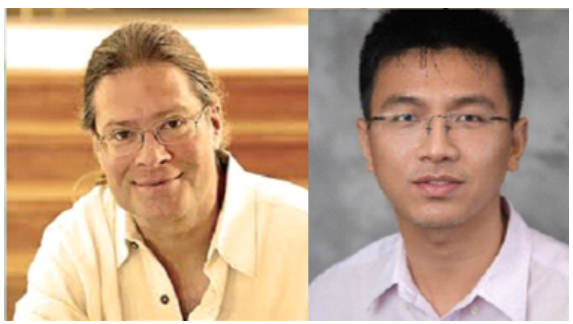
The Planted Matching Problem: Phase Transition and Exact Results

Mehrdad Moharrami

Collaborators: Cristopher Moore and Jiaming Xu



Planted Problems: good solution + noise

Constraint satisfaction, optimization

Planted Problems: good solution + noise

Constraint satisfaction, optimization

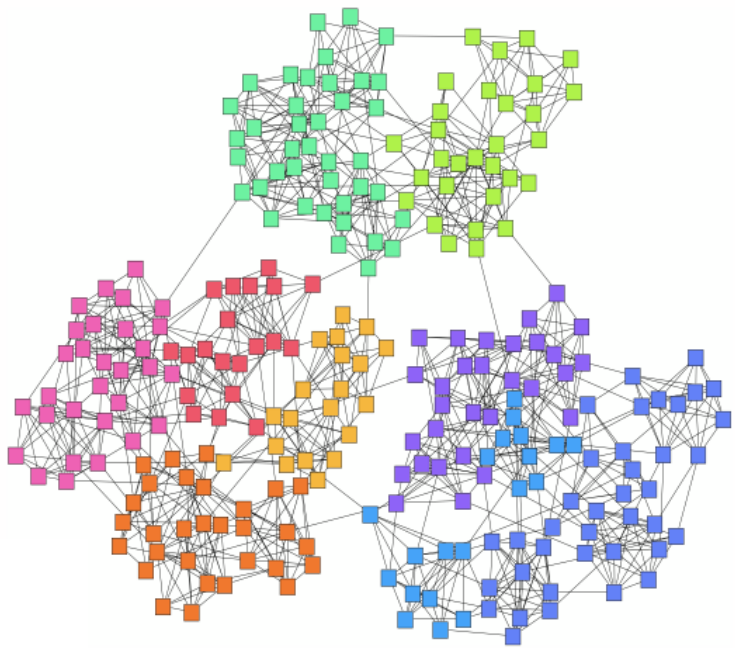
Generative model, Bayesian inference

Planted Problems: good solution + noise

Constraint satisfaction, optimization

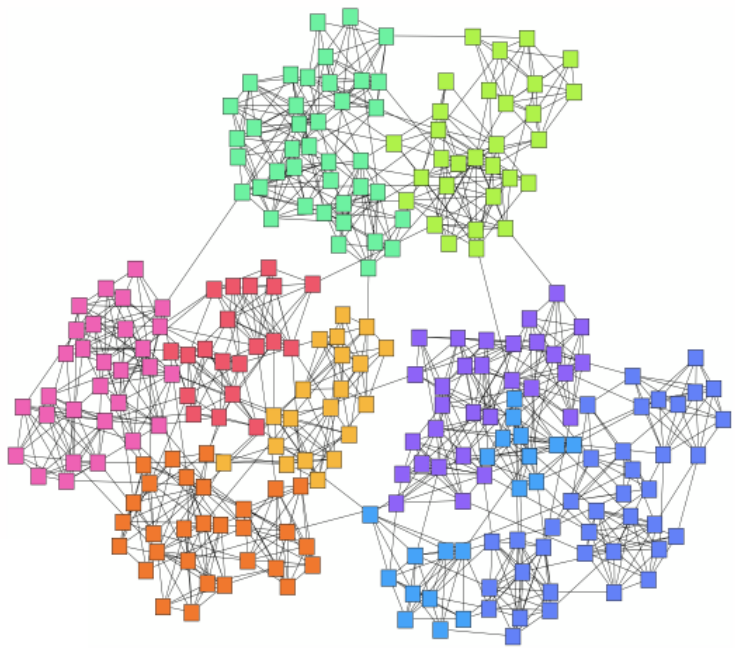
Generative model, Bayesian inference

Satisfying assignments, cliques, communities. . .



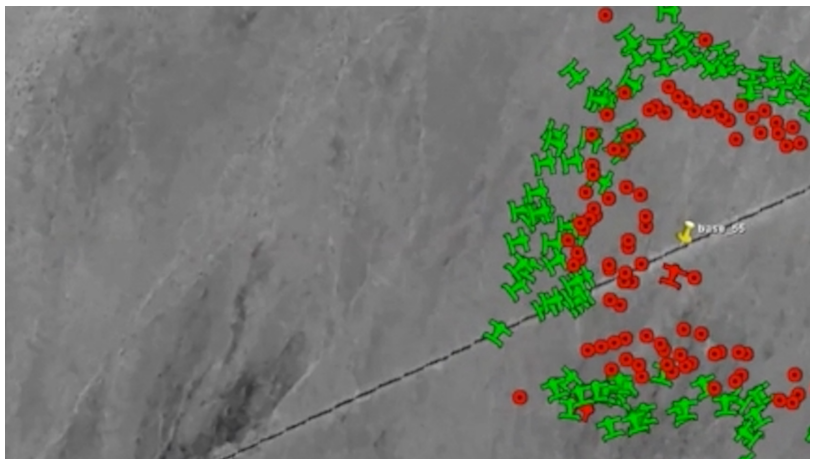
Planted Problems: good solution + noise

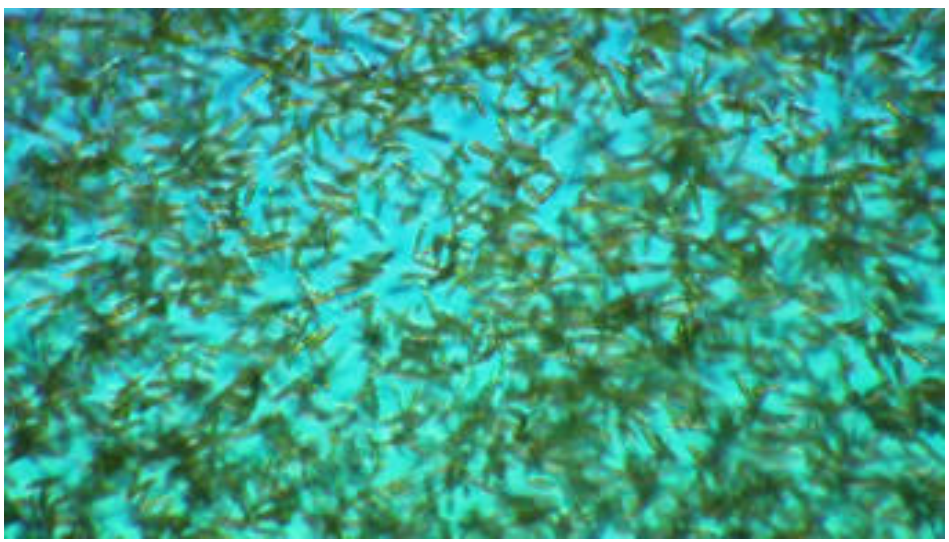
Constraint satisfaction, optimization
Generative model, Bayesian inference
Satisfying assignments, cliques, communities. . .



Information-theoretic (a.k.a. statistical) and computational barriers

Particle Tracking/Mobile Object





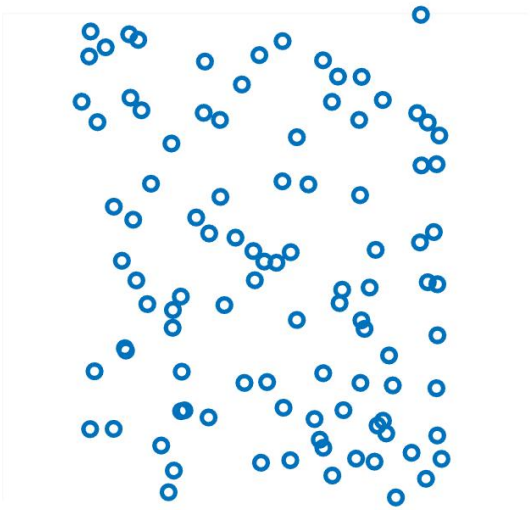
Particle Tracking/Mobile Object

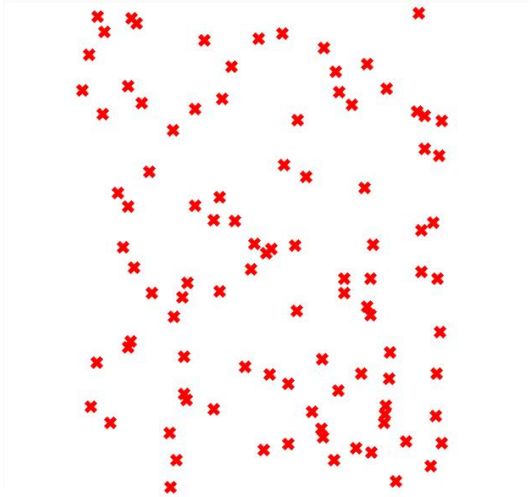
Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.

Particle Tracking/Mobile Object

Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.

Almost exact recovery is possible

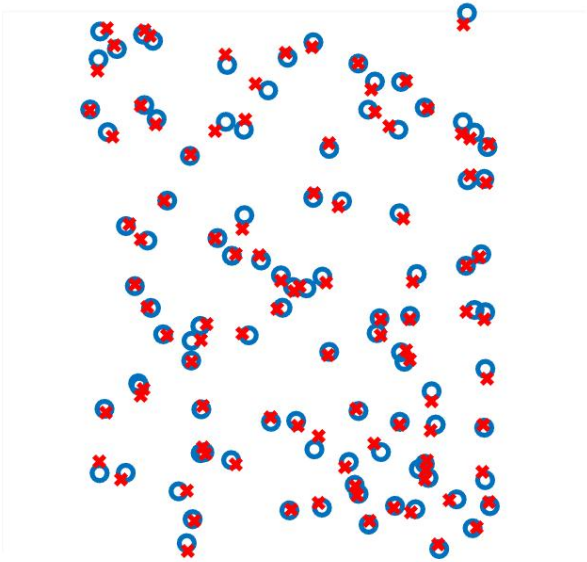




Particle Tracking/Mobile Object

Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.

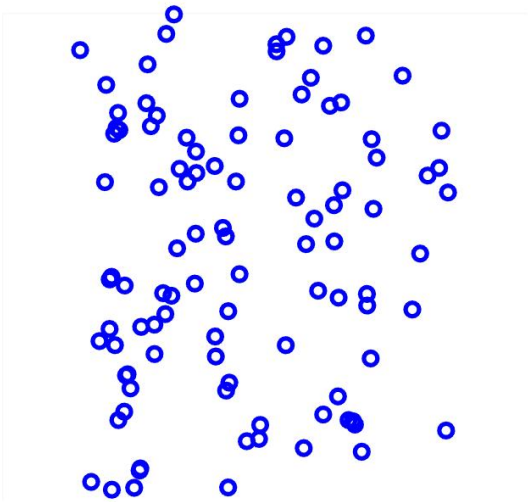
Almost exact recovery is possible

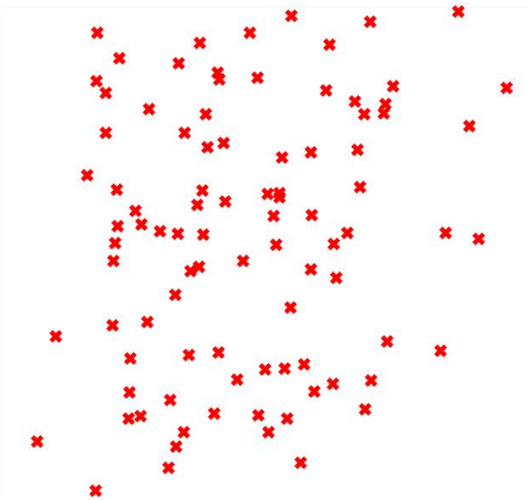


Particle Tracking/Mobile Object

Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.

