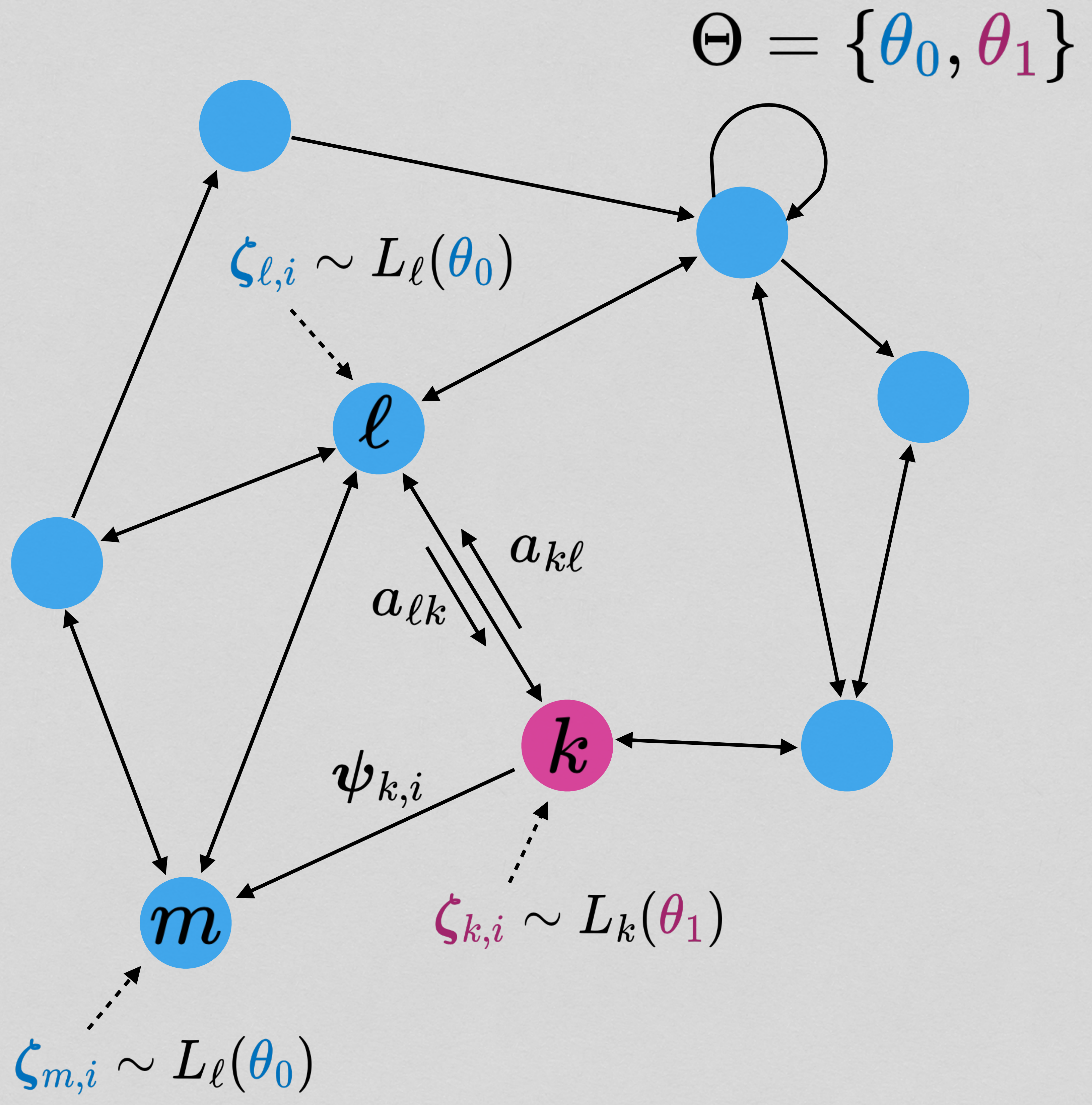
1589: Distributed Decision-Making for Community Structured Networks

Valentina Shumovskaia, Mert Kayaalp, and Ali H. Sayed

Social learning model

The social learning paradigm is a popular non-Bayesian formulation that enables a group of networked agents to learn and track the state of nature.

- **Network**: A set $\mathcal{N}$ of communicating agents receiving observations from the environment
- **Hypotheses**: A set of Θ possible states of environment
- Agents aim to identify the current true state of environment by
    - Receiving streaming private observations from environment
    - Exchanging beliefs with immediate neighbours
- **Beliefs**: Probability mass functions over the set Θ for each agent
- The main question in social learning is whether agents are able to learn the truth eventually.

