Social Learning in a Dynamic Environment

Krishna Dasaratha (Yale)
Benjamin Golub (Northwestern)
Nir Hak (Uber)

July 21, 2021
6th World Congress of the Game Theory Society

Question

External state, changing over time.

Question

External state, changing over time.

Each short-lived agent learns about it using:

- private information (private signal with individual-specific precision);
- past estimates of some of the others (e.g., by talking).

Question

External state, changing over time.

Each short-lived agent learns about it using:

- private information (private signal with individual-specific precision);
- past estimates of some of the others (e.g., by talking).

Can **decentralized communication** among **short-lived individuals** aggregate information quickly, keeping up with the changing environment?

Question

External state, changing over time.

Each short-lived agent learns about it using:

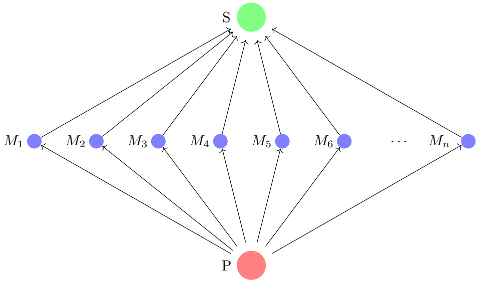
- private information (private signal with individual-specific precision);
- past estimates of some of the others (e.g., by talking).

Can **decentralized communication** among **short-lived individuals** aggregate information quickly, keeping up with the changing environment?

Key idea: Sufficient heterogeneity in signal distributions enables good filtering by Bayesians – whereas naive agents do very badly with or without it.

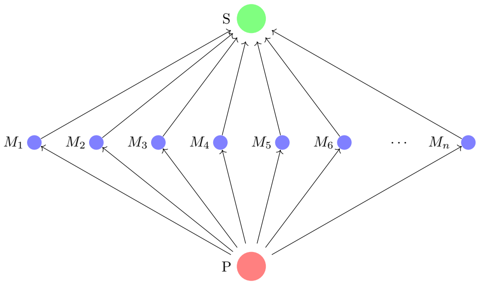
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic.



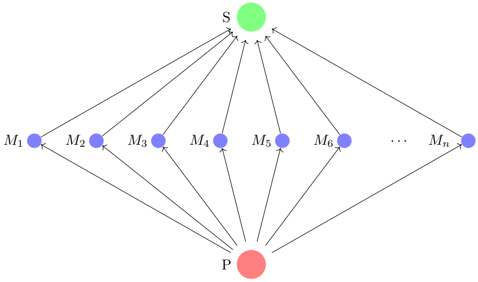
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes."



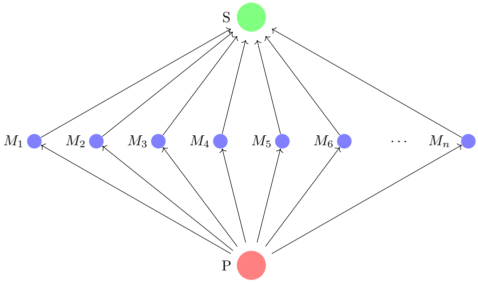
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian.



An illustrative example

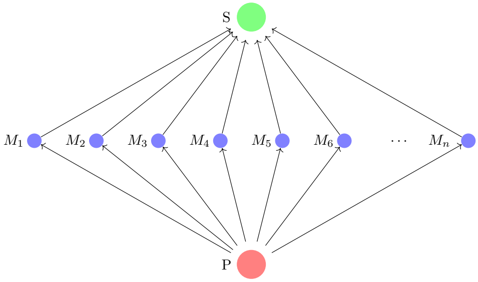
Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:



An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

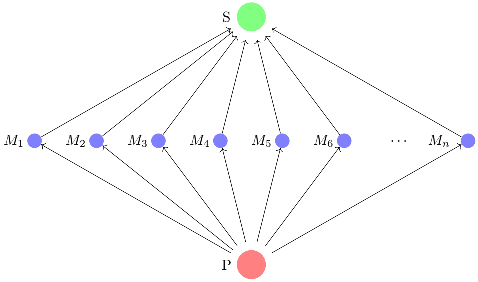
1. S sees $s_{\mathrm{S}} = \theta + \eta_{\mathrm{S}}$; takes an action equal to posterior mean of θ;



An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

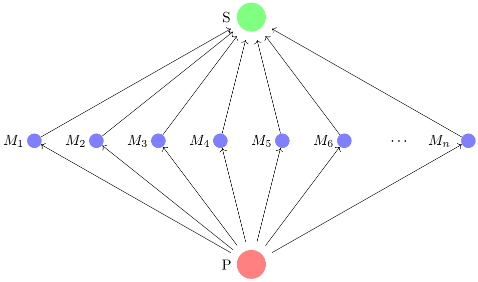
1. S sees $s_{\text{S}} = \theta + \eta_{\text{S}}$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;



An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_{\mathrm{S}} = \theta + \eta_{\mathrm{S}}$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

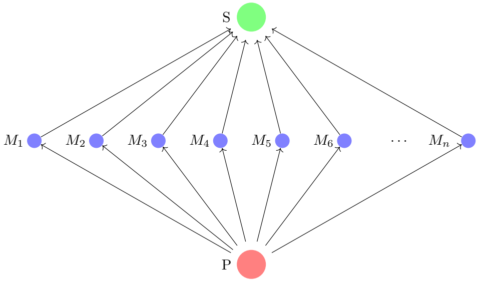


An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_{\mathrm{S}} = \theta + \eta_{\mathrm{S}}$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

All $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ independent.



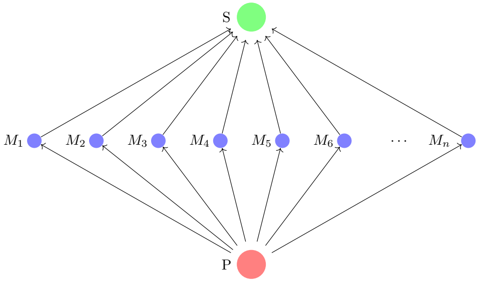
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_{\mathrm{S}} = \theta + \eta_{\mathrm{S}}$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

All $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ independent.

$a_i = w_i s_i + (1 - w_i) s_{\mathrm{S}}$



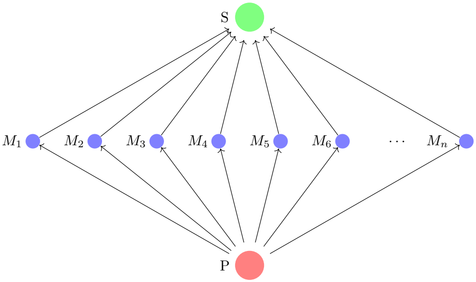
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_{\mathrm{S}} = \theta + \eta_{\mathrm{S}}$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

All $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ independent.

$a_i = w_i s_i + (1 - w_i) s_{\mathrm{S}}$



If σ_i^2 all equal, then w_i all equal $\Rightarrow$ P **cannot filter** out η_{S}.

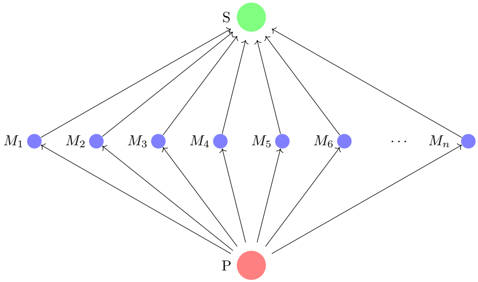
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_S = \theta + \eta_S$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

All $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ independent.

