Statistical Mechanics of Large Dense Random Graphs

Lorenzo Sadun

University of Texas

Avronfest, Jerusalem, July 3, 2013

Joint work with Charles Radin

• A Graph is a collection of n vertices and up to $\binom{n}{2}$ undirected edges.

э

- A Graph is a collection of n vertices and up to $\binom{n}{2}$ undirected edges.
- Points are distinguishable and abstract. There is no geometry.

- A Graph is a collection of n vertices and up to $\binom{n}{2}$ undirected edges.
- Points are distinguishable and abstract. There is no geometry.
- A Random Graph also has a measure on the $2^{\binom{n}{2}}$ possibilities.

- A Graph is a collection of n vertices and up to $\binom{n}{2}$ undirected edges.
- Points are distinguishable and abstract. There is no geometry.
- A Random Graph also has a measure on the $2^{\binom{n}{2}}$ possibilities.
- A Large Random Graph means taking the $n \to \infty$ limit.

- A Graph is a collection of n vertices and up to $\binom{n}{2}$ undirected edges.
- Points are distinguishable and abstract. There is no geometry.
- A Random Graph also has a measure on the $2^{\binom{n}{2}}$ possibilities.
- A Large Random Graph means taking the $n \to \infty$ limit.
- A Large Dense Random Graph means non-trivial edge density.

$$e := \frac{\# \text{ of edges}}{n^2/2} \not\rightarrow 0.$$

$$t := \frac{\# \text{ of triangles}}{n^3/6.}$$

• What values of (e, t) are possible in the $n \to \infty$ limit?

18 ►

э

• What values of (e, t) are possible in the $n \to \infty$ limit?

• How many graphs of given
$$(e, t)$$
 are possible?
 $Z_{e,t}^{\delta,n} =$ number of graphs with *n* vertices such that
 $\left|\frac{\# \text{ edges}}{n^2/2} - e\right| < \delta \text{ and } \left|\frac{\# \text{ triangles}}{n^3/6} - t\right| < \delta.$

18 ►

э

• What values of (e, t) are possible in the $n \to \infty$ limit?

• How many graphs of given
$$(e, t)$$
 are possible?
 $Z_{e,t}^{\delta,n} =$ number of graphs with n vertices such that
 $\left|\frac{\# \text{ edges}}{n^2/2} - e\right| < \delta$ and $\left|\frac{\# \text{ triangles}}{n^3/6} - t\right| < \delta$.
 $s(e, t) = \lim_{\delta \to 0} \lim_{n \to \infty} \frac{\ln Z_{e,t}^{\delta,n}}{n^2}$.

3

• What values of (e, t) are possible in the $n \to \infty$ limit?

• How many graphs of given
$$(e, t)$$
 are possible?
 $Z_{e,t}^{\delta,n} =$ number of graphs with n vertices such that
 $\left|\frac{\# \text{ edges}}{n^2/2} - e\right| < \delta$ and $\left|\frac{\# \text{ triangles}}{n^3/6} - t\right| < \delta$.
 $s(e, t) = \lim_{\delta \to 0} \lim_{n \to \infty} \frac{\ln Z_{e,t}^{\delta,n}}{n^2}$.

• What does a typical graph with a given (e, t) look like?

э

• What values of (e, t) are possible in the $n \to \infty$ limit?

• How many graphs of given
$$(e, t)$$
 are possible?
 $Z_{e,t}^{\delta,n} =$ number of graphs with n vertices such that
 $\left| \frac{\# \text{ edges}}{n^2/2} - e \right| < \delta$ and $\left| \frac{\# \text{ triangles}}{n^3/6} - t \right| < \delta$.
 $s(e, t) = \lim_{\delta \to 0} \lim_{n \to \infty} \frac{\ln Z_{e,t}^{\delta,n}}{n^2}$.

• What does a typical graph with a given (*e*, *t*) look like? Like microcanonical ensemble with

 $\begin{array}{rcl} \mathsf{edges} & \leftrightarrow & \mathsf{particles} \\ \mathsf{triangles} & \leftrightarrow & \mathsf{interaction\ energy} \end{array}$



Instead of using the microcanonical ensemble, most work has picked parameters β_1,β_2 and set

$$P(\text{graph } G) = \frac{1}{Z} \exp \left[\beta_1(\# \text{ edges}) + \frac{\beta_2}{n}(\# \text{ triangles}) \right].$$

B b 4

Instead of using the microcanonical ensemble, most work has picked parameters β_1,β_2 and set

$$P(\text{graph } G) = \frac{1}{Z} \exp \left[\beta_1(\# \text{ edges}) + \frac{\beta_2}{n}(\# \text{ triangles}) \right].$$

Like grand canonical ensemble.

