Aggregative Efficiency of Bayesian Learning in Networks

Krishna Dasaratha Kevin He

Simons Institute Economics of Networks Conference

December 1, 2022

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?
- Much of existing theoretical work on complete network

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?
- Much of existing theoretical work on complete network
- Less known about how rational social learning compares across networks, and existing results say agents **eventually** learn completely on **all** (reasonable) networks

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?
- Much of existing theoretical work on complete network
- Less known about how rational social learning compares across networks, and existing results say agents **eventually** learn completely on **all** (reasonable) networks
- Open question: impact of network on **how well signals are aggregated** — hence how quickly rational agents learn

Social-Learning Dynamics in Different Networks

- **Social learning**: info about unknown state dispersed among society of agents, agents act based on private signals and observations of social neighbors
- How does social network affect efficiency of info aggregation?
- Much of existing theoretical work on complete network
- Less known about how rational social learning compares across networks, and existing results say agents **eventually** learn completely on **all** (reasonable) networks
- Open question: impact of network on **how well signals are aggregated** — hence how quickly rational agents learn

Golub and Sadler (2016): "A significant gap in our knowledge concerns short-run dynamics and rates of learning in these models."

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior
- "intransitivity" that appears in almost all realistic observation networks can lead to **arbitrarily inefficient** social learning

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior
- "intransitivity" that appears in almost all realistic observation networks can lead to **arbitrarily inefficient** social learning

Generations network – observe subset of agents in previous gen

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior
- "intransitivity" that appears in almost all realistic observation networks can lead to **arbitrarily inefficient** social learning

Generations network – observe subset of agents in previous gen

- express learning rate as simple function of network parameters

Environment and Key Results

Introduce **tractable model of rational sequential learning** that lets us compare learning dynamics across different networks

- define measure that ranks networks wrt social-learning efficiency

Highlight **network-based informational confounds**

- suppose P2 and P3 see P1, but P4 sees only P2 and P3
- from P4's perspective, P1's action confounds the info content of P2 and P3's behavior
- "intransitivity" that appears in almost all realistic observation networks can lead to **arbitrarily inefficient** social learning

Generations network – observe subset of agents in previous gen

- express learning rate as simple function of network parameters
- extent of info loss: under a symmetry condition, learning aggregates **no more than 2 signals per gen** asymptotically

Related Social-Learning Literature

Sequential learning: Banerjee (1992), Bikhchandani, Hirshleifer, Welch (1992)

Network structure and Bayesian social learning

- Network does not matter (within "reasonable" class) for long-run learning: Acemoglu, Dahleh, Lobel, and Ozdaglar (2011), Lobel and Sadler (2015), Rosenberg and Vieille (2019)
- Examples and simulations: Sgroi (2002), Lobel, Acemoglu, Dahleh, and Ozdaglar (2009), Arieli and Mueller-Frank (2019)
- Adoption Dynamics: Board and Meyer-ter-Vehn (2021)
- This paper: **analytic ranking of networks on rate of learning**

Other obstructions to the efficient learning rate

- Coarse action space: Harel, Mossel, Strack, Tamuz (2020), Rosenberg and Vieille (2019), Hann-Caruthers, Martynov, Tamuz (2018)
- Endogenous info: Burguet and Vives (2000), Mueller-Frank and Pai (2016), Ali (2018), Lomys (2020), Liang and Mu (2020)
- This paper: **network-based** obstructions to efficient learning

Speed of learning under heuristics: Ellison and Fudenberg (1993), Golub and Jackson (2012), Molavi, Tahbaz-Salehi, Jadbabaie (2018). This paper: **rational learning**

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, \ldots$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, \ldots$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, \ldots i-1\}$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, ...i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, ... i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$
- Signals are Gaussian and conditionally i.i.d. given state,

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, \ldots$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, \ldots i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$
- Signals are Gaussian and conditionally i.i.d. given state, $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, ...i - 1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$
- Signals are Gaussian and conditionally i.i.d. given state, $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$ and $s_i \sim \mathcal{N}(-1, \sigma^2)$ when $\omega = 0$

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, ...$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, ...i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$
- Signals are Gaussian and conditionally i.i.d. given state, $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$ and $s_i \sim \mathcal{N}(-1, \sigma^2)$ when $\omega = 0$
- Neighborhoods $N(i)$ are common knowledge

Model and Notations

- Two equally likely states $\omega \in \{0, 1\}$
- Agents $i = 1, 2, 3, \ldots$ move in order, each acting once
  - i observes **private signal** $s_i \in \mathbb{R}$ and actions of **neighbors**, $N(i) \subseteq \{1, \ldots i-1\}$
  - picks **action** $a_i \in [0, 1]$ to maximize expectation of $-(a_i - \omega)^2$, so $a_i = \mathbb{P}[\omega = 1 \mid i\text{'s information}]$
- Signals are Gaussian and conditionally i.i.d. given state, $s_i \sim \mathcal{N}(1, \sigma^2)$ when $\omega = 1$ and $s_i \sim \mathcal{N}(-1, \sigma^2)$ when $\omega = 0$
- Neighborhoods $N(i)$ are common knowledge
- Agents are Bayesian and choose optimal actions (given observations and predecessors' play)

Log-Linearity of Actions

WLOG apply log-transformations and work with log-likelihoods

- **Log-signal** $\lambda_i := \ln\left(\frac{\mathbb{P}[\omega=1|s_i]}{\mathbb{P}[\omega=0|s_i]}\right)$, **log-actions** $\ell_i := \ln\left(\frac{a_i}{1-a_i}\right)$
- These changes are 1-to-1, so there is a (unique) map from i's log-signal and neighbors' log-actions to i's optimal log-action
- Our first result says this map is linear

Proposition 1

For each agent i with $N(i) = \{j(1), ..., j(n_i)\}$, there exist constants $(\beta_{i,j(k)})_{k=1}^{n_i}$ s.t.

$$\ell_i^* = \lambda_i + \sum_{k=1}^{n_i} \beta_{i,j(k)} \ell_{j(k)}^*.$$

- Proof gives explicit formula for coefficients $\beta_{i,j(k)}$

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if log-action $\ell_i^* \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if log-action $\ell_i^* \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

- When agents choose arbitrary actions (even if log-linear), need not hold for **any** $r \in \mathbb{R}$

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if log-action $\ell_i^* \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

- When agents choose arbitrary actions (even if log-linear), need not hold for **any** $r \in \mathbb{R}$
- But **rational** log-actions always admit this kind of signal-counting interpretation, suff. stat for rational accuracy

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if log-action $\ell_i^* \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

- When agents choose arbitrary actions (even if log-linear), need not hold for **any** $r \in \mathbb{R}$
- But **rational** log-actions always admit this kind of signal-counting interpretation, suff. stat for rational accuracy

Proposition 2

There exist $(r_i)_{i \geq 1}$ *so that social learning aggregates* r_i *signals by agent* i*. These* $(r_i)_{i \geq 1}$ *depend on the network, but not on* σ^2*.*

Signal-Counting Interpretation of Accuracy

If i's only info is $n \in \mathbb{N}_+$ indep signals, $\ell_i \sim \mathcal{N}\left(\pm n \cdot \frac{2}{\sigma^2}, n \cdot \frac{4}{\sigma^2}\right)$

Definition

Social learning **aggregates** $r \in \mathbb{R}_+$ **signals by agent** i if log-action $\ell_i^* \sim \mathcal{N}\left(\pm r \cdot \frac{2}{\sigma^2}, r \cdot \frac{4}{\sigma^2}\right)$ in two states.

