EPFL

Thèse n° 10 489

Sequential Reasoning with Socially Caused Beliefs

Présentée le 14 février 2025

Faculté des sciences et techniques de l'ingénieur
Laboratoire de Systèmes Adaptatifs
Programme doctoral en informatique et communications

pour l'obtention du grade de Docteur ès Sciences

par

Mert KAYAALP

Acceptée sur proposition du jury

Prof. M. C. Gastpar, président du jury
Prof. A. H. Sayed, directeur de thèse
Prof. P. M. Djuric, rapporteur
Prof. D. Shah, rapporteur
Prof. N. Kiyavash, rapporteuse

École polytechnique fédérale de Lausanne

2025

There's no sense in being precise when
you don't even know what you're talking about.
— John von Neumann

To my mother . . .

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Prof. Ali H. Sayed for giving me the opportunity to join his lab and for his guidance throughout my five-year journey. His exceptional work ethic and dedication have been a constant source of inspiration. Despite his numerous awards and career successes, he is the hardest-working person I have ever known; and even with his busy schedule, he always manages to find time for his students. I will always be indebted to him for all that I have learned during my PhD.

I would also like to thank Professors Emre Telatar and Visa Koivunen for their support and inspiring discussions, both in research and beyond. I feel privileged to have interacted with Prof. Telatar while collaborating with my colleague and friend Yunus. I was also extremely fortunate to meet Prof. Koivunen during his sabbatical at EPFL, and I feel honored to have collaborated with both of them.

I would like to extend my appreciation to the members of my thesis committee — Professors Negar Kiyavash, Devavrat Shah, Petar Djurić, and Michael Gastpar — for thoroughly evaluating my thesis and providing constructive feedback.

I would like to thank all my collaborators and labmates in the Adaptive Systems Laboratory — Konstantinos Ntemos, Stefan Vlaski, Virginia Bordignon, Lucas Cassano, Elsa Rizk, Ping Hu, Valentina Shumovskaia, Ying Cao, Flavio Pavan, Malek Khammassi, Haoyuan Cai, and Prof. Vincenzo Matta — for all the moments we have shared both inside and outside the lab. A special thanks goes to Yunus Inan for our research discussions that spanned two continents, countless coffee shops, Bilkent Center, Düveroglu, and the IPG Lounge. I was also extremely lucky to work with interns from whom I have learned a lot: Fatima Ghadieh, Ching-Fang Li, Mert Cemri, Elyas Benyamina, Eduardo Cabrera, and Vladyslav Shashkov. I am indebted to our lab secretary, Patricia Vonlanthen, for shielding us from the stress that comes with the bureaucratic side of the PhD journey.

During my time in Lausanne, I have met so many wonderful friends from all over the world. It would be impossible to list everyone individually without doing injustice to the importance of each connection, yet I am truly thankful to have met each and

every one of you. However, I must specifically mention my family in Lausanne — or more precisely, the kemik tayfa: Baran, Yunus, and Bahar — with whom I have shared many unforgettable adventures. I also want to thank my friends from Turkey: Mertcan, Lapaz, and Ali. I cannot believe we have been living in different cities since the end of high school, as I always feel their support by my side.

Last but not least, I would like to thank my family, whose love and support have been my greatest strength. This thesis is dedicated to my mother, Hatice, who always wanted me to be a doctor — though I know she meant a medical doctor, not this kind.

September 20, 2024 M.K.

Abstract

Machine learning and artificial intelligence methods have achieved remarkable success, matching and even surpassing human capabilities in various complex tasks. However, many demonstrations have generally neglected a critical part of the intelligence that is prevalent in the real world, namely, the one that emerges from the collective of interconnected individuals with diverse capabilities, perspectives and experiences.

To explore this fact, the current dissertation utilizes mathematical models of collaborative learning and reasoning. These models are based on the following two concepts: Bayesian inference, which is used to model how agents update their beliefs in the face of uncertain data, and graphs, which represent the communication links and information exchange among individuals.

Through these models, the current work examines the effect of dynamic models on learning, as well as the implications of causal interactions among agents on their decisions. In particular, this work is structured around (i) the effect of different information exchange procedures on learning, (ii) the need to adapt to changing environments, and (iii) the cause-and-effect relationships that arise among interacting agents over a graph. The net effect of our study is a collection of new results and design tools that strengthen our understanding of multi-agent networks.

A critical part of collaboration among agents is how information is exchanged among them. The first part of the dissertation examines how information is (i) fused and (ii) shared within a social network of interacting agents, and how these processes affect the learning capabilities of the network. In particular, the learning rates of the network are compared under both arithmetic and geometric fusion rules. The effect of network connectivity and information diversity is also clarified, in addition to the impact of random and partial information sharing.

The second part of the dissertation examines network behavior under changing environments, where the unknown state of nature is assumed to follow a hidden Markov model (HMM). The work examines the agents' ability to track the evolving state. It also considers the more challenging case of partially observable Markov decision processes

Abstract

(POMDP), where agents are able to take actions based on certain sequential policies. By acknowledging the uncertainties present in real-world scenarios and tackling them through cooperative state estimation, the methods devised in this work can facilitate the practical application of multi-agent reinforcement learning.

The first two parts of the dissertation address how procedural and environmental factors influence agent behavior. In the final part, the work focuses on the cause-and-effect relationships between agents. Specifically, causal inference tools are developed to determine how agents impact other agents' decisions and how this influence diffuses through the network. Expressions are derived for the total effects of agents over time in terms of the instantaneous direct effects, and an algorithm is devised to learn the causal effects from observational data.

Keywords: multi-agent networks, distributed Bayesian inference, information fusion, hidden Markov models, state tracking, partially observable Markov decision process, social influence, causal impact, social learning, multi-agent decision making

Résumé

Les méthodes d'apprentissage automatique et d'intelligence artificielle ont connu un succès remarquable, égalant et même surpassant les capacités humaines dans diverses tâches complexes. Cependant, de nombreuses démonstrations ont généralement négligé une partie essentielle de l'intelligence qui prévaut dans le monde réel, à savoir celle qui émerge du collectif d'individus interconnectés aux capacités, perspectives et expériences diversifiées.

Pour explorer cela, la présente thèse utilise des modèles mathématiques d'apprentissage collaboratif et de raisonnement. Ces modèles sont basés sur les deux concepts suivants : l'inférence bayésienne, utilisée pour modéliser la façon dont les agents mettent à jour leurs croyances face à des données incertaines, et les graphes, qui représentent les liens de communication et l'échange d'informations entre les individus.

À travers ces modèles, le présent travail examine l'effet des modèles dynamiques sur l'apprentissage, ainsi que les implications des interactions causales entre agents sur leurs décisions. En particulier, ce travail s'articule autour de (i) l'effet des différentes procédures d'échange d'informations sur l'apprentissage, (ii) la nécessité de s'adapter à des environnements changeants, et (iii) les relations de cause à effet qui surgissent entre les agents interagissant sur un graphe. Le résultat net de notre étude est une collection de nouveaux résultats et d'outils de conception qui renforcent notre compréhension des réseaux multi-agents.