Instead of using the microcanonical ensemble, most work has picked parameters β_1,β_2 and set

$$P(\text{graph } G) = \frac{1}{Z} \exp \left[\beta_1(\# \text{ edges}) + \frac{\beta_2}{n}(\# \text{ triangles}) \right].$$

Like grand canonical ensemble. Bad idea! (We'll see why soon) • s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.

同 ト イ ヨ ト イ ヨ ト

э

- s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.
- 2 There is a phase transition along the entire Erdös-Rényi (ER) curve $t = e^3$.

Image: A Image: A

- s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.
- Output: There is a phase transition along the entire Erdös-Rényi (ER) curve $t = e^3$.
- Survey Large regions of (e, t) space are invisible to the canonical or grand canonical ensembles.

• • = • • = •

- s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.
- 2 There is a phase transition along the entire Erdös-Rényi (ER) curve $t = e^3$.
- Survey Large regions of (e, t) space are invisible to the canonical or grand canonical ensembles.
- There are explicit formulas for s(e, t) on the boundary, on the ER curve, and on the line e = 1/2 below the ER curve.

伺 ト イ ヨ ト イ ヨ ト

- s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.
- 2 There is a phase transition along the entire Erdös-Rényi (ER) curve $t = e^3$.
- Survey Large regions of (e, t) space are invisible to the canonical or grand canonical ensembles.
- There are explicit formulas for s(e, t) on the boundary, on the ER curve, and on the line e = 1/2 below the ER curve.
- Analytic evidence that this last formula works for all $e \le 1/2$ below the ER curve.

通 と イ ヨ と イ ヨ と

- s(e, t) is well-defined and is the solution to an explicit minimization problem on functions $g : [0, 1]^2 \rightarrow [0, 1]$.
- 2 There is a phase transition along the entire Erdös-Rényi (ER) curve $t = e^3$.
- Survey Large regions of (e, t) space are invisible to the canonical or grand canonical ensembles.
- There are explicit formulas for s(e, t) on the boundary, on the ER curve, and on the line e = 1/2 below the ER curve.
- Analytic evidence that this last formula works for all $e \leq 1/2$ below the ER curve.
- Numerical evidence of other phase transitions, some 1st order and some 2nd.

直 ト イヨ ト イヨ ト

• Key idea: Embed all graphs in a single space of graphons. (Lovasz and collaborators)

伺 ト く ヨ ト く ヨ ト

3

- Key idea: Embed all graphs in a single space of graphons. (Lovasz and collaborators)
- A graphon is a (measurable) function $g : [0,1]^2 \rightarrow [0,1]$ with g(x,y) = g(y,x).

伺 ト く ヨ ト く ヨ ト

3

- Key idea: Embed all graphs in a single space of graphons. (Lovasz and collaborators)
- A graphon is a (measurable) function $g : [0,1]^2 \rightarrow [0,1]$ with g(x,y) = g(y,x).
- A graph G can be viewed as a checkerboard graphon. Break [0, 1] into n equal intervals, each corresponding to a vertex. This breaks [0, 1]² into n² squares, one for each possible edge. Let g(x, y) = 1 if the edge exists and 0 if it doesn't.

ゆ マ チョン ・ マン

- Key idea: Embed all graphs in a single space of graphons. (Lovasz and collaborators)
- A graphon is a (measurable) function $g : [0,1]^2 \rightarrow [0,1]$ with g(x,y) = g(y,x).
- A graph G can be viewed as a checkerboard graphon. Break [0, 1] into n equal intervals, each corresponding to a vertex. This breaks [0, 1]² into n² squares, one for each possible edge. Let g(x, y) = 1 if the edge exists and 0 if it doesn't.
- Topology of space of graphons is technical. Buzzword is "Cut metric". Also need to mod out by measure-preserving transformations of [0, 1]. (Unnecessary for today's talk.)

伺 ト イ ヨ ト イ ヨ ト

Generating random graphs from a graphon

Graphons are best viewed as recipes for generating random graphs:

Graphons are best viewed as recipes for generating random graphs:

• Pick *n* points (x_1, \ldots, x_n) uniformly on [0, 1].

-

Generating random graphs from a graphon

Graphons are best viewed as recipes for generating random graphs:

- Pick *n* points (x_1, \ldots, x_n) uniformly on [0, 1].
- Vertex between *i*-th and *j*-th vertices exists with probability $g(x_i, x_j)$.

