How Networked Agents Make Decisions:

Coordination with Local Information & The Value of Temporal Data for Learning Influence Networks

Ph.D. Thesis Defense

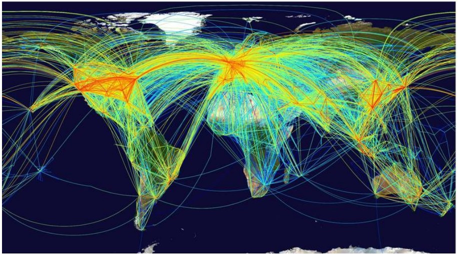
Spyros Zoumpoulis

Laboratory for Information and Decision Systems

Thesis Committee:
Marios Angeletos, Munther Dahleh, Devavrat Shah, John Tsitsiklis

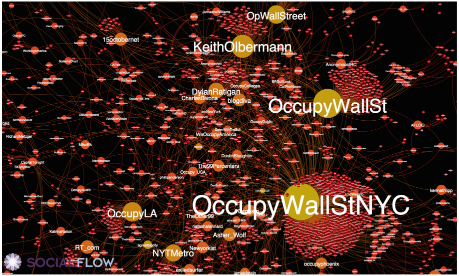
April 17, 2014

Motivation



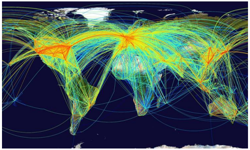
Motivation

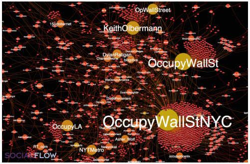
Motivation


OpWallStreet
KeithOlbermann
15octobernet
DylanRatigan
CharlesBivona
blogdiva
OccupyWallSt
WeOccupyAmerica
DustinSlaughter
The99Percenters
OccupyWallStNYC
OccupyLA
TheOther99
Asher_Wolf
RT_com
NYTMetro
Newyorkist
exiledsurfer
occupyphoenix
SOCIALFLOW

Motivation

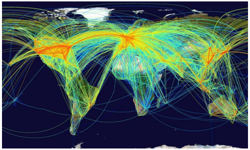
Motivation

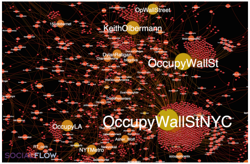




- Networks shape decisions/outcomes

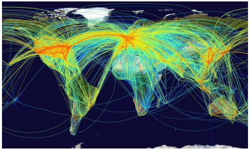
Motivation

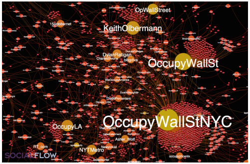




- Networks shape decisions/outcomes
- How do networked entities make decisions?

Motivation





- Networks shape decisions/outcomes
- How do networked entities make decisions?
    - What does knowledge of the network tell us about decision making of rational agents?
    - What do decisions/outcomes tell us about the network?

From Networks to Outcomes; from Outcomes to Networks

- With knowledge of the *network*, characterize the *outcomes*

- With knowledge of the *outcomes*, infer the *network*

Theme: Locality

- With knowledge of the *network*, characterize the *outcomes*

  - How do the *equilibria* of *coordination games* depend on *local* information sharing?

- With knowledge of the *outcomes*, infer the *network*

Theme: Locality

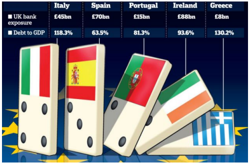
- With knowledge of the *network*, characterize the *outcomes*

  - How do the *equilibria* of *coordination games* depend on *local* information sharing?

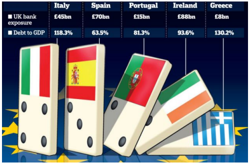
- With knowledge of the *outcomes*, infer the *network*

  - How can we *learn* the network of *local influence* between agents from records of their behavior?

Coordination with Local Information

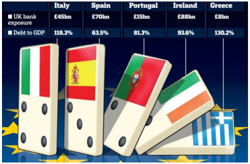


Coordination with Local Information



- Self-fulfilling crises
  - Debt crises, bank runs, currency attacks, social upheavals, ...

Coordination with Local Information



- Self-fulfilling crises
    - Debt crises, bank runs, currency attacks, social upheavals, ...
- *Local* information sharing enables coordination. How do outcomes depend on local information sharing?

Model — Agents and Payoffs

- Agents $1, \ldots, n$
- Actions: *risky* ($\alpha_i = 1$), *safe* ($\alpha_i = 0$)
- Payoffs:

$$u_i(a_i, a_{-i}, \theta) = \begin{cases} \pi(k, \theta) & \text{if } a_i = 1 \\ 0 & \text{if } a_i = 0, \end{cases}$$

where $k = \sum_{j=1}^{n} a_j$, $\theta \in \mathbb{R}$ *the fundamentals*.