Almost exact recovery seems impossible

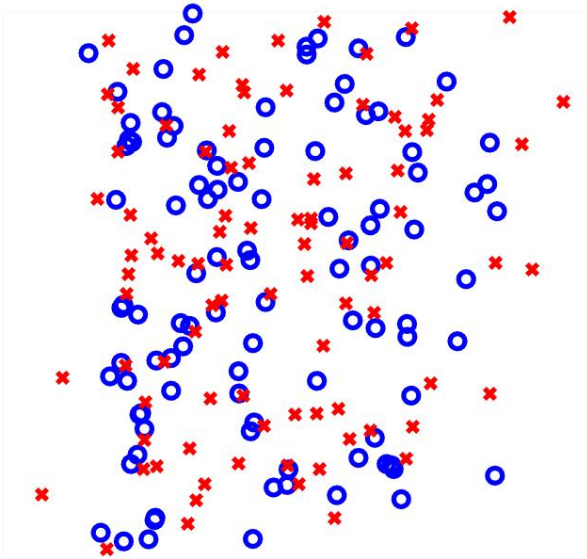




Particle Tracking/Mobile Object

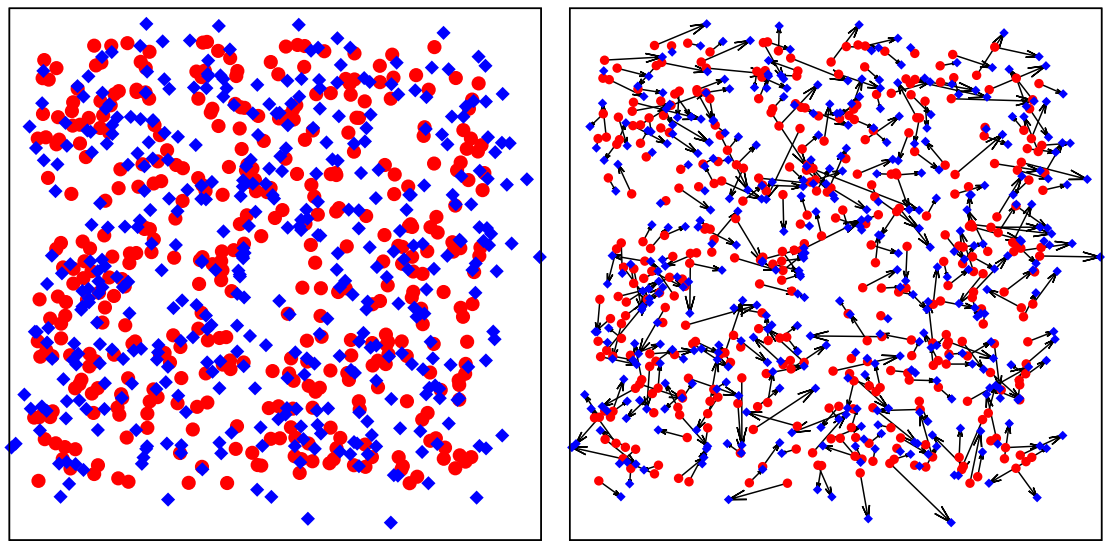
Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.

Almost exact recovery seems impossible



Particle Tracking/Mobile Object

Consider two successive images of a system. Depending on the sampling rate, it might be possible to track the mobile objects.



Goal: Recover the underlying true one-to-one mapping of the particles.

Particle Tracking/Mobile Object: Challenges

First attempt: movement of n particles in $\mathbb{R}^d$ according to some distribution

- Highly complex and intractable: correlations among the interparticle distances

Particle Tracking/Mobile Object: Challenges

First attempt: movement of n particles in $\mathbb{R}^d$ according to some distribution

- Highly complex and intractable: correlations among the interparticle distances

Simplification: decouple the n^2 distances, random distance model

Particle Tracking/Mobile Object: Challenges

First attempt: movement of n particles in $\mathbb{R}^d$ according to some distribution

- Highly complex and intractable: correlations among the interparticle distances

Simplification: decouple the n^2 distances, random distance model

Second attempt: n^2 independent distances d_i^j between particle i in first image and particle j in second image:

- $i = j$: distribution P that is fixed
- $i \neq j$: distribution Q that scales with n

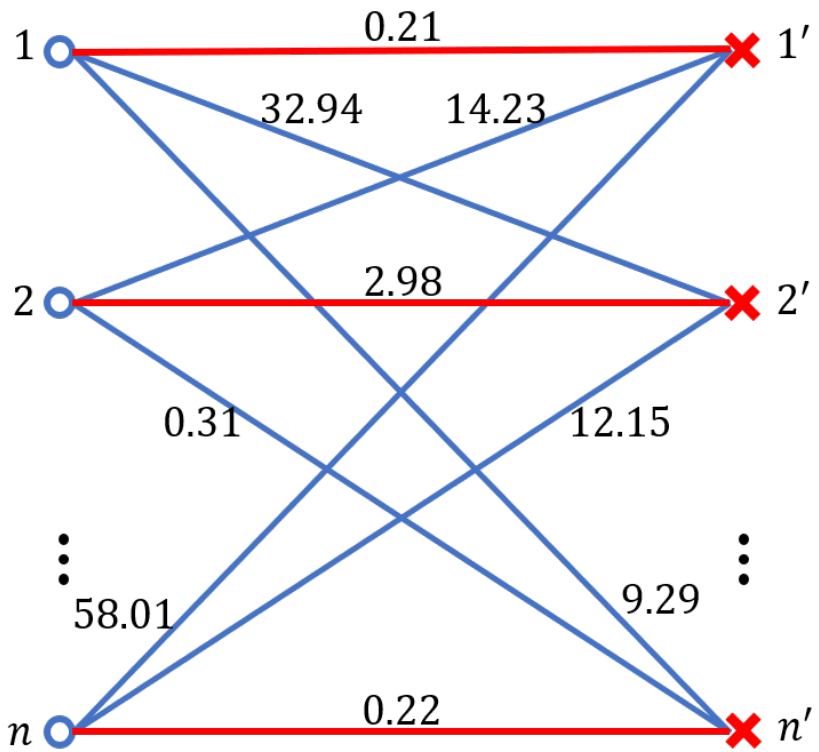
Drawback: triangular inequality fails

Model

- A complete bipartite graph
- A planted 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

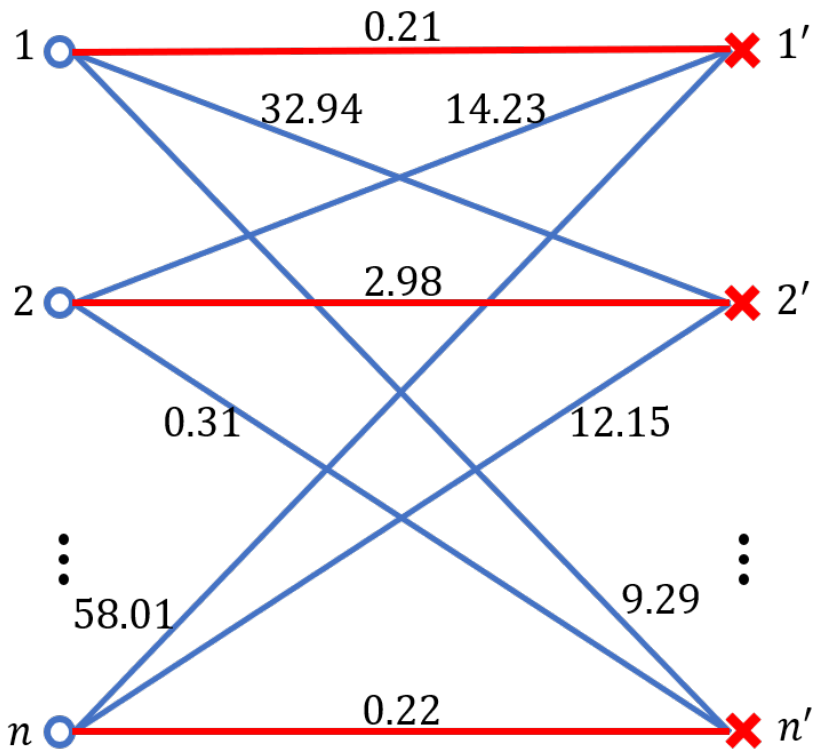


Model

- A complete bipartite graph
- A planted 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: $P = \exp(\lambda)$, $Q = \exp(1/n)$ (mean $1/\lambda$ vs. n)

Maximum Likelihood Estimator

The posterior probability for a given matching M', i.e., $\mathbb{P}(M^* = M' \mid G)$, is proportional to the density:

Maximum Likelihood Estimator

The posterior probability for a given matching M', i.e., $\mathbb{P}(M^* = M' \mid G)$, is proportional to the density:

$$\begin{aligned}\mathbb{P}(G \mid M')] &= \prod_{e \in M'} P(w_e) \prod_{e \notin M'} Q(w_e) \\ &\propto \prod_{e \in M'} \exp(-(\lambda - 1/n) w_e) \\ &= \exp\left(-(\lambda - 1/n) \sum_{e \in M'} w_e \right) .\end{aligned}$$

Maximum Likelihood Estimator

The posterior probability for a given matching M', i.e., $\mathbb{P}(M^* = M' \mid G)$, is proportional to the density:

$$\begin{aligned}\mathbb{P}(G \mid M')] &= \prod_{e \in M'} P(w_e) \prod_{e \notin M'} Q(w_e) \\ &\propto \prod_{e \in M'} \exp(-(\lambda - 1/n) w_e) \\ &= \exp\left(-(\lambda - 1/n) \sum_{e \in M'} w_e \right) .\end{aligned}$$

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

Maximum Likelihood Estimator

The posterior probability for a given matching M', i.e., $\mathbb{P}(M^* = M' \mid G)$, is proportional to the density:

$$\begin{aligned}\mathbb{P}(G \mid M')] &= \prod_{e \in M'} P(w_e) \prod_{e \notin M'} Q(w_e) \\ &\propto \prod_{e \in M'} \exp(-(\lambda - 1/n) w_e) \\ &= \exp\left(-(\lambda - 1/n) \sum_{e \in M'} w_e \right) .\end{aligned}$$

Minimum-weight matching $\widehat{M}$ is the Maximum Likelihood Estimator. Hungarian algorithm finds $\widehat{M}$ in time $O(n^3)$

Maximum Likelihood Estimator

The posterior probability for a given matching M', i.e., $\mathbb{P}(M^* = M' \mid G)$, is proportional to the density:

$$\begin{aligned}\mathbb{P}(G \mid M')] &= \prod_{e \in M'} P(w_e) \prod_{e \notin M'} Q(w_e) \\ &\propto \prod_{e \in M'} \exp(-(\lambda - 1/n) w_e) \\ &= \exp\left(-(\lambda - 1/n) \sum_{e \in M'} w_e \right) .\end{aligned}$$

Minimum-weight matching $\widehat{M}$ is the Maximum Likelihood Estimator.
Hungarian algorithm finds $\widehat{M}$ in time $O(n^3)$
Question: What is the overlap between $\widehat{M}$ and M^*?