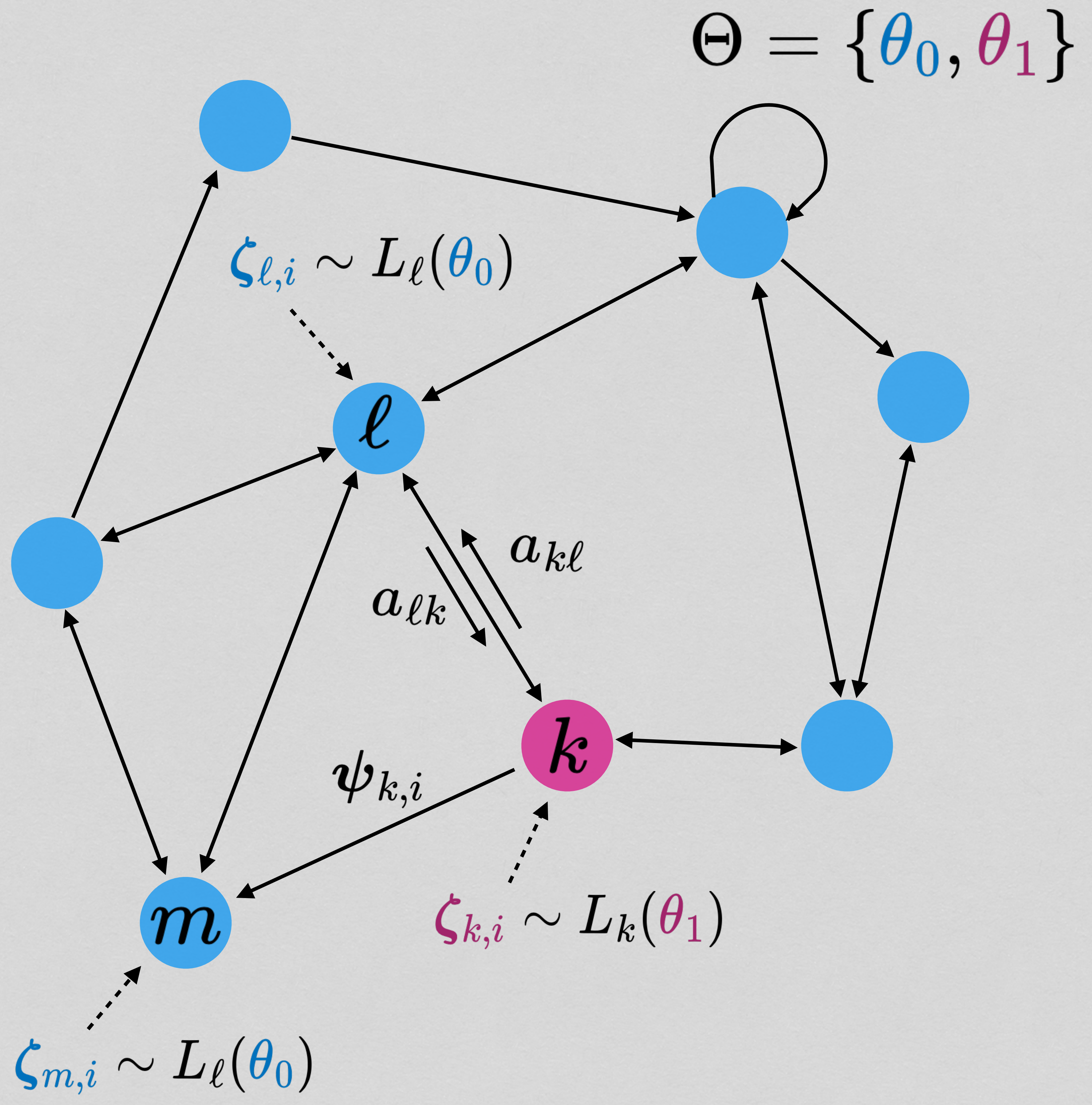


Social learning model

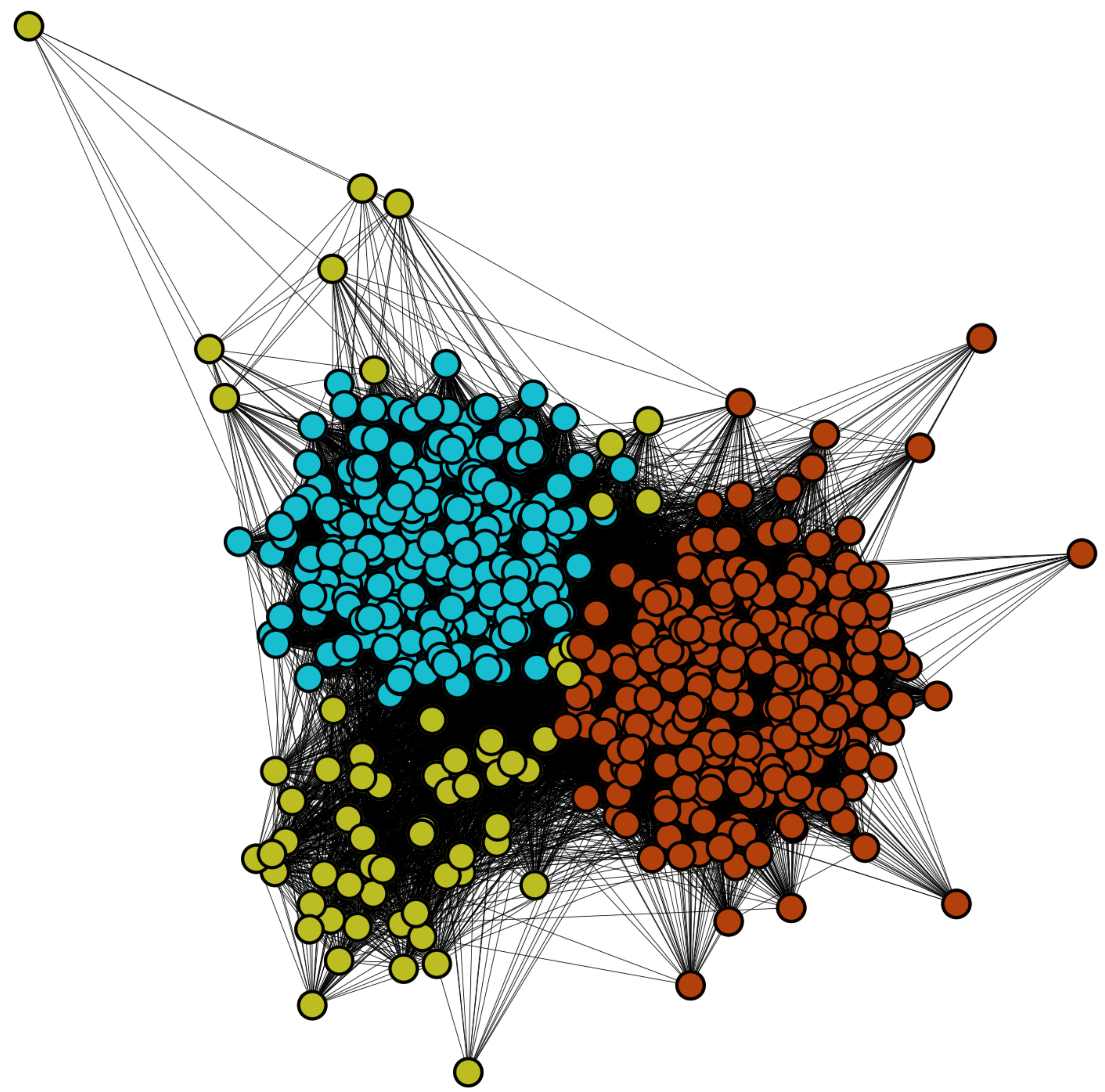
Applications:

- Distributed decision-making (e.g., sensors),
- Hypotheses testing,
- Opinion formation processes…

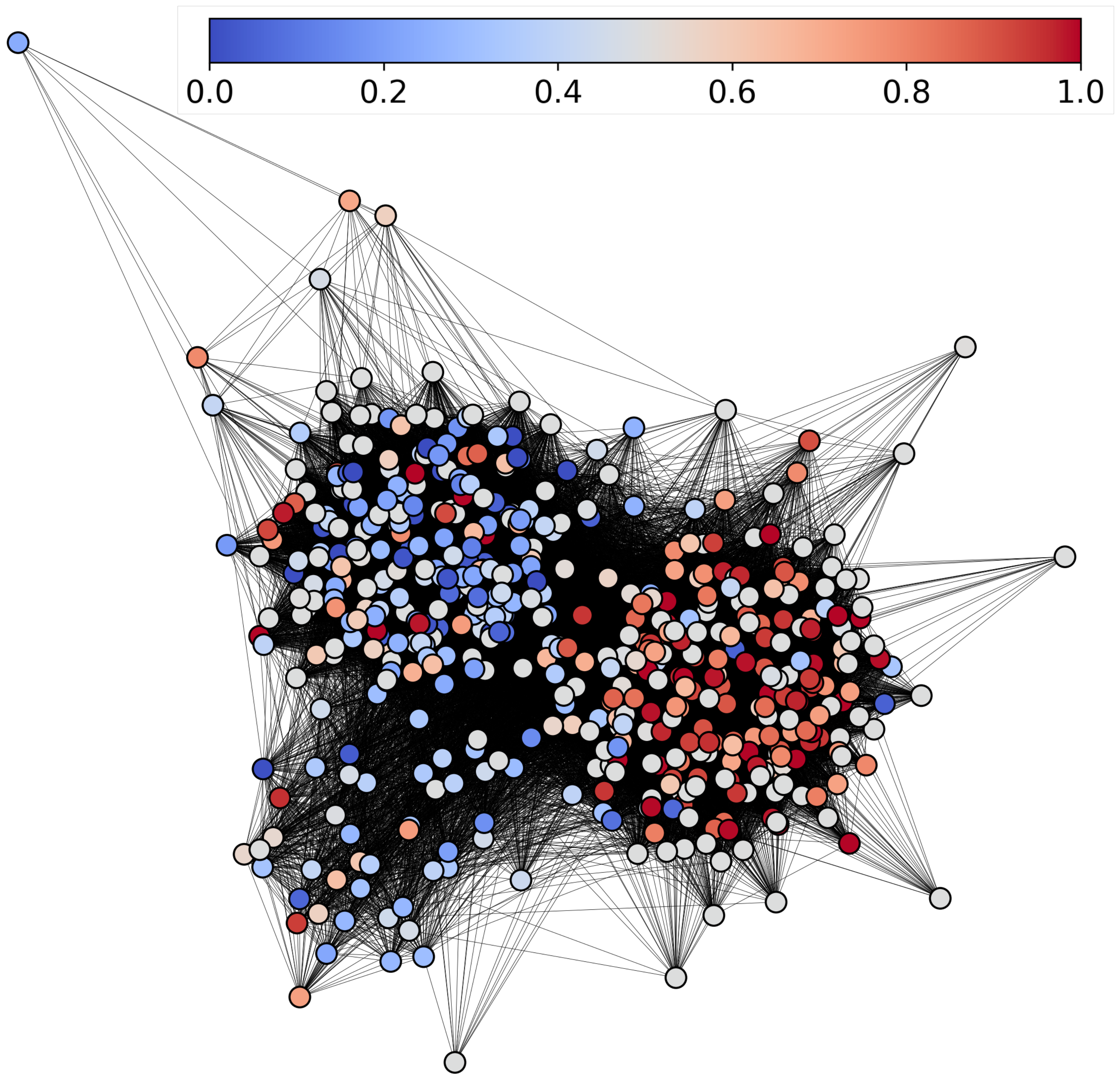
Typically, one single truth is assumed. But what if we consider a **clustered (or community structured) network**?



