrm{d}\mu \;\; \varrho(\mu|\mathbf{A})} \prod_{i=1}^{N} \delta_{3,\sum_j A_{ij}}$$

prediction:

$$\varrho(\mu) \;\; = \;\; \rho_0(\mu) + \frac{1}{N}\rho_1(\mu) + \sum_{\ell\geq 3} \frac{(q-1)^{\ell}}{2N} \left[\mathrm{e}^{\alpha\ell \int_{-1}^{1} \mathrm{d}\mu \;\; G_\ell(\mu)} - 1 \right] G_\ell(\mu),$$

$N = 1000$,
$\alpha = 0.5$



simulations non-trivial ...

Summary

- analytical approach to (processes on) loopy random graphs,
  max entropy ensembles with constrained *degrees and spectrum*

- imaginary replica formula to handle spectral constraint:

$$\mathrm{e}^{N\int \mathrm{d}\mu\;\hat{\varrho}(\mu)\varrho(\mu|\mathbf{A})} = \lim_{\varepsilon,\Delta\downarrow 0}\prod_{\mu}\Big[Z(\mu+\mathrm{i}\varepsilon|\mathbf{A})^{\mathrm{i}n(\mu)}\;\overline{Z(\mu+\mathrm{i}\varepsilon|\mathbf{A})}^{\;-\mathrm{i}n(\mu)}\Big]$$

$$Z(\mu|\mathbf{A}) = \int \mathrm{d}\boldsymbol{\phi}\;\mathrm{e}^{-\frac{1}{2}\mathrm{i}\boldsymbol{\phi}\cdot[\mathbf{A}-\mu\mathbf{1}]\boldsymbol{\phi}}, \qquad n(\mu) = \frac{\Delta}{\pi}\frac{\mathrm{d}}{\mathrm{d}\mu}\hat{\varrho}(\mu)$$

- RS eqns: reproduced tree-like formulae for $\hat{\varrho}(\mu)\to 0$,
  look natural (message passing + nontrivial acceptance probs),
  but disagreed with simulations of loopy ones ...

- simple solvable toy model: $p(k)=\delta_{k2}$, $\hat{\varrho}(\mu)=\sum_{\ell\geq 3}\alpha_\ell\mu^\ell$
  finite nonzero loop *densities* require $\alpha_\ell = O(\log N)$

- imaginary replica theory correctly describes *finite size corrections*
  to spectra of *weakly loopy* graphs

- finite nonzero loop *densities*: $\hat{\varrho}(\mu)=O(\log N)$,
  need to revisit steepest descent evaluation of path integrals ...

Papers, talks, seminars

https://nms.kcl.ac.uk/ton.coolen

Coolen, J. Phys. Conference Series 699 (2016)
Coolen, Annibale, Roberts, *Generating random networks and graphs* (2017)
Aguirre-Lopez, Barucca, Fekom, Coolen, J. Phys. A 51 (2018)
Aguirre-Lopez and Coolen, preprint arXiv:1907.06703 (2019)

January 2020

King's College London $\longrightarrow$ Radboud University