The Planted Matching Problem

Cristopher Moore, Santa Fe Institute

Joint work with Mehrdad Moharrami (Michigan) and Jiaming Xu (Duke)

December 8, 2019
Planted problems: good solution + noise

Constraint satisfaction, optimization
Planted problems: good solution + noise

Constraint satisfaction, optimization
Generative models, Bayesian inference
Planted problems: good solution + noise

Constraint satisfaction, optimization
Generative models, Bayesian inference
Satisfying assignments, cliques, communities…
Planted problems: good solution + noise

Constraint satisfaction, optimization
Generative models, Bayesian inference
Satisfying assignments, cliques, communities...

Information-theoretic (a.k.a. statistical) and computational barriers
Planted problems: good solution + noise

Constraint satisfaction, optimization
Generative models, Bayesian inference
Satisfying assignments, cliques, communities...

Information-theoretic (a.k.a. statistical) and computational barriers
Statistical physics $\Rightarrow$ conjectures, proofs, and algorithms
Planted matchings: particle tracking

Tracking particles advected by turbulent fluid flow

Goal: recover the underlying true one-to-one mapping of the particles
Flocks of birds, swimming microbes, . . .

[Chertkov-Kroc-Krzakala-Vergassola-Zdeborová PNAS'10]
The planted assignment model

- A complete bipartite graph
- A hidden perfect matching $M$
- Edge weight $W_{ij}$ ind.
  \[ W_{ij} \sim \begin{cases} P & e \in M \\ Q & e \notin M \end{cases} \]
- Goal: recover $M$ from $W$

\[ \text{Our work: } P = \exp(\lambda), \quad Q = \exp(1/n) \text{ (mean } 1/\lambda \text{ vs. } n) \]

Minimum-weight matching $\hat{M}$ is the Maximum Likelihood Estimator

- A phase transition in $\lambda$, and exact results
The planted assignment model

• A complete bipartite graph
• A hidden perfect matching $M$
• Edge weight

\[ W_{ij} \overset{\text{ind.}}{\sim} \begin{cases} P & e \in M \\ Q & e \notin M \end{cases} \]

• Goal: recover $M$ from $W$
• Our work: $P = \text{Exp}(\lambda)$, $Q = \text{Exp}(1/n)$ (mean $1/\lambda$ vs. $n$)
The planted assignment model

- A complete bipartite graph
- A hidden perfect matching $M$
- Edge weight $W_{ij}$ ind.

\[ W_{ij} \sim \begin{cases} P & e \in M \\ Q & e \notin M \end{cases} \]

- Goal: recover $M$ from $W$

- Our work: $P = \text{Exp}(\lambda)$, $Q = \text{Exp}(1/n)$ (mean $1/\lambda$ vs. $n$)
- Minimum-weight matching $\widehat{M}$ is the Maximum Likelihood Estimator
The planted assignment model

- A complete bipartite graph
- A hidden perfect matching $M$
- Edge weight $W_{ij} \overset{\text{ind.}}{\sim} \begin{cases} P & e \in M \\ Q & e \notin M \end{cases}$
- Goal: recover $M$ from $W$

- Our work: $P = \text{Exp}(\lambda), Q = \text{Exp}(1/n)$ (mean $1/\lambda$ vs. $n$)
- Minimum-weight matching $\hat{M}$ is the Maximum Likelihood Estimator
- How much does $\hat{M}$ have in common with $M$?
The planted assignment model

- A complete bipartite graph
- A hidden perfect matching $M$
- Edge weight $W_{ij} \overset{\text{ind.}}{\sim} \begin{cases} P & e \in M \\ Q & e \notin M \end{cases}$
- Goal: recover $M$ from $W$

- Our work: $P = \text{Exp}(\lambda)$, $Q = \text{Exp}(1/n)$ (mean $1/\lambda$ vs. $n$)
- Minimum-weight matching $\hat{M}$ is the Maximum Likelihood Estimator
- How much does $\hat{M}$ have in common with $M$?
- A phase transition in $\lambda$, and exact results
Main result: phase transition at $\lambda = 4$