Un aspect crucial de la collaboration entre agents est la façon dont l'information est échangée entre eux. La première partie de la thèse examine comment l'information est (i) fusionnée et (ii) partagée au sein d'un réseau social d'agents interagissants, et comment ces processus affectent les capacités d'apprentissage du réseau. En particulier, les taux d'apprentissage du réseau sont comparés selon des règles de fusion arithmétique et géométrique. L'effet de la connectivité du réseau et de la diversité de l'information est également clarifié, en plus de l'impact du partage d'informations aléatoire et partiel.

La deuxième partie de la thèse examine le comportement du réseau dans des environnements changeants, où l'état inconnu de la nature est supposé suivre un modèle de Markov caché (HMM). Le travail examine la capacité des agents à suivre l'évolution de

l'état. Il considère également le cas plus difficile des processus décisionnels de Markov partiellement observables (POMDP), où les agents peuvent prendre des actions basées sur certaines politiques séquentielles. En reconnaissant les incertitudes présentes dans les scénarios du monde réel et en les abordant par l'estimation coopérative d'état, les méthodes développées dans ce travail peuvent faciliter l'application pratique de l'apprentissage par renforcement multi-agents.

Les deux premières parties de la thèse abordent comment les facteurs procéduraux et environnementaux influencent le comportement des agents. Dans la dernière partie, le travail se concentre sur les relations de cause à effet entre agents. Plus précisément, des outils d'inférence causale sont développés pour déterminer comment les agents influencent les décisions des autres agents et comment cette influence se diffuse à travers le réseau. Des expressions sont dérivées pour les effets totaux des agents au fil du temps en termes d'effets directs instantanés, et un algorithme est conçu pour apprendre les effets causals à partir de données observationnelles.

Mots-clés : réseaux multi-agents, inférence bayésienne distribuée, fusion d'informations, modèles de Markov cachés, suivi d'état, processus décisionnels de Markov partiellement observables, influence sociale, impact causal, apprentissage social, prise de décision multi-agents

Notation

$\boldsymbol{x}$	Random variables are denoted in boldface.
$\mathbb{E}_x$	Expectation with respect to the distribution of $\boldsymbol{x}$.
$\mathcal{A}$	Sets and events in script-style letters
$\lvert\mathcal{A}\rvert$	Cardinality of set $\mathcal{A}$
Ω	Space of all data sequence realizations
ω	An element of Ω
$\mathcal{F}$	σ-field generated by the sequence of data
$\mathbb{P}$	Probability measure over sample paths $\omega \in \Omega$
$(\Omega, \mathcal{F}, \mathbb{P})$	Probability space
$D_{\mathrm{KL}}(p\|q)$	Kullback-Leibler (KL) divergence of distributions p and q
$\mathbb{1}_K$	All-ones vector of size K
$\mathbb{I}$	Indicator function: $\mathbb{I}\{B\} = 1$ if B is true, else 0
$\mu(\theta) \propto F(\theta)$	Normalization $\mu(\theta) = \dfrac{F(\theta)}{\sum_{\theta' \in \Theta} F(\theta')}$
A^{T}	Transpose of matrix A
A^{-1}	Inverse of matrix A
$A^{\dagger}$	Pseudo-inverse of matrix A
$\mathrm{rank}(A)$	Rank of matrix A
$\mathrm{RAN}(A)$	Range space of matrix A
$\mathrm{NULL}(A)$	Null space of matrix A
diag $\{$A,B$\}$	$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$

$\text{col}\{w_k\}_{k=1}^{K}$	$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_K \end{bmatrix}$
$v_k, [v]_k$	The k-th entry of vector v
$\otimes$	Kronecker product
i	Time index
$\lim_{i\to\infty} \boldsymbol{x}_i \overset{\text{a.s.}}{=} \boldsymbol{x}$	Almost sure convergence of a sequence $\{\boldsymbol{x}_i\}$ to $\boldsymbol{x}$
$\boldsymbol{x}_i \xrightarrow{\text{a.s.}} \boldsymbol{x}$	Simplified notation for almost sure convergence
$\boldsymbol{x}_i \xrightarrow{\text{dist.}} \boldsymbol{x}$	Convergence of a sequence of random variables $\{\boldsymbol{x}_i\}$ to $\boldsymbol{x}$ in distribution

Contents

1 Introduction

The real world is full of inherent uncertainties, where complete information about any given situation is often not available. Agents, whether they are humans or machines, need to make decisions based on their limited knowledge of their surrounding environment and act accordingly.

Consider, for instance, a trader in the foreign exchange market who considers whether to keep her money in dollars or convert it into euros or francs. The decision that would generate the highest profit in the long run depends on the current and future policies of central banks as well as global political developments. The trader cannot possibly know all the factors that influence market behavior beforehand.

Similarly, autonomous systems and machines face their own set of uncertainties. A self-driving car, for example, must confidently interpret traffic signals even in adverse conditions like fog, which reduces the intensity of the light that the car sensors receive. A misclassification here could have serious safety implications.

In response to these challenges, the trader might analyze past currency trends and market fluctuations to forecast future currency rates. The self-driving car might process data from a multitude of sensors to accurately estimate the traffic signals and make safe driving decisions. In general, rational agents, whether they are profit-seeking traders or autonomous vehicles, try to gather as much relevant information as possible and use it to minimize uncertainty. In other words, they try to form accurate *beliefs* about their environment.

In the following section, we introduce a classical mathematical framework for modeling rational belief formation in the face of uncertainty.

1.1 Bayesian Reasoning

In this section, we review Bayesian inference, which formalizes our earlier discussion on belief formation under uncertainty. This framework adopts a probabilistic approach to quantify uncertainties concerning the world. Here, an agent's belief is defined as a probability distribution over all possible states of the world.

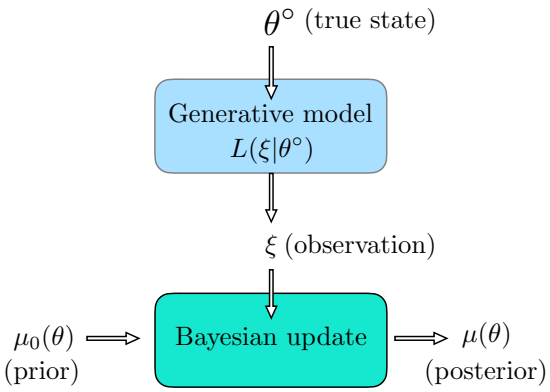


Figure 1.1: A block diagram representing a single Bayesian belief update.

More formally, a set of discrete hypotheses is represented as $\Theta = \{\theta_1, \theta_2, \ldots, \theta_H\}$, where each hypothesis $\theta \in \Theta$ represents a potential state of the world. The true state of the world, which is typically unknown, is denoted by θ° and is also a member of Θ.