Main Results: Phase Transition at $\lambda = 4$

Theorem

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

where $\alpha(\lambda) = 1 - 2\int_0^\infty (1 - F(x))\,(1 - G(x))\,V(x)W(x)\,dx < 1\,,$

Main Results: Phase Transition at $\lambda = 4$

Theorem

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

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:

$$\frac{\mathrm{d}F}{\mathrm{d}x} = (1-F(x))(1-G(x))V(x), \qquad \frac{\mathrm{d}V}{\mathrm{d}x} = \lambda(V(x)-F(x)),$$
$$\frac{\mathrm{d}G}{\mathrm{d}x} = -(1-F(x))(1-G(x))W(x), \qquad \frac{\mathrm{d}W}{\mathrm{d}x} = \lambda(G(x)-W(x)),$$

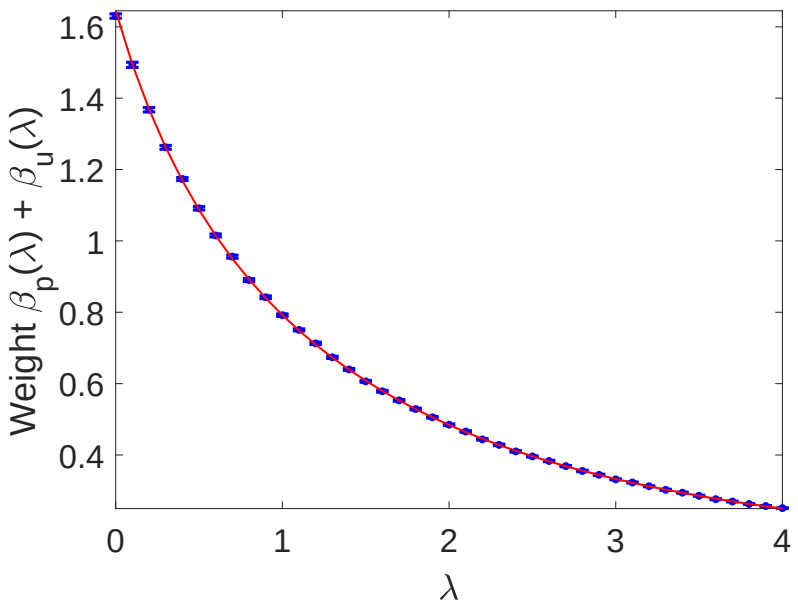
Boundary conditions: $F(x), V(x), G(-x), W(-x) \to \begin{cases} 1 & x \to +\infty \\ 0 & x \to -\infty \end{cases}$

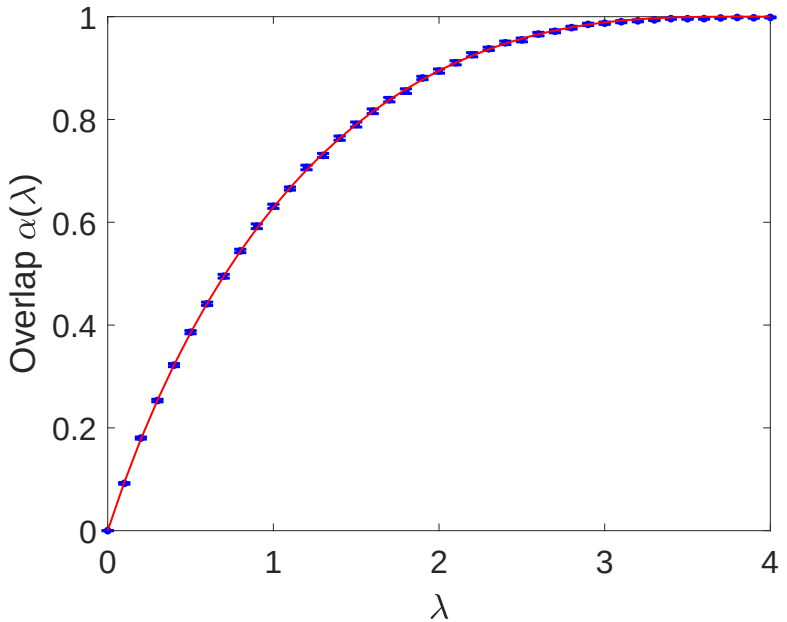
Main Results: Phase Transition at $\lambda = 4$

Theorem

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

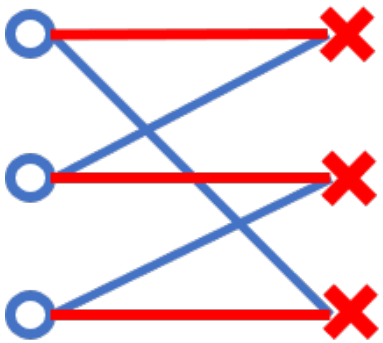
where $\alpha(\lambda) = 1 - 2\int_0^\infty (1-F(x))\,(1-G(x))\,V(x)W(x)\,dx < 1\,,$





When $\lambda \geq 4$: Count Augmenting Cycles

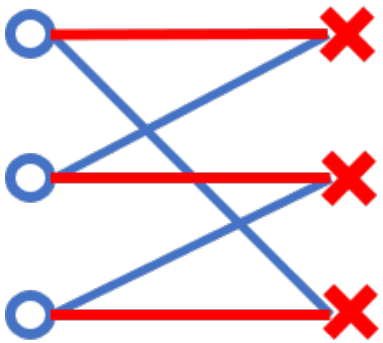
Probability that M' has lower total weight than M?



When $\lambda \geq 4$: Count Augmenting Cycles

Probability that M' has lower total weight than M?

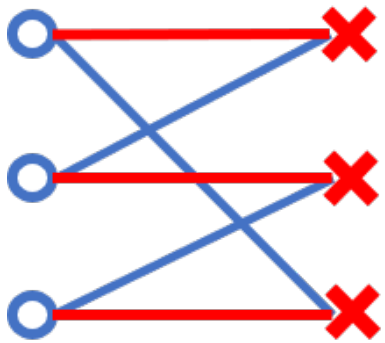
$$\mathbb{P}\left(\mathrm{Erlang}(\ell, \lambda) \geq \mathrm{Erlang}(\ell, 1/n)\right) \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$$



When $\lambda \geq 4$: Count Augmenting Cycles

Probability that M' has lower total weight than M?

$$\mathbb{P}\left(\mathrm{Erlang}(\ell, \lambda) \geq \mathrm{Erlang}(\ell, 1/n)\right) \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$$

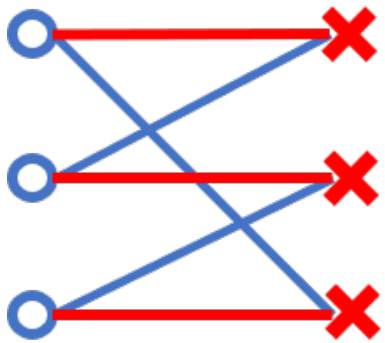


There are $\binom{n}{\ell}\ell! \leq n^{\ell} e^{-\ell^2/2n}$ matchings M' with $|M \triangle M'| = 2\ell$

When $\lambda \geq 4$: Count Augmenting Cycles

Probability that M' has lower total weight than M?

$$\mathbb{P}\left(\mathrm{Erlang}(\ell,\lambda) \geq \mathrm{Erlang}(\ell, 1/n)\right) \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$$



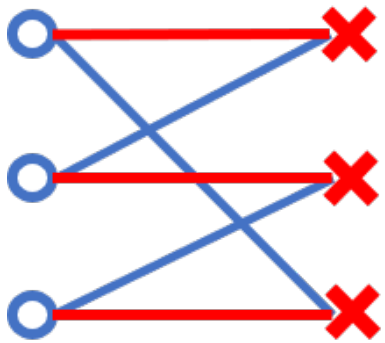
There are $\binom{n}{\ell}\ell! \leq n^{\ell} e^{-\ell^2/2n}$ matchings M' with $|M \triangle M'| = 2\ell$

$\Rightarrow$ 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?

$$\mathbb{P}\left(\mathrm{Erlang}(\ell, \lambda) \geq \mathrm{Erlang}(\ell, 1/n)\right) \leq \left(\frac{\lambda n}{4}\right)^{-\ell}$$



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

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

$\Rightarrow$ Sum over ℓ: total probability a planted edge is in an augmenting cycle is $o(1)$ if $\lambda \geq 4$

When $\lambda < 4$: Second Moment Method?

First attempt: Second moment method a.k.a. Paley-Zygmund inequality

– Dependency makes it hard.

When $\lambda < 4$: Second Moment Method?

First attempt: Second moment method a.k.a. Paley-Zygmund inequality

- Dependency makes it hard.

Second attempt: Do what a brave statistician does:

- Change the limit: take $n \to \infty$ and study the minimum matching problem on the asymptotic object

When $\lambda < 4$: Second Moment Method?

First attempt: Second moment method a.k.a. Paley-Zygmund inequality

- Dependency makes it hard.

Second attempt: Do what a brave statistician does:

- Change the limit: take $n \to \infty$ and study the minimum matching problem on the asymptotic object

Usually referred to as cavity method/mean-field approach.

When $\lambda < 4$: Second Moment Method?

First attempt: Second moment method a.k.a. Paley-Zygmund inequality

- Dependency makes it hard.

Second attempt: Do what a brave statistician does:

- Change the limit: take $n \to \infty$ and study the minimum matching problem on the asymptotic object

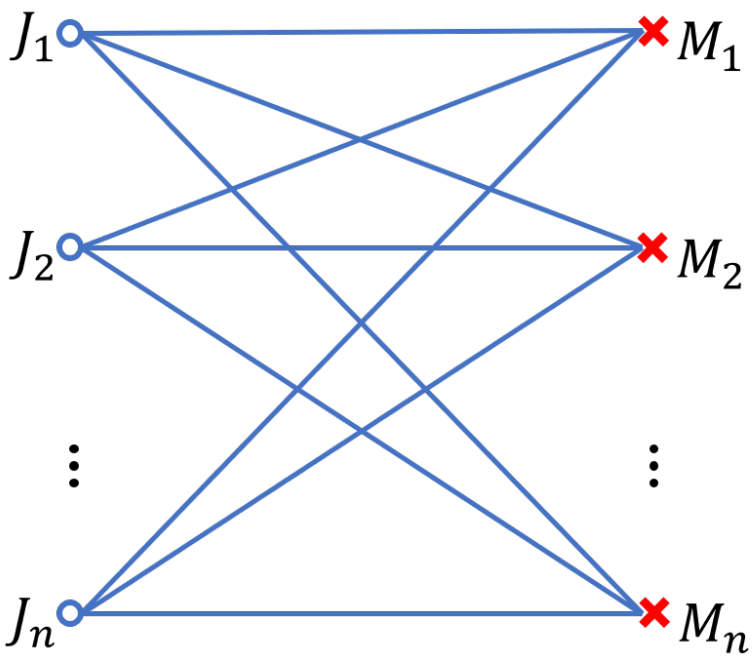
Usually referred to as cavity method/mean-field approach.

Challenges: What sort of infinite object we are looking at? How to justify the interchange of limit?