**Theorem (Moharrami-M.-Xu ’19)**

\[
\text{overlap: } \lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \left| \hat{M} \cap M \right| \right] = \begin{cases} 
1 & \text{if } \lambda \geq 4 \\
\alpha(\lambda) & \text{if } 0 < \lambda < 4
\end{cases}
\]

where $\alpha(\lambda) = 1 - 2 \int_0^\infty (1 - F(x)) (1 - G(x)) V(x) W(x) \, dx < 1$, 
Main result: phase transition at $\lambda = 4$

**Theorem (Moharrami-M.-Xu ’19)**

**overlap:**
\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \| \widehat{M} \cap M \| \right] = \begin{cases} 
1 & \text{if } \lambda \geq 4 \\
\alpha(\lambda) & \text{if } 0 < \lambda < 4
\end{cases}
\]

where $\alpha(\lambda) = 1 - 2 \int_{0}^{\infty} (1 - F(x))(1 - G(x)) V(x) W(x) \, dx < 1$,

and $F, G, V, W$ is the unique solution to a system of ODEs:

\[
\begin{align*}
\dot{F} &= (1 - F)(1 - G)V \\
\dot{G} &= -(1 - F)(1 - G)W \\
\dot{V} &= \lambda(V - F) \\
\dot{W} &= -\lambda(W - G)
\end{align*}
\]

Boundary conditions: $F(x), V(x), G(-x), W(-x) \to \begin{cases} 1 & x \to +\infty \\
0 & x \to -\infty \end{cases}$
Main result: phase transition at $\lambda = 4$

Theorem (Moharrami-M.-Xu ’19)

**overlap:** $\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ |\hat{M} \cap M| \right] = \begin{cases} 
1 & \text{if } \lambda \geq 4 \\
\alpha(\lambda) & \text{if } 0 < \lambda < 4
\end{cases}$

where $\alpha(\lambda) = 1 - 2 \int_{0}^{\infty} (1 - F(x)) (1 - G(x)) V(x)W(x) \, dx < 1$, 

![Graph showing the overlap $\alpha(\lambda)$ vs $\lambda$]
Main result: phase transition at $\lambda = 4$

**Theorem (Moharrami-M.-Xu ’19)**

The overlap is given by:

$$\lim_{n \to \infty} \frac{1}{n} \mathbb{E}[|\hat{M} \cap M|] = \begin{cases} 1 & \text{if } \lambda \geq 4 \\ \alpha(\lambda) & \text{if } 0 < \lambda < 4 \end{cases}$$

where $\alpha(\lambda) = 1 - 2 \int_0^\infty (1 - F(x))(1 - G(x)) V(x) W(x) \, dx < 1$, with $F(x)$ and $G(x)$ representing cumulative distribution functions, and $V(x)$ and $W(x)$ representing probability density functions, respectively.
When $\lambda \geq 4$: count augmenting cycles

- Probability that $M'$ has lower total weight than $M$ is $P[Erlang(\ell, \lambda) \geq Erlang(\ell, 1/n)] \leq \left(\frac{\lambda n}{4}\right)^{\ell - \ell/2n}$
- There are $(n\ell)!\leq ne^{-\ell/2n}$ matchings $M'$ with $|M \triangle M'| = 2\ell$
- Expected number of such $M'$ is at most $\left(\frac{\lambda}{4}\right)^{\ell - \ell/2n}$
- Sum over $\ell$: total probability a planted edge is in augmenting cycle is $o(1)$ if $\lambda \geq 4$
When $\lambda \geq 4$: count augmenting cycles

- Probability that $M'$ has lower total weight than $M$ is

$\mathbb{P}[\text{Erlang}(\ell, \lambda) \geq \text{Erlang}(\ell, 1/n)] \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$
When $\lambda \geq 4$: count augmenting cycles

- Probability that $M'$ has lower total weight than $M$ is
  \[ \mathbb{P}[\text{Erlang}(\ell, \lambda) \geq \text{Erlang}(\ell, 1/n)] \leq \left( \frac{\lambda n}{4} \right)^{-\ell} \]