Model — Agents and Payoffs

- Agents $1, \ldots, n$
- Actions: *risky* $(\alpha_i = 1)$, *safe* $(\alpha_i = 0)$
- Payoffs:

$$u_i(a_i, a_{-i}, \theta) = \begin{cases} \pi(k, \theta) & \text{if} \quad a_i = 1 \\ 0 & \text{if} \quad a_i = 0, \end{cases}$$

where $k = \sum_{j=1}^{n} a_j$, $\theta \in \mathbb{R}$ *the fundamentals*.

- Assumptions on π
  1. Strategic complementarities: $\pi(k, \theta) - \pi(k-1, \theta) > \rho > 0$
  2. State monotonicity: π strictly decreasing in θ
  3. Strict dominance regions: for sufficiently low (high) fundamentals, risky (safe) is strictly dominant

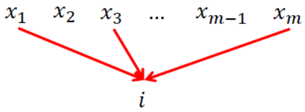
Model — Information

- θ is realized, agents hold improper prior over $\mathbb{R}$
- Conditional on θ, signals $(x_1, \ldots, x_m)$ generated: $x_r = \theta + \xi_r$
    - $(\xi_1, \ldots, \xi_m)$ independent of θ, drawn from continuous density with full support over $\mathbb{R}^m$

Model — Information

- θ is realized, agents hold improper prior over $\mathbb{R}$
- Conditional on θ, signals $(x_1, \ldots, x_m)$ generated: $x_r = \theta + \xi_r$
  - $(\xi_1, \ldots, \xi_m)$ independent of θ, drawn from continuous density with full support over $\mathbb{R}^m$

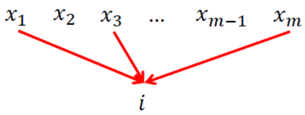
- Agent i: *observation set* $I_i \subseteq \{x_1, \ldots, x_m\}$
- $\{I_i\}_{i=1}^n$: *information structure*
- Strategy: $s_i : \mathbb{R}^{|I_i|} \to \{0, 1\}$



Model — Information

- θ is realized, agents hold improper prior over $\mathbb{R}$
- Conditional on θ, signals $(x_1, \ldots, x_m)$ generated: $x_r = \theta + \xi_r$
  - $(\xi_1, \ldots, \xi_m)$ independent of θ, drawn from continuous density with full support over $\mathbb{R}^m$

- Agent i: *observation set* $I_i \subseteq \{x_1, \ldots, x_m\}$
- $\{I_i\}_{i=1}^n$: *information structure*
- Strategy: $s_i : \mathbb{R}^{|I_i|} \to \{0, 1\}$



- We are interested in *information locality*
  - $x_r \in I_i \forall i$: public signal
  - $x_r \in I_i$ for only one i: private signal
  - all other cases: *local* signal

Problem and Contributions

- Positioning
  - Common knowledge of the fundamentals leads to multiple equilibria
  - Global-games framework leads to unique equilibrium selection
  - Endogeneity of information can restore multiplicity
  - Other natural, unstudied mechanism through which multiplicity can reemerge: *exogenous* information structure per se

Problem and Contributions

- Positioning
  - Common knowledge of the fundamentals leads to multiple equilibria
  - Global-games framework leads to unique equilibrium selection
  - Endogeneity of information can restore multiplicity
  - Other natural, unstudied mechanism through which multiplicity can reemerge: *exogenous* information structure per se

- Problem: How does the number of equilibria depend on the exogenous information structure?

Problem and Contributions

- Positioning
  - Common knowledge of the fundamentals leads to multiple equilibria
  - Global-games framework leads to unique equilibrium selection
  - Endogeneity of information can restore multiplicity
  - Other natural, unstudied mechanism through which multiplicity can reemerge: *exogenous* information structure per se
- Problem: How does the number of equilibria depend on the exogenous information structure?
- Contribution: The number of equilibria is highly sensitive to the details of information locality
  - Conditions for uniqueness vs. multiplicity that pertain solely to the details of information sharing
  - Characterization of width of multiplicity as a function of the information structure

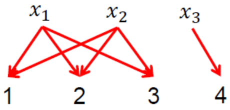
Sufficient Conditions for Multiplicity

Proposition

Information structure with $n \geq 2$ and collection C, $|C| \geq 2$, such that

- *all agents in C have the same observation set I and,*
- *for $i \notin C$, $I_i \cap I = \emptyset$.*

Then multiple Bayesian Nash equilibria.