How opinions evolve in communities: **Brexit** and British Parliament 𝕏



Louvain [2] community detection algorithm

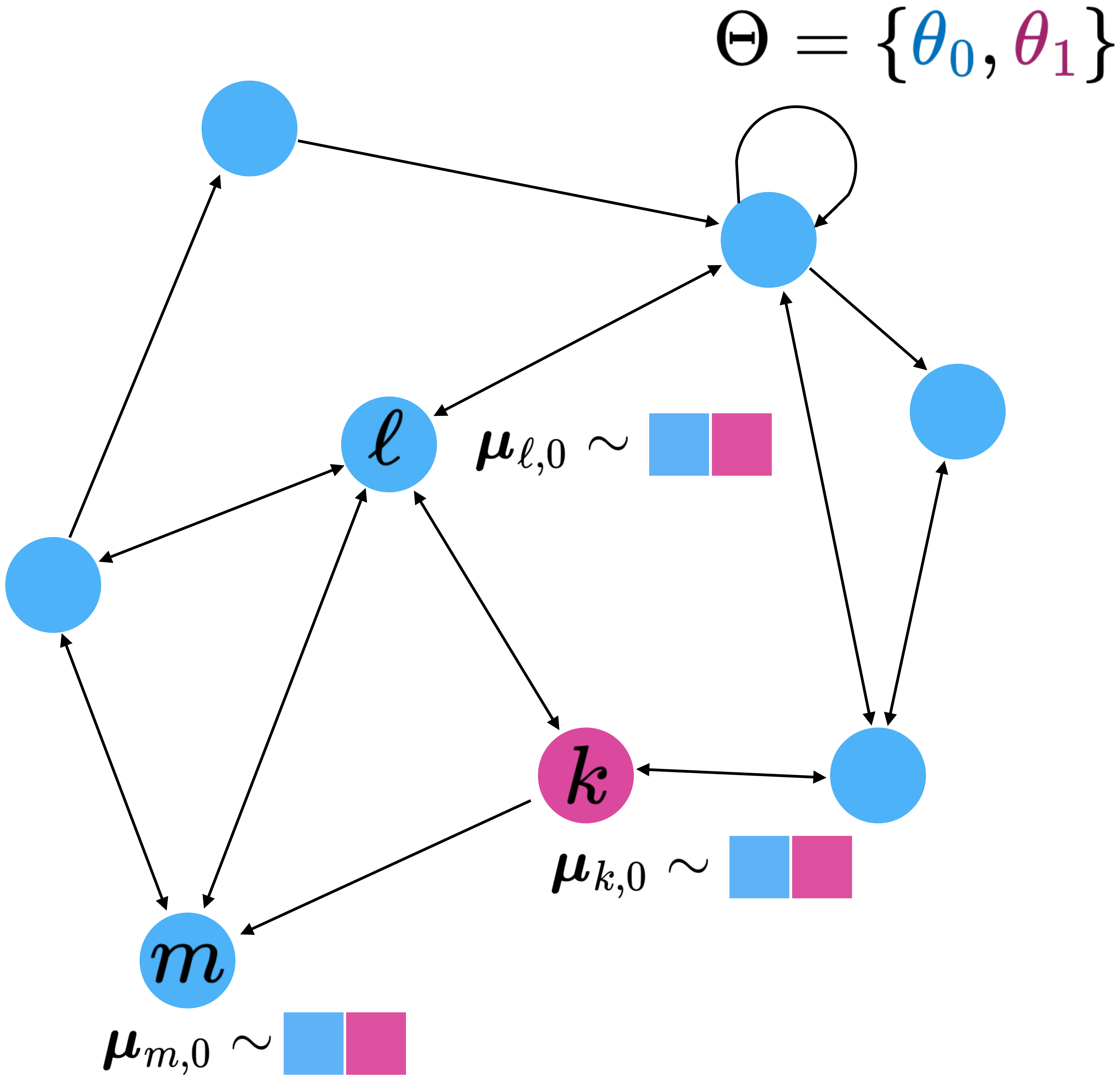


Average belief on Brexit from 01.01.2020 to 01.04.2020

[2] De Meo, Pasquale, et al. "Generalized Louvain method for community detection in large networks." *2011 11th international conference on intelligent systems design and applications*. IEEE, 2011.