- There are \( \binom{n}{\ell} \ell! \leq n^\ell e^{-\ell^2/2n} \) matchings $M'$ with $|M \triangle M'| = 2\ell$
When $\lambda \geq 4$: count augmenting cycles

- Probability that $M'$ has lower total weight than $M$ is

$$\Pr[\text{Erlang}(\ell, \lambda) \geq \text{Erlang}(\ell, 1/n)] \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$$

- There are $\binom{n}{\ell} \ell! \leq n^\ell e^{-\ell^2/2n}$ matchings $M'$ with $|M \triangle M'| = 2\ell$

$\Rightarrow$ Expected number of such $M'$ is at most $(\lambda/4)^{-\ell} e^{-\ell^2/2n}$
When $\lambda \geq 4$: count augmenting cycles

- Probability that $M'$ has lower total weight than $M$ is
  \[ P \left[ \text{Erlang}(\ell, \lambda) \geq \text{Erlang}(\ell, 1/n) \right] \leq \left( \frac{\lambda n}{4} \right)^{-\ell} \]

- There are $\binom{n}{\ell} \ell! \leq n^\ell e^{-\ell^2/2n}$ matchings $M'$ with $|M \triangle M'| = 2\ell$

\[ \Rightarrow \text{Expected number of such } M' \text{ is at most } (\lambda/4)^{-\ell} e^{-\ell^2/2n} \]

\[ \Rightarrow \text{Sum over } \ell: \text{total probability a planted edge is in augmenting cycle is } o(1) \text{ if } \lambda \geq 4 \]
Warmup: the (un-planted) random assignment problem

- A complete bipartite graph
- Weights uniform in $[0, n]$ or $\text{Exp}(1/n)$
- Cost of minimum matching?

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\min_\pi \sum_{i=1}^n W_i \pi(i) \right] = \pi = \frac{1}{2} + \frac{1}{4} + \frac{1}{9} + \cdots
\]
Warmup: the (un-planted) random assignment problem

- A complete bipartite graph
- Weights uniform in $[0, n]$ or $\text{Exp}(1/n)$
- Cost of minimum matching?

$\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)} \right] = \frac{\pi^2}{6} = 1 + \frac{1}{4} + \frac{1}{9} + \cdots$

[Walkup '79, Mézard-Parisi '87, Aldous '92, Steele '97, Aldous '01, ...]
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi '87, Aldous'00]
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi '87, Aldous'00]
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi '87, Aldous'00]
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi ’87, Aldous’00]

\[ X_v \triangleq \text{cost of min matching on } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi ’87, Aldous’00]

\[ X_v \triangleq \text{cost of min matching on } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]

sort edge weights \( W_{\emptyset,1}, W_{\emptyset,2}, \ldots \) from smallest to largest:
arrivals \( \zeta_1, \zeta_2, \ldots \) of a Poisson process with rate 1
Warmup: the (un-planted) random assignment problem

Cavity method: model as a tree [Mézard-Parisi ’87, Aldous’00]

\[ X_v \triangleq \text{cost of min matching on } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]

sort edge weights \( W_{\emptyset,1}, W_{\emptyset,2}, \ldots \) from smallest to largest: arrivals \( \zeta_1, \zeta_2, \ldots \) of a Poisson process with rate 1

\[ X \overset{d}{=} \min_{i \geq 1} \{\zeta_i - X_i\} \]
From distributional to differential equations

\[ X \overset{d}{=} \min \{ \zeta_i - X_i \} \text{ where } \zeta_i \text{ are Poisson arrivals} \]
\( X \overset{d}{=} \min \{ \zeta_i - X_i \} \) where \( \zeta_i \) are Poisson arrivals

Generate pairs \((\zeta, x)\): two-dimensional Poisson process with density \( f(x) \)
From distributional to differential equations

\[ X \overset{d}{=} \min \{ \zeta_i - X_i \} \text{ where } \zeta_i \text{ are Poisson arrivals} \]