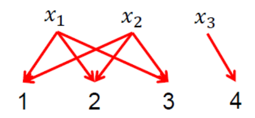
Sufficient Conditions for Multiplicity

Proposition

Information structure with $n \geq 2$ and collection C, $|C| \geq 2$, such that

- *all agents in C have the same observation set I and,*
- *for $i \notin C$, $I_i \cap I = \emptyset$.*

Then multiple Bayesian Nash equilibria.

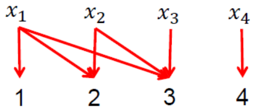


Proposition

Payoffs $\pi(k, \theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}(0, \sigma^2)$ i.i.d.
Information structure with $n \geq 2$, $\ell \geq 2$, such that

- $I_1 \subset I_2 \subset \ldots \subset I_\ell$
- *for $i > \ell$, $I_i \cap I_\ell = \emptyset$.*

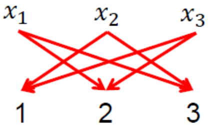
Then multiple Bayesian Nash equilibria.



May Have Uniqueness Despite Common Knowledge

May Have Uniqueness Despite Common Knowledge

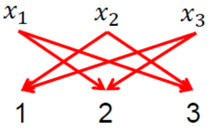
Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}\left(0, \sigma^2\right)$ i.i.d.



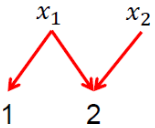
Uniqueness as long as $\sigma \geq \frac{1}{\sqrt{3\pi}}$

May Have Uniqueness Despite Common Knowledge

Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}(0, \sigma^2)$ i.i.d.



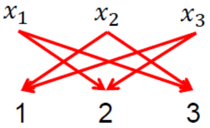
Uniqueness as long as $\sigma \geq \frac{1}{\sqrt{3\pi}}$



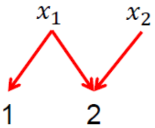
Multiplicity
(expected)

May Have Uniqueness Despite Common Knowledge

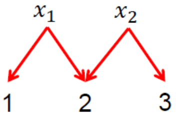
Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}\left(0, \sigma^2\right)$ i.i.d.



Uniqueness as long as $\sigma \geq \frac{1}{\sqrt{3\pi}}$



Multiplicity
(expected)



Uniqueness
(surprising)

Local Information and the Set of Equilibria

Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}\left(0,\sigma^2\right)$ i.i.d. Each agent observes only one signal. c_r: fraction of agents who observe x_r. $c_1 + \ldots + c_m = 1$.

Local Information and the Set of Equilibria

Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}(0,\sigma^2)$ i.i.d. Each agent observes only one signal. c_r: fraction of agents who observe x_r. $c_1 + \ldots + c_m = 1$.

Proposition

As $\sigma \to 0$, the strategy s_i of agent i is rationalizable if and only if

$$s_i(x) = \begin{cases} 1 & \text{if } x < \underline{\tau} \\ 0 & \text{if } x > \bar{\tau}, \end{cases}$$

where $\underline{\tau} = \frac{n}{2(n-1)}\left(1 - \|c\|_2^2\right)$ and $\bar{\tau} = 1 - \underline{\tau}$.

Local Information and the Set of Equilibria

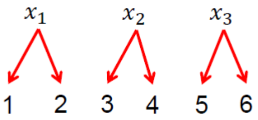
Payoffs $\pi(k,\theta) = \frac{k-1}{n-1} - \theta$, noise $\xi_r \sim \mathcal{N}(0,\sigma^2)$ i.i.d. Each agent observes only one signal. c_r: fraction of agents who observe x_r. $c_1 + \ldots + c_m = 1$.

Proposition

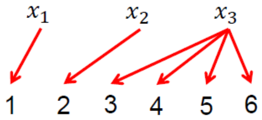
As $\sigma \to 0$, the strategy s_i of agent i is rationalizable if and only if

$$s_i(x) = \begin{cases} 1 & \textit{if } \; x < \underline{\tau} \\ 0 & \textit{if } \; x > \bar{\tau}, \end{cases}$$

where $\underline{\tau} = \frac{n}{2(n-1)}\left(1 - \|c\|_2^2\right)$ and $\bar{\tau} = 1 - \underline{\tau}$.



Narrower multiplicity



Wider multiplicity

Large Coordination Games

Proposition

As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

Large Coordination Games

Proposition

As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

- Application: networks of information exchange
  - The available signals are the idiosyncratic signals of the agents, and are exchanged through a social network $G = (\mathcal{V}, \mathcal{E})$
  - Agent i's observation set: $I_i = \{x_i, (x_j)_{j:(i,j) \in \mathcal{E}}\}$

Large Coordination Games

Proposition

As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

- Application: networks of information exchange
  - The available signals are the idiosyncratic signals of the agents, and are exchanged through a social network $G = (\mathcal{V}, \mathcal{E})$
  - Agent i's observation set: $I_i = \{x_i, (x_j)_{j:(i,j)\in\mathcal{E}}\}$
  - What networks induce unique/multiple equilibria?