Adaptive social learning model

0. Initialisation

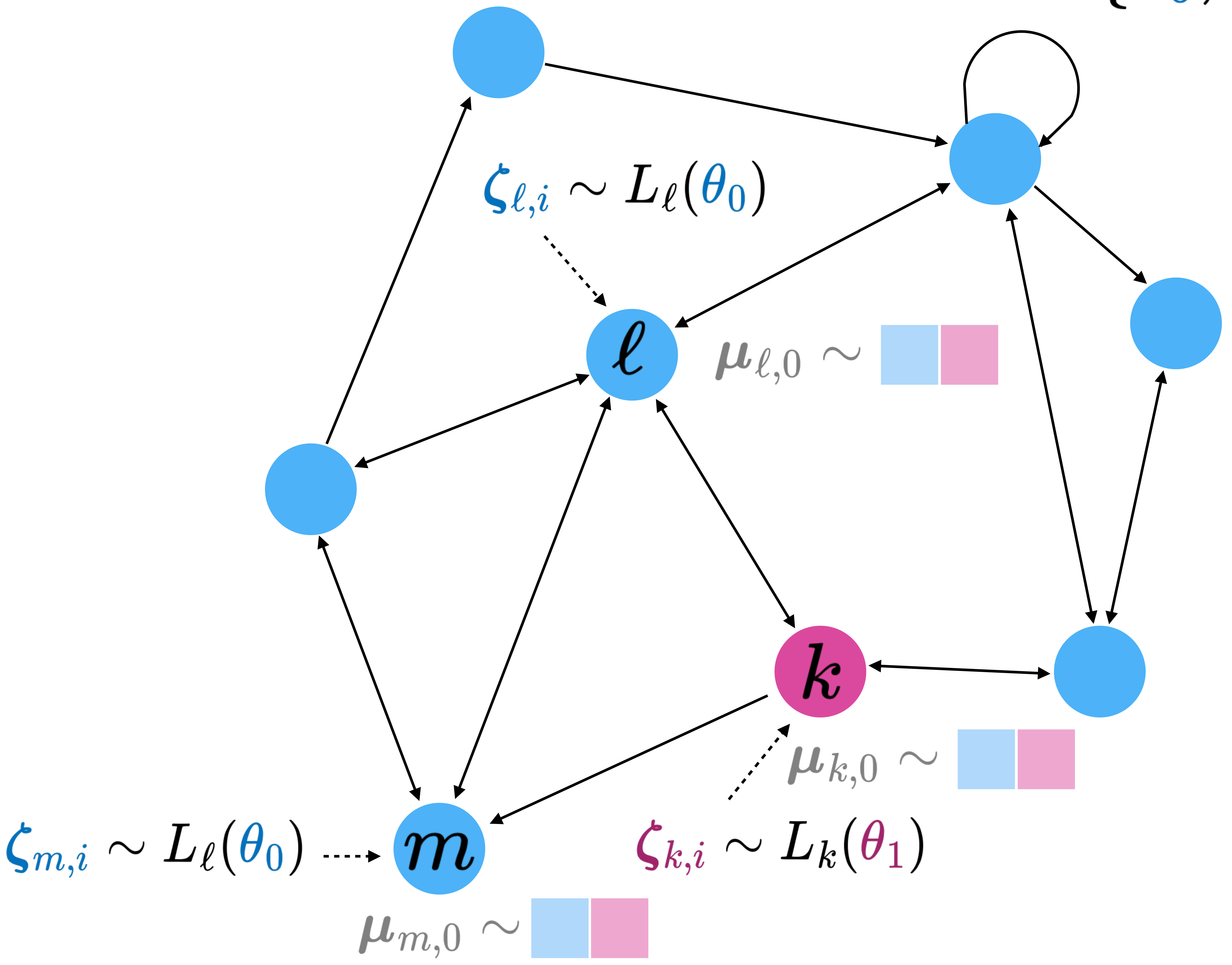


Adaptive social learning model

0. Initialisation

1. Observation

$\Theta = \{\theta_0, \theta_1\}$

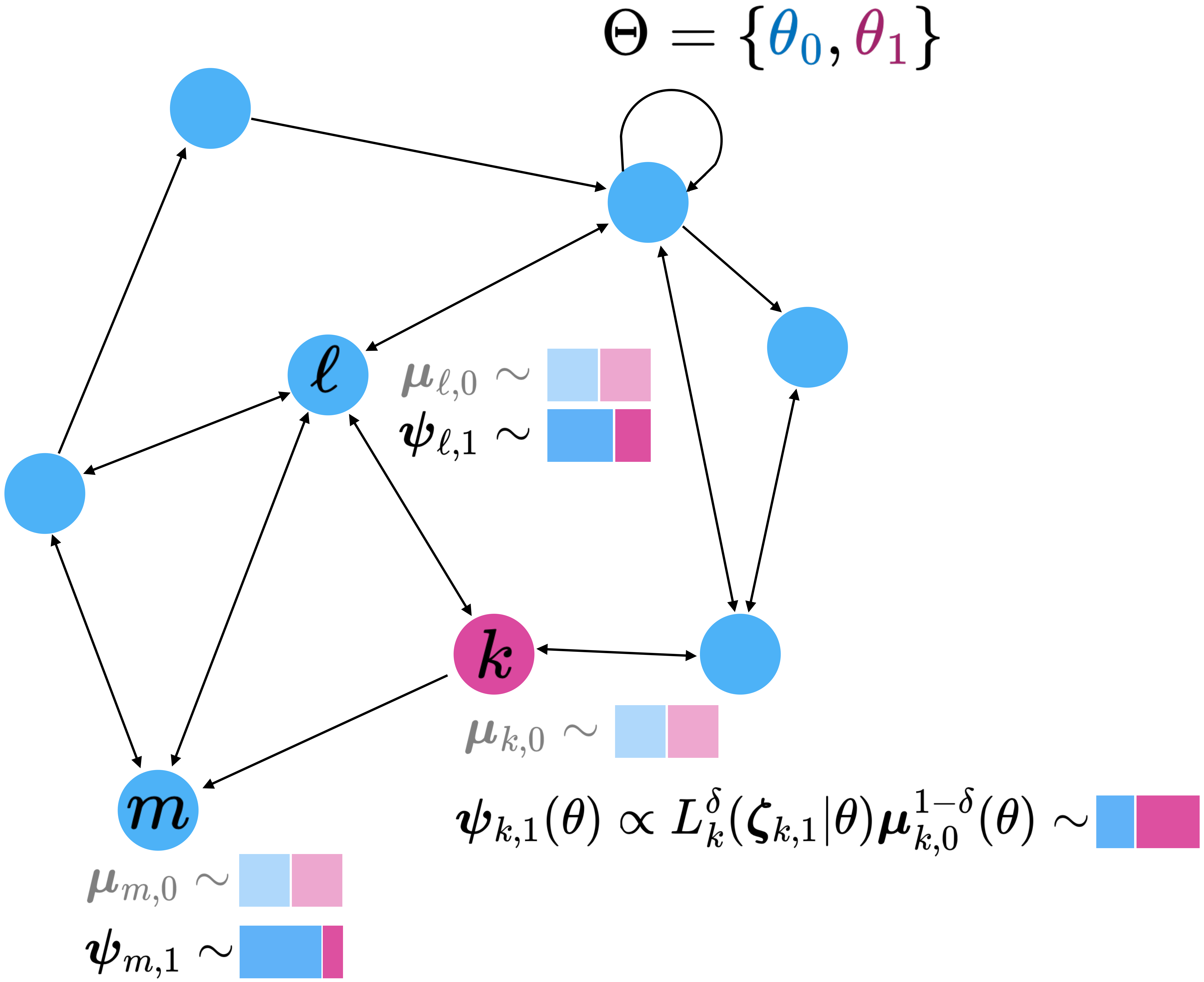


Adaptive social learning model

0. Initialisation

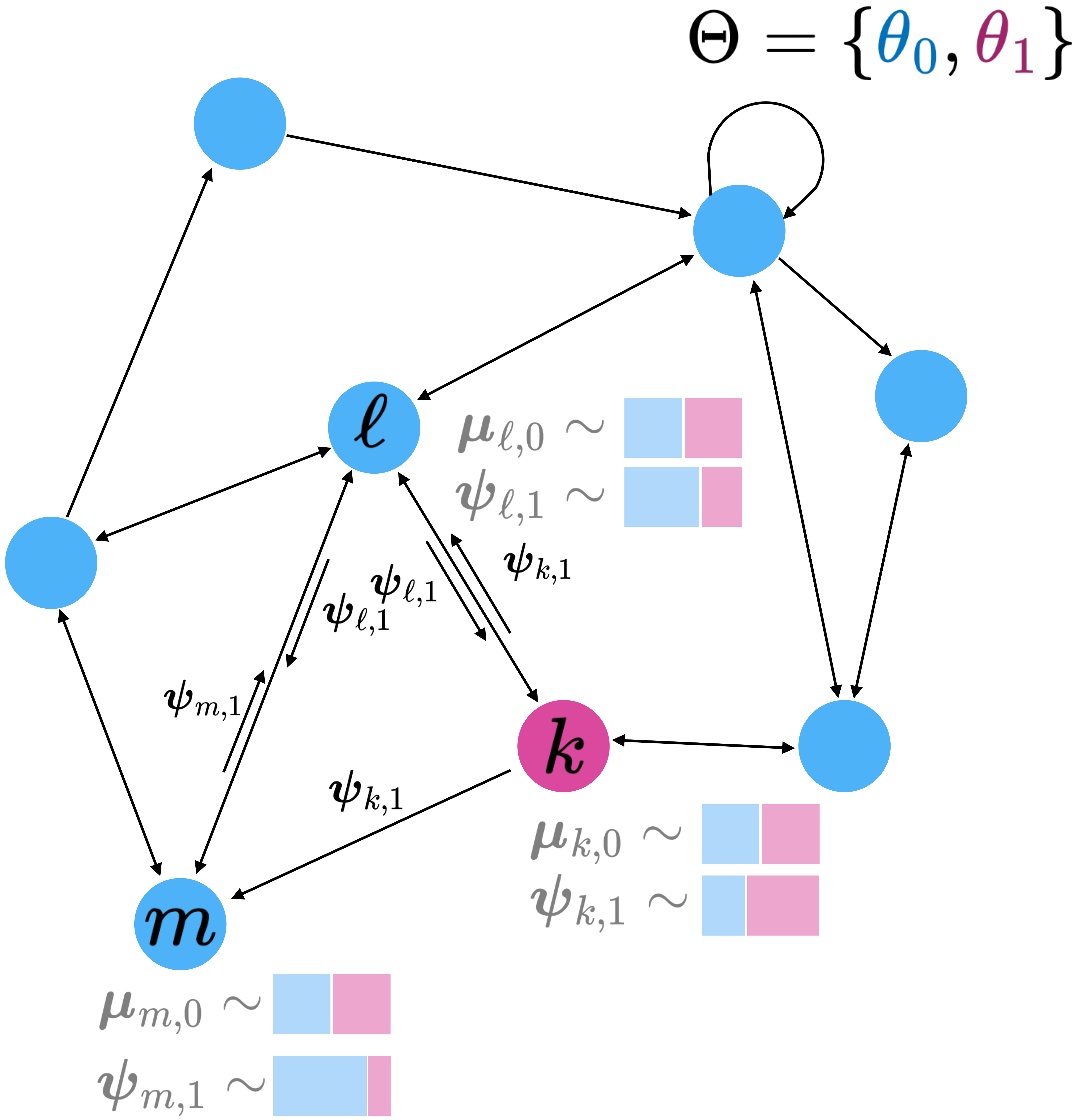
1. Observation

2. Local update



Adaptive social learning model

0. Initialisation
1. Observation
2. Local update
3. Communication



Adaptive social learning model

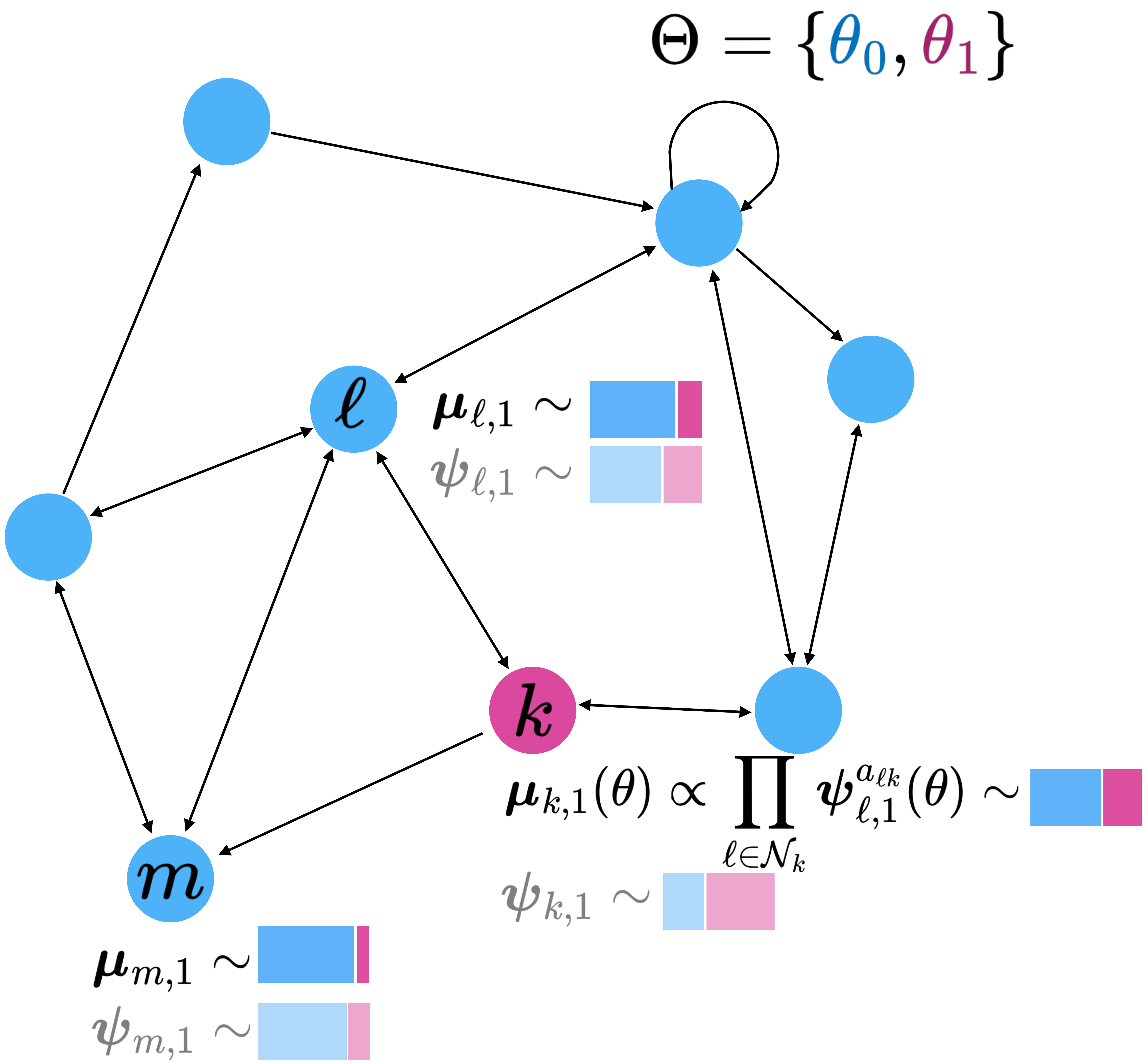
0. Initialisation

1. Observation

2. Local update

3. Communication

4. Fusion



Adaptive social learning model

0. Initialisation

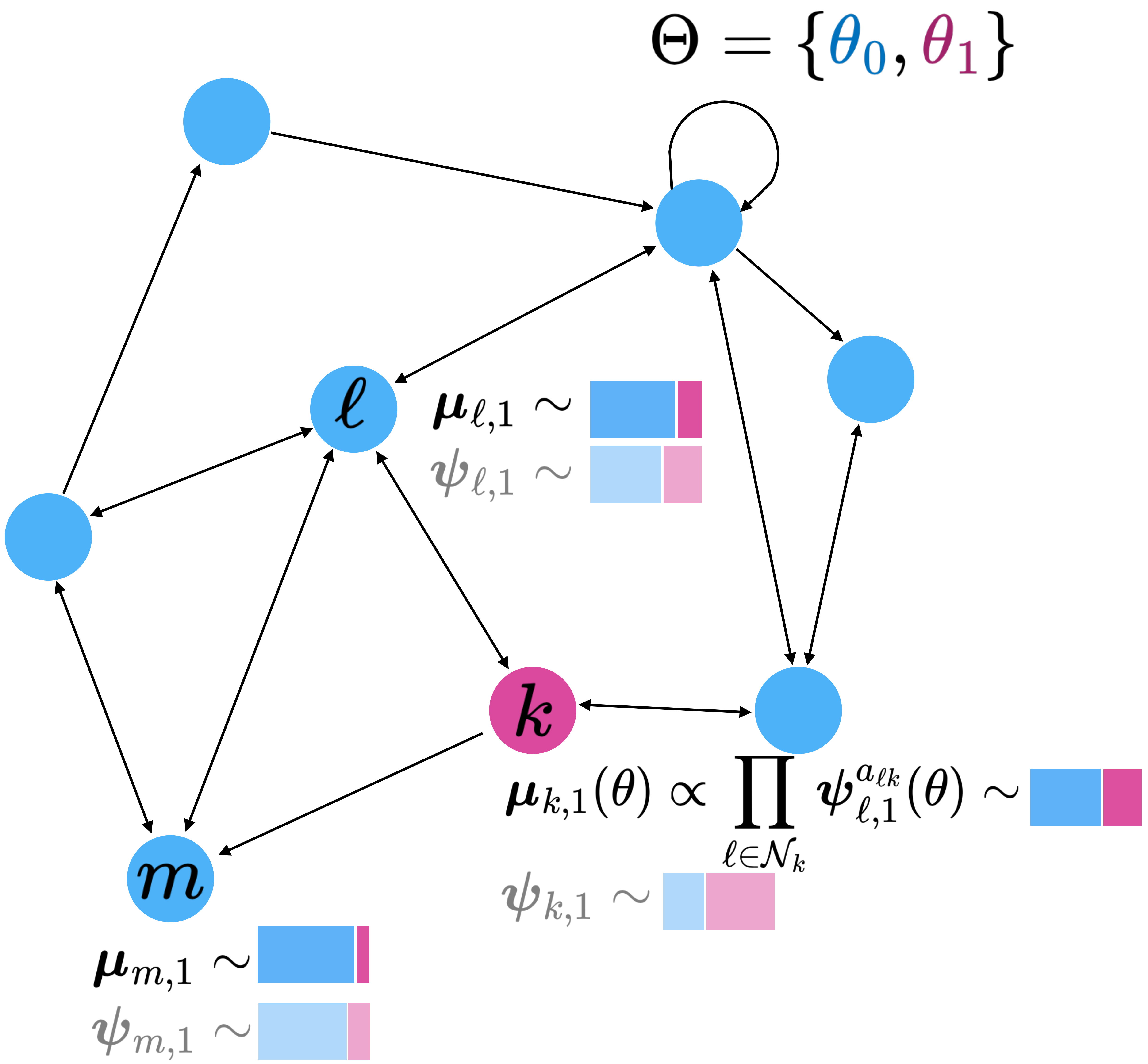
1. Observation

2. Local update

3. Communication

4. Fusion

Repeat steps 1-4



Adaptive social learning model

- At each iteration i, agent $k \in \mathcal{N}$ receives i.i.d. (w.r.t. iteration i) observation conditioned on **its own** truth