The prior belief of an agent is a probability mass function (pmf) over the hypotheses set Θ, denoted by μ_0. For each hypothesis θ, $\mu_0(\theta)$ quantifies agent's initial confidence that θ corresponds to the hidden true state of the world θ°.

Bayes' theorem provides a formula for the integration of new observations into one's prior beliefs. Specifically, a rational agent revises its initial belief $\mu_0(\theta)$ encountering a new observation $\xi \sim L(\xi|\theta^o)$ as follows:

$$\mu(\theta) = \frac{L(\xi|\theta)\mu_0(\theta)}{\sum_{\theta' \in \Theta} L(\xi|\theta')\mu_0(\theta')}. \tag{1.1}$$

Here, μ denotes the posterior belief and accordingly, $\mu(\theta)$ indicates the revised confidence in the proposition that "$\theta^\circ = \theta$" after the observation of new data ξ — see Fig. 1.1 for a visual representation.

In the Bayes' update (1.1), the quantity $L(\xi|\theta)$ represents the likelihood of observing ξ if the true state θ° were equal to θ. For each θ, the term $L(\cdot|\theta)$ is the "likelihood" function under θ. Obviously, given that the true state is θ°, the observations ξ are generated according to the distribution $L(\cdot|\theta^\circ)$.

Illustration with an application

To illustrate the use of Bayesian inference in a practical setting, let us consider an autonomous vehicle approaching traffic signals. The vehicle must determine the color of the light in order to decide on an appropriate action, e.g., slowing down.

The possible states of the world correspond to the colors of the traffic light. That is to say,

$$\Theta = \{\text{●}, \text{●}, \text{●}\}. \tag{1.2}$$

where these states represent the colors red, yellow, and green, respectively.

Let us assume that the prior belief of the vehicle reflects that the duration of the red and green traffic lights are typically longer compared to the duration of the yellow light. Therefore, the vehicle's initial belief might be that the traffic light is most likely red or green, with a lower chance of being yellow. Then, the prior pmf for these states can be, say,

$$\mu_0(\text{●}) = 0.4, \quad \mu_0(\text{●}) = 0.2, \quad \mu_0(\text{●}) = 0.4. \tag{1.3}$$

Next, suppose that the vehicle observes ξ, which could be a partial view of the traffic signal or the behavior of other vehicles in proximity. The likelihoods of this observation given each possible state of the traffic light might be:

$$L(\xi|\text{●}) = 0.7, \quad L(\xi|\text{●}) = 0.2, \quad L(\xi|\text{●}) = 0.1 \tag{1.4}$$

The vehicle can have processors on-board to compute (or estimate) these likelihood values with, say, neural networks that are trained offline. In light of (1.1), the posterior belief for the red light can be obtained by calculating

$$\mu(\text{●}) = \frac{L(\xi|\text{●})\mu_0(\text{●})}{L(\xi|\text{●})\mu_0(\text{●}) + L(\xi|\text{●})\mu_0(\text{●}) + L(\xi|\text{●})\mu_0(\text{●})}, \tag{1.5}$$

and similarly for the other lights. Inserting the numbers for the denominator yields:

$$\begin{aligned} L(\xi|\text{●})\mu_0(\text{●}) + L(\xi|\text{●})\mu_0(\text{●}) + L(\xi|\text{●})\mu_0(\text{●}) &= 0.7 \times 0.4 + 0.2 \times 0.2 + 0.1 \times 0.4 \\ &= 0.36. \end{aligned}$$

If we substitute this result into (1.5), we get:

$$\mu(\text{●}) = \frac{0.7 \times 0.4}{0.36} \approx 0.78. \tag{1.6}$$

Similarly, the posterior beliefs for the yellow and green lights are given by

$$\begin{aligned} \mu(\text{●}) &= \frac{L(\xi|\text{●})\mu_0(\text{●})}{0.36} = \frac{0.2 \times 0.2}{0.36} \approx 0.11 \\ \mu(\text{●}) &= \frac{L(\xi|\text{●})\mu_0(\text{●})}{0.36} = \frac{0.1 \times 0.4}{0.36} \approx 0.11. \end{aligned}$$

Therefore, the prior belief in (1.3) is adjusted to the posterior belief

$$\mu(\text{●}) = 0.78, \quad \mu(\text{●}) = 0.11, \quad \mu(\text{●}) = 0.11. \tag{1.7}$$

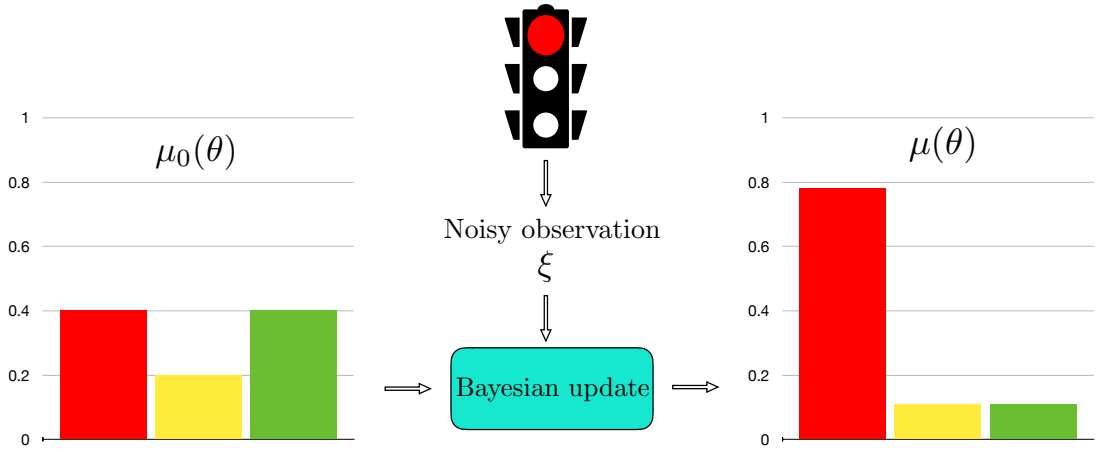


Figure 1.2: A single Bayesian update on the beliefs over traffic light colors.

Notice that the vehicle's belief on the traffic light being red has increased significantly from 0.4 to 0.78 after incorporating the new observation ξ. Figure 1.2 depicts this change in the belief.

1.2 Sequential Belief Updates

In real-world scenarios, observations or data often arrive sequentially over time rather than simultaneously. In such environments, the Bayesian framework discussed in the previous section can be applied iteratively, allowing beliefs to be continuously improved as new observations become available.

Formally, we consider a temporally ordered sequence of observations that are denoted by $\xi_1, \xi_2, \ldots, \xi_i$. Here, the index i denotes the particular time instant at which the observation ξ_i becomes available to the agent.