Large Coordination Games

Proposition

As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

- Application: networks of information exchange
  - The available signals are the idiosyncratic signals of the agents, and are exchanged through a social network $G = (\mathcal{V}, \mathcal{E})$
  - Agent i's observation set: $I_i = \{x_i, (x_j)_{j:(i,j)\in\mathcal{E}}\}$
  - What networks induce unique/multiple equilibria?
    - Network without edges: uniqueness. Add one edge: multiplicity

Large Coordination Games

Proposition

As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

- Application: networks of information exchange
    - The available signals are the idiosyncratic signals of the agents, and are exchanged through a social network $G = (\mathcal{V}, \mathcal{E})$
    - Agent i's observation set: $I_i = \{x_i, (x_j)_{j:(i,j)\in\mathcal{E}}\}$
    - What networks induce unique/multiple equilibria?
        - Network without edges: uniqueness. Add one edge: multiplicity
        - Rationalizable strategies easy to characterize for case of disconnected cliques

Large Coordination Games

Proposition

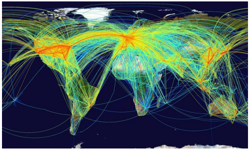
As $n \to \infty$, $\sigma \to 0$, uniqueness if and only if largest set of agents with common observation grows **sublinearly** *with n.*

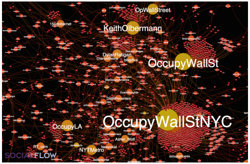
- Application: networks of information exchange
  - The available signals are the idiosyncratic signals of the agents, and are exchanged through a social network $G = (\mathcal{V}, \mathcal{E})$
  - Agent i's observation set: $I_i = \{x_i, (x_j)_{j:(i,j)\in\mathcal{E}}\}$
  - What networks induce unique/multiple equilibria?
    - Network without edges: uniqueness. Add one edge: multiplicity
    - Rationalizable strategies easy to characterize for case of disconnected cliques

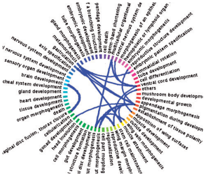
Proposition

Assume network is union of equally sized disconnected cliques. As $n \to \infty$, uniqueness if and only if cliques grow **sublinearly** *with n.*

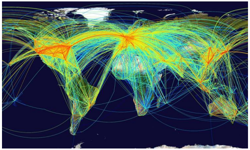
The Value of Temporal Data for Learning Influence Networks

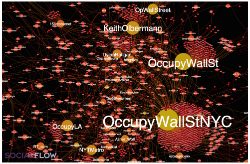


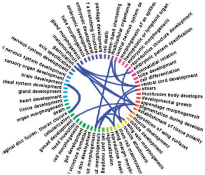




The Value of Temporal Data for Learning Influence Networks

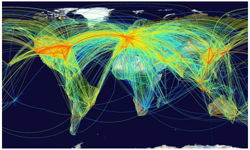


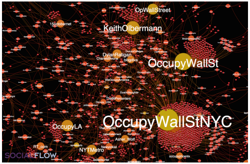


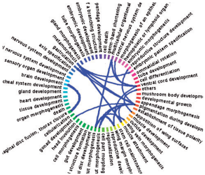


Outcomes are oftentimes observable (with time stamps), yet underlying network is hidden

The Value of Temporal Data for Learning Influence Networks





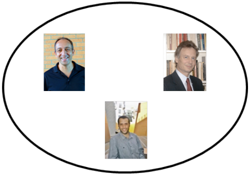


Outcomes are oftentimes observable (with time stamps), yet underlying network is hidden

- Can influences be untangled based on the outcomes in a principled manner?
- How much faster can we learn influences with access to increasingly informative temporal data?

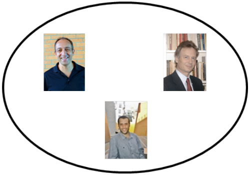
Three modes of data

- Learning with **sets** of decisions



Three modes of data

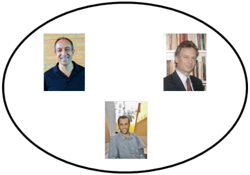
- Learning with **sets** of decisions



- Learning with **sequences** of decisions

Three modes of data

- Learning with **sets** of decisions



- Learning with **sequences** of decisions

- Learning with **times** of decisions

An example

- Agents $1, \ldots, 5$.

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5 →	$2, 1, 5$	$\{1, 2, 5\}$

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5 →	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4 →	$1, 5, 4$	$\{1, 4, 5\}$

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3	$5, 3$	$\{3, 5\}$

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

- How can we infer the large influence between $1 \longrightarrow 5$?