$$\zeta_{k,i} \sim L_k(\zeta_{k,i} | \theta_k^\star)$$

- And has an access to likelihood models

$$L_k(\zeta_{i,k} | \theta), \ \theta \in \Theta$$

- Start with randomly initialised $\mu_{k,0}$: $\mu_{k,0}(\theta) > 0$ and $\sum_\theta \mu_{k,0}(\theta) = 1$
- Local update:

public beliefs $\psi_{k,i}(\theta) \propto L_k^\delta(\zeta_{k,i} | \theta)\mu_{k,i-1}^{1-\delta}(\theta), \quad k \in \mathcal{N}$

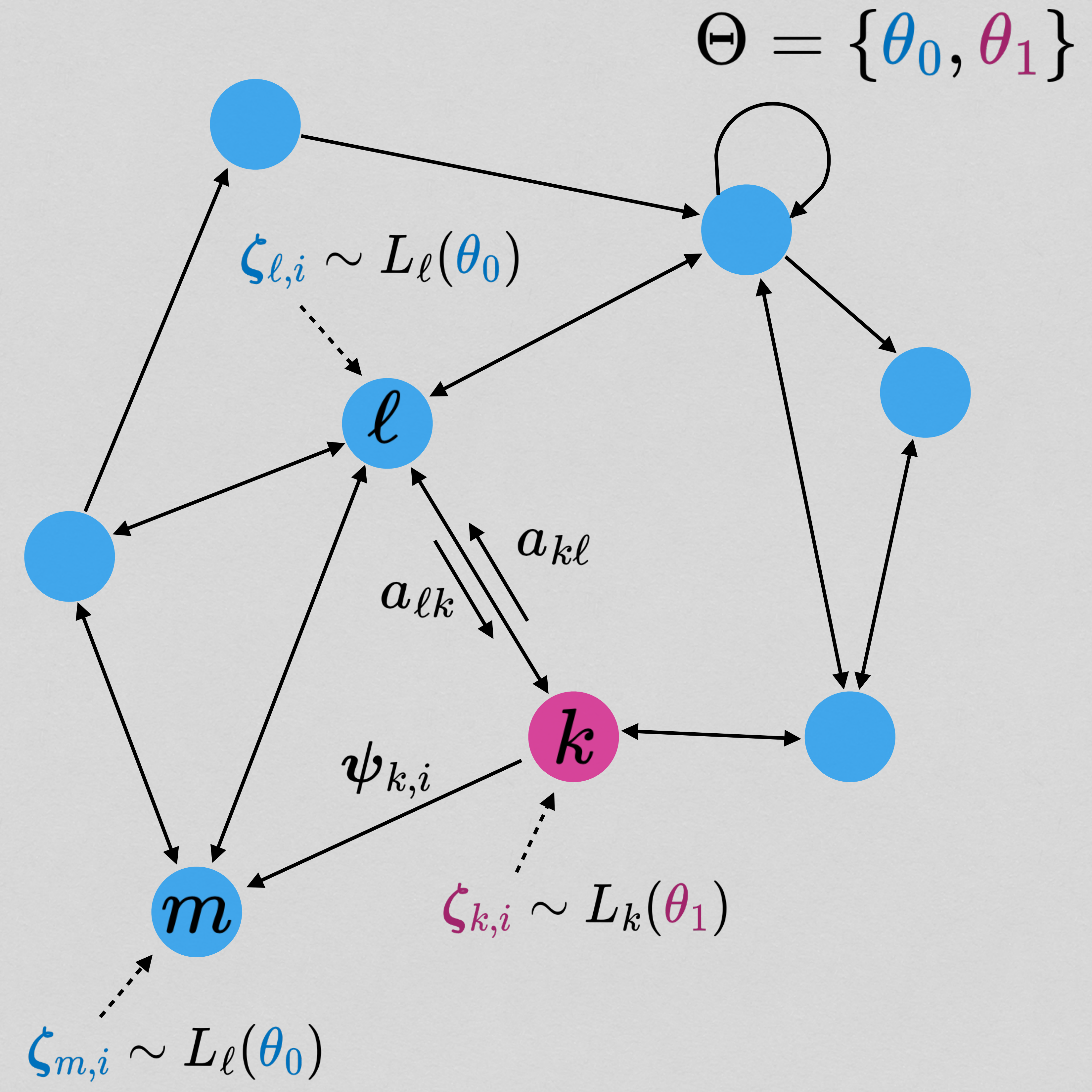
*Step-size parameter $\delta \in (0,1)$ controls the importance of newly received data

- Fusion:

private beliefs $\mu_{k,i}(\theta) \propto \prod_{\ell \in \mathcal{N}_k} \psi_{\ell,i}^{a_{\ell k}}(\theta), \quad k \in \mathcal{N}$

A: left-stochastic, strongly connected…

- Decision-making: $\hat{\theta} = \arg\max_{\theta \in \Theta} \mu_{k,i}$



[1] Bordignon, Virginia, Vincenzo Matta, and Ali H. Sayed. "Adaptive social learning." IEEE Transactions on Information Theory 67.9 (2021): 6053-6081.