In the sequential framework, the belief revision becomes a recursive application of Bayes update (1.1). At each time instant i, the posterior belief from the previous time $i-1$, denoted as μ_{i-1}, is taken as the prior for the current time i. In other words, the posterior μ_i at time i is obtained by using μ_{i-1} and ξ_i as follows (*cf.* (1.1))

$$\mu_i(\theta) = \frac{L(\xi_i|\theta)\mu_{i-1}(\theta)}{\sum_{\theta' \in \Theta} L(\xi_i|\theta')\mu_{i-1}(\theta')}, \tag{1.8}$$

according to (1.1) — see Fig. 1.3 for an illustration. The expectation is that as the observations accumulate more evidence over time, the belief should become more and more concentrated around the latent true state of the world.

This idea was first discussed in Thomas Bayes' seminal work [1] that was published posthumously by Richard Price in 1761. When Price published his late friend Bayes'

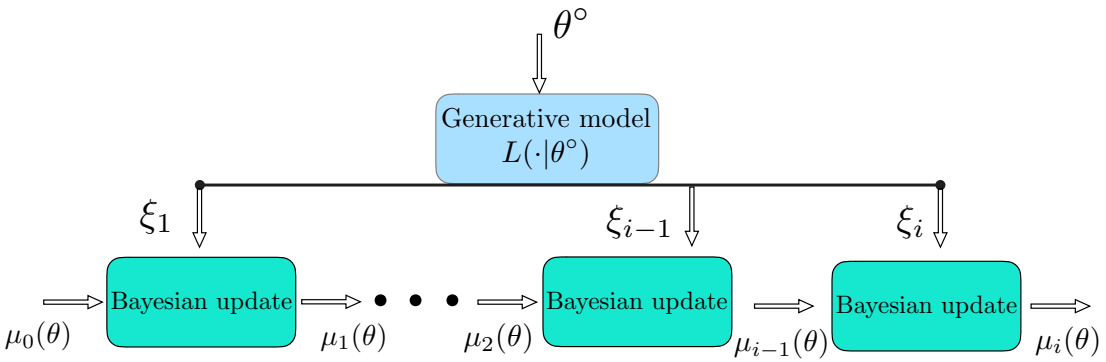


Figure 1.3: A block diagram representing sequential Bayesian belief updates over time.

ideas, he also incorporated a thought experiment to illustrate the sequential reasoning fundamental to Bayesian inference. Specifically, he imagined a scenario where an individual observes the sunrise multiple times and uses these observations to infer the chance of the sun rising again the next day, which exemplifies the process of sequential Bayesian reasoning discussed above. The original words of this example in [1] are as follows.

"... Let us imagine to ourselves the case of a person just brought forth into this, world and left to collect from his observations the order and course of events what powers and causes take place in it. The Sun would, probably, be the first object that would engage his attention; but after losing it the first night he would be entirely ignorant whether he should ever see it again. He would therefore be in the condition of a person making a first experiment about an event entirely unknown to him. But let him see a second appearance or one return of the Sun, and an expectation would be raised in him of a second return, and he might know that there was an odds of 3 to 1 for some probability of this. This odds would increase, as before represented, with the number of returns to which he was witness..."

In the context of our traffic light example, as the vehicle approaches the traffic signals, its sensors gather increasing amount of data — see Fig. 1.4 for an illustration. Each new observation can be used to update the vehicle's belief about the color of the traffic light. To illustrate the use of sequential updates in this practical setting, let us assume that after the first observation and the belief update in (1.3)–(1.7), the vehicle encounters another observation ξ_2. The likelihood values for ξ_2 can be, say,

$$L(\xi_2|\,\bullet) = 0.8, \quad L(\xi_2|\,\bullet) = 0.15, \quad L(\xi_2|\,\bullet) = 0.05. \tag{1.9}$$

If we incorporate ξ_2 to the posterior belief (1.7) of the first iteration that we now denote

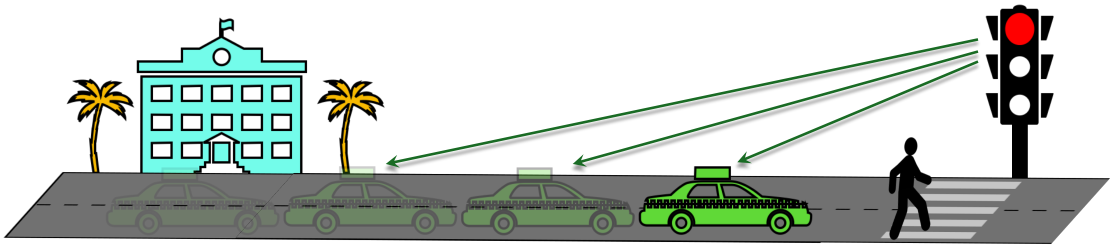


Figure 1.4: A vehicle getting more observations as it approaches to a traffic signal.

by μ_1, we arrive at the Bayes update

$$\begin{aligned}\mu_2(\color{red}\bullet) &= \frac{L(\xi_2|\color{red}\bullet)\mu_1(\color{red}\bullet)}{L(\xi_2|\color{red}\bullet)\mu_1(\color{red}\bullet) + L(\xi_2|\color{yellow}\bullet)\mu_1(\color{yellow}\bullet) + L(\xi_2|\color{green}\bullet)\mu_1(\color{green}\bullet)} \\ &= \frac{0.8 \times 0.78}{0.8 \times 0.78 + 0.15 \times 0.11 + 0.05 \times 0.11} \\ &\approx 0.96 \end{aligned} \tag{1.10}$$

Performing the same calculations for the other states as well, the posterior belief μ_2 can be found as

$$\mu_2(\color{red}\bullet) \approx 0.96, \quad \mu_2(\color{yellow}\bullet) = \frac{0.15 \times 0.11}{0.646} \approx 0.03, \quad \mu_2(\color{green}\bullet) = \frac{0.05 \times 0.11}{0.646} \approx 0.01. \tag{1.11}$$

Assume that another observation ξ_3 arrives for which the likelihood values again favor the red light, e.g.,

$$L(\xi_3|\color{red}\bullet) = 0.9, \quad L(\xi_3|\color{yellow}\bullet) = 0.05, \quad L(\xi_3|\color{green}\bullet) = 0.05. \tag{1.12}$$

The vehicle can update its belief similarly

$$\begin{aligned}\mu_3(\color{red}\bullet) &= \frac{L(\xi_3|\color{red}\bullet)\mu_2(\color{red}\bullet)}{L(\xi_3|\color{red}\bullet)\mu_2(\color{red}\bullet) + L(\xi_3|\color{yellow}\bullet)\mu_2(\color{yellow}\bullet) + L(\xi_3|\color{green}\bullet)\mu_2(\color{green}\bullet)} \\ &= \frac{0.9 \times 0.96}{0.9 \times 0.96 + 0.05 \times 0.03 + 0.05 \times 0.01} \\ &\approx 0.995, \end{aligned} \tag{1.13}$$

and arrives at the posterior belief

$$\mu_3(\color{red}\bullet) \approx 0.996 \quad \mu_3(\color{yellow}\bullet) \approx 0.003, \quad \mu_3(\color{green}\bullet) \approx 0.001. \tag{1.14}$$

Notice that with each observation ξ_i, the vehicle's belief μ_i converges towards the state of the traffic light being red. It reflects a higher confidence in this prediction as it accumulates more data supporting this hypothesis.