Warmup: The Job Assignment Problem

– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$

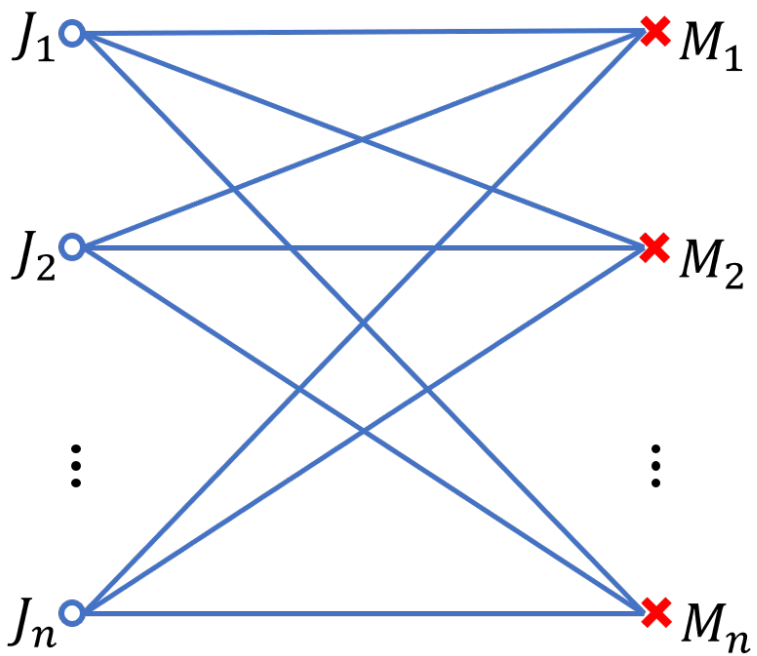


Warmup: The Job Assignment Problem

– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$

– Walkup [6]: It is bounded.



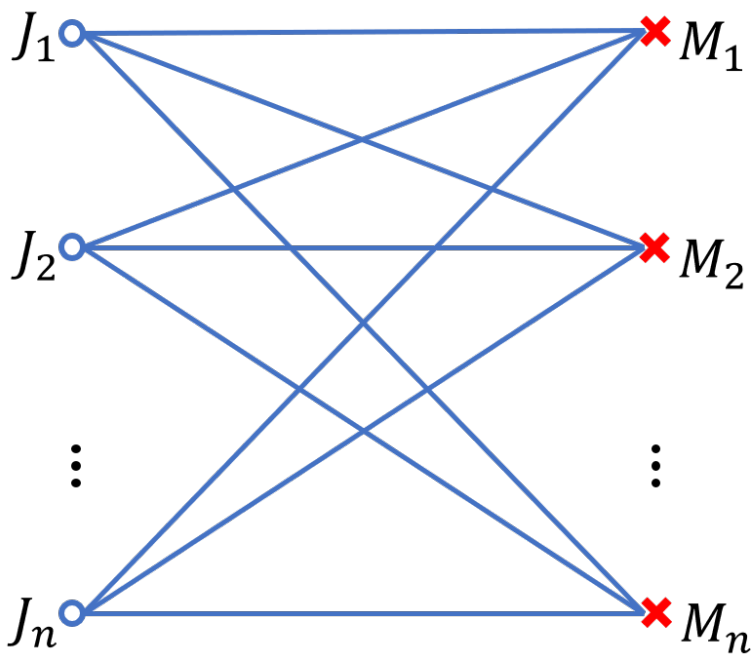
Warmup: The Job Assignment Problem

– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$

– Walkup [6]: It is bounded.

– Meźard & Parisi [5]: Conjectured $\pi^2/6$



Warmup: The Job Assignment Problem

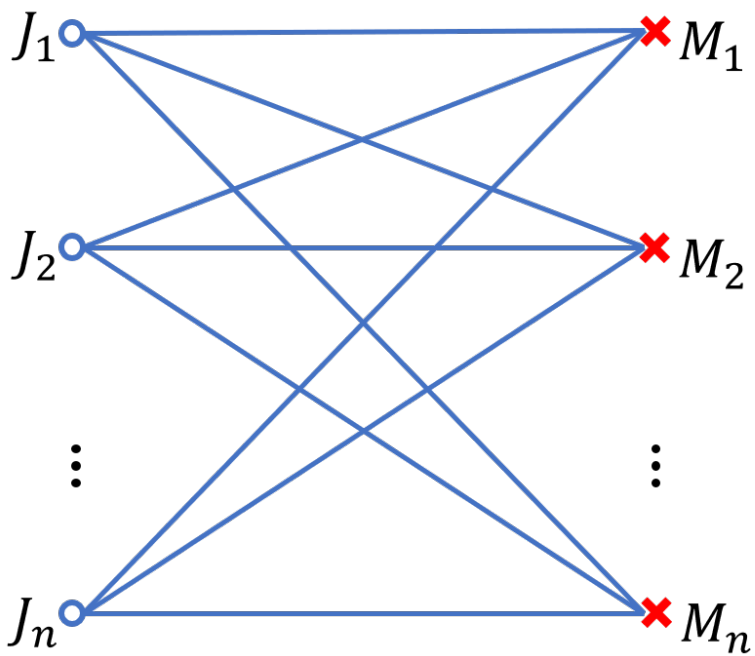
– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$

– Walkup [6]: It is bounded.

– Meźard & Parisi [5]: Conjectured $\pi^2/6$

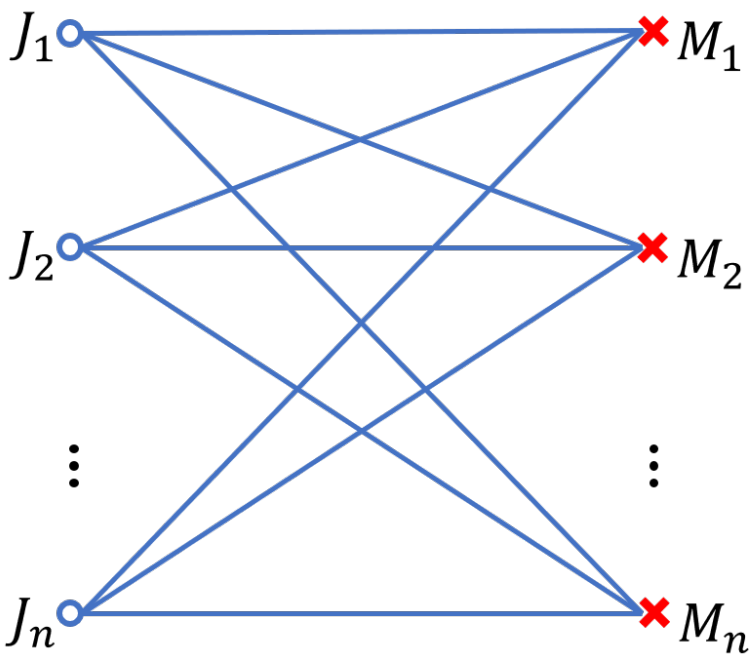
– Aldous [1]: Converges to something!



Warmup: The Job Assignment Problem

– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$



– Walkup [6]: It is bounded.

– Meźard & Parisi [5]: Conjectured $\pi^2/6$

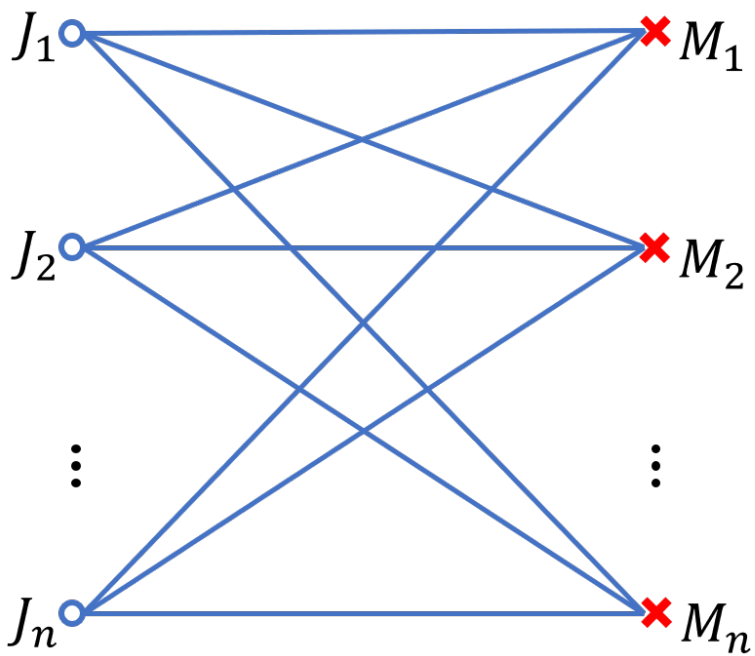
– Aldous [1]: Converges to something!

– Aldous [2]: Converges to $\pi^2/6$

Warmup: The Job Assignment Problem

– Kurtzberg [4]: Weights uniform in $[0, 1]$

$$\lim_{n\to\infty} \mathbb{E}\left[\min_{\pi} \sum_{i=1}^{n} W_{i\pi(i)}\right] < O(\log n)$$



– Walkup [6]: It is bounded.

– Meźard & Parisi [5]: Conjectured $\pi^2/6$

– Aldous [1]: Converges to something!

– Aldous [2]: Converges to $\pi^2/6$

Aldous and Steele [3]: The Objective Method

– Identify an infinite probabilistic object

– Solve the optimization problem on this infinite object

– Construct a near-optimal solution for finite problems

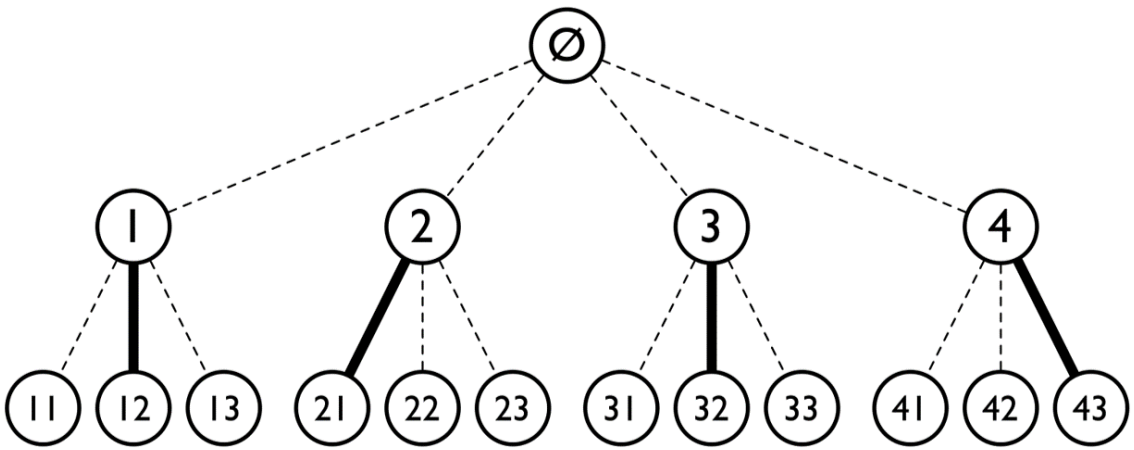
Warmup: The Job Assignment Problem

Converges locally weakly to Poisson Weighted Infinite Tree

Warmup: The Job Assignment Problem

Converges locally weakly to Poisson Weighted Infinite Tree
Solving minimum matching problem on a tree:

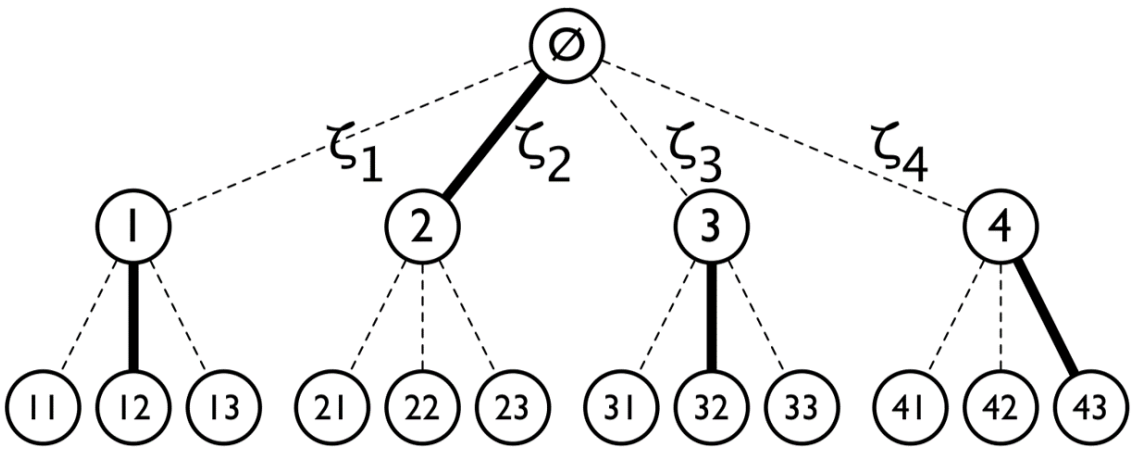
$$X_v := \text{ Cost between using and not using } v \text{ in a matching on } T_v$$