$a_i = w_i s_i + (1 - w_i) s_S$



If σ_i^2 all equal, then w_i all equal $\Rightarrow$ P **cannot filter** out η_S.

If σ_i^2 low for half the M_i and high for other half then...

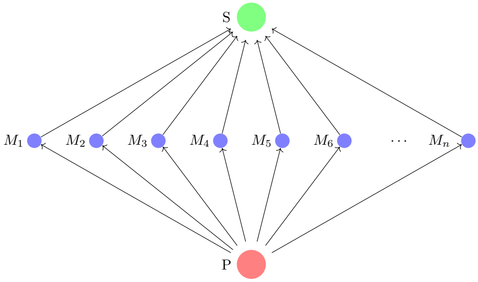
An illustrative example

Agents: single **S**ource, many **M**edia outlets, and a **P**ublic. Directed link means "observes." All r.v.'s will be Gaussian. Timing:

1. S sees $s_S = \theta + \eta_S$; takes an action equal to posterior mean of θ;
2. each M_i observes this and $s_i = \theta + \eta_i$; takes an action equal to posterior mean of θ;
3. P sees only these actions, estimates θ.

All $\eta_i \sim \mathcal{N}(0, \sigma_i^2)$ independent.

$a_i = w_i s_i + (1 - w_i) s_S$



If σ_i^2 all equal, then w_i all equal $\Rightarrow$ P **cannot filter** out η_S.

If σ_i^2 low for half the M_i and high for other half then...

P learns $\approx w\theta + (1 - w)s_S$ for two distinct values of w $\Rightarrow$ **learns** θ

Model

- Time is $t \in \mathcal{T} = \mathbb{Z}_{\geq 0}$ or $\mathbb{Z}$
- State θ evolves according to an AR(1) process:

$$\theta_t = \rho\theta_{t-1} + \nu_t,$$

where $0 < \rho \leq 1$ is a constant and $\nu_t \sim \mathcal{N}(0, 1)$ is normal

Model

- Time is $t \in \mathcal{T} = \mathbb{Z}_{\geq 0}$ or $\mathbb{Z}$
- State θ evolves according to an AR(1) process:

$$\theta_t = \rho\theta_{t-1} + \nu_t,$$

  where $0 < \rho \leq 1$ is a constant and $\nu_t \sim \mathcal{N}(0, 1)$ is normal
- Agents get signals about the state and also observe the actions of some others

Model

- Time is $t \in \mathcal{T} = \mathbb{Z}_{\geq 0}$ or $\mathbb{Z}$
- State θ evolves according to an AR(1) process:

$$\theta_t = \rho\theta_{t-1} + \nu_t,$$

where $0 < \rho \leq 1$ is a constant and $\nu_t \sim \mathcal{N}(0,1)$ is normal

- Agents get signals about the state and also observe the actions of some others
- There is a (directed or undirected) network of n nodes

Model

- Time is $t \in \mathcal{T} = \mathbb{Z}_{\geq 0}$ or $\mathbb{Z}$

- State θ evolves according to an AR(1) process:

$$\theta_t = \rho\theta_{t-1} + \nu_t,$$

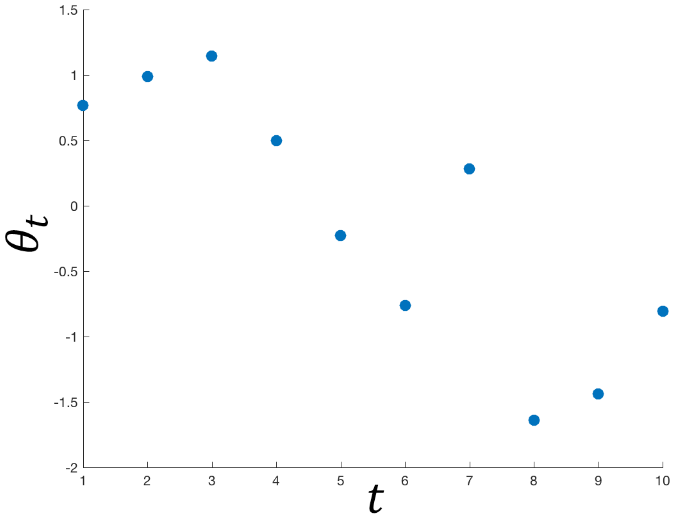
  where $0 < \rho \leq 1$ is a constant and $\nu_t \sim \mathcal{N}(0,1)$ is normal

- Agents get signals about the state and also observe the actions of some others

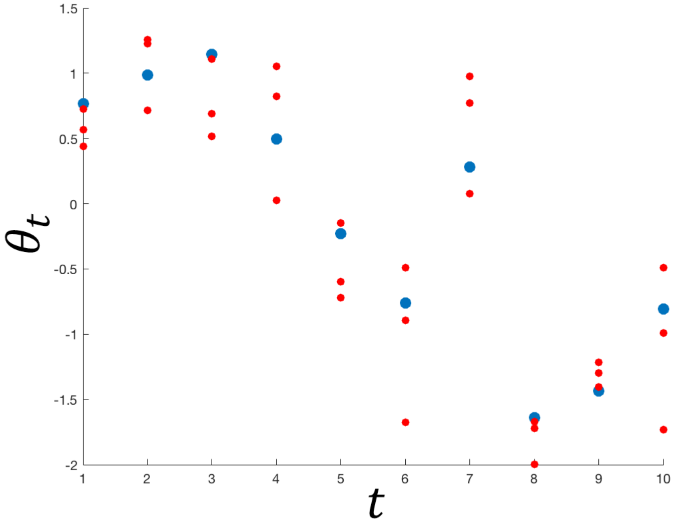
- There is a (directed or undirected) network of n nodes

- For each agent i, denote by N_i the neighbors of i (informally: people that i can observe)

Model



Model



- state $\theta_t = \rho\theta_{t-1} + \nu_t$; network with neighborhoods N_i;

- state $\theta_t = \rho\theta_{t-1} + \nu_t$; network with neighborhoods N_i;
- OLG social learning model (cf. Banerjee & Fudenberg 2004, Wolitzky 2018): agent (i, t) is born at $t - m$, observes:

  the estimates $a_{j,t-1}, \ldots, a_{j,t-m}$ of all neighbors $j \in N_i$ (including at own node) ;

  a private signal $s_{i,t} = \theta_t + \eta_{i,t}$ where $\eta_{i,t} \sim \mathcal{N}(0, \sigma_i^2)$

  all noise terms ν_t and $\eta_{i,t}$ are independent.