1.3 Social Reasoning

In the thought experiment from [1] discussed earlier, where a person just brought into the world forms beliefs by observing the sun, an important element is missing: social interactions. Humans do not exist in isolation; rather, they are continually and socially influenced by their environment and the people around them. In other words, *our beliefs are also shaped by others' beliefs*. Similarly, animals learn critical survival strategies, such as locating food or avoiding predators, from their elders.

These principles apply to autonomous systems as well. Consider our recurring example of a car attempting to predict the color of a traffic light. Typically, this scenario involves multiple vehicles on the road, as illustrated in Fig. 1.5, each equipped with advanced sensors and onboard processing capabilities. Despite these technological advances, each vehicle has inherently limited perception abilities confined by the physical and computational constraints of its sensors and algorithms. Through cooperation, these vehicles can achieve a collective understanding and predictive power far superior to that of any single vehicle. This form of social reasoning enables more effective synchronization of actions, which is essential for maintaining smooth traffic flow and preventing collisions. Similarly, the concept of collective reasoning extends beyond transportation to other socio-technical and cyber-physical systems such as wireless networks, power grids, and urban infrastructures, where communication and cooperation among various components are critical.

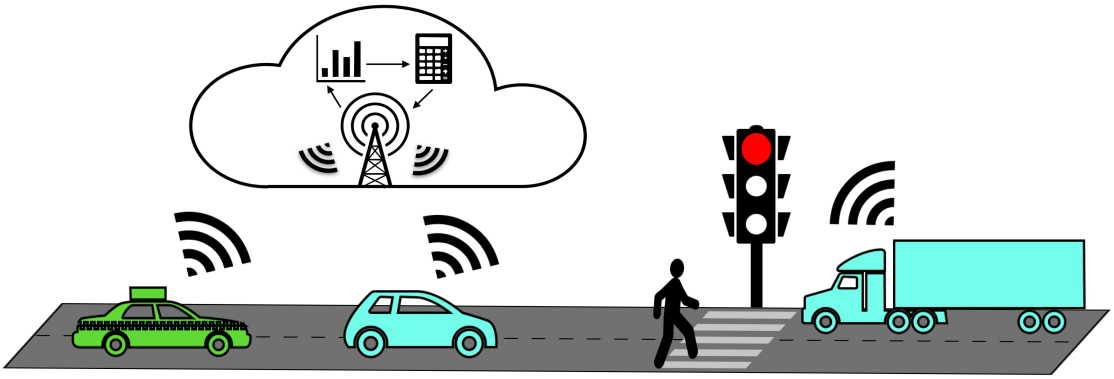


Figure 1.5: Intelligent vehicles and infrastructure can collaborate to enhance awareness of road conditions. Real-time and spontaneous cooperation is crucial in this context, as it allows for immediate responses to dynamic conditions, and hence improving the safety and efficiency of transportation.

In general, a rational agent uses the maximum available information to make the optimal decisions. In social settings, this information not only stems from one's own observations but also from the insights and beliefs of others. Building on this notion, in the current thesis, we will focus on belief formation in social settings, which is a topic of heightened interest in the literature. Namely, we will consider multi-agent sys-

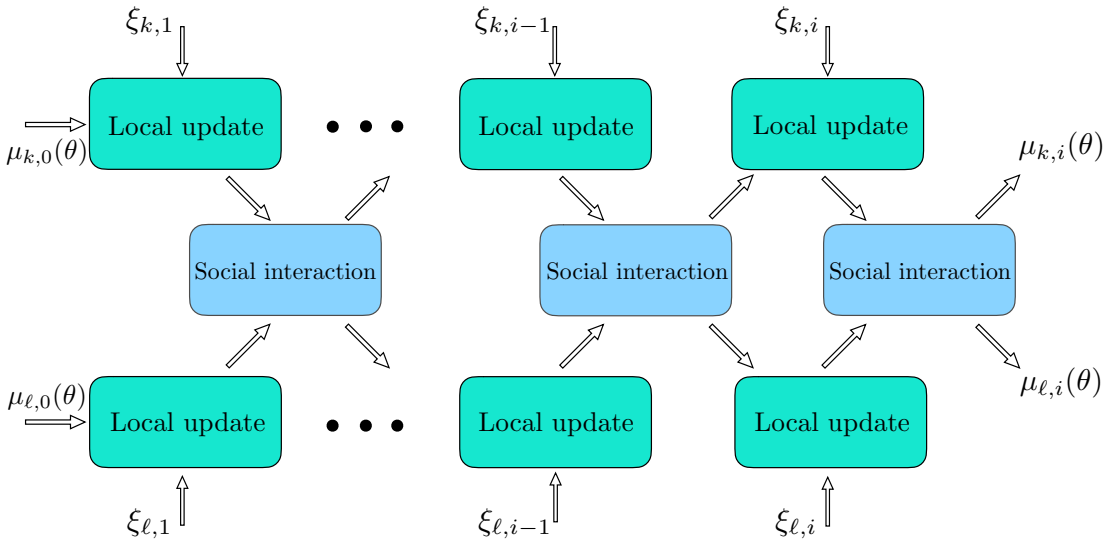


Figure 1.6: A block diagram representing belief updates and social interactions unrolled over time.

tems, where agents share a common world and a generative model of that world, and cooperate to form more precise beliefs about their environment, thereby enhancing inference and learning.

To better examine how social systems interact with each other, let us adjust our earlier notation to accommodate multiple agents. We consider a group of K agents. Each agent k maintains a belief $\mu_{k,i}$ about the common state of the world, with $\mu_{k,i}(\theta)$ representing the confidence agent k has at time i in θ being the true state θ°. We denote the personal (or private) observation agent k receives at time i by $\xi_{k,i}$, which is distributed according to the likelihood function $L_k(\cdot|\theta^\circ)$.

For visualization purposes, we refer to the block diagram in Fig. 1.6, which depicts consecutive local belief updates and interactions between two agents k and ℓ. While this diagram is useful for an initial understanding, in practice, multi-agent systems have complexities that far exceed those of two-agent models. As such, we dedicate Chapter 2 for describing our social learning model while introducing necessary graph and decision theoretical tools. But before proceeding to Chapter 2, we first discuss the thesis goal and contributions at a high level in the next two sections.

1.4 Revisiting the Title

After setting the stage in the previous sections, revisiting each word of our title, "Sequential Reasoning with Socially Caused Beliefs", can help clarify the goals of our thesis:

SEQUENTIAL refers to the process of continually updating beliefs based on streaming observations that arrive over time. In this thesis, we use the recursive Bayesian updates discussed in Section 1.2 as a model.

REASONING refers to making sense of data under uncertainty through likelihoods and beliefs. In this work, we consider probabilistic (or statistical) reasoning through Bayes' rule, as discussed in Section 1.1.

SOCIALLY refers to the multi-agent, cooperative, and collective dimensions of reasoning. As explored in Section 1.3, these interactions can enhance the capacity of individual agents to form more accurate beliefs of the world.