Generate pairs \((\zeta, x)\): two-dimensional Poisson process with density \(f(x)\)

Define the cdf \(\bar{F}(x) = 1 - F(x) = \mathbb{P}[X > x] = \mathbb{P}[\forall i : \zeta_i - x > X_i]\)

\[
\bar{F}(x) = \exp \left( - \int_{-x}^{\infty} \bar{F}(t) \, dt \right) \quad \Rightarrow \quad \frac{dF(x)}{dx} = F(x)F(-x)
\]
From distributional to differential equations, cont’d

$$\frac{dF(x)}{dx} = F(x)F(-x)$$
\[
\frac{dF(x)}{dx} = F(x)F(-x)
\]

\[
F(x) = \frac{e^x}{1 + e^x} \quad \text{or} \quad f(x) = \frac{1}{(e^{x/2} + e^{-x/2})^2}
\]
From distributional to differential equations, cont’d

\[ \frac{dF(x)}{dx} = F(x)F(-x) \]

\[ F(x) = \frac{e^x}{1 + e^x} \quad \text{or} \quad f(x) = \frac{1}{(e^{x/2} + e^{-x/2})^2} \]
\[ \frac{dF(x)}{dx} = F(x)F(-x) \]

\[ F(x) = \frac{e^x}{1 + e^x} \quad \text{or} \quad f(x) = \frac{1}{(e^{x/2} + e^{-x/2})^2} \]

**Contribution of a single edge:**

\[ E[W 1[W < X + X']] = \frac{1}{4} \text{Var}[X + X'] = \frac{1}{2} \text{Var}[X] = \frac{\pi^2}{6} \]
Now with planted edges

Partner in planted matching is either parent or child 0, other children sorted 1, 2, 3, …

\[ X_v \triangleq \text{cost of min matching in } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]
Now with planted edges

Partner in planted matching is either parent or child 0, other children sorted 1, 2, 3, …

\[ X_v \triangleq \text{cost of min matching in } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]

Recursion:

\[
X_\emptyset = \min \left\{ W_{\emptyset,0} - X_0, \min_{i \geq 1} \{ W_{\emptyset,i} - X_i \} \right\}
\]

\[
X_0 = \min_{i \geq 1} \{ W_{0,0i} - X_{0i} \}
\]
Now with planted edges

Partner in planted matching is either parent or child 0, other children sorted 1, 2, 3, …

\[ X_v \triangleq \text{cost of min matching in } T_v - \text{cost of min matching on } T_v \setminus \{v\} \]

Recursion:

\[ X_\emptyset = \min \left\{ W_{\emptyset,0} - X_0, \min_{i \geq 1} \{ W_{\emptyset,i} - X_i \} \right\} \]

\[ X_0 = \min_{i \geq 1} \{ W_{0,0i} - X_{0i} \} \]

\[ Y \overset{d}{=} \min \{ \eta - Z, Z' \} \]

\[ Z \overset{d}{=} \min_i \{ \zeta_i - Y_i \} \]
From distributional to differential equations, redux

\[
Y \overset{d}{=} \min \{ \eta - Z, Z' \}
\]
\[
Z \overset{d}{=} \min \{ \zeta_i - Y_i \}_{i=1}^\infty
\]

where \( \eta \sim \text{Exp}(\lambda) \) and \( \zeta_i \) are Poisson arrivals.
From distributional to differential equations, redux

\[
Y \overset{d}{=} \min \{\eta - Z, Z'\}
\]

\[
Z \overset{d}{=} \min \{\zeta_i - Y_i\}_{i=1}^{\infty}
\]

where \(\eta \sim \text{Exp}(\lambda)\) and \(\zeta_i\) are Poisson arrivals

\[
F(x) = \mathbb{P}[Z < x], \quad G(x) = F(-x), \quad V(x) = \mathbb{E}[F(x + \eta)], \quad W(x) = V(-x)
\]
From distributional to differential equations, redux

\[
Y = \min \{ \eta - Z, Z' \} \\
Z = \min \{ \zeta_i - Y_i \}_{i=1}^\infty
\]