An example

- Agents $1, \dots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5 →	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4 →	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3 →	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2 →	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

- How can we infer the large influence between $1 \longrightarrow 5$?
  - Times: agent 5 consistently adopts shortly after agent 1

An example

- Agents $1, \dots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

- How can we infer the large influence between $1 \longrightarrow 5$?
  - Times: agent 5 consistently adopts shortly after agent 1
  - Sequences: agent 5 is consistently the next to adopt following agent 1

An example

- Agents $1, \ldots, 5$.

	Times	Sequences	Sets
Record 1	2 1 5 $\longrightarrow$	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4 $\longrightarrow$	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3 $\longrightarrow$	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2 $\longrightarrow$	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

- How can we infer the large influence between $1 \longrightarrow 5$?
  - Times: agent 5 consistently adopts shortly after agent 1
  - Sequences: agent 5 is consistently the next to adopt following agent 1
  - Sets: whenever agent 1 adopts, agent 5 also adopts

An example

- Agents $1, \ldots, 5$.

	Times	**Sequences**	**Sets**
Record 1	2 1 5 $\longrightarrow$	$2, 1, 5$	$\{1, 2, 5\}$
Record 2	1 5 4 $\longrightarrow$	$1, 5, 4$	$\{1, 4, 5\}$
Record 3	5 3 $\longrightarrow$	$5, 3$	$\{3, 5\}$
Record 4	4 1 5 2 $\longrightarrow$	$4, 1, 5, 2$	$\{1, 2, 4, 5\}$

- How can we infer the large influence between $1 \longrightarrow 5$?
    - Times: agent 5 consistently adopts shortly after agent 1
    - Sequences: agent 5 is consistently the next to adopt following agent 1
    - Sets: whenever agent 1 adopts, agent 5 also adopts
- Richer data modes allow faster/better learning, but may require more effort/cost

Model

- A product becomes available at time $t = 0$

Model

- A product becomes available at time $t = 0$
- Each of $n + 1$ agents may adopt or not
  - Agent i adopts at a time $\sim Exp(\lambda_i \geq 0)$

Model

- A product becomes available at time $t = 0$
- Each of $n + 1$ agents may adopt or not
  - Agent i adopts at a time $\sim Exp(\lambda_i \geq 0)$
- After agent i adopts, $\lambda_j \leftarrow \lambda_j + \lambda_{ij}$, $j \neq i, \lambda_{ij} \geq 0$

Model

- A product becomes available at time $t = 0$
- Each of $n + 1$ agents may adopt or not
  - Agent i adopts at a time $\sim Exp(\lambda_i \geq 0)$
- After agent i adopts, $\lambda_j \leftarrow \lambda_j + \lambda_{ij}$, $j \neq i, \lambda_{ij} \geq 0$
- Duration of the horizon $\sim Exp(\lambda_{hor})$; no adoptions possible after the end of horizon

Model

- A product becomes available at time $t = 0$
- Each of $n + 1$ agents may adopt or not
  - Agent i adopts at a time $\sim Exp(\lambda_i \geq 0)$
- After agent i adopts, $\lambda_j \leftarrow \lambda_j + \lambda_{ij}$, $j \neq i, \lambda_{ij} \geq 0$
- Duration of the horizon $\sim Exp(\lambda_{hor})$; no adoptions possible after the end of horizon
- Collection of products, λ_i's, λ_{ij}'s are static across products, adoptions across products are independent

Model

- A product becomes available at time $t = 0$
- Each of $n + 1$ agents may adopt or not
  - Agent i adopts at a time $\sim Exp(\lambda_i \geq 0)$
- After agent i adopts, $\lambda_j \leftarrow \lambda_j + \lambda_{ij}$, $j \neq i, \lambda_{ij} \geq 0$
- Duration of the horizon $\sim Exp(\lambda_{hor})$; no adoptions possible after the end of horizon
- Collection of products, λ_i's, λ_{ij}'s are static across products, adoptions across products are independent
- We are after λ_{ij}'s, λ_i's

Roadmap of Results

- Theoretical formulations and results
  - KL divergence
  - Sample complexity
- Experiments

Roadmap of Results

- Theoretical formulations and results
  - KL divergence
  - Sample complexity
- Experiments

Problem Formulation

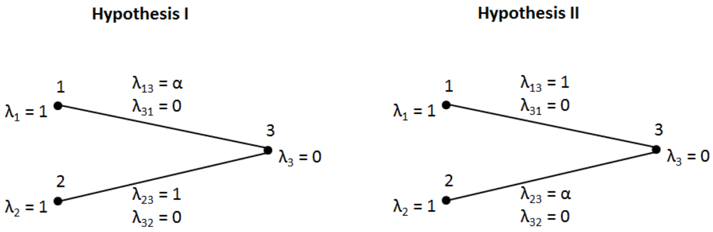
- We propose binary hypothesis testing problems

Problem Formulation

- We propose binary hypothesis testing problems
  - Which of two peers influences you crucially?