• • = • • = •

Generating random graphs from a graphon

Graphons are best viewed as recipes for generating random graphs:

- Pick *n* points (x_1, \ldots, x_n) uniformly on [0, 1].
- Vertex between *i*-th and *j*-th vertices exists with probability $g(x_i, x_j)$.
- Once (x_1, \ldots, x_n) are chosen, edges exist independently.

Graphons are best viewed as recipes for generating random graphs:

- Pick *n* points (x_1, \ldots, x_n) uniformly on [0, 1].
- Vertex between *i*-th and *j*-th vertices exists with probability $g(x_i, x_j)$.
- Once (x_1, \ldots, x_n) are chosen, edges exist independently.

The expected edge and triangle densities are:

$$e(g) = \iint g(x,y)dx dy, \quad t(g) = \iiint g(x,y)g(y,z)g(z,x)dxdydz.$$

• Entropy of one coin flip is $-[u \ln(u) + (1-u) \ln(1-u)]$.

直 マ イヨ マ イヨマ

- B

• Entropy of one coin flip is $-[u \ln(u) + (1-u) \ln(1-u)]$.

Let

$$I_o(u) = \frac{1}{2} (u \ln(u) + (1-u) \ln(1-u)) =$$
Rate function.

• Entropy of one coin flip is $-[u \ln(u) + (1-u) \ln(1-u)]$.

Let

$$I_o(u) = \frac{1}{2} (u \ln(u) + (1 - u) \ln(1 - u)) =$$
Rate function.

• Rate function of graphon g is

$$I(g) = \iint I_0(g(x,y))dx\,dy.$$

伺 ト く ヨ ト く ヨ ト

3

• Entropy of one coin flip is $-[u \ln(u) + (1-u) \ln(1-u)]$.

Let

$$I_o(u) = \frac{1}{2} (u \ln(u) + (1-u) \ln(1-u)) =$$
Rate function.

• Rate function of graphon g is

$$I(g) = \iint I_0(g(x,y)) dx \, dy.$$

- Entropy of random process described by graphon is $-n^2 I(g)$
- Entropy density is -I(g).

伺 ト イ ヨ ト イ ヨ ト

-

Theorem (Chatterjee-Varadhan)

Large deviation principle for ER graphs. The rate function controls everything.

3 N

Theorem (Chatterjee-Varadhan)

Large deviation principle for ER graphs. The rate function controls everything.

Theorem (Radin-S)

The entropy density s(e, t) is well-defined, and equals $-\inf I(g)$, where \inf is over graphons g with $\iint g(x, y)dx dy = e$ and $\iiint g(x, y)g(y, z)g(x, z)dx dy dz = t.$
Theorem (Chatterjee-Varadhan)

Large deviation principle for ER graphs. The rate function controls everything.

Theorem (Radin-S)

The entropy density s(e, t) is well-defined, and equals $-\inf I(g)$, where \inf is over graphons g with $\iint g(x, y)dx dy = e$ and $\iiint g(x, y)g(y, z)g(x, z)dx dy dz = t.$

This reduces all questions to minimizing I(g) for fixed e and t.

Theorem (Chatterjee-Varadhan)

Large deviation principle for ER graphs. The rate function controls everything.

Theorem (Radin-S)

The entropy density s(e, t) is well-defined, and equals $-\inf I(g)$, where \inf is over graphons g with $\iint g(x, y)dx dy = e$ and $\iiint g(x, y)g(y, z)g(x, z)dx dy dz = t.$

This reduces all questions to minimizing I(g) for fixed e and t. Goodbye statistical mechanics. Hello functional analysis.

/□ ▶ < 글 ▶ < 글

高 とう きょう うちょう しょう

This is the maximum entropy for fixed e and variable t.

A B + A B +

-

This is the maximum entropy for fixed *e* and variable *t*. How smooth is the maximum? (Answer: not very)

4 B K 4 B K

This is the maximum entropy for fixed *e* and variable *t*. How smooth is the maximum? (Answer: not very)

2) On bottom boundary, minimizing graphon is

$$g(x, y) = \begin{cases} 2e & x < \frac{1}{2} < y \text{ or } y < \frac{1}{2} < x \\ 0 & \text{otherwise.} \end{cases}$$

Entropy is $s(e, 0) = -\frac{1}{2}I_0(2e)$.

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

 $t = e^3 - b^3$. $I(g) = [I_0(e+b) + I_0(e-b)]/2$.

3 N 4 3 N

3

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

$$t = e^3 - b^3$$
. $I(g) = [I_0(e+b) + I_0(e-b)]/2$.

When $b \approx e$, describes almost-bipartite graphs, with clusters of equal size.

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

$$t = e^3 - b^3$$
. $I(g) = [I_0(e+b) + I_0(e-b)]/2$.