- When agents choose arbitrary actions (even if log-linear), need not hold for **any** $r \in \mathbb{R}$
- But **rational** log-actions always admit this kind of signal-counting interpretation, suff. stat for rational accuracy

Proposition 2

There exist $(r_i)_{i \geq 1}$ so that social learning aggregates r_i signals by agent i. These $(r_i)_{i \geq 1}$ depend on the network, but not on σ^2.

- Can measure each i's accuracy in units of private signals

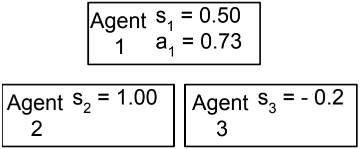
An Example of Informational Confound

Agent 1

$s_1 = 0.50$

$a_1 = 0.73$

An Example of Informational Confound



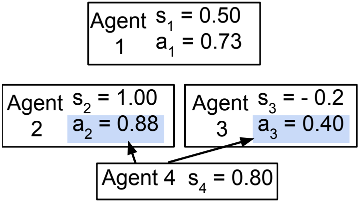
An Example of Informational Confound

Agent 1: $s_1 = 0.50$, $a_1 = 0.73$

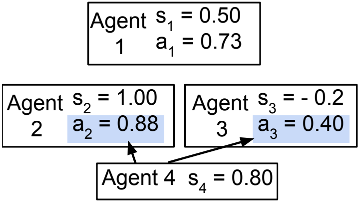
Agent 2: $s_2 = 1.00$, $a_2 = 0.88$

Agent 3: $s_3 = -0.2$, $a_3 = 0.40$

An Example of Informational Confound

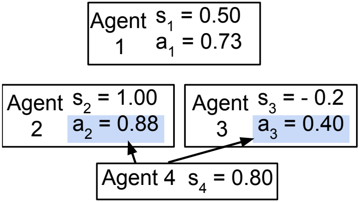


An Example of Informational Confound



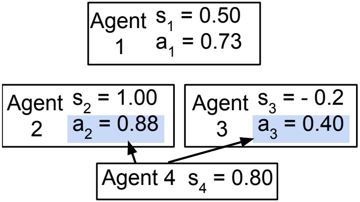
- P4 perfectly infers P2 and P3's signals from their actions

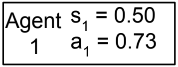
An Example of Informational Confound



- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

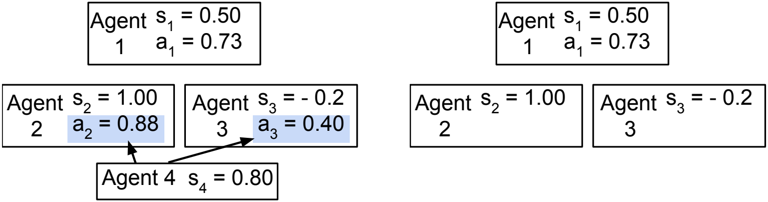
An Example of Informational Confound





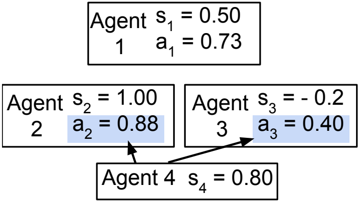
- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

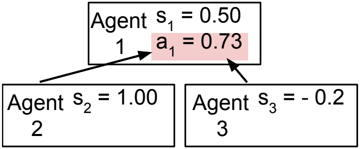
An Example of Informational Confound



- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

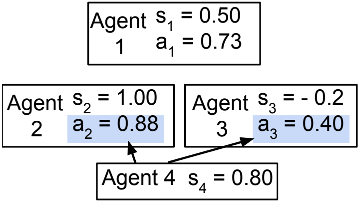
An Example of Informational Confound

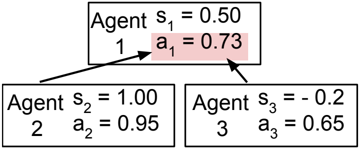




- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

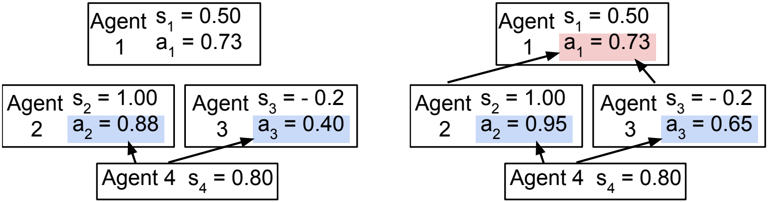
An Example of Informational Confound





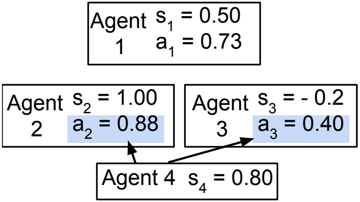
- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

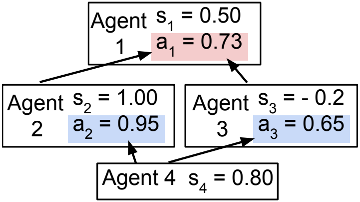
An Example of Informational Confound



- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

An Example of Informational Confound

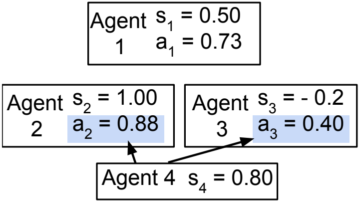




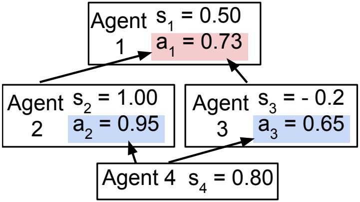
- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

- a_1 influences both a_2 and a_3, but is unobserved by P4

An Example of Informational Confound

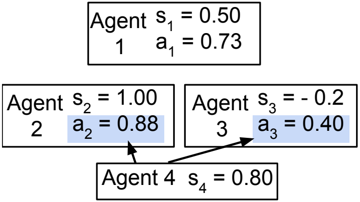


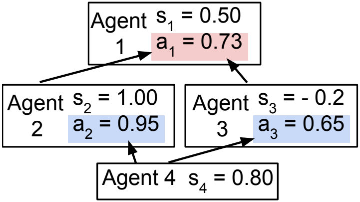
- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4



- a_1 influences both a_2 and a_3, but is unobserved by P4
- P4 cannot fully incorporate s_2 and s_3 without over-counting s_1

An Example of Informational Confound

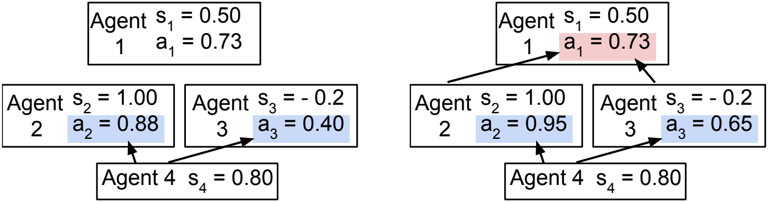




- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

- a_1 influences both a_2 and a_3, but is unobserved by P4
- P4 cannot fully incorporate s_2 and s_3 without over-counting s_1
- With optimal signal extraction, $r_4 =$

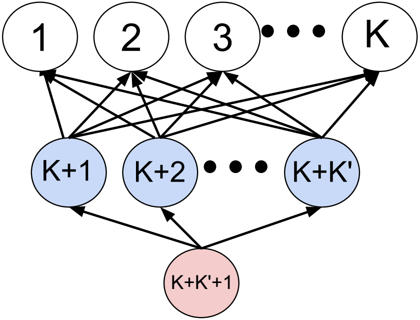
An Example of Informational Confound



- P4 perfectly infers P2 and P3's signals from their actions
- $r_4 = 3$ signals, fully incorporates info in s_2, s_3, and s_4