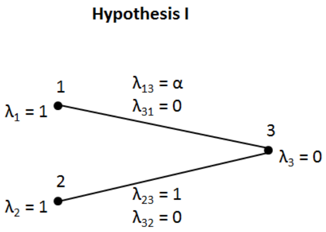
Problem Formulation

- We propose binary hypothesis testing problems
  - Which of two peers influences you crucially?



Problem Formulation

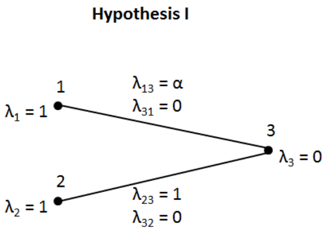
- We propose binary hypothesis testing problems
  - Which of two peers influences you crucially?



- Ability to tell hypotheses apart depends on how far apart distributions of outcomes are under hypotheses

Problem Formulation

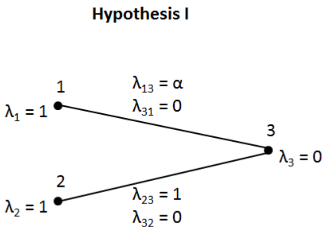
- We propose binary hypothesis testing problems
  - Which of two peers influences you crucially?



- Ability to tell hypotheses apart depends on how far apart distributions of outcomes are under hypotheses
- Best can do asymptotically: $Pr(\text{error}) \approx e^{-D(P_1||P_2)\cdot(\#\text{ of samples})}$

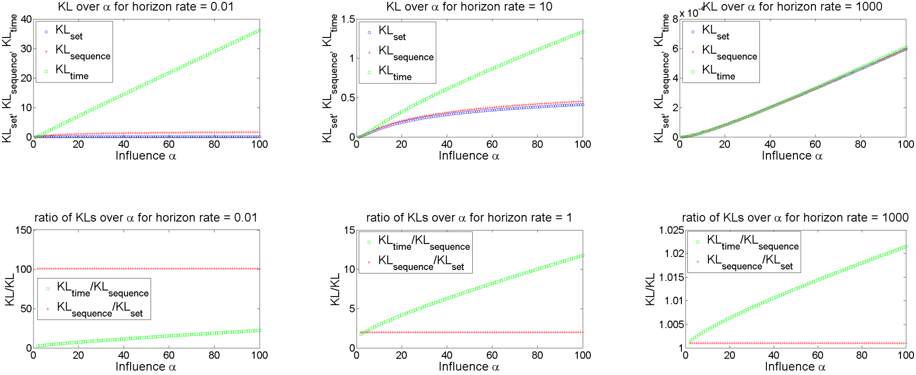
Problem Formulation

- We propose binary hypothesis testing problems
  - Which of two peers influences you crucially?



- Ability to tell hypotheses apart depends on how far apart distributions of outcomes are under hypotheses
- Best can do asymptotically: $Pr(\text{error}) \approx e^{-D(P_1||P_2)\cdot(\#\text{ of samples})}$
- Focus on KL divergence $D(P_1||P_2) \equiv \mathbb{E}_{P_1}\left[\log \frac{dP_1}{dP_2}\right]$ for cases of outcomes: sets, sequences, times.

Which of two peers influences you crucially?



Large horizon: sequences large gain over sets, times have gain but smaller for small α

Small horizon: sets provide almost all the information, times have value only for large α

Roadmap of Results

- Theoretical formulations and results
    - KL divergence
    - Sample complexity
- Experiments

Problem Formulation

- Prior knowledge: graph $G = (\mathcal{V}, \mathcal{E})$. $\lambda_{ij} = 0$ if edge $i - j$ is not in $\mathcal{E}$.

Problem Formulation

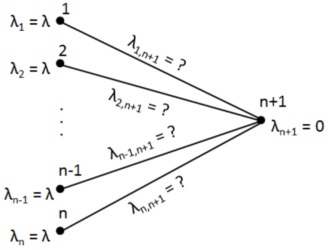
- Prior knowledge: graph $G = (\mathcal{V}, \mathcal{E})$. $\lambda_{ij} = 0$ if edge $i - j$ is not in $\mathcal{E}$.
- How many i.i.d. products required to learn correctly?