Warmup: The Job Assignment Problem

Converges locally weakly to Poisson Weighted Infinite Tree
Solving minimum matching problem on a tree:

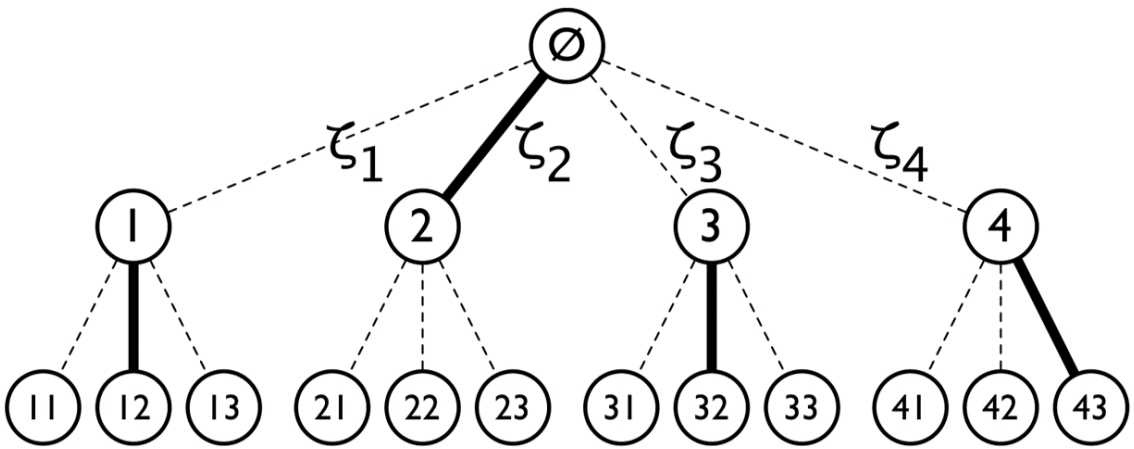
$$X_v := \text{ Cost between using and not using } v \text{ in a matching on } T_v$$



Warmup: The Job Assignment Problem

Converges locally weakly to Poisson Weighted Infinite Tree
Solving minimum matching problem on a tree:

$$X_v := \text{ Cost between using and not using } v \text{ in a matching on } T_v$$

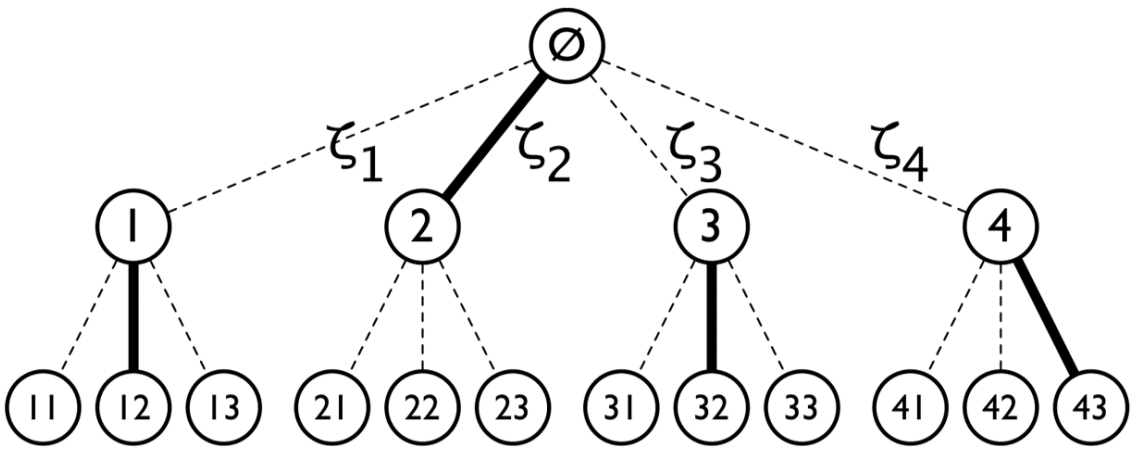


$$X_\emptyset = \text{cost of min matching on } T_\emptyset \; - \text{cost of min matching on } T_\emptyset \setminus \{\emptyset\}$$

Warmup: The Job Assignment Problem

Converges locally weakly to Poisson Weighted Infinite Tree
Solving minimum matching problem on a tree:

$$X_v := \text{ Cost between using and not using } v \text{ in a matching on } T_v$$



$$X_\emptyset = \text{cost of min matching on } T_\emptyset \; - \text{cost of min matching on } T_\emptyset \setminus \{\emptyset\}$$
$$= \min\{\zeta_i - X_i\}$$

Sorted weights: $\zeta_1, \zeta_2, \ldots$ are arrivals of a Poisson process with rate 1

Warmup: The Job Assignment Problem

Pretend everything is finite:

$$I(T_v) := \text{Cost of matching on the tree } T_v \text{ rooted at } v$$
$$I(T_v \setminus \{v\}) := \text{Cost of matching on the forest } T_v \setminus \{v\}$$

Warmup: The Job Assignment Problem

Pretend everything is finite:

$$I(T_v) := \text{Cost of matching on the tree } T_v \text{ rooted at } v$$
$$I(T_v \setminus \{v\}) := \text{Cost of matching on the forest } T_v \setminus \{v\}$$

We have:

$$\begin{aligned} I(T_\emptyset) &= \min_{i=1,2,\cdots} \{\zeta_i + I(T_i \setminus \{i\}) + \sum_{j \neq i} I(T_j)\} \\ &= \sum_{j \geq 1} I(T_j) + \min_{i=1,2,\cdots} \{\zeta_i + I(T_i \setminus \{i\}) - I(T_i)\} \end{aligned}$$

Warmup: The Job Assignment Problem

Pretend everything is finite:

$$I(T_v) := \text{Cost of matching on the tree } T_v \text{ rooted at } v$$
$$I(T_v \setminus \{v\}) := \text{Cost of matching on the forest } T_v \setminus \{v\}$$

We have:

$$\begin{aligned} I(T_\emptyset) &= \min_{i=1,2,\cdots} \{\zeta_i + I(T_i \setminus \{i\}) + \sum_{j \neq i} I(T_j)\} \\ &= \sum_{j \geq 1} I(T_j) + \min_{i=1,2,\cdots} \{\zeta_i + I(T_i \setminus \{i\}) - I(T_i)\} \end{aligned}$$

Hence,

$$X_\emptyset = I(T_\emptyset) - I(T_\emptyset \setminus \{\emptyset\}) = \min\{\zeta_i - (I(T_i) - I(T_i \setminus \{i\}))\} = \min\{\zeta_i - X_i\}$$

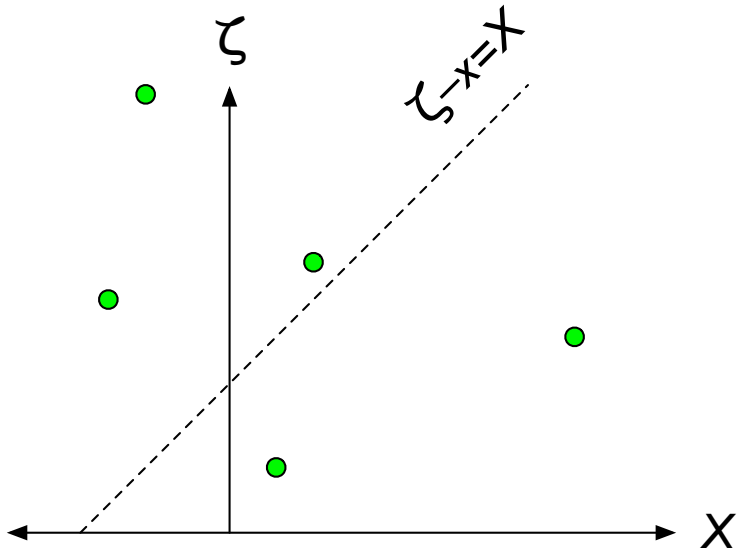
Warmup: From Distributional to Differential Equations

$X \stackrel{d}{=} \min\{\zeta_i - X_i\}$ where ζ_i are Poisson arrivals

Warmup: From Distributional to Differential Equations

$X \stackrel{d}{=} \min\{\zeta_i - X_i\}$ where ζ_i are Poisson arrivals

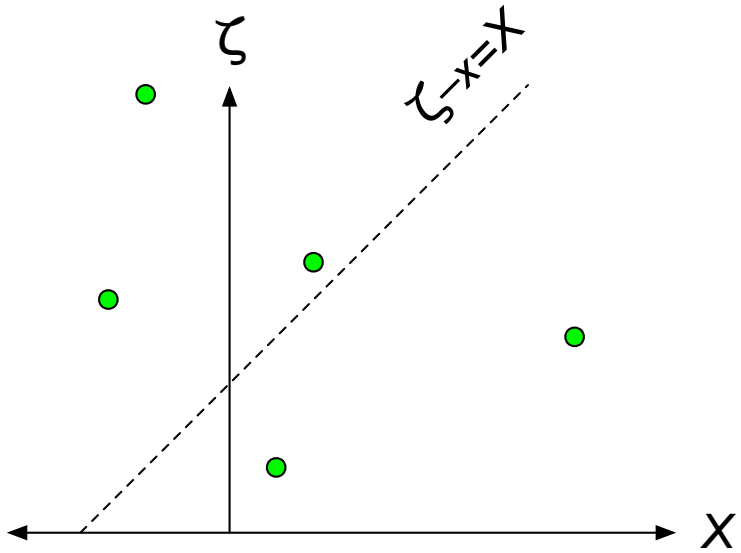
Generate pairs (ζ, x): two-dimensional Poisson process with density $f(x)$



Warmup: From Distributional to Differential Equations

$X \stackrel{d}{=} \min\{\zeta_i - X_i\}$ where ζ_i are Poisson arrivals

Generate pairs (ζ, x): two-dimensional Poisson process with density $f(x)$



Define the complementary cdf

$$\bar{F}(x) = 1 - F(x) = \mathbb{P}(X > x) = \mathbb{P}(\forall i : \zeta_i - x > X_i)$$

$$\bar{F}(x) = \exp\left(-\int_{-x}^{\infty} \bar{F}(t)\, dt\right) \quad \Rightarrow \quad \frac{\mathrm{d}F(x)}{\mathrm{d}x} = F(x)F(-x)$$

Warmup: Solving the Differential Equations

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = F(x)F(-x)$$

Warmup: Solving the Differential Equations

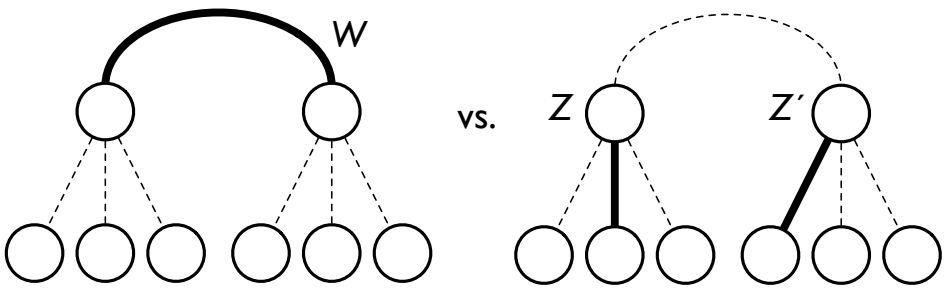
$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = F(x)F(-x)$$