When $b \approx e$, describes almost-bipartite graphs, with clusters of equal size.

Appears to minimize I(g). When $b \approx e$, 2nd variation w.r.t. L^{∞} -small changes is positive operator. But L^{1} -small changes needn't be L^{∞} -small.

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

$$t = e^3 - b^3$$
. $I(g) = [I_0(e+b) + I_0(e-b)]/2$.

When $b \approx e$, describes almost-bipartite graphs, with clusters of equal size.

Appears to minimize I(g). When $b \approx e$, 2nd variation w.r.t. L^{∞} -small changes is positive operator. But L^{1} -small changes needn't be L^{∞} -small.

Numerical evidence that this minimizes I(g) for all $e \leq 1/2$, $t \leq e^3$.

Consider the graphon

$$g(x,y) = egin{cases} e+b & x < rac{1}{2} < y ext{ or } y < rac{1}{2} < x \ e-b & ext{otherwise.} \end{cases}$$

$$t = e^3 - b^3$$
. $I(g) = [I_0(e+b) + I_0(e-b)]/2$.

When $b \approx e$, describes almost-bipartite graphs, with clusters of equal size.

Appears to minimize I(g). When $b \approx e$, 2nd variation w.r.t. L^{∞} -small changes is positive operator. But L^{1} -small changes needn't be L^{∞} -small.

Numerical evidence that this minimizes I(g) for all $e \leq 1/2$, $t \leq e^3$.

Rigorous proof that this minimizes I(g) for e = 1/2, $t \le e^3$. (Later slide)

Theorem (Radin-S)

There exist positive constants c_1 , c_2 such that: If $t < e^3$, $s(e, t) \le s(e, e^3) - c_1(e^3 - t)^{2/3}$, If $t > e^3$, $s(e, t) \le s(e, e^3) - c_2(t - e^3)$.

Theorem (Radin-S)

There exist positive constants c_1 , c_2 such that: If $t < e^3$, $s(e, t) \le s(e, e^3) - c_1(e^3 - t)^{2/3}$, If $t > e^3$, $s(e, t) \le s(e, e^3) - c_2(t - e^3)$.

s(e, t) has at most a 1-sided derivative w.r.t. t at $t = e^3$.

Theorem (Radin-S)

There exist positive constants c_1 , c_2 such that: If $t < e^3$, $s(e, t) \le s(e, e^3) - c_1(e^3 - t)^{2/3}$, If $t > e^3$, $s(e, t) \le s(e, e^3) - c_2(t - e^3)$.

s(e, t) has at most a 1-sided derivative w.r.t. t at $t = e^3$. s(e, t) is concave up for t slightly less than e^3 . Invisible to Legendre transform.

Write
$$g(x, y) = e + \delta g(x, y)$$
, with $\int \int \delta g(x, y) dx dy = 0$.

æ

□ ▶ ▲ 臣 ▶ ▲ 臣

Write
$$g(x, y) = e + \delta g(x, y)$$
, with $\int \int \delta g(x, y) dx dy = 0$.
Expand $t = \int \int \int g(x, y)g(y, z)g(z, x)$:
 $t = e^3 + 3e^2 \int \int \delta g(x, y) + 3e \int \int \int \delta g(x, z)\delta g(y, z)$
 $+ \int \int \int \delta g(x, y)\delta g(y, z)\delta g(z, x)$.

æ

□ ▶ ▲ 臣 ▶ ▲ 臣

Write
$$g(x, y) = e + \delta g(x, y)$$
, with $\int \int \delta g(x, y) dx dy = 0$.
Expand $t = \int \int \int g(x, y)g(y, z)g(z, x)$:
 $t = e^3 + 3e^2 \int \int \delta g(x, y) + 3e \int \int \int \delta g(x, z)\delta g(y, z)$
 $+ \int \int \int \delta g(x, y)\delta g(y, z)\delta g(z, x)$.

First integral is zero.

∃ → < ∃</p>

э

Write
$$g(x, y) = e + \delta g(x, y)$$
, with $\int \int \delta g(x, y) dx dy = 0$.
Expand $t = \int \int \int g(x, y)g(y, z)g(z, x)$:
 $t = e^3 + 3e^2 \int \int \delta g(x, y) + 3e \int \int \int \delta g(x, z)\delta g(y, z) + \int \int \int \delta g(x, y)\delta g(y, z)\delta g(z, x)$.

First integral is zero.

Second integral is positive definite. If $\alpha(z) = \int \delta g(x, z) dx$, 2nd integral is $3e \int \alpha(z)^2 dz$.