- state $\theta_t = \rho\theta_{t-1} + \nu_t$; network with neighborhoods N_i;
- OLG social learning model (cf. Banerjee & Fudenberg 2004, Wolitzky 2018): agent (i, t) is born at $t - m$, observes:

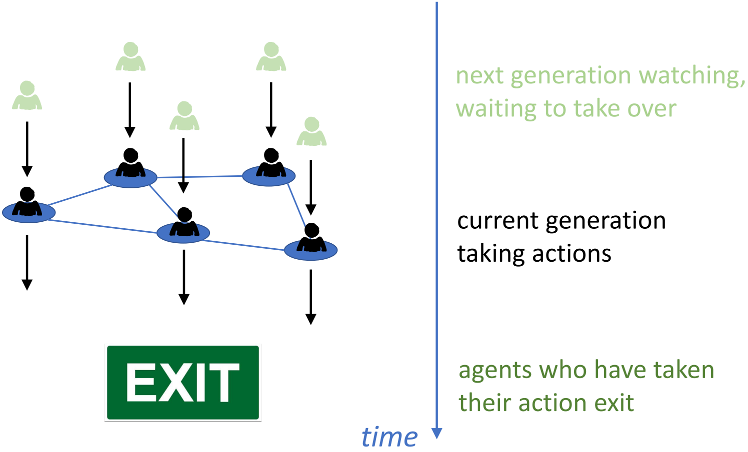
  the estimates $a_{j,t-1}, \ldots, a_{j,t-m}$ of all neighbors $j \in N_i$ (including at own node) ;

  a private signal $s_{i,t} = \theta_t + \eta_{i,t}$ where $\eta_{i,t} \sim \mathcal{N}(0, \sigma_i^2)$

  all noise terms ν_t and $\eta_{i,t}$ are independent.

- Makes an estimate $a_{i,t}$ to maximize the expectation of $-(a_{i,t} - \theta_t)^2$ so

$$a_{i,t} = \mathbb{E}\left[\theta_t \mid i\text{'s observations}\right].$$


next generation watching,
waiting to take over
current generation
taking actions
EXIT
agents who have taken
their action exit
time

Context

Old question: when do decentralized systems aggregate information well enough to facilitate efficient adaptation?

cf. Hayek, 1945: "the economic problem of society is mainly one of rapid adaptation to changes"; central in RBC models, e.g., Molavi 18

Context

Old question: when do decentralized systems aggregate information well enough to facilitate efficient adaptation?

> cf. Hayek, 1945: "the economic problem of society is mainly one of rapid adaptation to changes"; central in RBC models, e.g., Molavi 18

Despite huge social learning literature, surprisingly little on a **moving target** and the question of **responsiveness**.

> **Sequential** soc. learning: Moscarini, Ottaviani, and Smith 98

Context

Old question: when do decentralized systems aggregate information well enough to facilitate efficient adaptation?

> cf. Hayek, 1945: "the economic problem of society is mainly one of rapid adaptation to changes"; central in RBC models, e.g., Molavi 18

Despite huge social learning literature, surprisingly little on a **moving target** and the question of **responsiveness**.

> **Sequential** soc. learning: Moscarini, Ottaviani, and Smith 98

Moving states and network – distributed Kalman filtering:

- Olfati-Saber 07; Shahrampour, Rakhlin and Jadbabaie 13; Frongillo, Schoenebeck, and Tamuz 11

Very recently: Kabos and Meyer (WP 21), Levy, Marcin Peski, Vieille (WP 21)

Main contributions in context

1. **Methodological: stationary model of learning in a network about a dynamic state.**

Main contributions in context

1. **Methodological: stationary model of learning in a network about a dynamic state.**

2. **Substantive: Conditions for fast aggregation.**
    - Bayesians can use diversity of information endowments to learn (and need it).
    - Naive agents are much worse off than in a fixed-state model.

Existence of a stationary equilibrium

Definition: A **stationary equilibrium in linear strategies** is one where strategies are linear and time-invariant.

Existence of a stationary equilibrium

Definition: A **stationary equilibrium in linear strategies** is one where strategies are linear and time-invariant.

Proposition

There exists a stationary equilibrium in linear strategies. proof

Existence of a stationary equilibrium

Definition: A **stationary equilibrium in linear strategies** is one where strategies are linear and time-invariant.

Proposition

There exists a stationary equilibrium in linear strategies. proof

- As in DeGroot learning, at our equilibrium agents add up their observations with constant weights.

Existence of a stationary equilibrium

Definition: A **stationary equilibrium in linear strategies** is one where strategies are linear and time-invariant.

Proposition

There exists a stationary equilibrium in linear strategies. proof

- As in DeGroot learning, at our equilibrium agents add up their observations with constant weights.
- Studied in engineering literature mainly with exogenous weights; we consider Bayesian equilibrium.
- Can bring your own behavioral model of learning, define analogous fixed point.

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

If $-i$ play admissible linear strategies, the vector

$$\mathbf{\Delta}_t = \left(a_{i,t} - \theta_t\right)_{1 \le i \le n}$$

of last-period agents' errors is multivariate normal (we take $m = 1$)

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

If $-i$ play admissible linear strategies, the vector

$$\mathbf{\Delta}_t = \Big(a_{i,t} - \theta_t\Big)_{1 \leq i \leq n}$$

of last-period agents' errors is multivariate normal (we take $m = 1$)

Let $\boldsymbol{V}_t$ be the covariance matrix of $\mathbf{\Delta}_t$: records how **accurate** and how **correlated** observations are.

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

If $-i$ play admissible linear strategies, the vector

$$\boldsymbol{\Delta}_t = \Big(a_{i,t} - \theta_t\Big)_{1 \le i \le n}$$

of last-period agents' errors is multivariate normal (we take $m = 1$)

Let $\boldsymbol{V}_t$ be the covariance matrix of $\boldsymbol{\Delta}_t$: records how **accurate** and how **correlated** observations are.

Writing

$$a_{i,t+1} = \sum_j w_{ij,t} a_{j,t},$$

$\boldsymbol{V}_t$ determines how to weight others' past actions $(w_{ij,t})$. Also, $\boldsymbol{V}_t$ says how those actions are distributed. $\Rightarrow$ determines $\boldsymbol{V}_{t+1}$.

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

If $-i$ play admissible linear strategies, the vector

$$\boldsymbol{\Delta}_t = \Big(a_{i,t} - \theta_t\Big)_{1 \le i \le n}$$

of last-period agents' errors is multivariate normal (we take $m = 1$)

Let $\boldsymbol{V}_t$ be the covariance matrix of $\boldsymbol{\Delta}_t$: records how **accurate** and how **correlated** observations are.

Writing

$$a_{i,t+1} = \sum_j w_{ij,t} a_{j,t},$$

$\boldsymbol{V}_t$ determines how to weight others' past actions $(w_{ij,t})$. Also, $\boldsymbol{V}_t$ says how those actions are distributed. $\Rightarrow$ determines $\boldsymbol{V}_{t+1}$.

Thus, can define $\Phi(\boldsymbol{V}_t)$, a (deterministic) map $\boldsymbol{V}_t \mapsto \boldsymbol{V}_{t+1}$.

One-step linkage between distributions of behavior

Agents at $t+1$ want to estimate θ_t, which is a sufficient statistic for past information, from observed actions

If $-i$ play admissible linear strategies, the vector

$$\boldsymbol{\Delta}_t = \Big(a_{i,t} - \theta_t\Big)_{1 \le i \le n}$$

of last-period agents' errors is multivariate normal (we take $m = 1$)

Let $\boldsymbol{V}_t$ be the covariance matrix of $\boldsymbol{\Delta}_t$: records how **accurate** and how **correlated** observations are.