where \( \eta \sim \text{Exp}(\lambda) \) and \( \zeta_i \) are Poisson arrivals

\[
F(x) = \mathbb{P}[Z < x], \ G(x) = F(-x), \ V(x) = \mathbb{E}[F(x + \eta)], \ W(x) = V(-x)
\]

\[
\dot{F} = (1 - F)(1 - G)V \\
\dot{G} = -(1 - F)(1 - G)W \\
\dot{V} = \lambda(V - F) \\
\dot{W} = -\lambda(W - G)
\]
From distributional to differential equations, redux

\[ \begin{align*}
Y & \overset{d}{=} \min \{ \eta - Z, Z' \} \\
Z & \overset{d}{=} \min \{ \zeta_i - Y_i \}_{i=1}^\infty
\end{align*} \]

where \( \eta \sim \text{Exp}(\lambda) \) and \( \zeta_i \) are Poisson arrivals

\[ \begin{align*}
F(x) & = \mathbb{P}[Z < x],
G(x) & = F(-x),
V(x) & = \mathbb{E}[F(x + \eta)],
W(x) & = V(-x)
\end{align*} \]

\[ \begin{align*}
\dot{F} & = (1 - F)(1 - G)V \\
\dot{G} & = -(1 - F)(1 - G)W \\
\dot{V} & = \lambda(V - F) \\
\dot{W} & = -\lambda(W - G)
\end{align*} \]

\( \dot{V} \) and \( \dot{W} \) from \( \eta \sim \text{Exp}(\lambda) \), integration by parts
From distributional to differential equations, redux

\[ Y \stackrel{d}{=} \min \{ \eta - Z, Z' \} \]
\[ Z \stackrel{d}{=} \min \{ \zeta_i - Y_i \}_{i=1}^\infty \]

where \( \eta \sim \text{Exp}(\lambda) \) and \( \zeta_i \) are Poisson arrivals

\[ F(x) = \mathbb{P}[Z < x], \quad G(x) = F(-x), \quad V(x) = \mathbb{E}[F(x + \eta)], \quad W(x) = V(-x) \]

\[
\begin{align*}
\dot{F} &= (1 - F)(1 - G)V \\
\dot{G} &= -(1 - F)(1 - G)W \\
\dot{V} &= \lambda(V - F) \\
\dot{W} &= -\lambda(W - G)
\end{align*}
\]

\( \dot{V} \) and \( \dot{W} \) from \( \eta \sim \text{Exp}(\lambda) \), integration by parts

Boundary conditions: \( F(\infty) = V(\infty) = 1, \quad F(-\infty) = V(-\infty) = 0 \)
No solution for $\lambda \geq 4$

At least no sensible one...
No solution for $\lambda \geq 4$

At least no sensible one...

Want $F(+\infty) = V(+\infty) = 1$. But...

$F(x), V(x)$

![Graph showing $F(x)$ and $V(x)$ as a function of $x$.]
No solution for $\lambda \geq 4$

Conservation law: $FW + GV - VW = 0 \Rightarrow V(0) = 2F(0)$
No solution for $\lambda \geq 4$

Conservation law: $FW + GV - VW = 0 \Rightarrow V(0) = 2F(0)$

Let $U(x) = F(x)/V(x)$. Then $U(0) = 1/2$, want $U(+\infty) = 1$...

\[ \dot{U} = -\lambda U(1-U) + (1-F)(1-G) \leq -\lambda U(1-U) + 1 \]

If $\lambda \geq 4$, $\dot{U}(1/2) \leq 0$
No solution for $\lambda \geq 4$

Conservation law: $FW + GV - VW = 0 \Rightarrow V(0) = 2F(0)$

Let $U(x) = F(x)/V(x)$. Then $U(0) = 1/2$, want $U(+\infty) = 1$.

\[ \dot{U} = -\lambda U(1 - U) + (1 - F)(1 - G) \leq -\lambda U(1 - U) + 1 \]
No solution for $\lambda \geq 4$