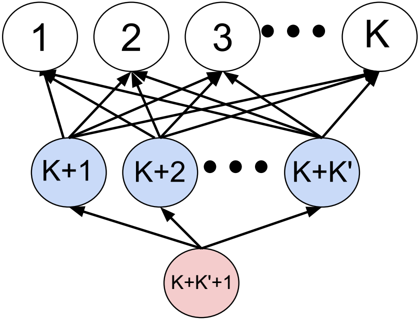
- a_1 influences both a_2 and a_3, but is unobserved by P4
- P4 cannot fully incorporate s_2 and s_3 without over-counting s_1
- With optimal signal extraction, $r_4 =$ "$3.\overline{66}$ signals" (to be formalized soon)

A More Severe Confound



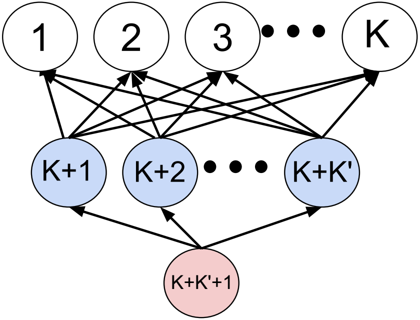
- Three generations of differing sizes, each observe all members of previous generation

A More Severe Confound



- Three generations of differing sizes, each observe all members of previous generation
- Change in accuracy between generations 2 and 3 is equivalent to getting $\frac{(K+1)(K'-1)}{KK'+1} + 1 < 3$ additional signals

A More Severe Confound



- Three generations of differing sizes, each observe all members of previous generation
- Change in accuracy between generations 2 and 3 is equivalent to getting $\frac{(K+1)(K'-1)}{KK'+1} + 1 < 3$ additional signals
- Little change in accuracy–even if K small, so little confounding information!
  - Even if K' is also large–many new signals in generation 2, but almost all information lost

Condition for Long-Run Learning

- Society learns completely in the long run if actions $a_i^* \to \omega$ in probability (equivalent to $r_i \to \infty$)

Condition for Long-Run Learning

- Society learns completely in the long run if actions $a_i^* \to \omega$ in probability (equivalent to $r_i \to \infty$)

Proposition 3

Society learns completely in the long run if and only if

$$\lim_{i \to \infty} \left[\max_{j \in N(i)} j \right] = \infty.$$

Condition for Long-Run Learning

- Society learns completely in the long run if actions $a_i^* \to \omega$ in probability (equivalent to $r_i \to \infty$)

Proposition 3

Society learns completely in the long run if and only if

$$\lim_{i \to \infty} \left[\max_{j \in N(i)} j \right] = \infty.$$

- Analog of Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s **expanding observations** property for deterministic network

Condition for Long-Run Learning

- Society learns completely in the long run if actions $a_i^* \to \omega$ in probability (equivalent to $r_i \to \infty$)

Proposition 3

Society learns completely in the long run if and only if

$$\lim_{i \to \infty} \left[\max_{j \in N(i)} j \right] = \infty.$$

- Analog of Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s **expanding observations** property for deterministic network
- Mild and clearly necessary: else for some $C < \infty$, infinitely many i cannot access the signal of any $j > C$ except their own

Condition for Long-Run Learning

- Society learns completely in the long run if actions $a_i^* \to \omega$ in probability (equivalent to $r_i \to \infty$)

Proposition 3

Society learns completely in the long run if and only if

$$\lim_{i \to \infty} \left[\max_{j \in N(i)} j \right] = \infty.$$

- Analog of Acemoglu, Dahleh, Lobel, and Ozdaglar (2011)'s **expanding observations** property for deterministic network
- Mild and clearly necessary: else for some $C < \infty$, infinitely many i cannot access the signal of any $j > C$ except their own
- Satisfied in all "reasonable" networks, not useful for ranking

Aggregative Efficiency

Definition

$\lim_{i \to \infty}(r_i/i)$ is the **aggregative efficiency** of the network

Aggregative Efficiency

Definition

$\lim_{i\to\infty}(r_i/i)$ is the **aggregative efficiency** of the network

- What fraction of signals in the entire society do individuals aggregate under social learning?

Aggregative Efficiency

Definition

$\lim_{i\to\infty}(r_i/i)$ is the **aggregative efficiency** of the network

- What fraction of signals in the entire society do individuals aggregate under social learning?
- Can have $r_i \to \infty$ but $\lim_{i\to\infty}(r_i/i)$ near 0: complete long-run learning, but get there very slowly

Aggregative Efficiency

Definition

$\lim_{i\to\infty}(r_i/i)$ is the **aggregative efficiency** of the network

- What fraction of signals in the entire society do individuals aggregate under social learning?
- Can have $r_i \to \infty$ but $\lim_{i\to\infty}(r_i/i)$ near 0: complete long-run learning, but get there very slowly
- Higher aggregative efficiency $\Rightarrow$ higher welfare if signals are not too precise and welfare function is patient

Aggregative Efficiency

Definition

$\lim_{i \to \infty}(r_i/i)$ is the **aggregative efficiency** of the network

- What fraction of signals in the entire society do individuals aggregate under social learning?
- Can have $r_i \to \infty$ but $\lim_{i \to \infty}(r_i/i)$ near 0: complete long-run learning, but get there very slowly
- Higher aggregative efficiency $\Rightarrow$ higher welfare if signals are not too precise and welfare function is patient
- Rest of the talk: compare networks for social learning by comparing their aggregative efficiency

Maximal Generations Networks

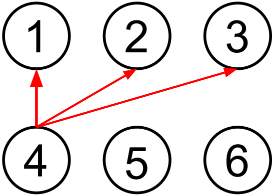
- $K \geq 1$ agents per generation

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$

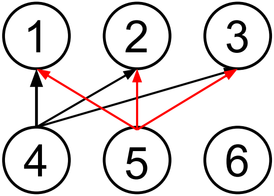
Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



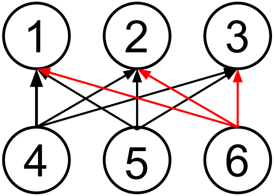
Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



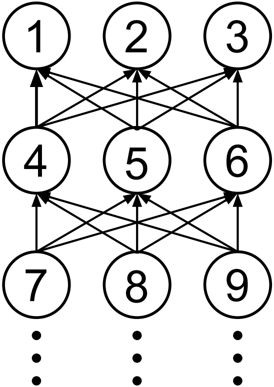
Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t - 1$



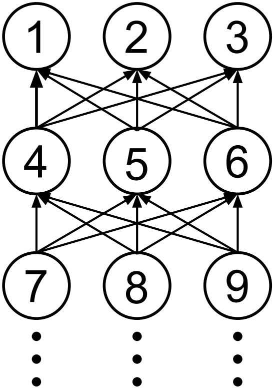
Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$

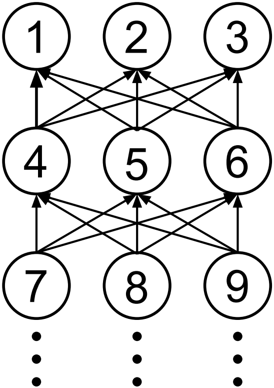


Proposition 4

In maximal generations networks:

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



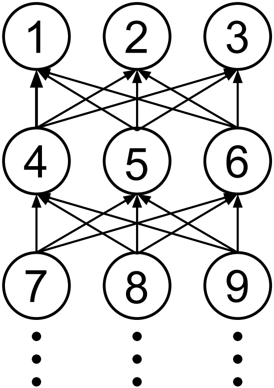
Proposition 4

In maximal generations networks:

- $\lim_{i\to\infty}(r_i/i) = \frac{(2K-1)}{K^2}$

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



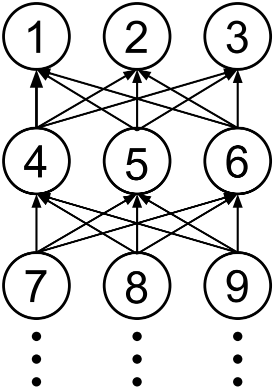
Proposition 4

In maximal generations networks:

- $\lim_{i\to\infty}(r_i/i) = \frac{(2K-1)}{K^2}$
- *In the long run, social learning aggregates...*

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



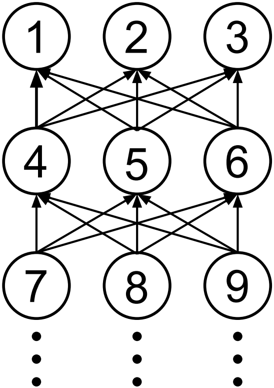
Proposition 4

In maximal generations networks:

- $\lim_{i \to \infty}(r_i/i) = \frac{(2K-1)}{K^2}$
- *In the long run, social learning aggregates...*
  - *fewer than 2 signals per generation with any K*

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



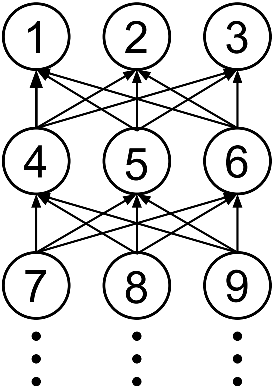
Proposition 4

In maximal generations networks:

- $\lim_{i\to\infty}(r_i/i) = \frac{(2K-1)}{K^2}$
- *In the long run, social learning aggregates...*
  - *fewer than 2 signals per generation with any K*
  - *fewer signals per agent with larger K*

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



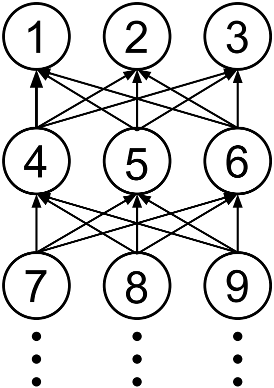
Proposition 4

In maximal generations networks:

- $\lim_{i \to \infty}(r_i/i) = \frac{(2K-1)}{K^2}$
- *In the long run, social learning aggregates...*
  - *fewer than 2 signals per generation with any K*
  - *fewer signals per agent with larger K*
  - *more signals per generation with larger K*

Maximal Generations Networks

- $K \geq 1$ agents per generation
- Agents in gen t observe all agents in gen $t-1$



Proposition 4

In maximal generations networks:

- $\lim_{i \to \infty}(r_i/i) = \frac{(2K-1)}{K^2}$
- *In the long run, social learning aggregates...*
  - *fewer than 2 signals per generation with any K*
  - *fewer signals per agent with larger K*
  - *more signals per generation with larger K*
- *After generation 2, social learning aggregates fewer than 3 signals per generation with any K*

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)
- This lower-bound not too far from the actual learning rate:

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)

- This lower-bound not too far from the actual learning rate:

$$r_i \ / \underbrace{\lceil i/K \rceil}_{\text{gen of } i} = \underbrace{\frac{(2K-1)}{K}}_{<2} + o(1)$$

(No more than **2 signals** per gen in long-run, for any K)

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)

- This lower-bound not too far from the actual learning rate:

$$r_i \, / \underbrace{\lceil i/K \rceil}_{\text{gen of } i} = \underbrace{\frac{(2K-1)}{K}}_{<2} + o(1)$$

(No more than **2 signals** per gen in long-run, for any K)

$$r_i - r_{i'} \leq 3, \qquad \text{for } i, i' \text{ in gen } t, t-1 \text{ where } t \geq 3$$

(No more than **3 signals** per gen **starting with gen 3**, for any K)

Bounds on Signals Aggregated Per Generation

- Social learning must aggregate at least 1 signal per gen (improvement by combining own signal with social obs)

- This lower-bound not too far from the actual learning rate:

$$r_i \;/ \underbrace{\lceil i/K \rceil}_{\text{gen of } i} = \underbrace{\frac{(2K-1)}{K}}_{<2} + o(1)$$

(No more than **2 signals** per gen in long-run, for any K)

$$r_i - r_{i'} \leq 3, \qquad \text{for } i, i' \text{ in gen } t, t-1 \text{ where } t \geq 3$$

(No more than **3 signals** per gen **starting with gen 3**, for any K)

- For K large, aggregate only an unboundedly small fraction of the private signals

Intuition for Inefficient Learning

- Someone in gen $t+1$ finds it hard to figure out gen t's private signals due to **info confounding**
  - Which part of neighbors' actions come from their signals, and which part from their own social observations?

Intuition for Inefficient Learning

- Someone in gen $t+1$ finds it hard to figure out gen t's private signals due to **info confounding**
  - Which part of neighbors' actions come from their signals, and which part from their own social observations?

- Must trade off overweighting gen t's private signals and underweighting gen t's common social information

Intuition for Inefficient Learning

- Someone in gen $t+1$ finds it hard to figure out gen t's private signals due to **info confounding**
  - Which part of neighbors' actions come from their signals, and which part from their own social observations?

- Must trade off overweighting gen t's private signals and underweighting gen t's common social information

- When generations are large, optimal action severely underweights private signals from generation t
  - Will see later that total weight on private signals from the generation t is close to 1 for t large
  - Without confounding, would place weight 1 on **each** private signal from generation t

Generation Size and Rate of Learning

1. Per-agent rate of learning is slower with larger generations:

Generation Size and Rate of Learning

1. Per-agent rate of learning is slower with larger generations:
    - If $K = 1$, every agent perfectly incorporates all past private signals $\Rightarrow$ fastest possible speed of social learning
    - Prop 4 says aggregative efficiency strictly decreases in K
    - Worse learning with larger K also holds numerically starting from agent $i = 16$ when comparing among $K \in \{2, 3, 4, 5\}$

Generation Size and Rate of Learning

1. Per-agent rate of learning is slower with larger generations:
   - If $K = 1$, every agent perfectly incorporates all past private signals $\Rightarrow$ fastest possible speed of social learning
   - Prop 4 says aggregative efficiency strictly decreases in K
   - Worse learning with larger K also holds numerically starting from agent $i = 16$ when comparing among $K \in \{2, 3, 4, 5\}$

2. Per-generation rate of learning is faster with larger generations:

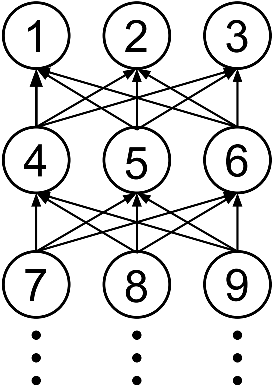
Generation Size and Rate of Learning

1. Per-agent rate of learning is slower with larger generations:
   - If $K = 1$, every agent perfectly incorporates all past private signals $\Rightarrow$ fastest possible speed of social learning
   - Prop 4 says aggregative efficiency strictly decreases in K
   - Worse learning with larger K also holds numerically starting from agent $i = 16$ when comparing among $K \in \{2, 3, 4, 5\}$

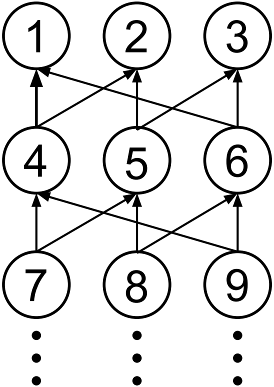
2. Per-generation rate of learning is faster with larger generations:
   - On the other hand larger $K \Rightarrow$ more learning per generation
   - But differences are small, and per-generation rate of learning is bounded above by 2 signals

Which Network Leads to Faster Learning?