Writing

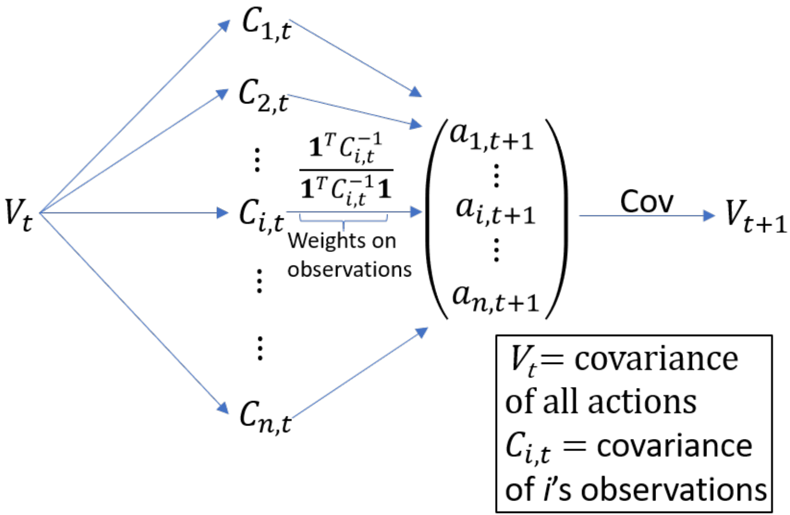
$$a_{i,t+1} = \sum_j w_{ij,t} a_{j,t},$$

$\boldsymbol{V}_t$ determines how to weight others' past actions $(w_{ij,t})$. Also, $\boldsymbol{V}_t$ says how those actions are distributed. $\Rightarrow$ determines $\boldsymbol{V}_{t+1}$.

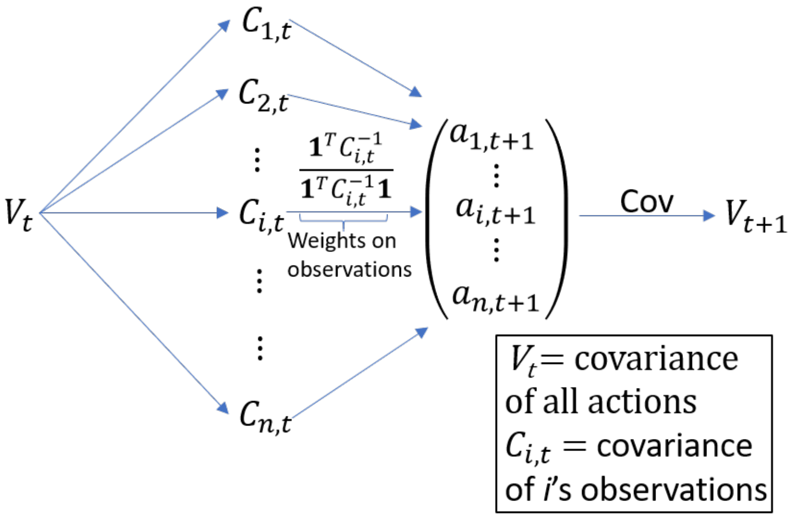
Thus, can define $\Phi(\boldsymbol{V}_t)$, a (deterministic) map $\boldsymbol{V}_t \mapsto \boldsymbol{V}_{t+1}$.

A fixed point of Φ; exists by Brouwer (define compact C s.t. $\boldsymbol{V}_t \in C$).

Distribution of past determines distribution of present



Distribution of past determines distribution of present



Putting these together gives the map Φ. The behavior of the map Φ is key to understanding learning outcomes over time.

Homogeneous signals: Can be far from benchmark

- Learning very well: learn θ_{t-1} exactly (it's the most you could hope to learn from social information).
  Learning well: within ϵ of this, in payoffs.

Homogeneous signals: Can be far from benchmark

- Learning very well: learn θ_{t-1} exactly (it's the most you could hope to learn from social information).
  Learning well: within ϵ of this, in payoffs.

- Results:
  1. Even for Bayesians, diversity of information can be necessary to learn well.
  2. Diversity in a suitable sense is sufficient for Bayesians to learn well.
  3. Naive agents cannot do well even with diversity.

Learning benchmark: What does it mean to learn well?

- i at $t+1$ achieves the **perfect aggregation** benchmark if he learns as well as if he knows θ_t and own private signal

  eq'm action has variance $(\sigma_i^{-2}+1)^{-1}$

Learning benchmark: What does it mean to learn well?

- i at $t+1$ achieves the **perfect aggregation** benchmark if he learns as well as if he knows θ_t and own private signal

  eq'm action has variance $(\sigma_i^{-2}+1)^{-1}$

  a best case; cannot be achieved exactly.

Learning benchmark: What does it mean to learn well?

- i at $t+1$ achieves the **perfect aggregation** benchmark if he learns as well as if he knows θ_t and own private signal

    eq'm action has variance $(\sigma_i^{-2}+1)^{-1}$

    a best case; cannot be achieved exactly.

- We ask whether agents achieve this benchmark in large networks with many observations
    - Hope: many signals are helpful for learning.
    - Challenge: neighbors incorporate realizations i doesn't know $\Rightarrow$ correlated noise $\Rightarrow$ LLN doesn't apply.

Learning benchmark: What does it mean to learn well?

- i at $t+1$ achieves the **perfect aggregation** benchmark if he learns as well as if he knows θ_t and own private signal

  eq'm action has variance $(\sigma_i^{-2}+1)^{-1}$

  a best case; cannot be achieved exactly.

- We ask whether agents achieve this benchmark in large networks with many observations
  - Hope: many signals are helpful for learning.
  - Challenge: neighbors incorporate realizations i doesn't know $\Rightarrow$ correlated noise $\Rightarrow$ LLN doesn't apply.

- Results:
  1. Even for Bayesians, diversity of information can be necessary to learn well.
  2. Diversity in a suitable sense is sufficient for Bayesians to learn well.
  3. Naive agents cannot do well even with diversity.

Homogeneous signals: Can be far from benchmark

- **Result 1:** Learning well is not guaranteed, despite rational agents and abundant information.

Homogeneous signals: Can be far from benchmark

- **Result 1:** Learning well is not guaranteed, despite rational agents and abundant information.
- Suppose all agents have the same private signal variance σ^2.

Homogeneous signals: Can be far from benchmark

- **Result 1:** Learning well is not guaranteed, despite rational agents and abundant information.
- Suppose all agents have the same private signal variance σ^2.
- Take, for example, the complete graph and $m = 1$ (generalizations in paper).

Proposition

There is a constant $c > 0$ such that for the complete graph on n nodes

- there is a unique stationary linear equilibrium;
- and in it all agents have variance exceeding the perfect aggregation benchmark by at least c.

Homogeneous signals: Can be far from benchmark

- **Result 1:** Learning well is not guaranteed, despite rational agents and abundant information.
- Suppose all agents have the same private signal variance σ^2.
- Take, for example, the complete graph and $m = 1$ (generalizations in paper).

Proposition

There is a constant $c > 0$ such that for the complete graph on n nodes

- there is a unique stationary linear equilibrium;
- and in it all agents have variance exceeding the perfect aggregation benchmark by at least c.

- Without signal heterogeneity, agents learn imperfectly.
- Same result in graphs with *symmetric neighbors*, Erdos-Renyi random graph.

Homogeneous signals: Can be far from benchmark

Consider an agent at time $t+1$.

All observed actions $a_{i,t}$ are exchangeable, so the period $t+1$ social signal is just the (unweighted) average of period t actions $a_{i,t}$

Homogeneous signals: Can be far from benchmark

Consider an agent at time $t+1$.