CAUSED refers to how agents' beliefs are formed and adjusted. In the context of this thesis, "socially caused" refers to the beliefs that are influenced by the interactions within a social network.

BELIEFS refer to the confidence agents have on possible hypotheses being true about their environment. In this thesis, the terms "opinion" and "soft-decision" are used interchangeably with the term "belief". The former is borrowed from behavioral contexts, while the latter is used to highlight decisions involving a range of possibilities, and not just binary outcomes.

Overall, the primary goal of this thesis is to understand and enhance collective decision-making in uncertain environments. We aim to merge sequential Bayesian reasoning methods with the networked interactions. Our approach extends previous studies by incorporating and exploiting Markovian models for the evolution of the state, and by examining the causal effects that arise among agents. The thesis also examines enhanced policies for fusing information and opens up applications in the context of reinforcement learning.

1.5 Thesis Contributions

In particular, the dissertation makes contributions in three key areas of sequential reasoning with socially caused beliefs: information exchange, dynamic environments, and causal influence. Accordingly, we have structured our thesis into three parts, each focusing on one of these areas.

1.5.1 Part I: Information Exchange

The first part of the thesis addresses the information fusion and sharing aspects within networked interactions.

Chapter 3: Information Fusion

A foundational element of social systems is the aggregation of diverse information sources. In Chapter 3, we provide a theoretical framework that compares the effectiveness of arithmetic and geometric averaging based fusion rules, and evaluates them based on their learning rates. We establish quantitative bounds in terms of network connectivity and the diversity of information across the network. These bounds can help practitioners select the appropriate fusion strategy for various applications.

Chapter 4: Information Sharing

The communication patterns that networked systems exhibit can significantly influence the behavior of the system. In Chapter 4, we investigate a pattern where the nature of information sharing is random. Under this pattern, we provide conditions under which agents can successfully learn the truth or risk mislearning. The theoretical results yield useful insights into the impact of communication protocols and frequency on the speed of collective inference.

1.5.2 Part II: Dynamic Environments

The second part of the thesis focuses on dynamic environments, where state of nature evolves over time. Here, the goal is not to learn a static hidden parameter but to track its changes over time.

Chapter 5: Markovian States

In Chapter 5, we use hidden Markov models (HMMs) to represent the changes in the true state of nature. We extend previous works on networked belief formation for inferring static states to tracking dynamic states. Additionally, we provide theoretical performance comparisons between an information theoretically optimal strategy and the decentralized strategy we propose.

Chapter 6: Policy Evaluation in Dec-POMDPs

In Chapter 6, we extend the framework of Chapter 5 to include partially observable Markov decision processes (POMDPs), where agents perform actions based on their beliefs. This extension is particularly relevant for real-world applications of multi-agent reinforcement learning, where the complete knowledge of the state is rarely available.

1.5.3 Part III: Causal Influence

The final part of the thesis explores how agents' beliefs are causally influenced by others.

Chapter 7: Causality in Social Networks

In Chapter 7, we establish a theoretical framework that reveals the causal relations between agents, distinguishing it from merely correlational associations. Additionally, we introduce algorithms that rank agents based on their influence and propose methods for learning causal relationships from observational data.

Chapter 8: Causality under Asynchronicity

Motivated by the fact that real-world applications exhibit significant asynchronous behavior, in Chapter 8, we extend our discussions from Chapter 7 to understand how different communication participation patterns and fusion policies change causal influence.

2 Social Reasoning

In the previous chapter, we focused on single-agent Bayesian reasoning and discussed its extension to multi-agent scenarios. In this chapter, we build on the concepts introduced previously and review a standard graph-theoretical approach to model the dynamics of social interactions. This chapter sets the stage for our contributions in the following sections by introducing the necessary notation and outlining the modeling conditions.

2.1 Network Topology

In this section, we describe the graphical models [2–8] used to represent the network architectures considered in this thesis. Graphs are ideal for modeling networks; for example, in a social network, graph nodes can represent users, while edges represent the relationships and communication links between them. This facilitates a spatio-temporal representation of the network.

There are two main architectures of interest depending on the communication topology of agents.

2.1.1 Decentralized Peer-to-Peer Networks

Peer-to-peer networks are based on localized interactions and are fully decentralized, meaning there is no central unit. In these networks, agents perform local adaptation and integrate information from their immediate neighbors [7, 9].

In particular, we consider strongly-connected networks [7,8] — see Fig. 2.1. This means that there exists a path linking any agent ℓ to any other agent k, which starts at ℓ and ends at k. Moreover, there exists at least one agent k° in the graph with a self-loop, i.e., an agent that does not discard its own observations and has positive self-reliance.

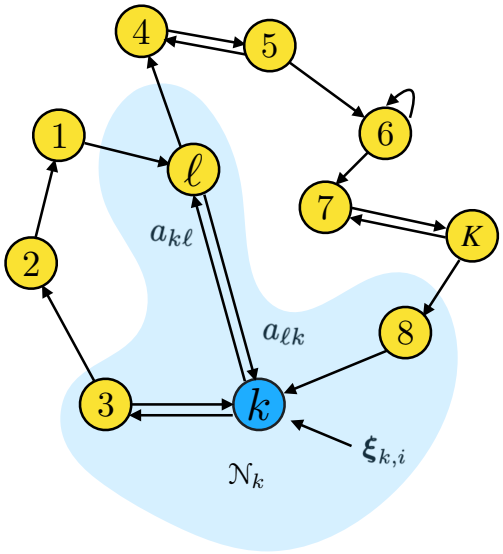


Figure 2.1: An illustration of a strongly connected decentralized network topology.

These conditions allow the information to diffuse across the network thoroughly.

We associate a non-negative combination coefficient $a_{\ell k}$ with the link from ℓ to k. This entry denotes the confidence score that agent k assigns to the information received from agent ℓ. The score $a_{\ell k}$ is positive if, and only if, agent ℓ is an immediate neighbor of agent k. Otherwise it is equal to zero. We denote the neighborhood of agent k by $\mathcal{N}_k$. In other words, $a_{\ell k} > 0$ if, and only if, $\ell \in \mathcal{N}_k$. Assuming K agents, we collect the weights over the graph into the $K \times K$ combination matrix $A = [a_{\ell k}]$.

The graph underlying the network is directed and hence the combination matrix A is not necessarily symmetric, i.e., in general $a_{\ell k} \neq a_{k\ell}$. Nevertheless, the confidence scores $a_{\ell k}$ that agent k assigns to its neighbors in $\mathcal{N}_k$ should add up to one. This means that the matrix A is a left-stochastic matrix, meaning the entries on each of its columns add up to 1, i.e.,

$$A^{\mathsf{T}}\mathbb{1}_K = \mathbb{1}_K. \tag{2.1}$$

The strong-connectedness of the graph translates into A being a primitive matrix [7]. In terms of Markov chain terminology [10], this means that A is an irreducible and aperiodic Markov kernel. Therefore, in view of the Perron-Frobenius theorem [11], the matrix A will have a unique eigenvalue at 1, and the corresponding eigenvector v satisfies:

$$Av = v, \quad \mathbb{1}_K^{\mathsf{T}} v = 1, \quad v_k > 0. \tag{2.2}$$

The eigenvector v is called the Perron vector of A and all its entries are positive and normalized to add up to 1. The k-th entry of v, denoted by v_k, measures the relative network centrality of agent k in the network.