Problem Formulation

- Prior knowledge: graph $G = (\mathcal{V}, \mathcal{E})$. $\lambda_{ij} = 0$ if edge $i - j$ is not in $\mathcal{E}$.
- How many i.i.d. products required to learn correctly?
  - Learn one edge vs. learn all edges
  - Different prior knowledge
  - Moderate horizon vs. small horizon
  - Sets vs. sequences vs. times

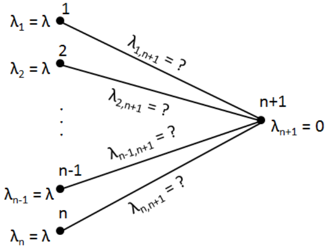
Problem Formulation

- Prior knowledge: graph $G = (\mathcal{V}, \mathcal{E})$. $\lambda_{ij} = 0$ if edge $i - j$ is not in $\mathcal{E}$.
- How many i.i.d. products required to learn correctly?
  - Learn one edge vs. learn all edges
  - Different prior knowledge
  - Moderate horizon vs. small horizon
  - Sets vs. sequences vs. times



Problem Formulation

- Prior knowledge: graph $G = (\mathcal{V}, \mathcal{E})$. $\lambda_{ij} = 0$ if edge $i - j$ is not in $\mathcal{E}$.
- How many i.i.d. products required to learn correctly?
  - Learn one edge vs. learn all edges
  - Different prior knowledge
  - Moderate horizon vs. small horizon
  - Sets vs. sequences vs. times



- Consider $\lambda_{ij} = 0$ or ∞

Sample Complexity Results

- Bayesian setting, $p = \frac{1}{2}$ of ∞ for each edge

	$\lambda_{hor} = \lambda$		$\lambda_{hor} = n\lambda$	
	Sets	Sequences	Sets	Sequences
Learn one	$\Theta(n^2)$	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$
Learn all	$\Theta(n^2 \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$

Sample Complexity Results

- Bayesian setting, $p = \frac{1}{2}$ of ∞ for each edge

	$\lambda_{hor} = \lambda$		$\lambda_{hor} = n\lambda$	
	Sets	Sequences	Sets	Sequences
Learn one	$\Theta(n^2)$	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$
Learn all	$\Theta(n^2 \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$

- Known scaling ℓ of agents with influence rate ∞ to $n+1$

	$\lambda_{hor} = \lambda$		$\lambda_{hor} = n\lambda$	
	Sets	Sequences	Sets	Sequences
$\ell = 1$	$\Theta(\log n)$	$\Theta(1)$	$\Theta(n)$	$\Theta(n)$
$\ell = \alpha n, \alpha \in (0,1)$	$\Theta(n^2 \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$	$\Theta(n \log n)$
$\ell = n-1$	$\Theta(n^2)$	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$

Roadmap of Results

- Theoretical formulations and results
  - KL divergence
  - Sample complexity
- Experiments

Real Data — Mobile Apps

- Mobile applications dataset (courtesy of Sandy Pentland, MIT Media Lab)
  - Installations of mobile apps by 55 users during the experimental period of four months
  - Call logs
  - Bluetooth hits
  - Networks of declared affiliation and friendship among the participants

Real Data — Mobile Apps

- Mobile applications dataset (courtesy of Sandy Pentland, MIT Media Lab)
  - Installations of mobile apps by 55 users during the experimental period of four months
  - Call logs
  - Bluetooth hits
  - Networks of declared affiliation and friendship among the participants
- Infer influence network from data on actions; validate with social data

Real Data — Mobile Apps

- Mobile applications dataset (courtesy of Sandy Pentland, MIT Media Lab)
  - Installations of mobile apps by 55 users during the experimental period of four months
  - Call logs
  - Bluetooth hits
  - Networks of declared affiliation and friendship among the participants

- Infer influence network from data on actions; validate with social data

- A friendship edge exists between two randomly selected nodes in the dataset with probability 0.3508

Real Data — Mobile Apps

- Mobile applications dataset (courtesy of Sandy Pentland, MIT Media Lab)
  - Installations of mobile apps by 55 users during the experimental period of four months
  - Call logs
  - Bluetooth hits
  - Networks of declared affiliation and friendship among the participants

- Infer influence network from data on actions; validate with social data

- A friendship edge exists between two randomly selected nodes in the dataset with probability 0.3508

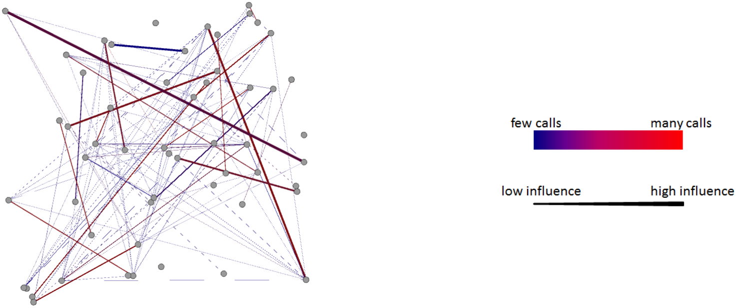
Out of top	10	joint influence edges, friendship exists in	70%
	20		65%
	50		54%
	100		38%