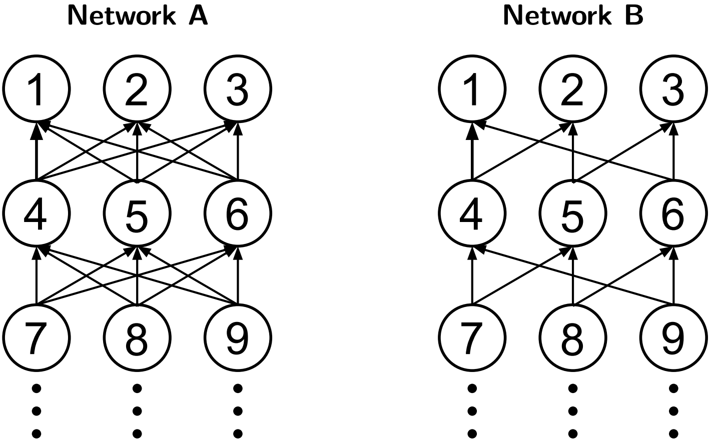
Network A



Network B

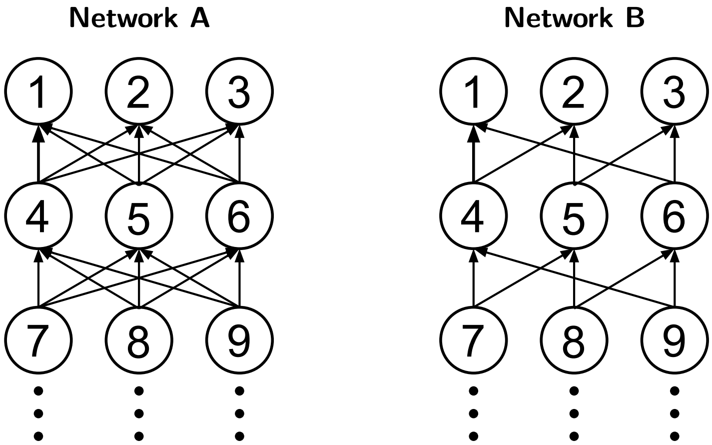


Which Network Leads to Faster Learning?



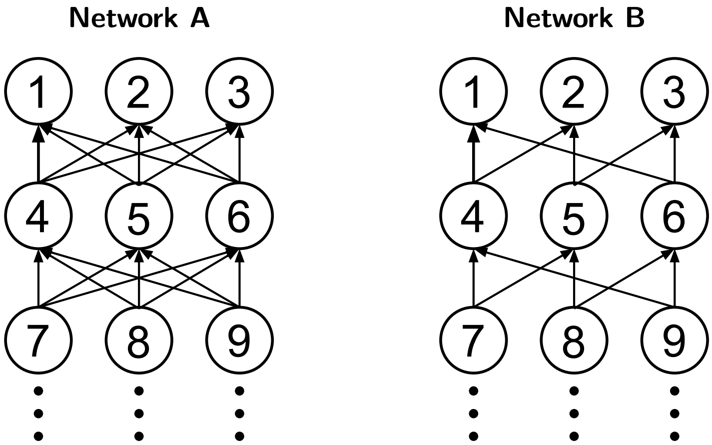
- **Network A** is the maximal generations network with $K = 3$

Which Network Leads to Faster Learning?



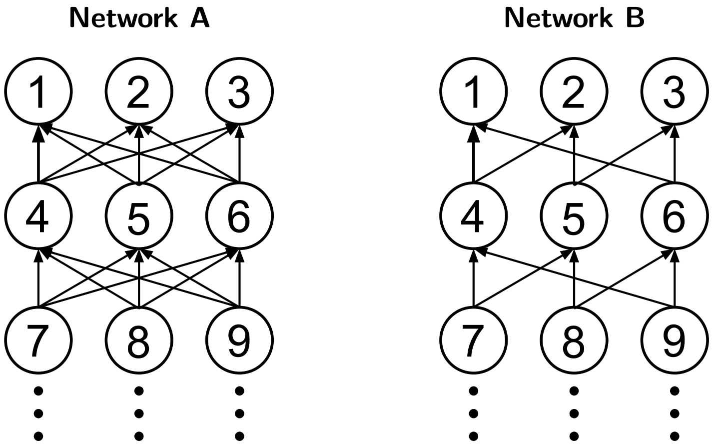
- **Network A** is the maximal generations network with $K = 3$
- **Network B** puts agents in each gen into 3 slots, $k \in \{1, 2, 3\}$.
  $k = 1$ sees 1 and 2, $k = 2$ sees 2 and 3, $k = 3$ sees 3 and 1.

Which Network Leads to Faster Learning?



- **Network A** is the maximal generations network with $K = 3$
- **Network B** puts agents in each gen into 3 slots, $k \in \{1, 2, 3\}$. $k = 1$ sees 1 and 2, $k = 2$ sees 2 and 3, $k = 3$ sees 3 and 1.
- Fewer social observations, but also less info confounding

Which Network Leads to Faster Learning?



- **Network A** is the maximal generations network with $K = 3$
- **Network B** puts agents in each gen into 3 slots, $k \in \{1, 2, 3\}$. $k = 1$ sees 1 and 2, $k = 2$ sees 2 and 3, $k = 3$ sees 3 and 1.
- Fewer social observations, but also less info confounding
- Need: aggregative efficiency on more general networks

Generations Network with Partial Observations

Generations Network with Partial Observations

- Generations network with K agents per gen

Generations Network with Partial Observations

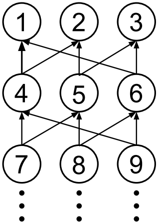
- Generations network with K agents per gen
- $\Psi_k \subseteq \{1, ..., K\}$, **observation set**, define which gen $t-1$ slots are observed by a gen t agent in slot k

Generations Network with Partial Observations

- Generations network with K agents per gen
- $\Psi_k \subseteq \{1, ..., K\}$, **observation set**, define which gen $t-1$ slots are observed by a gen t agent in slot k
- Maximal generations network is the case of $\Psi_k = \{1, ..., K\}$

Generations Network with Partial Observations

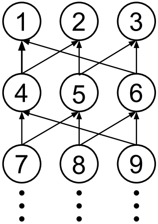
- Generations network with K agents per gen
- $\Psi_k \subseteq \{1, ..., K\}$, **observation set**, define which gen $t-1$ slots are observed by a gen t agent in slot k
- Maximal generations network is the case of $\Psi_k = \{1, ..., K\}$



$$\Psi_1 = \{1, 2\},$$
$$\Psi_2 = \{2, 3\},$$
$$\Psi_3 = \{1, 3\}.$$

Generations Network with Partial Observations

- Generations network with K agents per gen
- $\Psi_k \subseteq \{1, ..., K\}$, **observation set**, define which gen $t-1$ slots are observed by a gen t agent in slot k
- Maximal generations network is the case of $\Psi_k = \{1, ..., K\}$



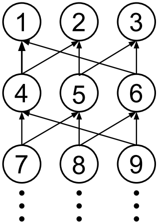
$$\Psi_1 = \{1,2\},$$
$$\Psi_2 = \{2,3\},$$
$$\Psi_3 = \{1,3\}.$$

Definition

The network is **symmetric** if all agents observe $d \geq 1$ neighbors and all pairs of agents in the same generation share c common neighbors.