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

Warmup: Solving the Differential Equations

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = F(x)F(-x)$$

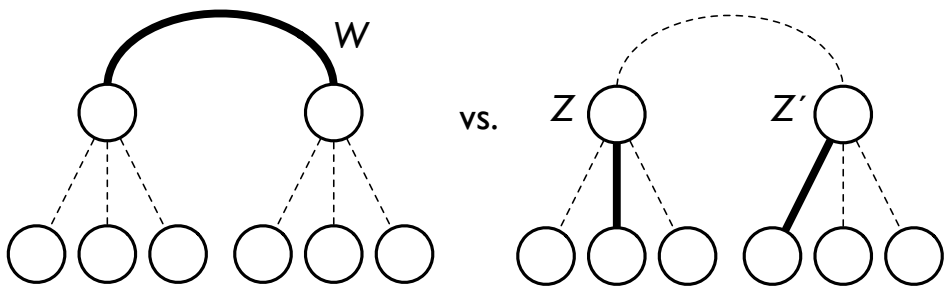
$$F(x) = \frac{e^x}{1+e^x} \quad \text{or} \quad f(x) = \frac{1}{(e^{x/2}+e^{-x/2})^2}$$



Warmup: Solving the Differential Equations

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = F(x)F(-x)$$

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

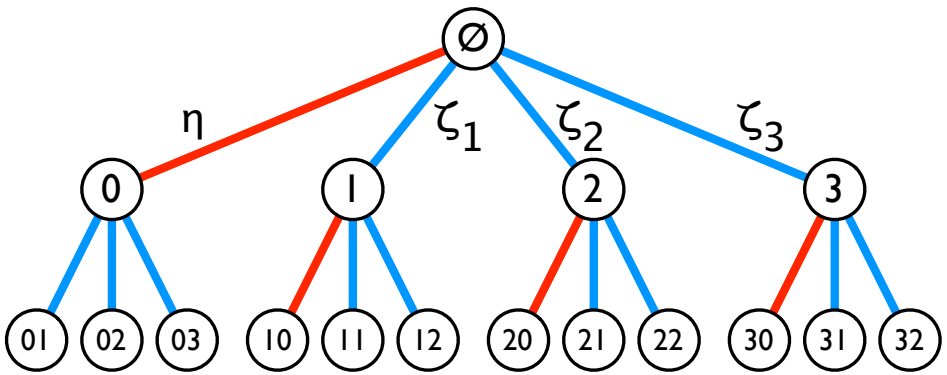


Contribution of a single edge: $\mathbb{E}\big[W\,1[W < X + X']\big]$

$$= \frac{1}{4}\,\mathrm{Var}[X+X'] = \frac{1}{2}\,\mathrm{Var}[X] = \frac{\pi^2}{6}$$

Planted Matching Problem

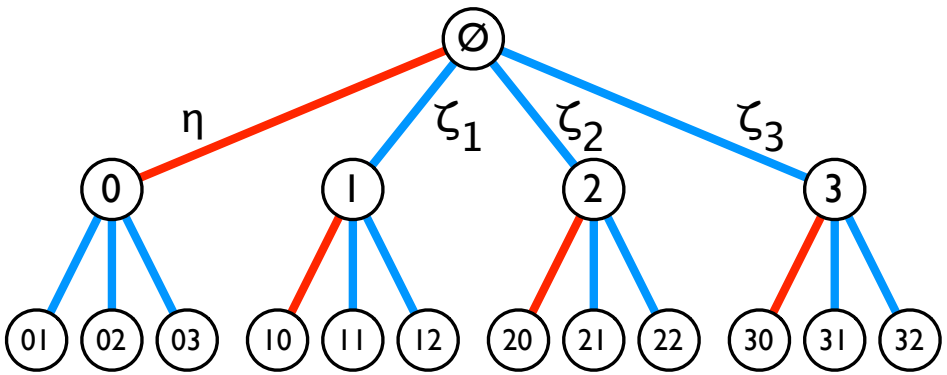
Partner in planted matching is either parent or child 0, other children sorted $1, 2, 3, \ldots$



$X_v \triangleq$ cost of min matching in T_v $-$ cost of min matching on $T_v \setminus \{v\}$

Planted Matching Problem

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



$X_v \triangleq$ cost of min matching in T_v $-$ cost of min matching on $T_v \setminus \{v\}$

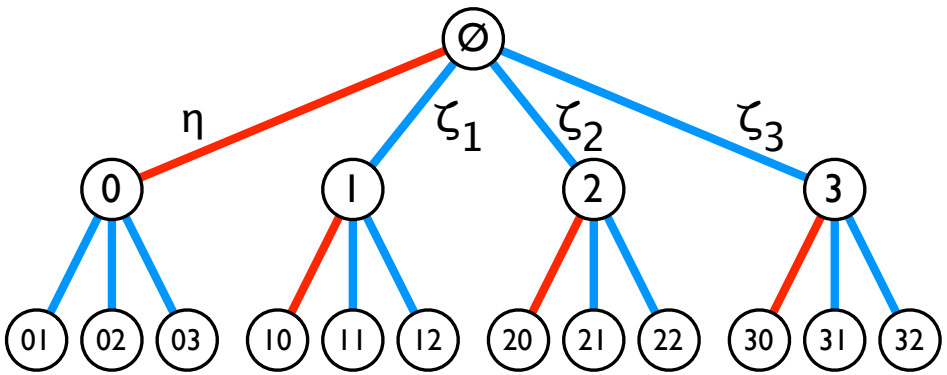
Recursion:

$$X_\varnothing = \min\left\{ W_{\varnothing,0} - X_0,\ \min_{i\geq 1}\{W_{\varnothing,i} - X_i\}\right\}$$

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

Planted Matching Problem

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



$X_v \triangleq$ cost of min matching in T_v $-$ cost of min matching on $T_v \setminus \{v\}$

Recursion:

$$X_\varnothing = \min\left\{ W_{\varnothing,0} - X_0,\ \min_{i\geq 1}\{W_{\varnothing,i} - X_i\}\right\}$$

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

$$Y \stackrel{d}{=} \min\{\eta - Z, Z'\}$$

$$Z \stackrel{d}{=} \min_i\{\zeta_i - Y_i\}$$

From Distributional to Differential Equations

$$\boxed{\begin{aligned} Y &\stackrel{d}{=} \min\left\{\eta - Z, Z'\right\} \\ Z &\stackrel{d}{=} \min\left\{\zeta_i - Y_i\right\}_{i=1}^{\infty} \end{aligned}}$$

where $\eta \sim \exp(\lambda)$ and ζ_i are Poisson arrivals

From Distributional to Differential Equations

$$\boxed{\begin{aligned} Y &\stackrel{d}{=} \min\left\{\eta - Z, Z'\right\} \\ Z &\stackrel{d}{=} \min\left\{\zeta_i - Y_i\right\}_{i=1}^{\infty} \end{aligned}}$$

where $\eta \sim \exp(\lambda)$ and ζ_i are Poisson arrivals

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

From Distributional to Differential Equations

$$\boxed{\begin{aligned} Y &\stackrel{d}{=} \min\left\{\eta - Z, Z'\right\} \\ Z &\stackrel{d}{=} \min\left\{\zeta_i - Y_i\right\}_{i=1}^{\infty} \end{aligned}}$$

where $\eta \sim \exp(\lambda)$ and ζ_i are Poisson arrivals

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

$$\boxed{\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}x} &= (1-F)(1-G)V & \frac{\mathrm{d}G}{\mathrm{d}x} &= -(1-F)(1-G)W \\ \frac{\mathrm{d}V}{\mathrm{d}x} &= \lambda(V-F) & \frac{\mathrm{d}W}{\mathrm{d}x} &= -\lambda(W-G) \end{aligned}}$$

From Distributional to Differential Equations

$$\boxed{\begin{aligned} Y &\stackrel{d}{=} \min\left\{\eta - Z, Z'\right\} \\ Z &\stackrel{d}{=} \min\left\{\zeta_i - Y_i\right\}_{i=1}^{\infty} \end{aligned}}$$

where $\eta \sim \exp(\lambda)$ and ζ_i are Poisson arrivals

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

$$\boxed{\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}x} &= (1-F)(1-G)V & \frac{\mathrm{d}G}{\mathrm{d}x} &= -(1-F)(1-G)W \\ \frac{\mathrm{d}V}{\mathrm{d}x} &= \lambda(V-F) & \frac{\mathrm{d}W}{\mathrm{d}x} &= -\lambda(W-G) \end{aligned}}$$

Boundary conditions: $F(+\infty) = V(+\infty) = 1,\ F(-\infty) = V(-\infty) = 0$

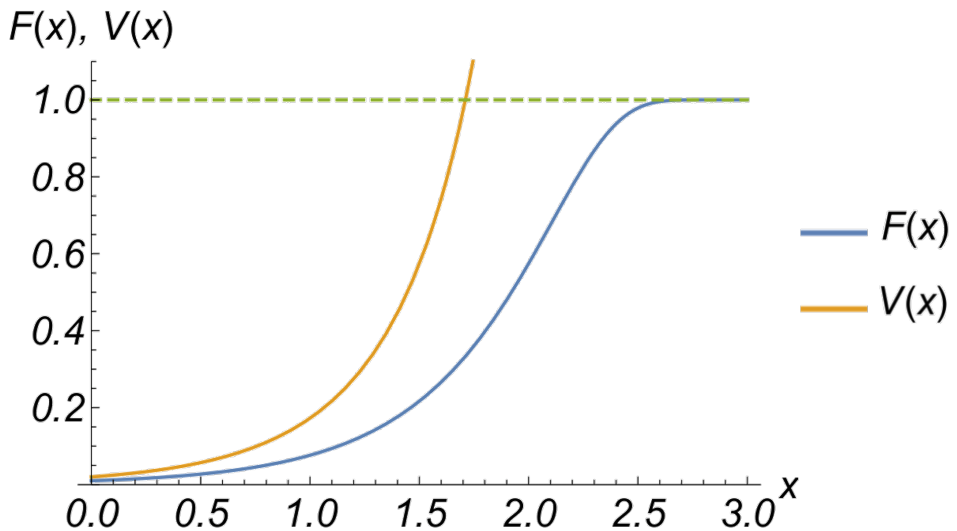
No Solution for $\lambda \geq 4$

At least no sensible one. . .

No Solution for $\lambda \geq 4$

At least no sensible one...