All observed actions $a_{i,t}$ are exchangeable, so the period $t+1$ social signal is just the (unweighted) average of period t actions $a_{i,t}$

Because these actions $a_{i,t}$ in turn place substantial weight on period $t-1$ actions, correlated error from the change in state $\nu_t = \theta_t - \rho\theta_{t-1}$ prevents perfect aggregation

More generally: dimensionality of relevant state updates exceeds identification power afforded by your social neighborhood.

Homogeneous signals: Can be far from benchmark

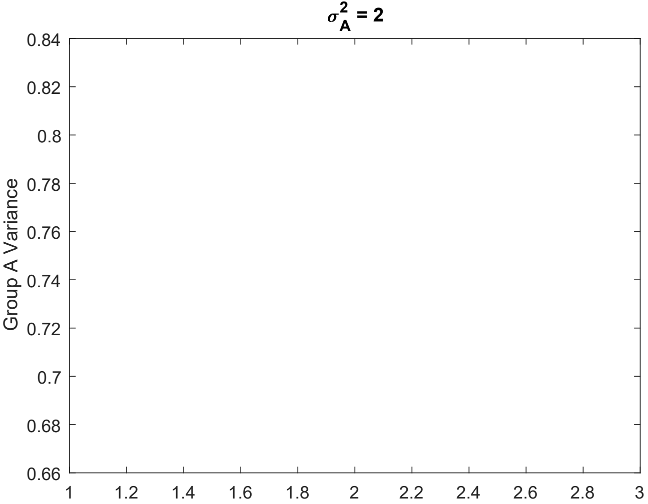
Consider an agent at time $t+1$.

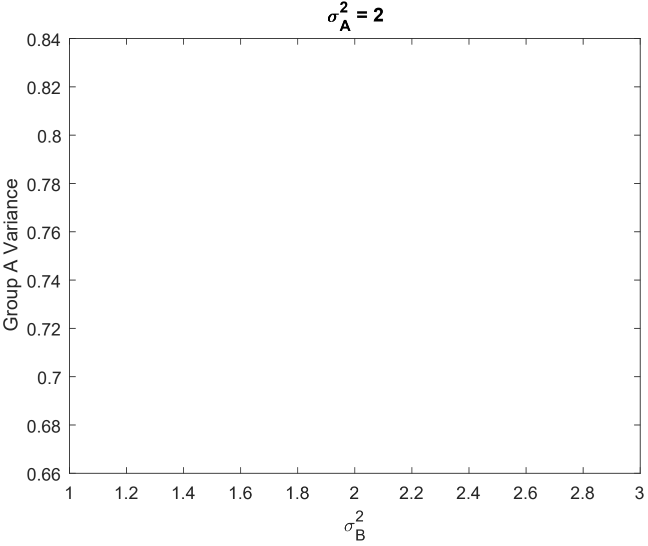
All observed actions $a_{i,t}$ are exchangeable, so the period $t+1$ social signal is just the (unweighted) average of period t actions $a_{i,t}$

Because these actions $a_{i,t}$ in turn place substantial weight on period $t-1$ actions, correlated error from the change in state $\nu_t = \theta_t - \rho\theta_{t-1}$ prevents perfect aggregation

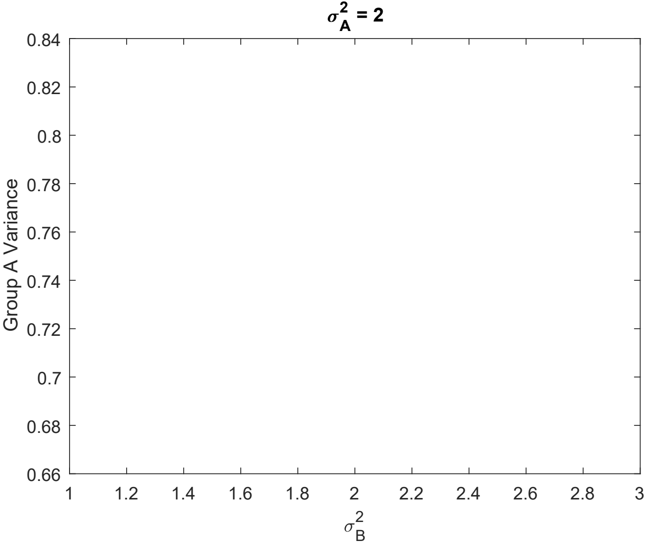
More generally: dimensionality of relevant state updates exceeds identification power afforded by your social neighborhood.

Figure: Tumbleweed: Picks up the dust along its way, rolls along with it

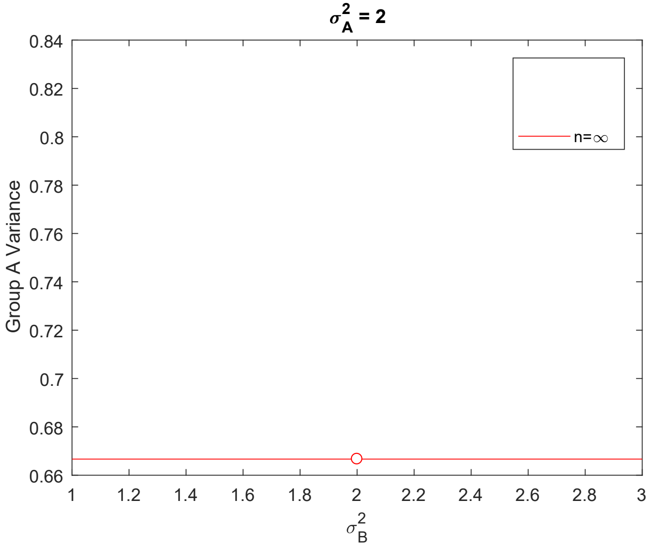

σ_A^2 = 2
Group A Variance
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3



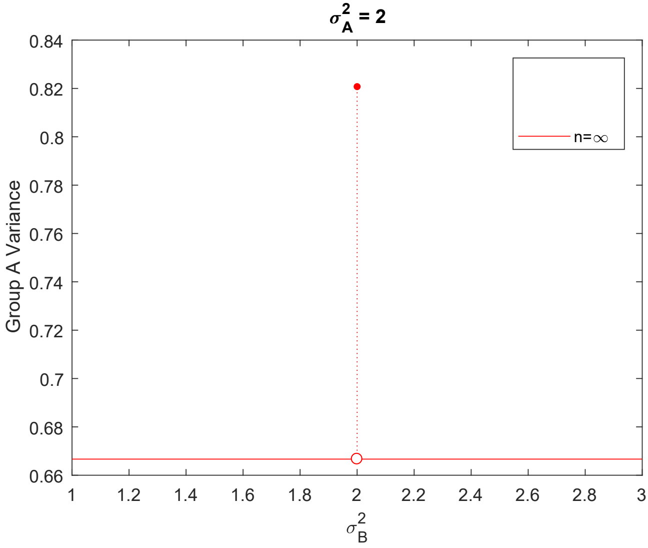
σ²_A = 2
Group A Variance
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ²_B