2.1.2 Architectures with Fusion Center

In this architecture, agents communicate with and through a central node instead of direct communication with other agents. The agents send their information such as raw data, opinion or belief about the environment, to a fusion center. The fusion center combines the information received from peripheral agents and may broadcast the aggregated information back to agents — see Fig. 2.2 for an illustration.

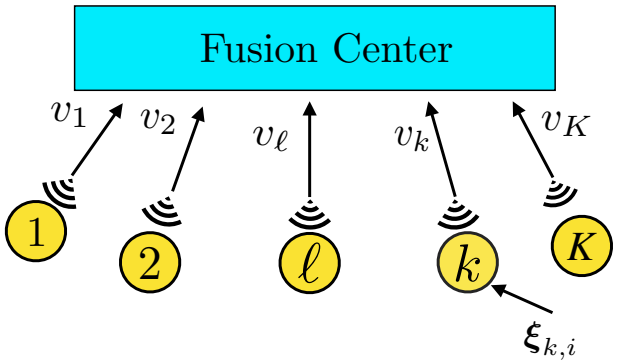


Figure 2.2: An illustration of a federated architecture.

In the figure, we display confidence weights denoted by v_k, which are used by the central unit to fuse the information arriving from the agents. We collect these weights into the vector of confidence weights with $v \triangleq [v_1, \ldots, v_K]^{\mathsf{T}}$. The k-th entry $v_k \in (0,1)$, which is assigned by the fusion center to agent k [7, 12], is usually constructed based on the previous interactions with the agents. These weights are assumed to be positive constants that add up to 1.

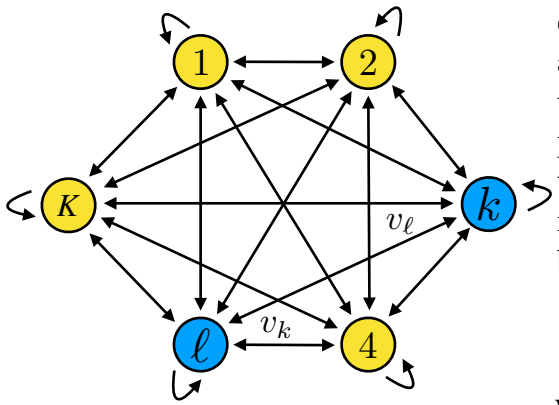


Figure 2.3: A fully connected network topology.

Notice the parallel between the definition of the confidence vector v for the federated architecture and the Perron eigenvector for the decentralized architecture in (2.2). In particular, note that, the federated architecture is equivalent to a fully-connected network as in Fig. 2.3 with a rank-one combination matrix:

$$A = v\mathbb{1}_K^{\mathsf{T}}, \tag{2.3}$$

whose Perron vector coincides with the confidence vector, v. Therefore, in the sequel, we focus on general combination matrices A described for decentralized networks, which also cover federated architectures as a special case. Whenever it yields additional results and insights on top of the general case, we constrain A to be a

rank-one matrix.

2.2 Inference Framework

In this section, we explain how the Bayesian reasoning framework presented in the previous chapter has been extended to distributed architectures. By inference, we refer to the challenge of predicting hidden variables from partially informative observations, data, or features.

Thus, consider a group $\mathcal{N}$ of K agents (e.g., social media users, sensors, or machines) that are trying to discover the hypothesis that best explains their observations of the world. More formally, these agents aim to distinguish the true state of nature θ° from a finite set of H hypotheses or states, denoted as $\Theta = \{\theta_1, \theta_2, \ldots, \theta_H\}$. In this thesis, we focus on categorical variables that belong to sets with finite cardinality, as opposed to estimating continuous-valued random variables.

Illustratively, in a behavioral context, voters might aim to reach a consensus on the best political representative from a set of candidates {A, B, C, D}, influenced by personal biases and their interactions over a social network. Similarly, in engineering systems, a group of vehicles might cooperate to determine the color of a traffic light among the set of possibilities $\Theta = \{$🔴, 🟡, 🟢$\}$.

At each time instant i, each agent k receives a personal and private observation $\xi_{k,i}$, which we assume to encapsulate all out-of-network information available to k at time i. It is distributed according to some marginal likelihood function $L_k(\xi|\theta^\circ)$ dependent on the true state θ°. The social interactions and cooperations are of interest for agents. This is because observations typically provide partial information, which may be limited or distorted by each agent's potentially restricted or noisy perspective on the overall phenomenon.

In general, the observations $\boldsymbol{\xi}_{k,i}$ are assumed independent and identically distributed (i.i.d.) over time, and the likelihood functions $L_k(\xi|\theta^\circ)$ are assumed to be time-invariant. However, the observations are not necessarily independent across the agents. This is common in applications where ensuring spatial independence of observations is impractical. In such environments, there might be confounding factors that affect multiple agents simultaneously and induce correlation among the observations. In order to emphasize the random nature of the observations, we will use the boldface notation $\boldsymbol{\xi}_{k,i}$ from now on. This randomness is also reflected on the belief vectors, which will similarly denote in boldface and write as $\boldsymbol{\mu}_{k,i}$.

Each agent k knows the likelihood model $L_k(\xi|\theta)$ for every possible hypothesis $\theta \in \Theta$. If the distribution $L_k(\xi|\theta)$ for a hypothesis $\theta \neq \theta^\circ$ is sufficiently different from $L_k(\xi|\theta^\circ)$, then it is easier for agent k to distinguish θ° from θ. This "distinguishing power" can be

quantified using the Kullback-Leibler divergence [13] between the likelihood models, namely,

$$d_k(\theta) \triangleq D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)||L_k(\cdot|\theta)\Big). \tag{2.4}$$

The larger this quantity is, the more *informative* agent k's observations are for distinguishing a wrong hypothesis θ from the true hypothesis θ°.

Based on the local observations and on interactions with other agents, each agent k forms a belief vector, denoted by $\boldsymbol{\mu}_{k,i}$, which is a probability mass function defined over the set of hypotheses Θ. Here, the entry $\boldsymbol{\mu}_{k,i}(\theta)$ quantifies the confidence k has about $\theta \in \Theta$ being the true hypothesis θ° at time i. Agents exchange their beliefs with each other according to some communication topology from the previous section. The way we model their information exchange will be detailed in Section 2.3.

In this work, we are interested in whether individual agents can discern (learn) the true state θ° from other hypotheses. We distinguish three possible scenarios on this matter.