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



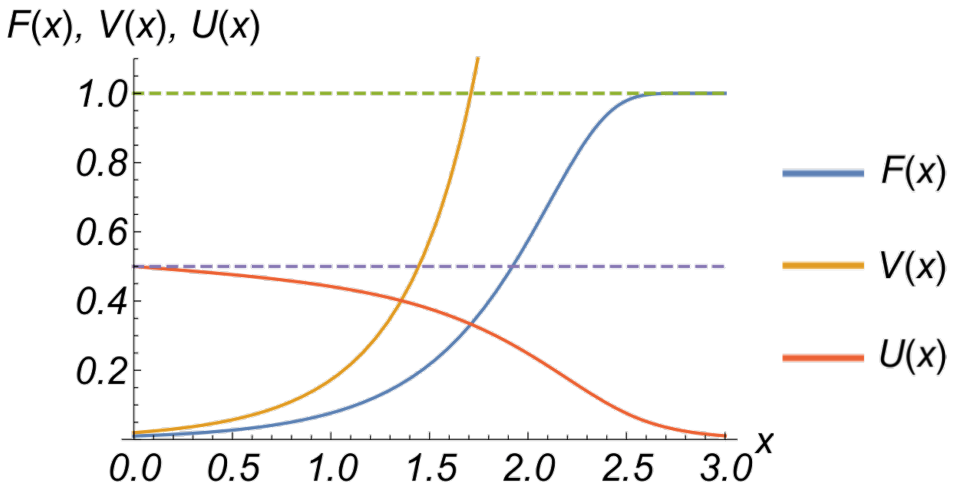
No Solution for $\lambda \geq 4$

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

No Solution for $\lambda \geq 4$

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

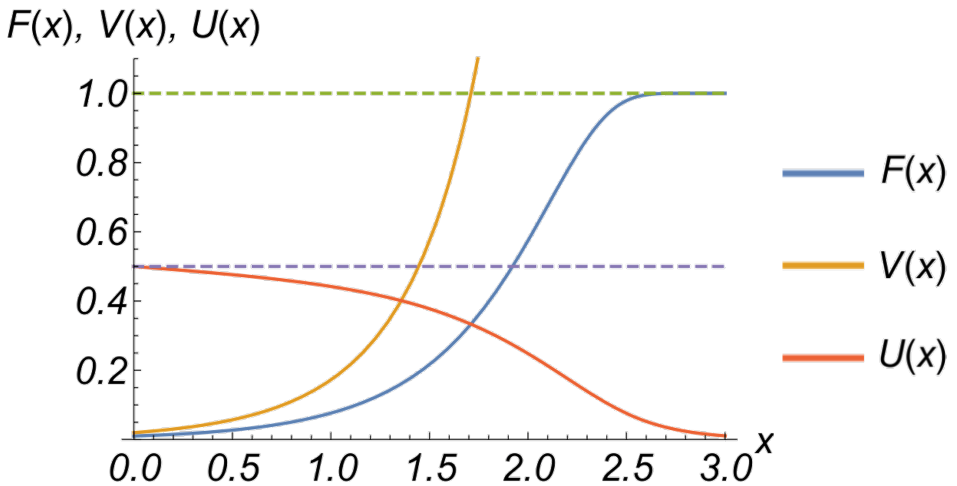
Let $U(x) = F(x)/V(x)$. Then $U(0) = 1/2$, want $U(+\infty) = 1\ldots$



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 $U(+\infty) = 1\ldots$

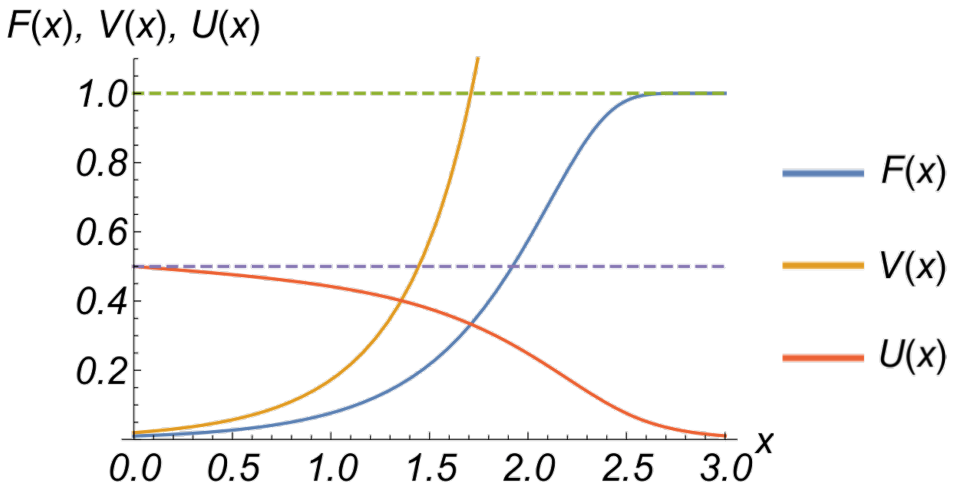


$$\dot{U} = -\lambda U(1-U) + (1-F)(1-G) \leq -\lambda U(1-U) + 1$$

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 $U(+\infty) = 1\ldots$



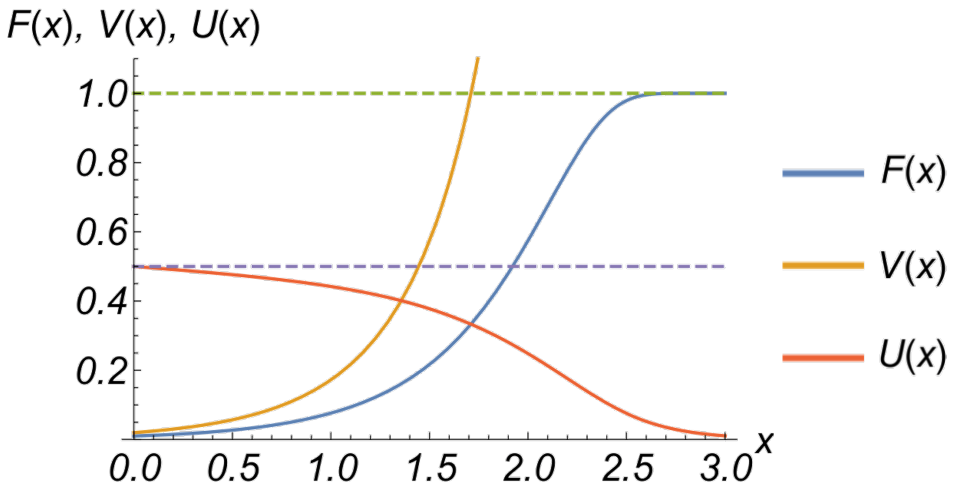
$\dot{U} = -\lambda U(1-U) + (1-F)(1-G) \leq -\lambda U(1-U) + 1$

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

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 $U(+\infty) = 1\ldots$



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

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

No fixed distribution on finite values: cost of un-planted edge is $+\infty$

$\Rightarrow$ suggests almost perfect recovery

A Unique Solution when $\lambda < 4$

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

$$\dot{U} = -\lambda U(1-U) + (1-UV)\,(1-(1-U)W)$$
$$\dot{V} = \lambda V(1-U)$$
$$\dot{W} = \lambda WU$$

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

A Unique Solution when $\lambda < 4$

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

$$\dot{U} = -\lambda U(1-U) + (1-UV)\,(1-(1-U)W)$$
$$\dot{V} = \lambda V(1-U)$$
$$\dot{W} = \lambda WU$$
$$\text{Initial conditions: } U(0) = \frac{1}{2},\ V(0) = W(0) = \epsilon$$

Lemma

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

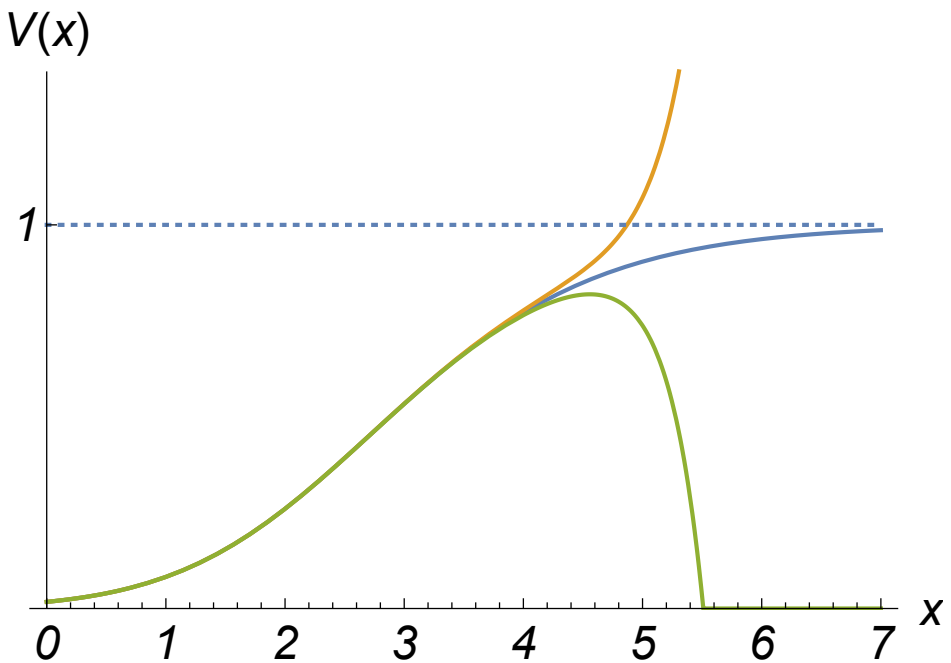
If $\epsilon \in [0, \epsilon_0)$, $U(x) \to +\infty$

If $\epsilon = \epsilon_0$, $U(x) \to 1$ and $V(x) \to 1$

If $\epsilon \in (\epsilon_0, 1]$, $V(x) \to +\infty$

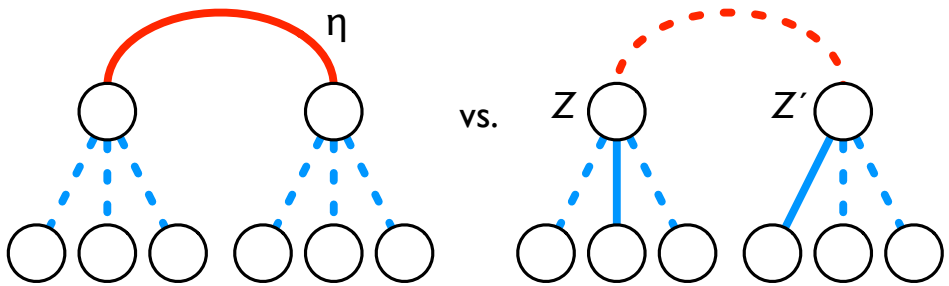
A Unique Solution when $\lambda < 4$

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

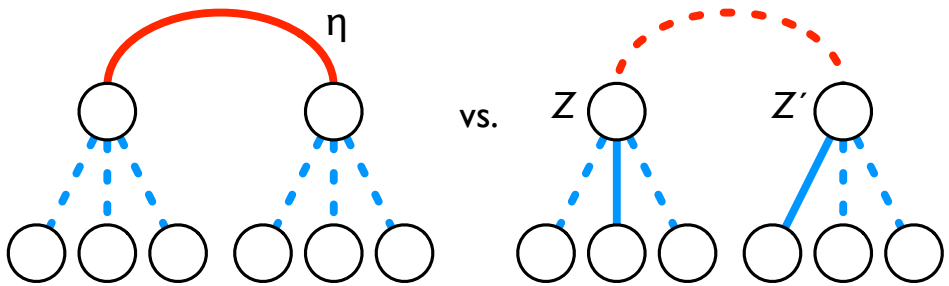


This gives cdfs $F, V \to 1$ of the unique fixed point distribution

Computing the Overlap for $\lambda < 4$



Computing the Overlap for $\lambda < 4$



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

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

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

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

Open Problems

– Order of the phase transition?

Open Problems

- Order of the phase transition?

- Concentration of the overlap?

Open Problems

- Order of the phase transition?
- Concentration of the overlap?
- Information-theoretically optimal recovery?

Open Problems

- Order of the phase transition?
- Concentration of the overlap?
- Information-theoretically optimal recovery?
- Distributions other than $\eta \sim \exp(\lambda)$?

Open Problems

- Order of the phase transition?
- Concentration of the overlap?
- Information-theoretically optimal recovery?
- Distributions other than $\eta \sim \exp(\lambda)$?
- Spatial structure (particle tracking)?

Open Problems

- Order of the phase transition?
- Concentration of the overlap?
- Information-theoretically optimal recovery?
- Distributions other than $\eta \sim \exp(\lambda)$?
- Spatial structure (particle tracking)?
- Similar arguments for the uniqueness of the solution of differential equations?