Conservation law: $FW + GV - VW = 0 \Rightarrow V(0) = 2F(0)$

Let $U(x) = F(x)/V(x)$. Then $U(0) = 1/2$, want $U(+\infty) = 1$.

\[
\dot{U} = -\lambda U(1 - U) + (1 - F)(1 - G) \leq -\lambda U(1 - U) + 1
\]

If $\lambda \geq 4$, $\dot{U}(1/2) \leq 0$
No solution for $\lambda \geq 4$

Conservation law: $FW + GV - VW = 0 \Rightarrow V(0) = 2F(0)$

Let $U(x) = F(x)/V(x)$. Then $U(0) = 1/2$, want $U(\infty) = 1$…

$$\dot{U} = -\lambda U(1 - U) + (1 - F)(1 - G) \leq -\lambda U(1 - U) + 1$$

If $\lambda \geq 4$, $\dot{U}(1/2) \leq 0$

No fixed distribution on finite values: cost of un-planted edge is $+\infty$ \Rightarrow almost perfect recovery
A unique solution when $\lambda < 4$

$(F, G, V, W) \iff (U, V, W)$: three-dimensional dynamical system

\[
\begin{align*}
\dot{U} &= -\lambda U(1 - U) + (1 - UV)(1 - (1 - U)W) \\
\dot{V} &= \lambda V(1 - U) \\
\dot{W} &= \lambda WU
\end{align*}
\]

Initial conditions: $U(0) = \frac{1}{2}$, $V(0) = W(0) = \epsilon$
A unique solution when $\lambda < 4$

$(F, G, V, W) \iff (U, V, W)$: three-dimensional dynamical system

\[
\begin{align*}
\dot{U} &= -\lambda U(1 - U) + (1 - UV)(1 - (1 - U)W) \\
\dot{V} &= \lambda V(1 - U) \\
\dot{W} &= \lambda WU
\end{align*}
\]

Initial conditions: $U(0) = \frac{1}{2}$, $V(0) = W(0) = \epsilon$

Lemma

If $\lambda < 4$ there is a unique $\epsilon_0 \in (0, 1)$ such that

- If $\epsilon \in [0, \epsilon_0)$, $U(x) \to +\infty$
- If $\epsilon = \epsilon_0$, $U(x) \to 1$ and $V(x) \to 1$
- If $\epsilon \in (\epsilon_0, 1]$, $V(x) \to +\infty$
A unique solution when $\lambda < 4$

Geometric interpretation: $(U = 1, V = 1, W = 0)$ is a saddle point. If $V(0) = W(0) = \epsilon_0$ we approach the saddle along its unstable manifold.

This gives cdfs $F, V \to 1$ of the unique fixed point distribution.
A numerical experiment

$\lambda = 2.5$, population dynamics with $N = 10^6$

CDFs for $Y, Z$
Finally, computing the overlap for $\lambda < 4$
Finally, computing the overlap for $\lambda < 4$

$$\alpha(\lambda) = \mathbb{P}[\eta < Z + Z'] = 1 - \mathbb{E}_\eta \int_{-\infty}^{+\infty} f(x)F(\eta - x) \, dx$$

$$= 1 - \int_{-\infty}^{+\infty} f(x) \mathbb{E}_\eta F(\eta - x) \, dx$$

$$= 1 - \int_{-\infty}^{+\infty} (1 - F(x))(1 - G(x))V(x)W(x) \, dx$$

$$= 1 - 2 \int_{0}^{+\infty} (1 - F(x))(1 - G(x))V(x)W(x) \, dx$$
• \( \hat{M} \) only depends on weights \( \Rightarrow \) symmetry in the joint distribution of weights and matching

• Vertex-transitive involutions on \( K_{n,n} \) or infinite tree \( T_\infty \)

• A random matching is *involution invariant* if it has these symmetries

• We have constructed an involution invariant \( M_{opt} \) on \( T_\infty \) and computed its cost and overlap
• Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)

\( \text{Cristopher Moore} \)
Proving it: Local weak convergence (Aldous 1992, 2001)