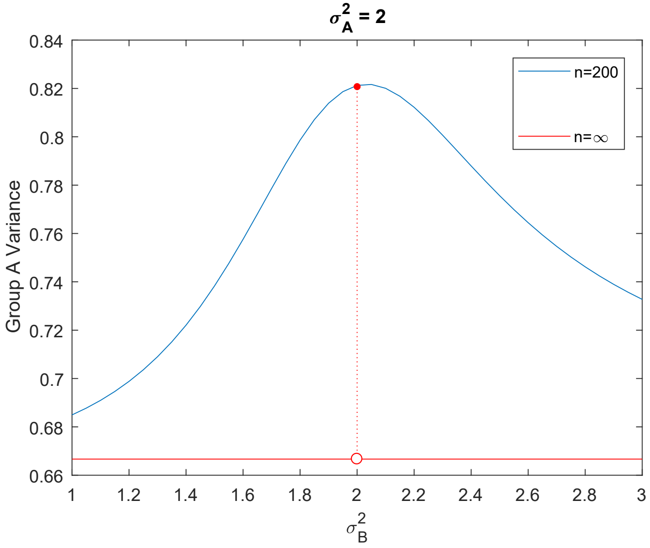
$\sigma_A^2 = 2$
Group A Variance
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ_B^2



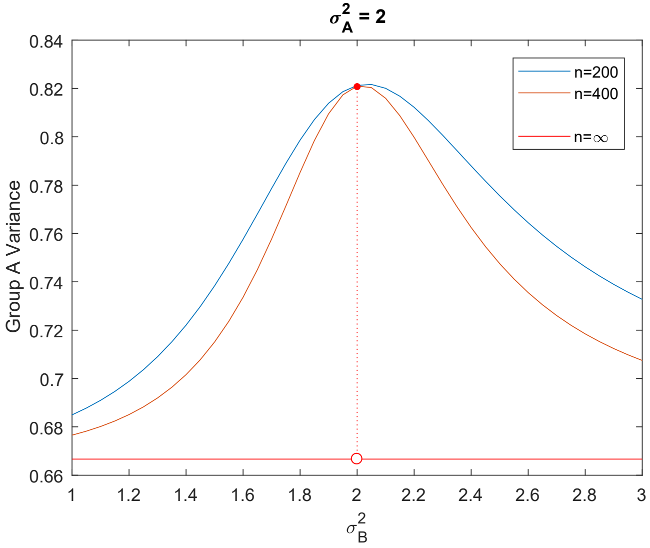
$\sigma_A^2 = 2$
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
Group A Variance
n=∞
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ_B^2



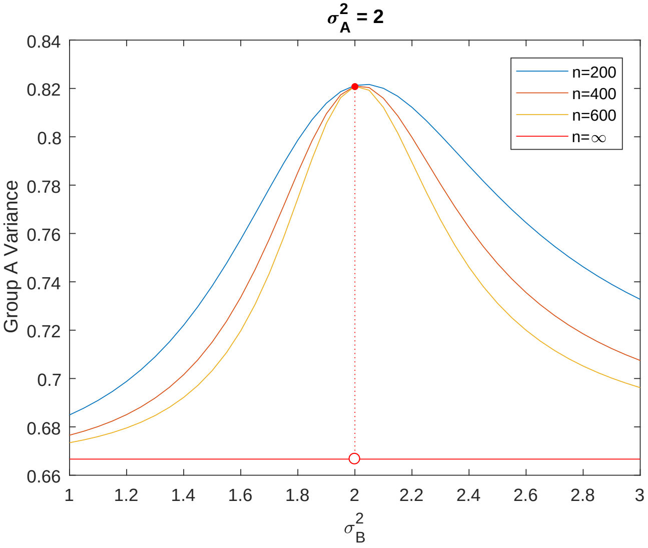
σ_A^2 = 2
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
Group A Variance
n=∞
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ_B^2



σ_A^2 = 2
n=200
n=∞
Group A Variance
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ_B^2



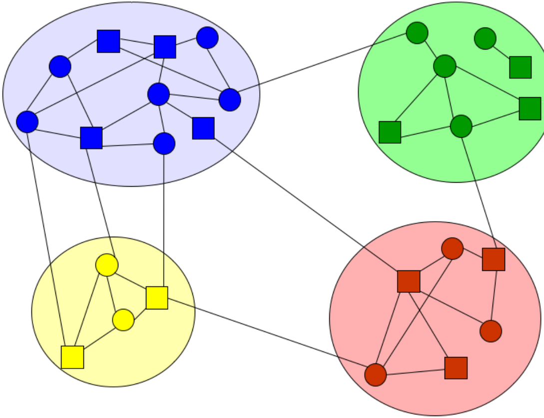
$\sigma_A^2 = 2$
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
Group A Variance
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
σ_B^2
n=200
n=400
n=∞



σ_A^2 = 2
Group A Variance
σ_B^2
0.84
0.82
0.8
0.78
0.76
0.74
0.72
0.7
0.68
0.66
1
1.2
1.4
1.6
1.8
2
2.2
2.4
2.6
2.8
3
n=200
n=400
n=600
n=∞

Heterogeneous signals, flexible networks

Stochastic block model: finitely many types; probabilities of linking between types given (depend on n) different signal types within network types.



Assume each neighborhood has **many** individuals of each of **at least two** signal types.

Heterogeneous signals, flexible networks

1. **Networks**
   - Large random network: n agents of finitely many network types comprising fixed population shares
   - types k and k' linked with probability $p_{kk'}$; links drawn independently; no isolated types

Heterogeneous signals, flexible networks

1. **Networks**
   - Large random network: n agents of finitely many network types comprising fixed population shares
   - types k and k' linked with probability $p_{kk'}$; links drawn independently; no isolated types
2. **Signals**
   - Each agent has one of many possible signal variances
   - Each network type contains a given share of agents with each private signal variance

Heterogeneous signals, flexible networks

1. **Networks**
   - Large random network: n agents of finitely many network types comprising fixed population shares
   - types k and k' linked with probability $p_{kk'}$; links drawn independently; no isolated types
2. **Signals**
   - Each agent has one of many possible signal variances
   - Each network type contains a given share of agents with each private signal variance
3. **Example:** Complete network with equal shares of agents with each signal quality

Heterogeneous signals, flexible networks

Condition. We say **signal diversity** holds if each network type has positive shares of agents with at least two distinct signal variances.

Heterogeneous signals, flexible networks

Condition. We say **signal diversity** holds if each network type has positive shares of agents with at least two distinct signal variances.

Theorem

Assume signal diversity. Let $\epsilon > 0$. If n is large enough, with probability $1 - \epsilon$ there is a stationary equilibrium where all agents have variances within ϵ of the perfect aggregation benchmark.

Heterogeneous signals, flexible networks

Condition. We say **signal diversity** holds if each network type has positive shares of agents with at least two distinct signal variances.

Theorem

Assume signal diversity. Let $\epsilon > 0$. If n is large enough, with probability $1 - \epsilon$ there is a stationary equilibrium where all agents have variances within ϵ of the perfect aggregation benchmark.

- With signal heterogeneity, Bayesian agents in stationary linear equilibrium achieve perfect aggregation on a broad class of networks

Heterogeneous signals, flexible networks

Condition. We say **signal diversity** holds if each network type has positive shares of agents with at least two distinct signal variances.

Theorem

Assume signal diversity. Let $\epsilon > 0$. If n is large enough, with probability $1 - \epsilon$ there is a stationary equilibrium where all agents have variances within ϵ of the perfect aggregation benchmark.