Open Problems

- Order of the phase transition?
- Concentration of the overlap?
- Information-theoretically optimal recovery?
- Distributions other than $\eta \sim \exp(\lambda)$?
- Spatial structure (particle tracking)?
- Similar arguments for the uniqueness of the solution of differential equations?
- Other planted structures: spanning trees, traveling salespeople?

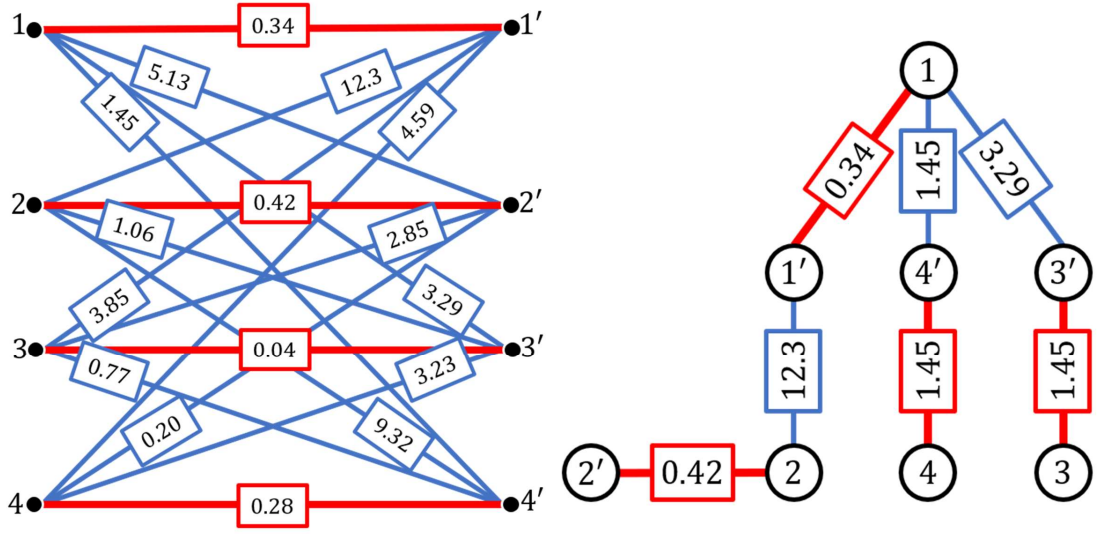
High Level Takeaway

– The Objective Method: applicable as long as there is a correlation decay, useful in combinatorial optimization problems

Thanks for your Attention

Proof Idea: Local Weak Convergence (Aldous [1, 2])

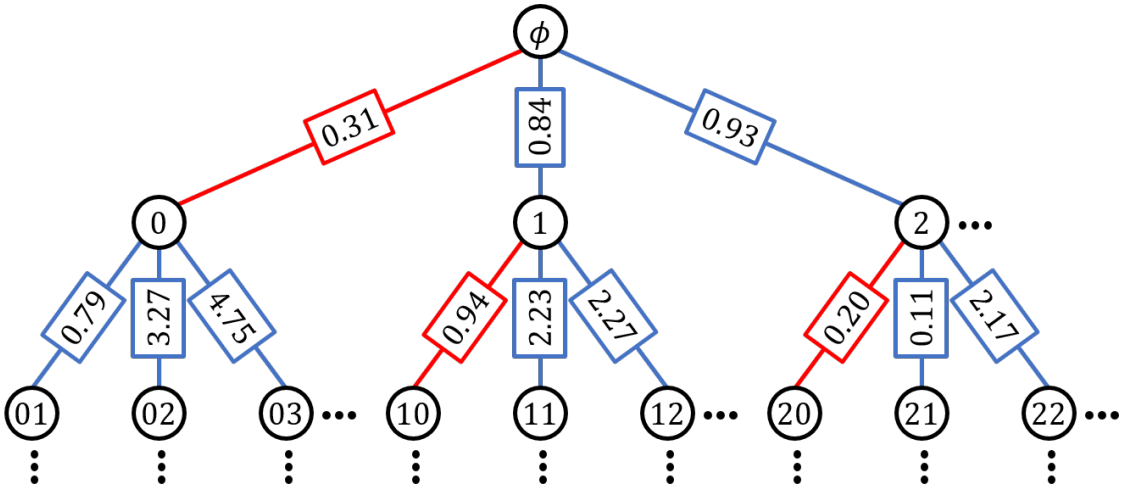
Unfolding Map



– Choose v uniformly at random and unfold the bipartite graph

Proof Idea: Local Weak Convergence (Aldous [1, 2])

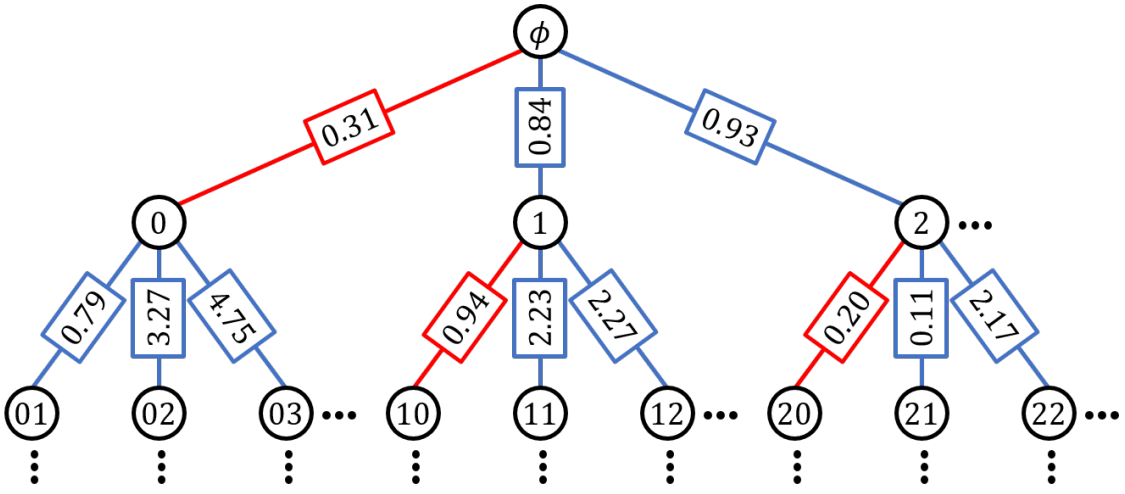
Unfolding Map



- Choose v uniformly at random and unfold the bipartite graph
- The resulted tree converges to planted PWIT locally weakly

Proof Idea: Local Weak Convergence (Aldous [1, 2])

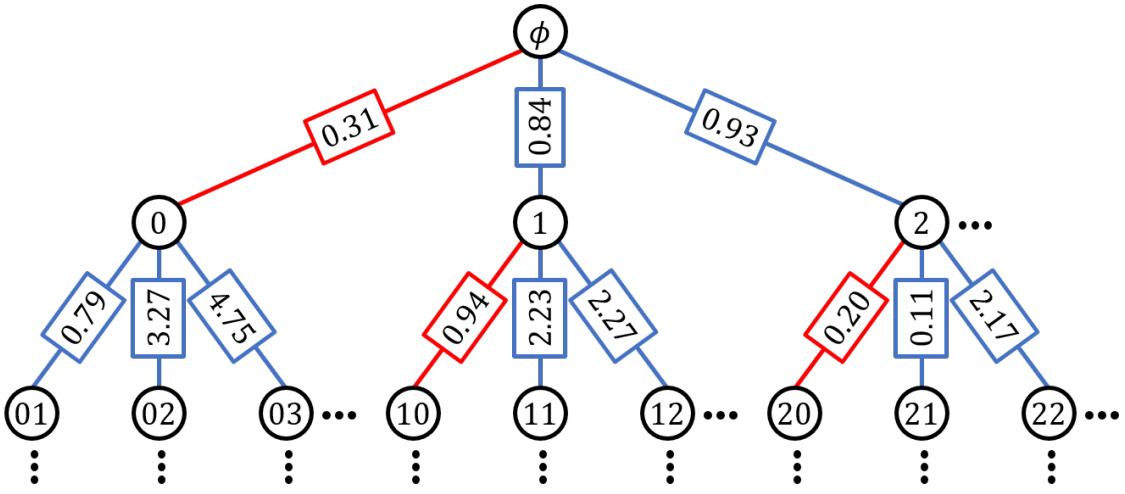
Unfolding Map



- Choose v uniformly at random and unfold the bipartite graph
- The resulted tree converges to planted PWIT locally weakly
- Uniform choice of node implies a modest symmetry property

Proof Idea: Local Weak Convergence (Aldous [1, 2])

Unfolding Map



- Choose v uniformly at random and unfold the bipartite graph
- The resulted tree converges to planted PWIT locally weakly
- Uniform choice of node implies a modest symmetry property
- The corresponding matching on infinite tree satisfies this property: *involution invariant*

Proof Idea: Local Weak Convergence (Aldous [1, 2])

Minimum Involution Invariant Matching on T_∞

Lemma

If the system of recursive distributional equation has a solution, then this solution corresponds to the minimum involution invariance matching $\mathcal{M}_{opt}$ on T_∞. Moreover, this matching is unique.

Corollary

If the system of recursive distributional equation has a solution, then this solution is unique.

We have constructed the minimum involution invariant $\mathcal{M}_{\text{opt}}$ on T_∞ and computed its cost and overlap when $\lambda < 4$

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Combine all the above

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Combine all the above

- By invariance, M_∞^* and $\mathcal{M}_{\text{opt}}$ differ at the root

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Combine all the above

- By invariance, M^*_∞ and $\mathcal{M}_{\text{opt}}$ differ at the root
- M^*_∞ often chooses the wrong partner for $\varnothing$

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Combine all the above

- By invariance, M^*_∞ and $\mathcal{M}_{\text{opt}}$ differ at the root
- M^*_∞ often chooses the wrong partner for $\varnothing$
- Right partner given by recursion $\Rightarrow$ differential equations

Proof Idea: Local Weak Convergence (Aldous [1, 2])

From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$ and Vise-Versa

Easy Half: From $\widehat{M}$ to $\mathcal{M}_{\text{opt}}$

- Any sequence of $\{M_n\}$ of matchings on $K_{n,n}$ has a subsequence $\{M_{n_j}\}$ that converges to a (possibly random) involution invariant matching on T_∞, in the local weak sense
- Local weak convergence, compactness, and Skorokhod's theorem

Hard Half: From $\mathcal{M}_{\text{opt}}$ to $\widehat{M}$

- Martingale convergence
- Almost-doubly-stochastic matrix
- Almost-perfect matching on $K_{n,n}$
- A perfect matching on $K_{n,n}$

Combine all the above

- By invariance, M^*_∞ and $\mathcal{M}_{\text{opt}}$ differ at the root
- M^*_∞ often chooses the wrong partner for $\varnothing$
- Right partner given by recursion $\Rightarrow$ differential equations

Together these imply $\lim_{n\to\infty} \text{overlap}(\widehat{M}_n) = \text{overlap}(\mathcal{M}_{\text{opt}})$

References I

David Aldous.
Asymptotics in the random assignment problem.
Probability Theory and Related Fields, 93(4):507–534, Dec 1992.

David Aldous.
The $\zeta(2)$ limit in the random assignment problem.
Random Structures & Algorithms, 18(4):381–418, 2001.

David Aldous and Michael J. Steele.
The Objective Method: Probabilistic Combinatorial Optimization and Local Weak Convergence, pages 1–72.
Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

Jerome M. Kurtzberg.
On approximation methods for the assignment problem.
J. ACM, 9(4):419–439, oct 1962.

References II

Marc Mézard and Giorgio Parisi.
On the solution of the random link matching problems.
J. Phys. France, 48(9):1451–1459, 1987.

David W. Walkup.
On the expected value of a random assignment problem.
SIAM Journal on Computing, 8(3):440–442, 1979.