Generations Network with Partial Observations

- Generations network with K agents per gen
- $\Psi_k \subseteq \{1, ..., K\}$, **observation set**, define which gen $t-1$ slots are observed by a gen t agent in slot k
- Maximal generations network is the case of $\Psi_k = \{1, ..., K\}$



$$\Psi_1 = \{1, 2\},$$
$$\Psi_2 = \{2, 3\},$$
$$\Psi_3 = \{1, 3\}.$$

Definition

The network is **symmetric** if all agents observe $d \geq 1$ neighbors and all pairs of agents in the same generation share c common neighbors.

For example, "**Network B**" is symmetric with $d = 2$, $c = 1$

Speed of Learning with Partial Observations

- Each agent observes d neighbors and pairs of agents in the same generation have c common neighbors

Theorem 1

In symmetric generations networks,

$$\lim_{i \to \infty} (r_i/i) = \left(1 + \frac{d^2 - d}{d^2 - d + c}\right) \frac{1}{K}.$$

Speed of Learning with Partial Observations

- Each agent observes d neighbors and pairs of agents in the same generation have c common neighbors

Theorem 1

In symmetric generations networks,

$$\lim_{i\to\infty}(r_i/i) = \left(1 + \frac{d^2 - d}{d^2 - d + c}\right)\frac{1}{K}.$$

- Exact expression of aggregative efficiency for a broader class of generations networks Proof

Speed of Learning with Partial Observations

- Each agent observes d neighbors and pairs of agents in the same generation have c common neighbors

Theorem 1

In symmetric generations networks,

$$\lim_{i\to\infty}(r_i/i) = \left(1 + \frac{d^2 - d}{d^2 - d + c}\right)\frac{1}{K}.$$

- Exact expression of aggregative efficiency for a broader class of generations networks Proof
- Theorem 1 shows asymptotic bound of 2 signals per gen applies to **all** such networks, strengthening Proposition 4

Speed of Learning with Partial Observations

- Each agent observes d neighbors and pairs of agents in the same generation have c common neighbors

Theorem 1

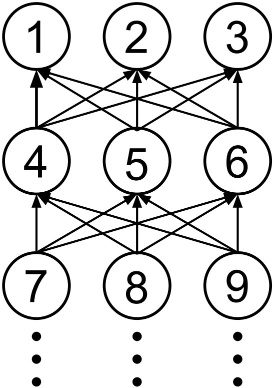
In symmetric generations networks,

$$\lim_{i \to \infty} (r_i / i) = \left(1 + \frac{d^2 - d}{d^2 - d + c}\right) \frac{1}{K}.$$

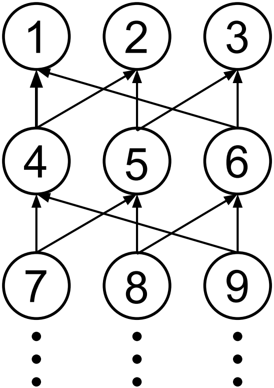
- Exact expression of aggregative efficiency for a broader class of generations networks Proof
- Theorem 1 shows asymptotic bound of 2 signals per gen applies to **all** such networks, strengthening Proposition 4
- Term in parenthesis increases in d and decreases in c — more obs speeds up rate of learning per gen but more confounding slows it down, all else equal

Which Network Leads to Faster Learning?

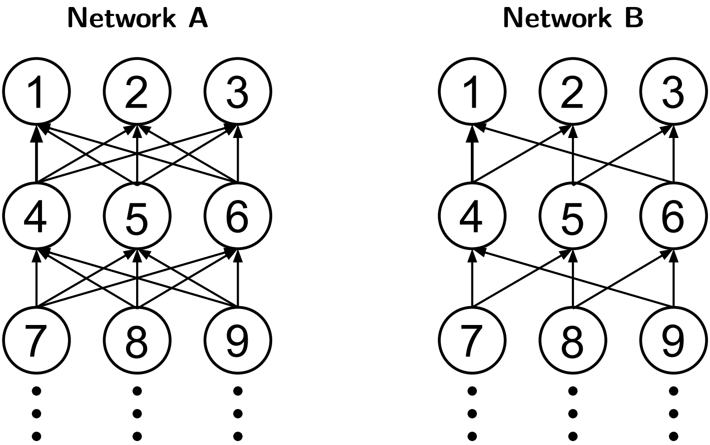
Network A



Network B

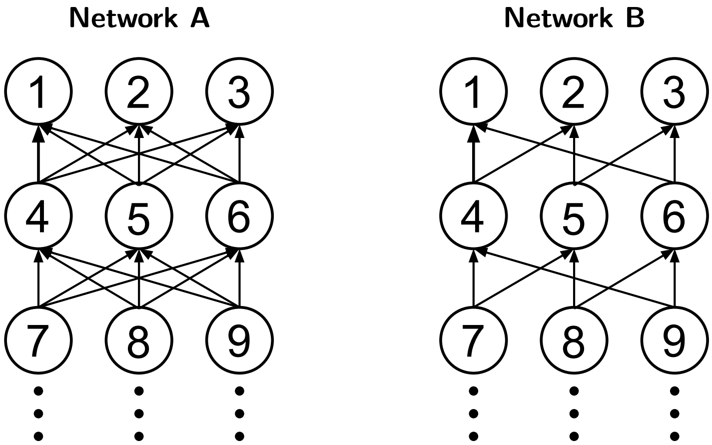


Which Network Leads to Faster Learning?



- Applying Theorem 1, aggregative efficiency is the same in **Network A** ($d = 3, c = 3$) and **Network B** ($d = 2, c = 1$)!

Which Network Leads to Faster Learning?



- Applying Theorem 1, aggregative efficiency is the same in **Network A** ($d = 3, c = 3$) and **Network B** ($d = 2, c = 1$)!
- Extra social obs exactly cancel out reduced info content of each obs Conclusion

Amplifying Private Signals

- Inefficiency can be corrected if agents place more weight on their private signals

Amplifying Private Signals

- Inefficiency can be corrected if agents place more weight on their private signals
- Suppose all agents in generations $t \geq 2$ in a maximal generations network place weight $1 + x$ on private signal

Amplifying Private Signals

- Inefficiency can be corrected if agents place more weight on their private signals
- Suppose all agents in generations $t \geq 2$ in a maximal generations network place weight $1 + x$ on private signal
- Choose weight on previous generation so there is a signal counting interpretation

Amplifying Private Signals

- Inefficiency can be corrected if agents place more weight on their private signals
- Suppose all agents in generations $t \geq 2$ in a maximal generations network place weight $1 + x$ on private signal
- Choose weight on previous generation so there is a signal counting interpretation

Proposition 6

Suppose each agent i aggregates $r_i(x)$ signals for each $x \geq 0$. Then $r_i'(0) = 0$ for any i in generation 2 while $r_i'(0) > 0$ for any i after generation 2.