- With signal heterogeneity, Bayesian agents in stationary linear equilibrium achieve perfect aggregation on a broad class of networks
- The uncertainty is over the network: with small probability we could get a network that prevents learning

Naive agents

- Consider agents who incorrectly believe that their neighbors choose actions equal to their private signals, but are otherwise Bayesian (as in Eyster and Rabin, 2010)

Naive agents

- Consider agents who incorrectly believe that their neighbors choose actions equal to their private signals, but are otherwise Bayesian (as in Eyster and Rabin, 2010)
- Take a sequence of complete (or Erdös-Rényi) networks G_n as n grows large, two signal variances σ_A^2 and σ_B^2

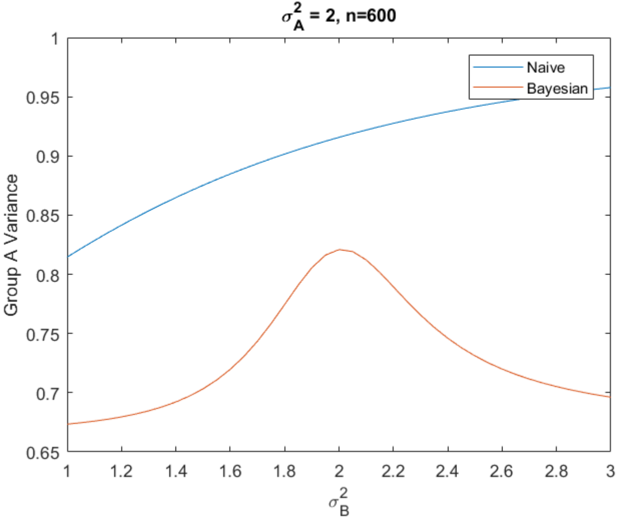
Naive agents

- Consider agents who incorrectly believe that their neighbors choose actions equal to their private signals, but are otherwise Bayesian (as in Eyster and Rabin, 2010)
- Take a sequence of complete (or Erdös-Rényi) networks G_n as n grows large, two signal variances σ_A^2 and σ_B^2
- The naive agents' equilibrium variances converge to values far from the equilibrium benchmark.

Naive agents

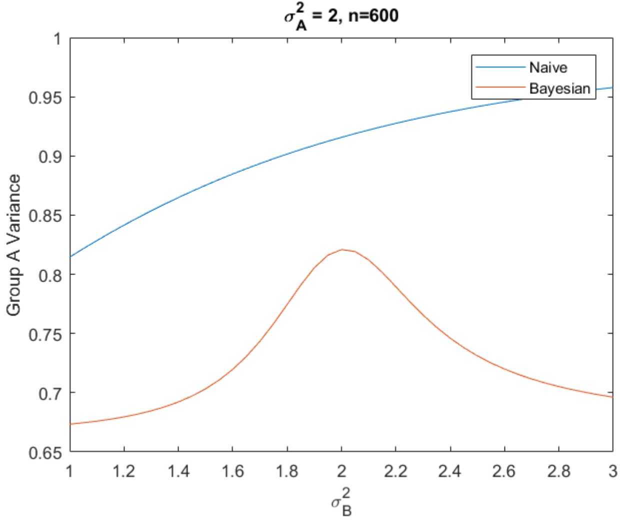
- Consider agents who incorrectly believe that their neighbors choose actions equal to their private signals, but are otherwise Bayesian (as in Eyster and Rabin, 2010)
- Take a sequence of complete (or Erdös-Rényi) networks G_n as n grows large, two signal variances σ_A^2 and σ_B^2
- The naive agents' equilibrium variances converge to values far from the equilibrium benchmark.
- Perfect aggregation requires a sophisticated response to correlation, while naive agents completely ignore correlation.

Comparing naive and Bayesian agents



Complete graph with two signal variances

Comparing naive and Bayesian agents



Complete graph with two signal variances

Failure to achieve benchmark with naive agents

Proposition

Assume all updating weights are positive and agents put total weight $\geq \delta > 0$ on neighbors and on own signal.

Then in any sequence of weight matrices, there is a constant $c > 0$ s.t. at all times $t \geq 1$ all agents have variance exceeding the perfect aggregation benchmark by at least c.

Failure to achieve benchmark with naive agents

Proposition

Assume all updating weights are positive and agents put total weight $\geq \delta > 0$ on neighbors and on own signal.

Then in any sequence of weight matrices, there is a constant $c > 0$ s.t. at all times $t \geq 1$ all agents have variance exceeding the perfect aggregation benchmark by at least c.

"Tumbleweed" intuition: pick up old noise even though it's irrelevant.

Failure to achieve benchmark with naive agents

Proposition

Assume all updating weights are positive and agents put total weight $\geq \delta > 0$ on neighbors and on own signal.

Then in any sequence of weight matrices, there is a constant $c > 0$ s.t. at all times $t \geq 1$ all agents have variance exceeding the perfect aggregation benchmark by at least c.

"Tumbleweed" intuition: pick up old noise even though it's irrelevant.

Compare with "wisdom of crowds" in fixed-state environments – e.g., Jadbabaie, Molavi, Sandroni, Tahbaz-Salehi 12.

Conclusion

- Introduced a model of social learning with a moving target.
- Key idea: diversity of signal distributions in one's neighborhood helps one to filter. **A (distinctive) reason to have specialized expertise.**
- Methodology: study action of Φ: fixed points (stationary equilibrium, which is a DeGroot-type behavior) or dynamics starting from initial time.
- Sophistication is crucial.
- Diversity helps rational agents even in real-world, small networks.

Numerical results

- Social networks of 43 villages in rural India (Banerjee et al. 2013);

  mean size 212 (s.d. 53.5); mean degree 19 (s.d. 7.5).

Numerical results

- Social networks of 43 villages in rural India (Banerjee et al. 2013);

  mean size 212 (s.d. 53.5); mean degree 19 (s.d. 7.5).

- Two cases:
  1. **homogeneous**: all signal variances equal to 2;

Numerical results

- Social networks of 43 villages in rural India (Banerjee et al. 2013);

  mean size 212 (s.d. 53.5); mean degree 19 (s.d. 7.5).

- Two cases:
  1. **homogeneous**: all signal variances equal to 2;
  2. **heterogeneous**: majority (92%) has the same signal distribution as in the first case, but a minority (people lacking electricity) has a substantially worse signal.

Numerical results

- Social networks of 43 villages in rural India (Banerjee et al. 2013);

  mean size 212 (s.d. 53.5); mean degree 19 (s.d. 7.5).

- Two cases:
  1. **homogeneous**: all signal variances equal to 2;
  2. **heterogeneous**: majority (92%) has the same signal distribution as in the first case, but a minority (people lacking electricity) has a substantially worse signal.

- In eq'm, median agent in terms of learning quality has more precise estimates of the state in heterogeneous case.

Numerical results

- Social networks of 43 villages in rural India (Banerjee et al. 2013);

  mean size 212 (s.d. 53.5); mean degree 19 (s.d. 7.5).