Real Data — Mobile Apps

- The inferred influence rates are highly correlated with the realized social networks when learning with *sequences*

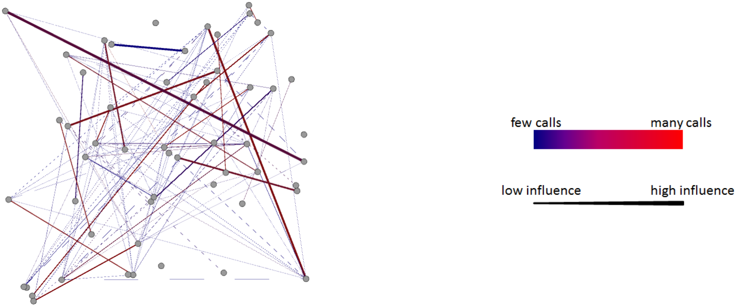
Real Data — Mobile Apps

- The inferred influence rates are highly correlated with the realized social networks when learning with *sequences*



Real Data — Mobile Apps

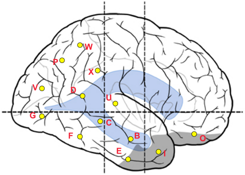
- The inferred influence rates are highly correlated with the realized social networks when learning with *sequences*



- Correlation only slightly lower when learning with *sets*

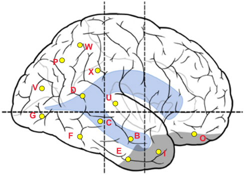
Real Data — Epileptic Seizures

- EEG data obtained from epileptic patients
  (courtesy of Sri Sarma, Johns Hopkins Inst. for Comp. Medicine)
  - 10-15 electrodes per patient, 10 channels per electrode
  - Different electrodes/channels monitor different regions of the brain
  - 3-6 seizure events per patient
  - For each seizure event, data is voltage measurements per millisecond across all channels, over 4-7 minutes



Real Data — Epileptic Seizures

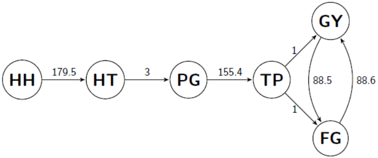
- EEG data obtained from epileptic patients
  (courtesy of Sri Sarma, Johns Hopkins Inst. for Comp. Medicine)
  - 10-15 electrodes per patient, 10 channels per electrode
  - Different electrodes/channels monitor different regions of the brain
  - 3-6 seizure events per patient
  - For each seizure event, data is voltage measurements per millisecond across all channels, over 4-7 minutes



- Infer influence network from data on seizures; generate hypotheses:
  - What regions are more likely to start the seizure?
  - How likely is it that a region that is under seizure infects another region?

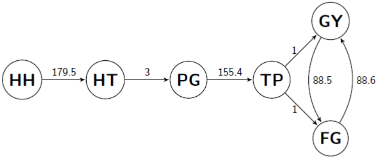
Real Data — Epileptic Seizures

- Unconstrained optimization of the likelihood, ℓ_1-regularized. Sparsest inferred network that is connected:

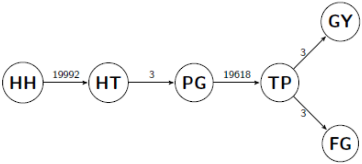


Real Data — Epileptic Seizures

- Unconstrained optimization of the likelihood, ℓ_1-regularized. Sparsest inferred network that is connected:



- Constrained optimization of the likelihood over trees whose nodes have at most two children



Summary

Networks

Game theory

Estimation
& Learning

Outcomes

Current and Future Work, Open Questions

- Coordination with local information
  - What happens with information asymmetry about the information structure?
  - Prove sufficient condition for multiplicity is not necessary
  - Open: complete characterization of sufficient and necessary conditions
  - Open: dependence of width of multiplicity on the information structure in the general case

Current and Future Work, Open Questions

- Coordination with local information
  - What happens with information asymmetry about the information structure?
  - Prove sufficient condition for multiplicity is not necessary
  - Open: complete characterization of sufficient and necessary conditions
  - Open: dependence of width of multiplicity on the information structure in the general case

- The value of temporal data for learning influence networks
  - Sample complexity results for broader families of networks
  - Reconstructing trees with latent nodes
  - Open: other influence/observation window models

Thank You!

- John, Munzer

Thank You!

- John, Munzer
- Devavrat, Marios

Thank You!

- John, Munzer
- Devavrat, Marios
- LIDS

Thank You!

- John, Munzer
- Devavrat, Marios
- LIDS
- Come visit in Paris!