Amplifying Private Signals

- Inefficiency can be corrected if agents place more weight on their private signals
- Suppose all agents in generations $t \geq 2$ in a maximal generations network place weight $1 + x$ on private signal
- Choose weight on previous generation so there is a signal counting interpretation

Proposition 6

Suppose each agent i aggregates $r_i(x)$ signals for each $x \geq 0$. Then $r_i'(0) = 0$ for any i in generation 2 while $r_i'(0) > 0$ for any i after generation 2.

- Slightly more weight on private signals is 'almost' Pareto-improving

Amplifying Private Signals

Proposition 6

Suppose each agent i aggregates $r_i(x)$ signals for each $x \geq 0$. Then $r_i'(0) = 0$ for any i in generation 2 while $r_i'(0) > 0$ for any i after generation 2.

- Slightly more weight on private signals is ‘almost’ Pareto-improving

Amplifying Private Signals

Proposition 6

Suppose each agent i aggregates $r_i(x)$ signals for each $x \geq 0$. Then $r_i'(0) = 0$ for any i in generation 2 while $r_i'(0) > 0$ for any i after generation 2.

- Slightly more weight on private signals is 'almost' Pareto-improving
- Suggests approach for policy to correct inefficiencies

Amplifying Private Signals

Proposition 6

Suppose each agent i aggregates $r_i(x)$ signals for each $x \geq 0$. Then $r_i'(0) = 0$ for any i in generation 2 while $r_i'(0) > 0$ for any i after generation 2.

- Slightly more weight on private signals is ‘almost’ Pareto-improving
- Suggests approach for policy to correct inefficiencies
- With much more weight on private signals, can attain aggregative efficiency of 1

Summary

- A tractable model of rational sequential learning that focuses on how the social network affects aggregative efficiency
- Generally, network confounds info content of neighbors' behavior and leads to info loss
- Exact aggregative efficiency in all generations networks with symmetric observation sets
- Significant info loss due to confounding: in any such network, each generation eventually aggregates no more than 2 signals
- Tractability of framework extends beyond generations networks, e.g., canonical random networks

Thank you!

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\mathrm{Var}[\lambda_i \mid \omega] = \mathrm{Var}_t$ for all i in generation t

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\text{Var}[\lambda_i \mid \omega] = \text{Var}_t$ for all i in generation t
- $\text{Cov}[\lambda_i, \lambda_{i'} \mid \omega] = \text{Cov}_t$ for all $i \neq i'$ in generation t (same covariance across all pairs)

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\text{Var}[\lambda_i \mid \omega] = \text{Var}_t$ for all i in generation t
- $\text{Cov}[\lambda_i, \lambda_{i'} \mid \omega] = \text{Cov}_t$ for all $i \neq i'$ in generation t (same covariance across all pairs)
- Each i in generation t puts weight β_t on each $j \in N(i)$

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\text{Var}[\lambda_i \mid \omega] = \text{Var}_t$ for all i in generation t
- $\text{Cov}[\lambda_i, \lambda_{i'} \mid \omega] = \text{Cov}_t$ for all $i \neq i'$ in generation t (same covariance across all pairs)
- Each i in generation t puts weight β_t on each $j \in N(i)$

From signal-counting interpretation, for i in generation t, r_i is proportional to Var_t, so can compute Var_t

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\text{Var}[\lambda_i \mid \omega] = \text{Var}_t$ for all i in generation t
- $\text{Cov}[\lambda_i, \lambda_{i'} \mid \omega] = \text{Cov}_t$ for all $i \neq i'$ in generation t (same covariance across all pairs)
- Each i in generation t puts weight β_t on each $j \in N(i)$

From signal-counting interpretation, for i in generation t, r_i is proportional to Var_t, so can compute Var_t

The optimal action from Prop 1 implies expressions for Var_{t+1} and Cov_{t+1} in terms of $\text{Var}_t, \text{Cov}_t$, and β_{t+1}

Proof Sketch of Theorem 1

Using symmetry, can show there exist numbers Var_t, Cov_t, β_t s.t.

- $\mathrm{Var}[\lambda_i \mid \omega] = \mathrm{Var}_t$ for all i in generation t
- $\mathrm{Cov}[\lambda_i, \lambda_{i'} \mid \omega] = \mathrm{Cov}_t$ for all $i \neq i'$ in generation t (same covariance across all pairs)
- Each i in generation t puts weight β_t on each $j \in N(i)$

From signal-counting interpretation, for i in generation t, r_i is proportional to Var_t, so can compute Var_t

The optimal action from Prop 1 implies expressions for Var_{t+1} and Cov_{t+1} in terms of $\mathrm{Var}_t, \mathrm{Cov}_t$, and β_{t+1}

Will sketch proof of key lemma:

Lemma 1

$$\lim_{t \to \infty} \beta_t = 1/d.$$

$\beta_t \to 1/d$ Using a Mixing Argument

- Showing $\beta_t \to 1/d$ amounts to showing $\text{Corr}(\ell_i, \ell_{i'} \mid \omega) \to 1$ for $i \neq i'$ in generation t, as $t \to \infty$
  - Observe almost perfectly correlated actions $\approx$ observe only one action $\Rightarrow$ total weight is close to one

$\beta_t \to 1/d$ Using a Mixing Argument

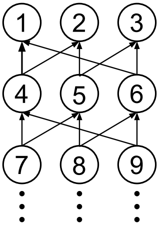
- Showing $\beta_t \to 1/d$ amounts to showing $\mathrm{Corr}(\ell_i, \ell_{i'} \mid \omega) \to 1$ for $i \neq i'$ in generation t, as $t \to \infty$
  - Observe almost perfectly correlated actions $\approx$ observe only one action $\Rightarrow$ total weight is close to one
- Weight that i puts on j's action is proportional to number of paths in the network from i to j

$\beta_t \to 1/d$ Using a Mixing Argument

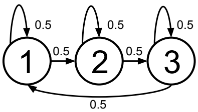
- Showing $\beta_t \to 1/d$ amounts to showing $\text{Corr}(\ell_i, \ell_{i'} \mid \omega) \to 1$ for $i \neq i'$ in generation t, as $t \to \infty$
  - Observe almost perfectly correlated actions $\approx$ observe only one action $\Rightarrow$ total weight is close to one
- Weight that i puts on j's action is proportional to number of paths in the network from i to j
- Equivalent to a Markov chain with state space $\{1, ..., K\}$ and transitions to each neighbor with probability $1/d$

$\beta_t \to 1/d$ Using a Mixing Argument

- Showing $\beta_t \to 1/d$ amounts to showing $\text{Corr}(\ell_i, \ell_{i'} \mid \omega) \to 1$ for $i \neq i'$ in generation t, as $t \to \infty$
  - Observe almost perfectly correlated actions $\approx$ observe only one action $\Rightarrow$ total weight is close to one
- Weight that i puts on j's action is proportional to number of paths in the network from i to j
- Equivalent to a Markov chain with state space $\{1, ..., K\}$ and transitions to each neighbor with probability $1/d$

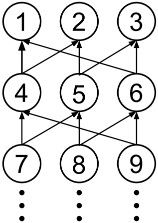


induces the stochastic process

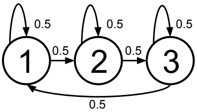


$\beta_t \to 1/d$ Using a Mixing Argument

- Showing $\beta_t \to 1/d$ amounts to showing $\text{Corr}(\ell_i, \ell_{i'} \mid \omega) \to 1$ for $i \neq i'$ in generation t, as $t \to \infty$
  - Observe almost perfectly correlated actions $\approx$ observe only one action $\Rightarrow$ total weight is close to one
- Weight that i puts on j's action is proportional to number of paths in the network from i to j
- Equivalent to a Markov chain with state space $\{1, ..., K\}$ and transitions to each neighbor with probability $1/d$



induces the stochastic process



- Markov chain mixing theorem implies steady-state distribution that does not depend on starting state $\Rightarrow$ number of paths to distant j almost independent of i Back