Adaptive social learning model

- Theoretical guarantees for learning the state $\theta^\star$:

$$\textcolor{red}{\lim_{\delta\to 0}} \lim_{i\to\infty} \mathbb{P}(\arg\max_{\theta\in\Theta} \boldsymbol{\mu}_{k,i}(\theta) \neq \theta^\star) = 0, \forall k \in \mathcal{N}$$

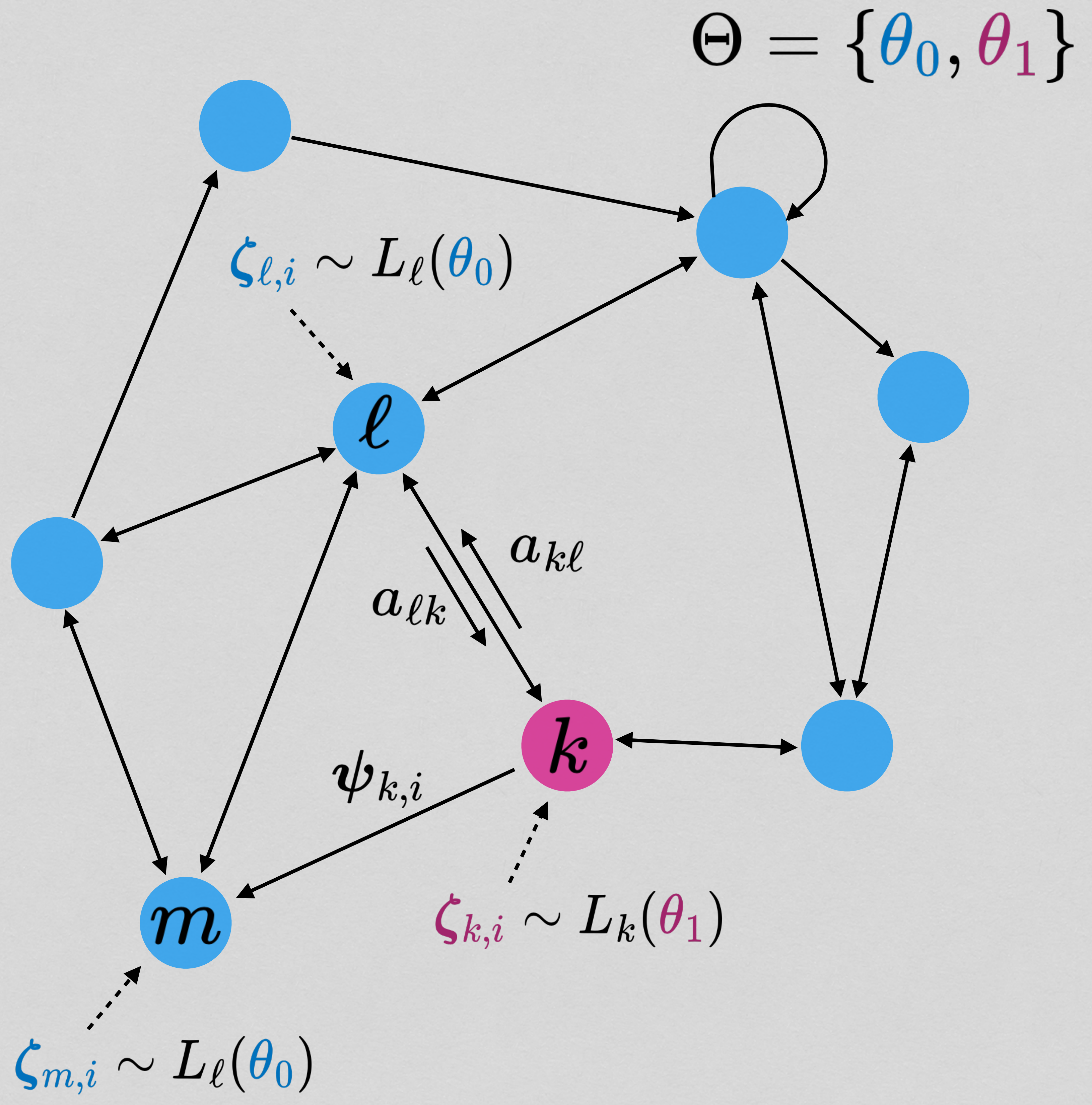
- Where the state $\theta^\star$ is the optimal solution of:

$$\min_{\theta} \sum_{k\in\mathcal{N}} u_k D_{\mathsf{KL}}\big(L_k(\theta_k^\star)\,||\,L_k(\theta)\big)$$

  - D_{KL} is Kullback-Leibler divergence between two distributions:

$$D_{\mathsf{KL}}\big(L_k(\theta)\,||\,L_k(\theta')\big) \triangleq \mathbb{E}_{\zeta\sim L_k(\zeta|\theta)} \log \frac{L_k(\zeta|\theta)}{L_k(\zeta|\theta')}$$

  - u_k is Perron entry of combination matrix A



Adaptive social learning model

- Theoretical guarantees for learning the state $\theta^\star$:

$$\lim_{\delta\to 0}\lim_{i\to\infty}\mathbb{P}(\arg\max_{\theta\in\Theta}\boldsymbol{\mu}_{k,i}(\theta)\neq\theta^\star)=0,\ \forall k\in\mathcal{N}$$

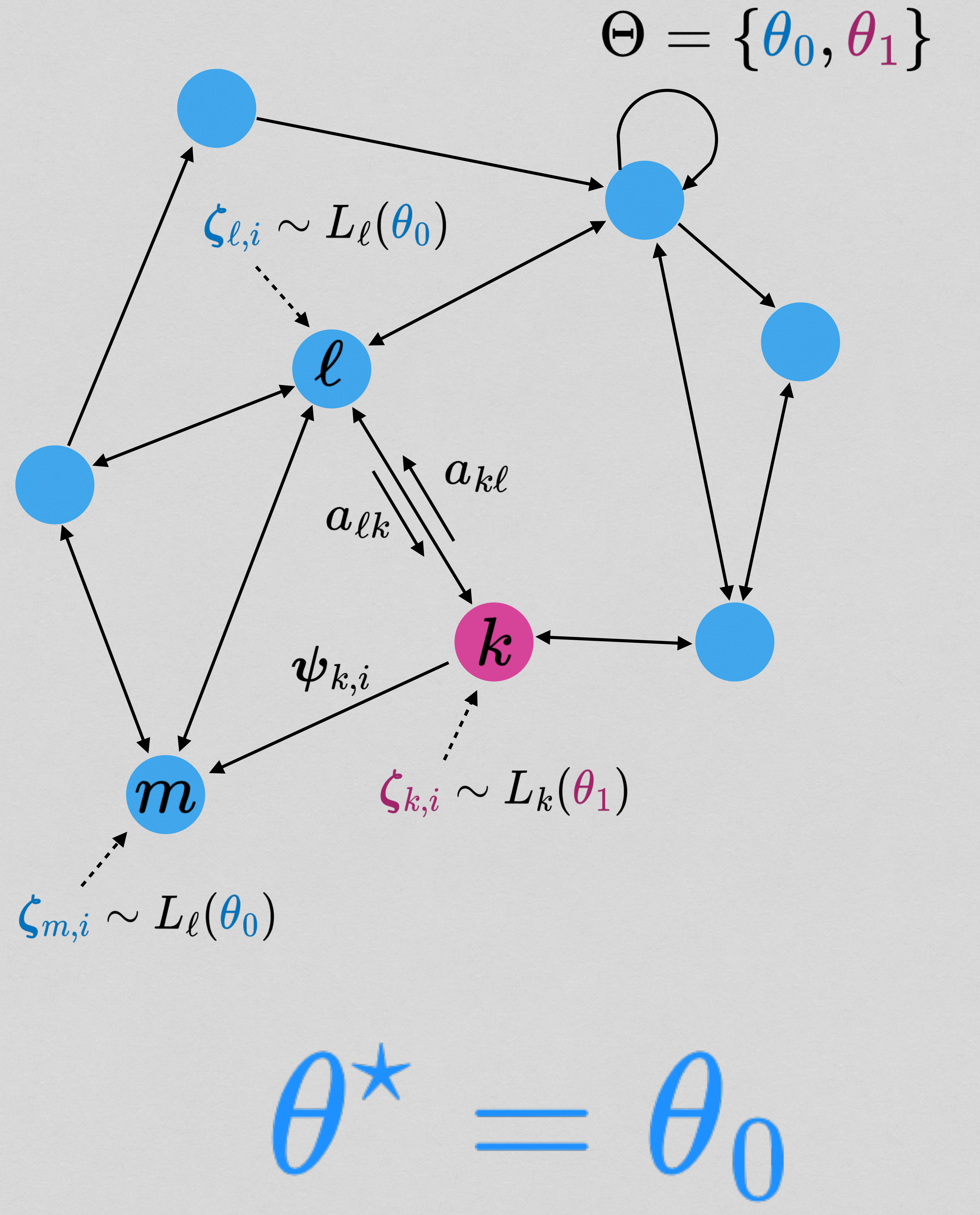
- Where the state $\theta^\star$ is the optimal solution of:

$$\min_{\theta}\sum_{k\in\mathcal{N}}u_k D_{\mathsf{KL}}\big(L_k(\theta_k^\star)\,||\,L_k(\theta)\big)$$

  - D_{KL} is Kullback-Leibler divergence between two distributions:

$$D_{\mathsf{KL}}\big(L_k(\theta)\,||\,L_k(\theta')\big)\triangleq\mathbb{E}_{\zeta\sim L_k(\zeta|\theta)}\log\frac{L_k(\zeta|\theta)}{L_k(\zeta|\theta')}$$

  - u_k is Perron entry of combination matrix A



Adaptive parameter δ

- Allows for **more diverse behavior** of social learning: the agents are able to quickly react to the true state change

- Recap:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k^{\delta}(\boldsymbol{\zeta}_{k,i}\,|\,\theta)\boldsymbol{\mu}_{k,i-1}^{1-\delta}(\theta), \quad k \in \mathcal{N}$$

$$\lim_{\delta\to 0}\lim_{i\to\infty} \mathbb{P}(\arg\max_{\theta\in\Theta} \boldsymbol{\mu}_{k,i}(\theta) \neq \theta^{\star}) = 0, \forall k \in \mathcal{N}$$

- The algorithm is studied in **homogeneous environments**, i.e. when all agents observations are conditioned on one single $\theta^{\star}$

- What happens if we let $\delta \gg 0$?

Log-beliefs ratio

- It can be shown that:

$$\log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta')} \xrightarrow[i\to\infty]{d} \lim_{i\to\infty} \delta \sum_{\ell\in\mathcal{N}} \sum_{t=0}^{i} (1-\delta)^t [A^{t+1}]_{\ell,k} \times \log \frac{L_\ell(\boldsymbol{\zeta}_{\ell,i} \mid \theta)}{L_\ell(\boldsymbol{\zeta}_{\ell,i} \mid \theta')}$$

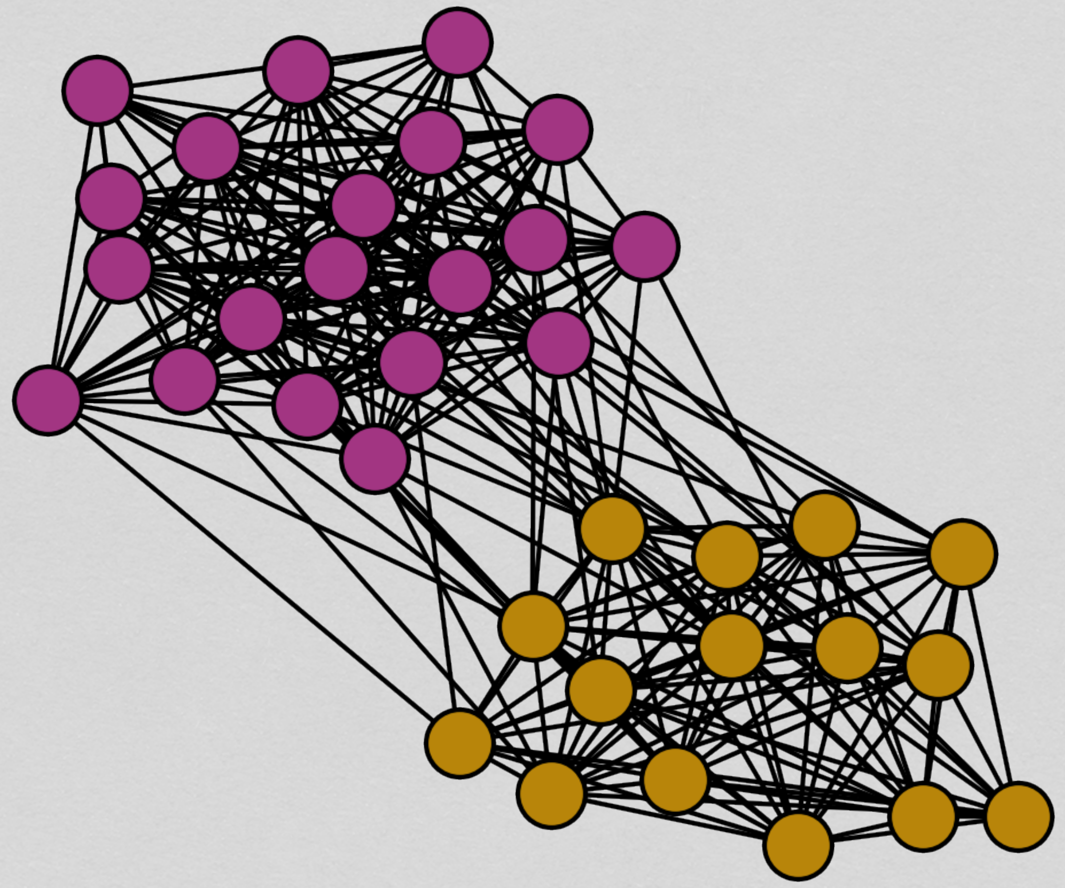
- Each agent gives **higher importance** to its **local topology**
  - $(1-\delta)$ scales the immediate one-hop neighbors
  - $(1-\delta)^2$ scales the agents from the 2-hop neighborhood
  - …
- What if we consider clustered graphs?

Stochastic Block Model (2 communities)