э

Write
$$g(x, y) = e + \delta g(x, y)$$
, with $\int \int \delta g(x, y) dx dy = 0$.
Expand $t = \int \int \int g(x, y)g(y, z)g(z, x)$:
 $t = e^3 + 3e^2 \int \int \delta g(x, y) + 3e \int \int \int \delta g(x, z)\delta g(y, z) + \int \int \int \delta g(x, y)\delta g(y, z)\delta g(z, x)$.

First integral is zero.

Second integral is positive definite. If $\alpha(z) = \int \delta g(x, z) dx$, 2nd integral is $3e \int \alpha(z)^2 dz$. Last integral is $O(\delta g^3)$, so if $t < e^3$, $e^3 - t < c \|\delta g\|_{L^2}^{3/2}$.

$$s(e, e^3) - s(e, t) = \iint (I_0(e) - I_0(g(x, y)) dx dy = O(\delta g^2).$$

< ∃ >

æ

$$s(e, e^{3}) - s(e, t) = \iint (I_{0}(e) - I_{0}(g(x, y))dx dy = O(\delta g^{2}).$$

Since
$$\delta t = \begin{cases} O(\delta g^{2}) & \delta t > 0;\\ O(\delta g^{3}) & \delta t < 0, \end{cases}$$

.⊒ . ►

æ

$$s(e, e^{3}) - s(e, t) = \iint (I_{0}(e) - I_{0}(g(x, y))dx dy = O(\delta g^{2}).$$

Since
$$\delta t = \begin{cases} O(\delta g^{2}) & \delta t > 0; \\ O(\delta g^{3}) & \delta t < 0, \\ \delta s = \begin{cases} O(\delta t) & \delta t > 0; \\ O(|\delta t|^{2/3}) & \delta t < 0. \end{cases}$$

< ∃ >

э

$$s(e, e^{3}) - s(e, t) = \iint (I_{0}(e) - I_{0}(g(x, y))dx dy = O(\delta g^{2}).$$

Since
$$\delta t = \begin{cases} O(\delta g^{2}) & \delta t > 0; \\ O(\delta g^{3}) & \delta t < 0, \\ \\ \delta s = \begin{cases} O(\delta t) & \delta t > 0; \\ O(|\delta t|^{2/3}) & \delta t < 0. \end{cases}$$

When e = 1/2 and $\delta t < 0$, all estimates are saturated by symmetric bipartite graphon, so we know it's a minimizer.

• • = • • = •

Phase I: Symmetric bipartite.

-

э

Phase I: Symmetric bipartite.

Phase II: Asymmetric bipartite. Like symmetric bipartite, but two subintervals have unequal size and unequal probabilities of internal edges. There is a constant $c \neq 1/2$ such that

$$g(x, y) = \begin{cases} a & x, y < c \\ b & x, y > c \\ d & x < c < y \text{ or } y < c < x \end{cases}$$

Phase I: Symmetric bipartite.

Phase II: Asymmetric bipartite. Like symmetric bipartite, but two subintervals have unequal size and unequal probabilities of internal edges. There is a constant $c \neq 1/2$ such that

$$g(x, y) = \begin{cases} a & x, y < c \\ b & x, y > c \\ d & x < c < y \text{ or } y < c < x \end{cases}$$

Transition from Phase I to Phase II is 2nd-order.

Phase III: Tripartite. Like bipartite, but interval [c, 1] divided into equal sub-intervals. When x, y > c, g(x, y) depends on whether x and y are in same or different sub-intervals.

Phase III: Tripartite. Like bipartite, but interval [c, 1] divided into equal sub-intervals. When x, y > c, g(x, y) depends on whether x and y are in same or different sub-intervals. Applies near first "scallop".

Phase III: Tripartite. Like bipartite, but interval [c, 1] divided into equal sub-intervals. When x, y > c, g(x, y) depends on whether xand y are in same or different sub-intervals. Applies near first "scallop". Transition from Phase II to Phase III is first-order. (Partial derivatives of s diverge).

Bottom line: Lots of phase transitions, and lots of interesting behavior, but only in the microcanonical ensemble.

Bottom line: Lots of phase transitions, and lots of interesting behavior, but only in the microcanonical ensemble. Canonical ensemble (*e* fixed, *t* variable) misses most of the fun.

Bottom line: Lots of phase transitions, and lots of interesting behavior, but only in the microcanonical ensemble. Canonical ensemble (*e* fixed, *t* variable) misses most of the fun. Grand canonical ensemble misses almost all of the fun.



Happy Birthday, Yosi!

3



Ein Hagigah Bli Uga

Lorenzo Sadun

Statistical Mechanics of Large Dense Random Graphs

э