- Easy: any invariant sequence $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) invariant matching on $T_\infty$
  - Local treelikeness of light edges, compactness

- Hard: for any invariant matching $M_\infty$ there is a sequence $\{M_n: n \to \infty\}$ that converges to $M_\infty$
- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$, can fix to make a perfect matching

- Uniqueness: any invariant matching $M'_\infty$ that differs from $M_{\text{opt}}$ with positive probability has strictly greater cost
  - By invariance, $M'_\infty$ and $M_{\text{opt}}$ differ at the root
  - $M'_\infty$ often chooses the wrong partner for $\emptyset$
  - Right partner given by recursion ⇒ differential equations

- Together these imply $\lim_{n \to \infty} \text{overlap}(\hat{M}_n) = \text{overlap}(\hat{M}_\infty)$
Proving it: Local weak convergence (Aldous 1992, 2001)

- Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  - Local treelikeness of light edges, compactness

- Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
Proving it: Local weak convergence (Aldous 1992, 2001)

- Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  - Local treelikeness of light edges, compactness

- Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  - Martingale convergence
Proving it: Local weak convergence (Aldous 1992, 2001)

- **Easy**: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  - Local treelikeness of light edges, compactness

- **Hard**: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  - Martingale convergence
  - Almost-doubly-stochastic matrix

\[ \lim_{n \to \infty} \text{overlap}(\hat{M}_n) = \text{overlap}(\hat{M}_\infty) \]
Proving it: Local weak convergence (Aldous 1992, 2001)

• Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  ▶ Local treelikeness of light edges, compactness

• Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  ▶ Martingale convergence
  ▶ Almost-doubly-stochastic matrix
  ▶ Almost-perfect matching on \( K_{n,n} \), can fix to make a perfect matching
Proving it: Local weak convergence (Aldous 1992, 2001)

- Easy: any invariant sequence $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) invariant matching on $T_\infty$
  - Local treelikeness of light edges, compactness

- Hard: for any invariant matching $M_\infty$ there is a sequence $\{M_n : n \to \infty\}$ that converges to $M_\infty$
  - Martingale convergence
  - Almost-doubly-stochastic matrix
  - Almost-perfect matching on $K_{n,n}$, can fix to make a perfect matching

- Uniqueness: any invariant matching $M'_\infty$ that differs from $M_{opt}$ with positive probability has strictly greater cost
Proving it: Local weak convergence (Aldous 1992, 2001)

- Easy: any invariant sequence \( \{ M_n \} \) of matchings on \( K_{n,n} \) has a subsequence \( \{ M_{n_j} \} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  - Local treelikeness of light edges, compactness

- Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{ M_n : n \to \infty \} \) that converges to \( M_\infty \)
  - Martingale convergence
  - Almost-doubly-stochastic matrix
  - Almost-perfect matching on \( K_{n,n} \), can fix to make a perfect matching

- Uniqueness: any invariant matching \( M'_\infty \) that differs from \( M_{\text{opt}} \) with positive probability has strictly greater cost
  - By invariance, \( M'_\infty \) and \( M_{\text{opt}} \) differ at the root
• Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  ▶ Local treelikeness of light edges, compactness

• Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  ▶ Martingale convergence
  ▶ Almost-doubly-stochastic matrix
  ▶ Almost-perfect matching on \( K_{n,n} \), can fix to make a perfect matching

• Uniqueness: any invariant matching \( M'_\infty \) that differs from \( M_{\text{opt}} \) with positive probability has strictly greater cost
  ▶ By invariance, \( M'_\infty \) and \( M_{\text{opt}} \) differ at the root
  ▶ \( M'_\infty \) often chooses the wrong partner for \( \emptyset \)
Proving it: Local weak convergence (Aldous 1992, 2001)

• Easy: any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  ▶ Local treelikeness of light edges, compactness

• Hard: for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  ▶ Martingale convergence
  ▶ Almost-doubly-stochastic matrix
  ▶ Almost-perfect matching on \( K_{n,n} \), can fix to make a perfect matching