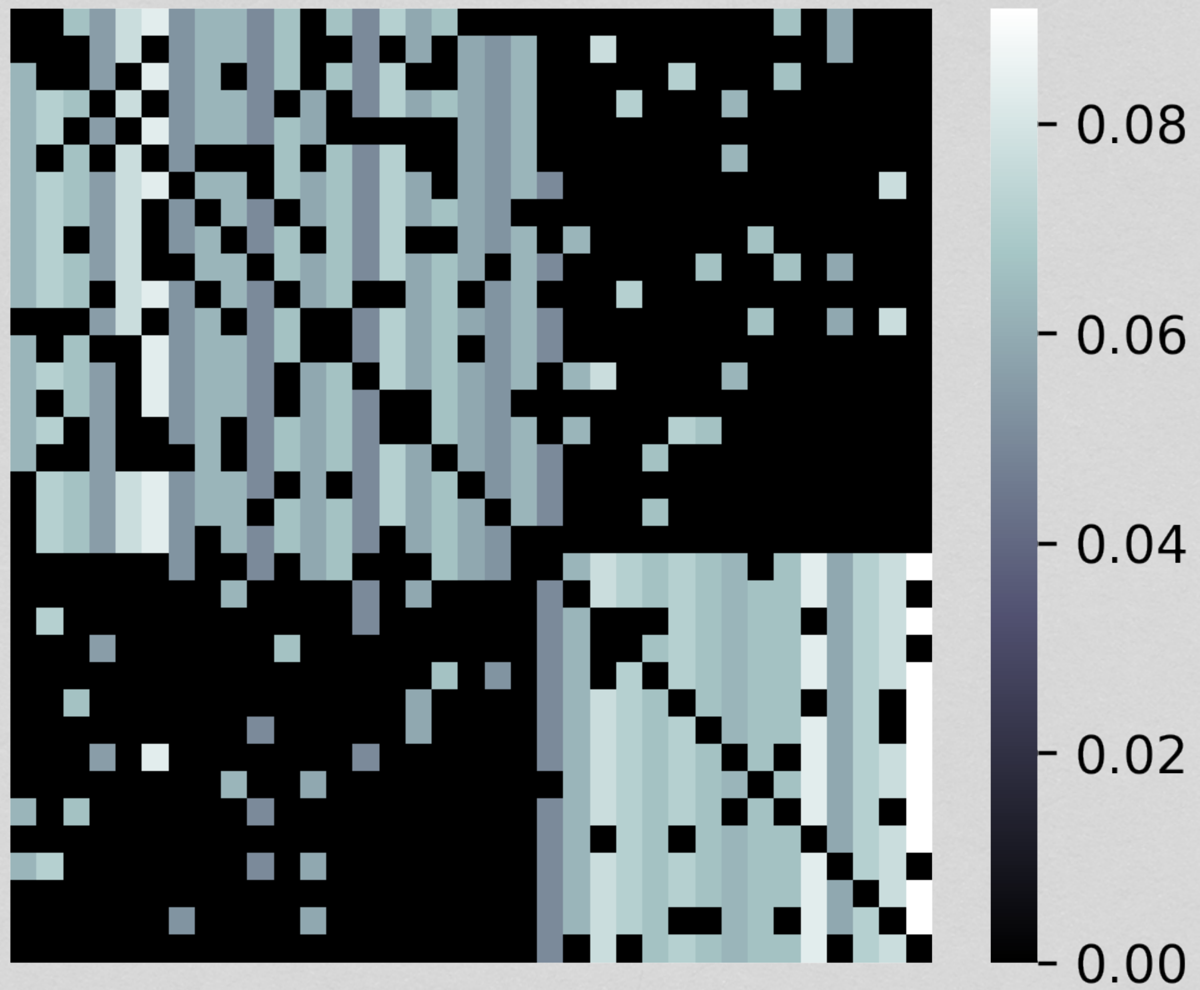
- Edges are generated as: $\boldsymbol{E} \sim Bernoulli(P)$
- Where P is probability matrix of a block form:
  - $$P \triangleq \left[\begin{array}{c|c} p_0 \mathbb{1}_{n_0} \mathbb{1}_{n_0}^\top & q_0 \mathbb{1}_{n_0} \mathbb{1}_{n_1}^\top \\ \hline q_1 \mathbb{1}_{n_1} \mathbb{1}_{n_0}^\top & p_1 \mathbb{1}_{n_1} \mathbb{1}_{n_1}^\top \end{array}\right]$$
  - $q_0, q_1 < \min\{p_0, p_1\}$, $n_0 + n_1 = |\mathcal{N}|$ are communities sizes
- Agents communicate according to left-stochastic **combination matrix**

$\boldsymbol{A}$ follows averaging rule $[\boldsymbol{A}]_{\ell,k} = \boldsymbol{E}_{\ell,k} / \sum_\ell \boldsymbol{E}_{\ell,k}$

Network illustration with $n_1 = 15$, $n_0 = 20$, $p_0 = 0.8$, $p_1 = 0.9$, $q_0 = q_1 = 0.1$.



SBM graph model with two communities



Combination matrix

Truth Learning theorem

- **Assumption:** within each cluster, agents have the same level of informativeness. For any $k \in [0, n_0]$ and $\ell \in [n_0+1, n_0+n_1]$:

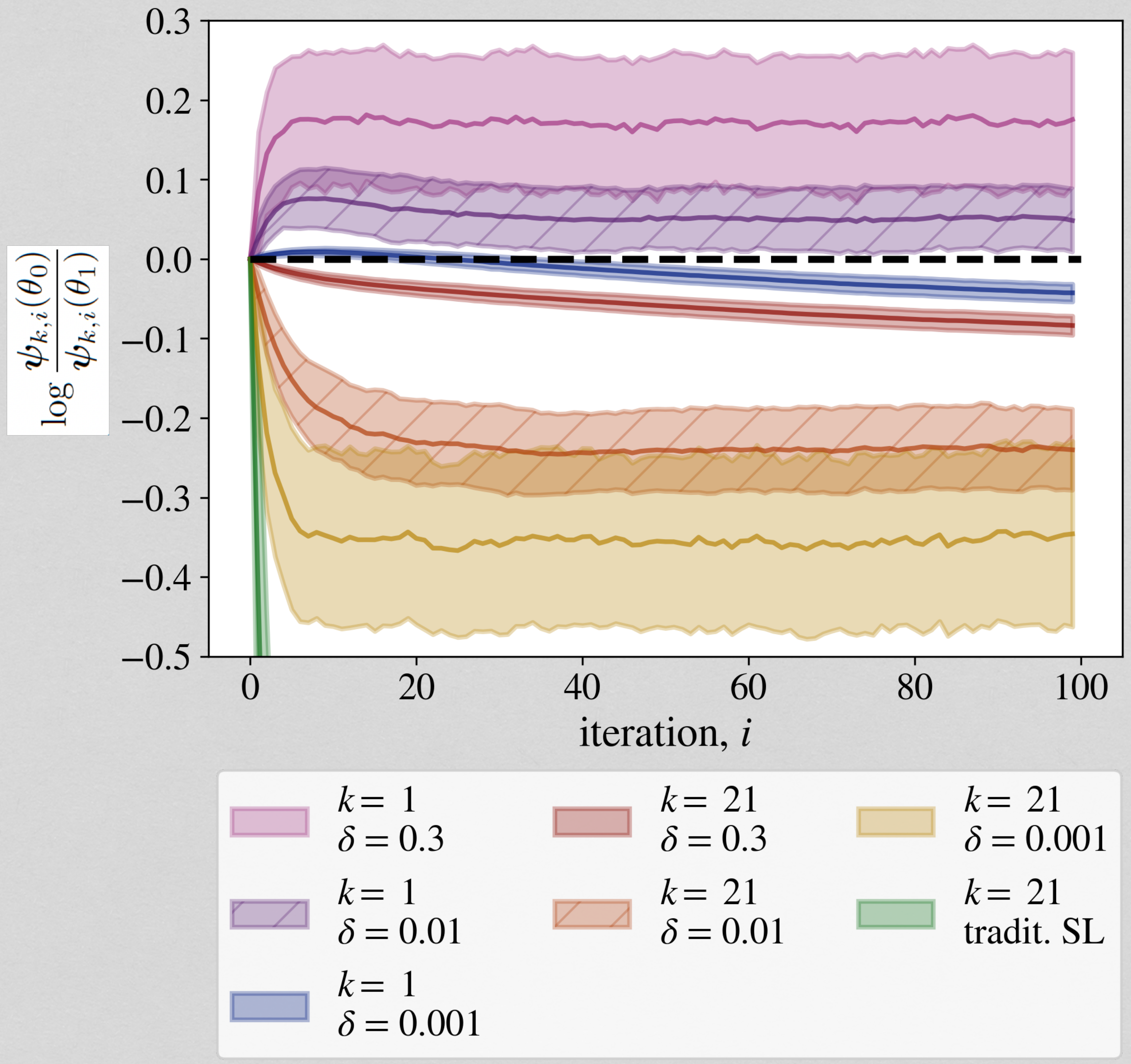
$$d_0 \triangleq D_{\mathrm{KL}}\left(L_k(\theta_0)||L_k(\theta_1)\right), \quad d_1 \triangleq D_{\mathrm{KL}}\left(L_\ell(\theta_1)||L_\ell(\theta_0)\right)$$

- **Theorem:** If the probabilities between clusters are sufficiently smaller than the probabilities inside the clusters, more specifically:

$$p_0 n_0 d_0 - q_1 n_1 d_1 > 0, \quad p_1 n_1 d_1 - q_0 n_0 d_0 > 0$$

Then, there exist a $\delta_0 \in (0,1)$, that for any $\delta > \delta_0$, on average, each cluster converges to its own hypothesis, i.e. both $\lim_{i\to\infty} \mathbb{E} \log \frac{\boldsymbol{\mu}_{k,i}(\theta_0)}{\boldsymbol{\mu}_{k,i}(\theta_1)}$ and $\lim_{i\to\infty} \mathbb{E} \log \frac{\boldsymbol{\psi}_{k,i}(\theta_0)}{\boldsymbol{\psi}_{k,i}(\theta_1)}$ are strictly positive or strictly negative depending on the cluster.

Network illustration with $n_0 = n_1 = 15$, $p_0 = p_1 = 0.8$, $q_0 = q_1 = 0.1$, $0.37 = d_0 < d_1 = 0.51$



Agent $k = 1$: cluster #1, θ_0
Agent $k = 21$: cluster #2, θ_1

mean and standard deviations over 500 algorithm runs

Conclusions

- **Traditional social learning** techniques behave conservatively and the whole network converges to a **consensus solution**, which is not necessarily optimal at the individual or cluster level.

- **Adaptive social learning** strategies

  - Behave similarly to traditional strategies when the adaptation hyperparameter $\delta \to 0$.

  - For sufficiently large $\delta > 0$, the ASL strategy is the **preferred** choice for **graphs with community structure** (such as SBM).

- Thus δ is not only introducing the adaptivity to the state changes in homogeneous environments, but also plays a **role of adaptivity** to the local neighbourhood of each individual agent.