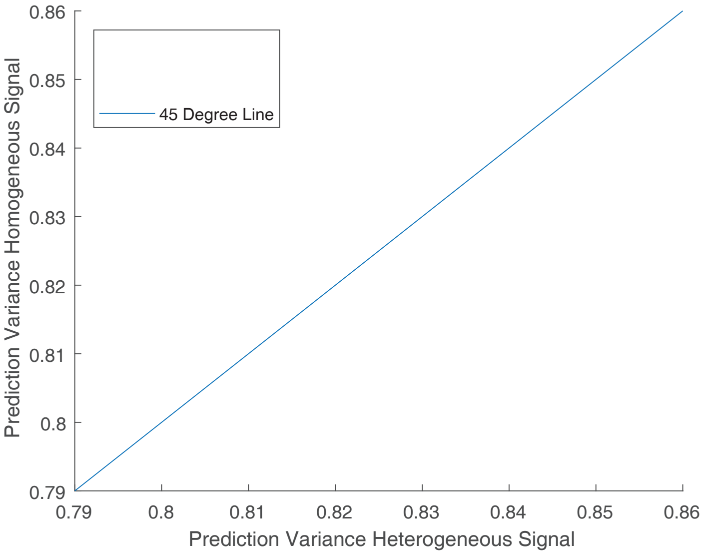
- Two cases:
  1. **homogeneous**: all signal variances equal to 2;
  2. **heterogeneous**: majority (92%) has the same signal distribution as in the first case, but a minority (people lacking electricity) has a substantially worse signal.

- In eq'm, median agent in terms of learning quality has more precise estimates of the state in heterogeneous case.

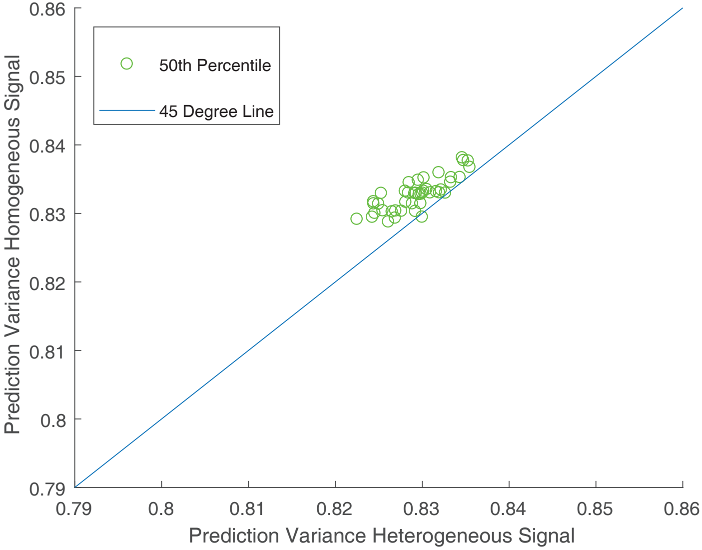
- Also consider an agent who estimates the state better than 75 percent of agents); advantage of these agents in the heterogeneous case is even more pronounced.

Social influence: A classic networks question

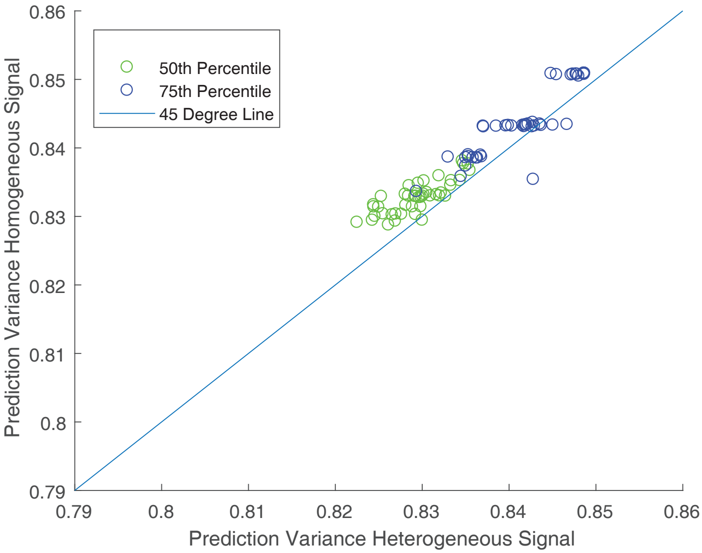
- Let an agent's **social influence** be the effect of changing her time-t private signal by 1 unit on the average beliefs of all agents, summed across all times.
- Focusing on the positive-weights case, we analyze social influence and how it depends on the network and signal qualities.
- Two equal groups with similar signal variances σ_A, σ_B. Either complete or random with average degrees d_A and d_B
- Suppose we "improve" A's position in some way (higher σ_A, d_A).
  - Ratio [A influence]/[B influence] $> \frac{\sigma_A}{\sigma_B}$.
  - Ratio [A influence]/[B influence] $< \frac{d_A}{d_B}$



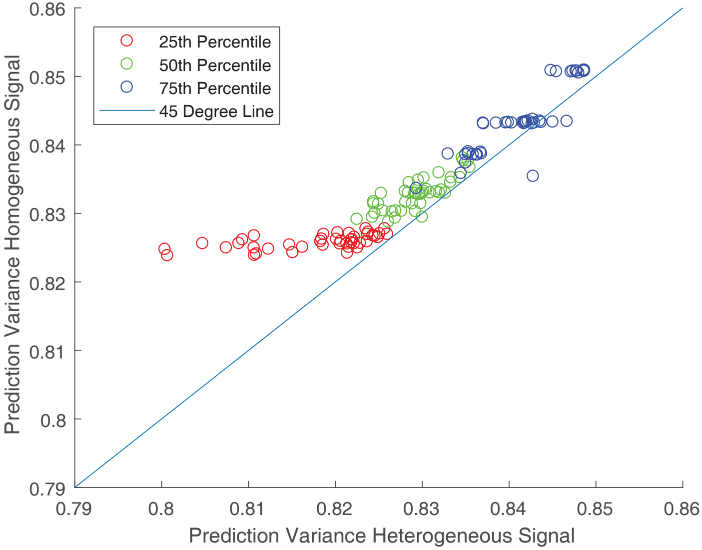
Village networks with homogeneous and heterogeneous signal variances.



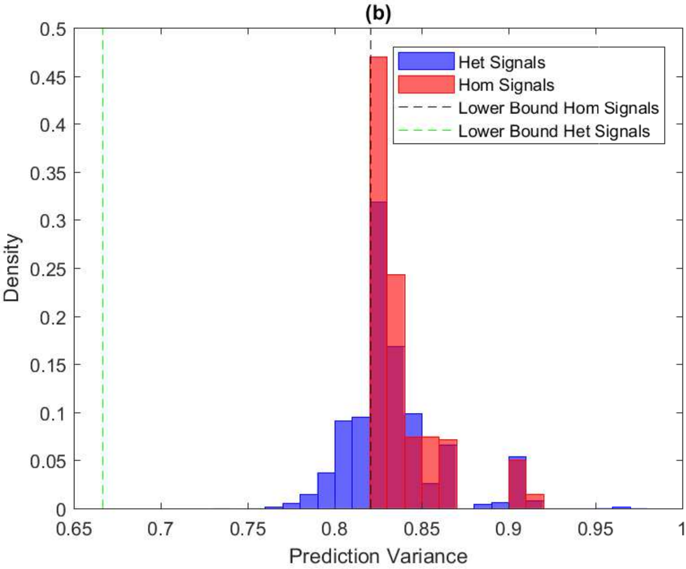
Village networks with homogeneous and heterogeneous signal variances.



Village networks with homogeneous and heterogeneous signal variances.



Village networks with homogeneous and heterogeneous signal variances.



Village networks with homogeneous and heterogeneous signal variances.