Calculating Aggregative Efficiency

- Expressions for Var_{t+1} and Cov_{t+1} give

$$\mathsf{Var}_{t+1} - \mathsf{Cov}_{t+1} = \frac{4}{\sigma^2} + \beta_{t+1}^2 (d - c)(\mathsf{Var}_t - \mathsf{Cov}_t)$$

Calculating Aggregative Efficiency

- Expressions for Var_{t+1} and Cov_{t+1} give

$$\text{Var}_{t+1} - \text{Cov}_{t+1} = \frac{4}{\sigma^2} + \beta_{t+1}^2 (d - c)(\text{Var}_t - \text{Cov}_t)$$

- For β close enough to $1/d$,

$$\phi(x) = \frac{4}{\sigma^2} + \beta^2 (d - c)x$$

  is a contraction mapping
  - Difference $\text{Var}_t - \text{Cov}_t$ converges to the unique fixed point with $\beta = 1/d$

Calculating Aggregative Efficiency

- Expressions for Var_{t+1} and Cov_{t+1} give

$$\text{Var}_{t+1} - \text{Cov}_{t+1} = \frac{4}{\sigma^2} + \beta_{t+1}^2 (d - c)(\text{Var}_t - \text{Cov}_t)$$

- For β close enough to $1/d$,

$$\phi(x) = \frac{4}{\sigma^2} + \beta^2 (d - c) x$$

  is a contraction mapping
  - Difference $\text{Var}_t - \text{Cov}_t$ converges to the unique fixed point with $\beta = 1/d$

- From this fixed point, can compute the growth rate of Var_t and therefore the aggregative efficiency Back

Finite Populations

Proposition 7

Let $\epsilon > 0$. There exists a constant $C > 0$ such that for any symmetric generations network and any generation $t \geq CK\log(K)$, at most $K\lim_i(r_i/i) + \epsilon$ signals are aggregated between generations t and $t+1$.

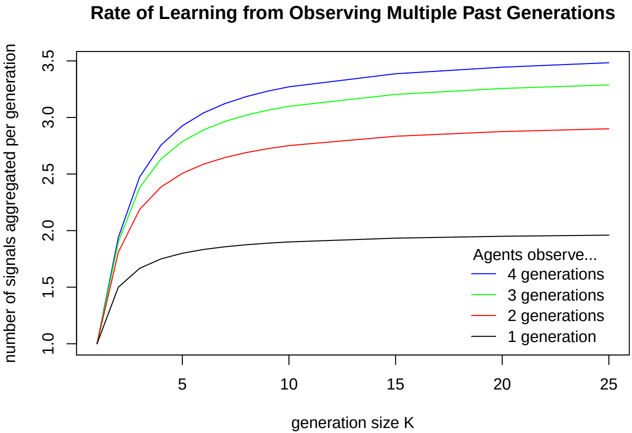
- Gives an upper bound on how long it takes for Theorem 1 to apply

Simulation: Observing Multiple Past Generations

Each agent observes all predecessors from past $\tau \geq 1$ generations

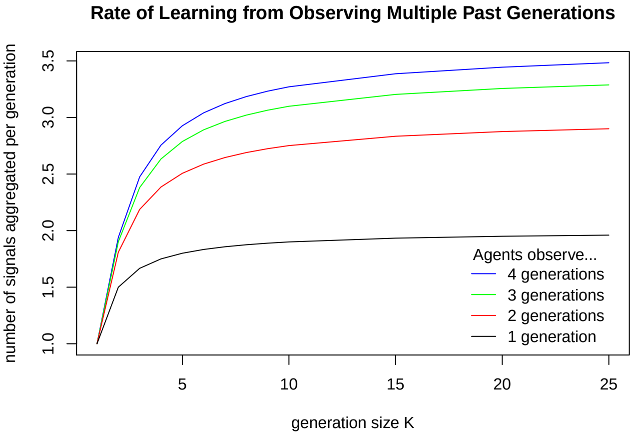
Simulation: Observing Multiple Past Generations

Each agent observes all predecessors from past $\tau \geq 1$ generations



Simulation: Observing Multiple Past Generations

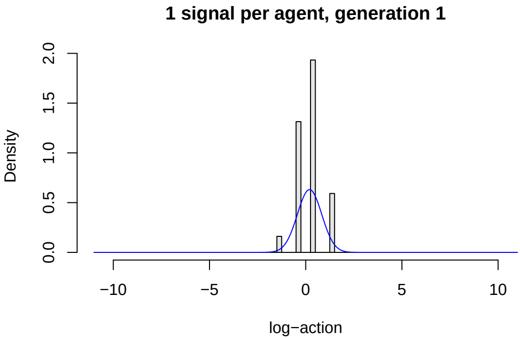
Each agent observes all predecessors from past $\tau \geq 1$ generations



- Limited improvement in aggregative efficiency: removes some confounds but creates new ones

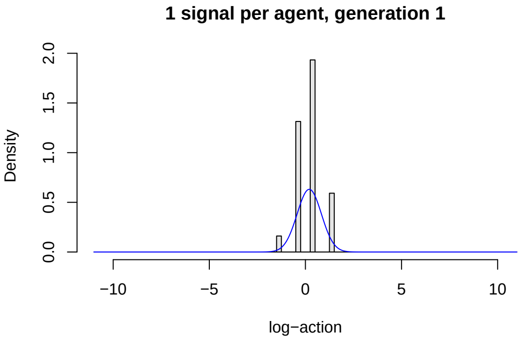
Simulation: General Signal Structures

- Each signal is finitely supported



Simulation: General Signal Structures

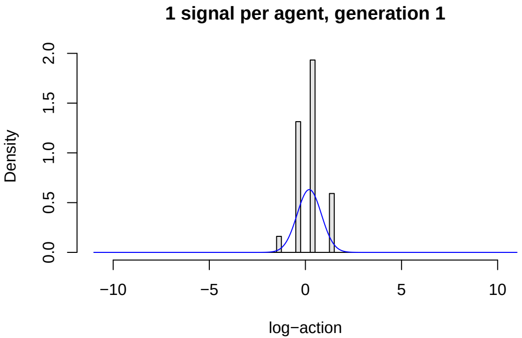
- Each signal is finitely supported



- Each agent has not 1, but *n* conditionally i.i.d. signals

Simulation: General Signal Structures

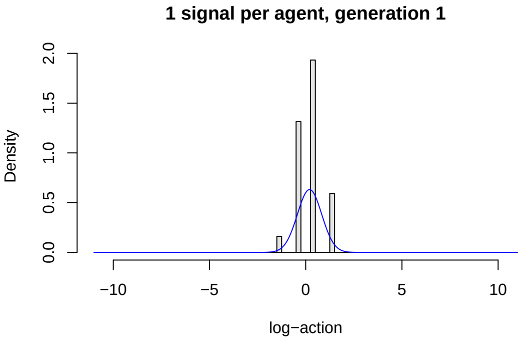
- Each signal is finitely supported



- Each agent has not 1, but *n* conditionally i.i.d. signals
- Think of agents who gather info over a period of time

Simulation: General Signal Structures

- Each signal is finitely supported



- Each agent has not 1, but n conditionally i.i.d. signals
- Think of agents who gather info over a period of time
- Increase n and scale down informativeness of each signal, fixing mean and SD of private log-belief (based on all n private signals) to match the Gaussian case