• Uniqueness: any invariant matching \( M'_\infty \) that differs from \( M_{\text{opt}} \) with positive probability has strictly greater cost
  ▶ By invariance, \( M'_\infty \) and \( M_{\text{opt}} \) differ at the root
  ▶ \( M'_\infty \) often chooses the wrong partner for \( \emptyset \)
  ▶ Right partner given by recursion \( \Rightarrow \) differential equations
Proving it: Local weak convergence (Aldous 1992, 2001)

- **Easy:** any invariant sequence \( \{M_n\} \) of matchings on \( K_{n,n} \) has a subsequence \( \{M_{n_j}\} \) that converges to a (possibly random) invariant matching on \( T_\infty \)
  - Local treelikeness of light edges, compactness

- **Hard:** for any invariant matching \( M_\infty \) there is a sequence \( \{M_n : n \to \infty\} \) that converges to \( M_\infty \)
  - Martingale convergence
  - Almost-doubly-stochastic matrix
  - Almost-perfect matching on \( K_{n,n} \), can fix to make a perfect matching

- **Uniqueness:** any invariant matching \( M'_\infty \) that differs from \( M_{\text{opt}} \) with positive probability has strictly greater cost
  - By invariance, \( M'_\infty \) and \( M_{\text{opt}} \) differ at the root
  - \( M'_\infty \) often chooses the wrong partner for \( \emptyset \)
  - Right partner given by recursion \( \Rightarrow \) differential equations

- Together these imply \( \lim_{n \to \infty} \text{overlap}(\hat{M}_n) = \text{overlap}(\hat{M}_\infty) \)
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher

2. Concentration of the overlap?
   - We computed its expectation

3. Information-theoretically optimal recovery?
   - Gibbs sampling, posterior marginals

4. Distributions other than $\eta \sim \text{Exp}(\lambda)$?
   - Distributional equations rarely collapse to ODEs

5. Spatial structure (particle tracking)?

6. Other planted structures: spanning trees, traveling salespeople?
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher

2. Concentration of the overlap?
   - We computed its expectation
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher

2. Concentration of the overlap?
   - We computed its expectation

3. Information-theoretically optimal recovery?
   - Gibbs sampling, posterior marginals
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher

2. Concentration of the overlap?
   - We computed its expectation

3. Information-theoretically optimal recovery?
   - Gibbs sampling, posterior marginals

4. Distributions other than \( \eta \sim \text{Exp}(\lambda) \)?
   - Distributional equations rarely collapse to ODEs
Open questions

1. Order of the phase transition?
   - Overlap is continuous, and so is its derivative
   - Appears to be third or higher

2. Concentration of the overlap?
   - We computed its expectation

3. Information-theoretically optimal recovery?
   - Gibbs sampling, posterior marginals

4. Distributions other than $\eta \sim \text{Exp}(\lambda)$?
   - Distributional equations rarely collapse to ODEs

5. Spatial structure (particle tracking)?
Open questions

1. Order of the phase transition?
   ▶ Overlap is continuous, and so is its derivative
   ▶ Appears to be third or higher

2. Concentration of the overlap?
   ▶ We computed its expectation

3. Information-theoretically optimal recovery?
   ▶ Gibbs sampling, posterior marginals

4. Distributions other than $\eta \sim \text{Exp}(\lambda)$?
   ▶ Distributional equations rarely collapse to ODEs

5. Spatial structure (particle tracking)?

6. Other planted structures: spanning trees, traveling salespeople?
Shameless Plug

To put it bluntly: this book rocks! It somehow manages to combine the fun of a popular book with the intellectual heft of a textbook.

Scott Aaronson, UT Austin

This is, simply put, the best-written book on the theory of computation I have ever read; one of the best-written mathematical books I have ever read, period.

Cosma Shalizi, Carnegie Mellon

A creative, insightful, and accessible introduction to the theory of computing, written with a keen eye toward the frontiers of the field and a vivid enthusiasm for the subject matter.

Jon Kleinberg, Cornell

www.nature-of-computation.org