Definition 2.1 (**Learning and mislearning**). *We say that truth learning occurs whenever:*

$$\boldsymbol{\mu}_{k,i}(\theta^\circ) \xrightarrow{\text{a.s.}} 1, \qquad \forall k \in \mathcal{N} \tag{2.5}$$

In other words, truth learning means that agents become fully confident on the true hypothesis with probability one. In any other scenario we say that there is no learning. Among the cases where there is no learning, we define total mislearning as corresponding to the case in which for some $\theta \in \Theta \setminus \{\theta^\circ\}$:

$$\boldsymbol{\mu}_{k,i}(\theta) \xrightarrow{\text{a.s.}} 1, \qquad \forall k \in \mathcal{N} \tag{2.6}$$

In this case, agents become fully confident on a wrong hypothesis, with probability one.

We emphasize that according to Definition 2.1, learning and mislearning occur whenever agents become fully confident on some hypothesis. In traditional Bayesian learning, the belief incorporates knowledge from an increasing number of measurements, with the expectation that the belief is not only maximized around the truth but also concentrates its mass on it. Any other outcome would imply a system malfunction. The same asymptotic learning outcome is expected in our definition as well. An infinite amount of evidence should lead to certainty. An alternative definition of learning could be to require that the agents' beliefs are maximized at the true hypothesis [14]. We adopt the former definition for consistency with other traditional works in the field.

2.3 Social Learning

In this section, we review social learning models, which are mathematical models designed to analyze how agents form opinions and make decisions within a network defined by a graph topology [14–26].

In particular, our focus is on *locally* Bayesian social learning strategies [21–25], which are also known as non-Bayesian strategies in the literature. While there exist *fully* Bayesian social learning strategies seeking to form the global Bayesian posterior in a decentralized manner [20], they nevertheless necessitate extensive knowledge about other agents, such as their likelihood functions and network topology. Even under this extensive knowledge, achieving the global Bayesian posterior is known to be NP-hard [27]. In contrast, non-Bayesian social learning strategies rely solely on the localized processing of data and on localized interactions, and have been shown to allow the inference of the true state of nature [21, 22]. Furthermore, these strategies are better suited to real-world scenarios compared to fully Bayesian strategies for at least two reasons. First, from a behavioral perspective, non-Bayesian strategies agree with the theory of bounded rationality in human decision making [28, 29]. Second, from an engineering perspective, they allow fully decentralized designs with moderate complexity and efficient memory.

2.3.1 Non-Bayesian Social Learning

Non-Bayesian social learning algorithms [21–24] are based on consecutive local update and information exchange steps, which are detailed next.

Local Update

At each time instant i, each agent first incorporates its personal observation $\boldsymbol{\xi}_{k,i}$ to its belief in a *locally Bayesian* fashion to obtain its intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta). \tag{2.7}$$

Here, the proportionality sign $\propto$ means that the entries of the resulting vector are normalized to add up to 1, as befits a true probability mass function. That is, expression (2.7) is a compact notation for

$$\boldsymbol{\psi}_{k,i}(\theta) = \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta)}{\sum_{\theta'\in\Theta} L_k(\boldsymbol{\xi}_{k,i}|\theta')\boldsymbol{\mu}_{k,i-1}(\theta')}. \tag{2.8}$$

The motivation for the local Bayesian update step (2.7) is at least two-fold. From a behavioral point of view, Bayes's rule is used to model human reasoning under uncertainty in neuroscience [30] and the social sciences [4, 31]. From a system design

perspective, Bayes's rule is known to be an optimal information processing rule [32].

Information Exchange

In the next step, the intermediate beliefs are shared with other agents, which may average them in a geometric manner to form the updated belief using the confidence scores they assign to their neighbors as follows:

$$\boldsymbol{\mu}_{k,i}(\theta) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(\theta)\Big)^{a_{\ell k}}. \tag{2.9}$$

Again, the proportionality sign $\propto$ implies that (2.9) is equivalent to

$$\boldsymbol{\mu}_{k,i}(\theta) = \frac{\prod_{\ell \in \mathcal{N}_k} (\boldsymbol{\psi}_{\ell,i}(\theta))^{a_{\ell k}}}{\sum_{\theta' \in \Theta} \prod_{\ell \in \mathcal{N}_k} (\boldsymbol{\psi}_{\ell,i}(\theta'))^{a_{\ell k}}}. \tag{2.10}$$

In (2.10), the term in the numerator is a weighted geometric average of the received intermediate beliefs, and the denominator is a normalization term that makes sure the belief vector $\boldsymbol{\mu}_{k,i}$ is a pmf over Θ.

The combination step (2.9) is a non-Bayesian way of combining beliefs and is inspired by the fact that humans are boundedly rational [29]. In the above implementation, the agents are combining their neighbors' *instantaneous* opinions, as opposed to behaving in a *fully* Bayesian manner [20], which would require global information (such as the graph topology and access to all observations) and is computationally intractable in general [27].

In geometric social learning, steps (2.7) and (2.9) are repeated at each time instant. We summarize the resulting procedure in Algorithm 2.1. Nevertheless, there are other variations of non-Bayesian social learning algorithms which we describe in the next section.

2.3.2 Model Variations

In the literature, there are other non-Bayesian social learning algorithms with different information pooling strategies or local adaptation steps. In this section, we introduce three of them that are commonly encountered in the literature.

Arithmetic social learning

In (2.9), the fusion of beliefs is based on geometric averaging (GA), which is also known as logarithmic pooling [33]. There are variations where the combination step is based

Algorithm 2.1 Geometric social learning

while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta) \tag{2.11}$$

  end for
  agents **exchange** intermediate beliefs with their neighbors
  for each agent $k = 1, 2, \ldots, K$ **do**
    combine received beliefs with **geometric averaging** (GA) rule:

$$\boldsymbol{\mu}_{k,i}(\theta) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(\theta)\Big)^{a_{\ell k}} \tag{2.12}$$

  end for
  $i \leftarrow i + 1$
end while

on an arithmetic averaging operation [22]. Namely, step (2.9) is replaced by

$$\boldsymbol{\mu}_{k,i}(\theta) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k}\boldsymbol{\psi}_{\ell,i}(\theta) \tag{2.13}$$

where arithmetic averaging (AA) [34] replaces GA — see Algorithm 2.2. These two averaging strategies possess distinct properties [35] and the choice of which to use for social learning is dependent upon the specific application under consideration. Here, we comment briefly on some features of AA and GA; the reader may refer to [35] for other useful characteristics.

While GA is externally Bayesian, namely, the Bayesian updates on the beliefs and the aggregation step are commutative, it nevertheless gives agents a veto power. In other words, if one agent has zero belief on some hypothesis, then the composite belief will also be zero on that same hypothesis regardless of the beliefs by the other agents. The AA rule, on the other hand, does not give this much power to individual agents and hence can be more robust against adversarial attacks. For this same reason, the support of the averaged belief vectors is the union of the supports in the AA case, while it is the intersection of the supports in the GA case. Hence, in GA, all agents need to have positive initial beliefs for all hypotheses in order not to discard any hypothesis. In contrast, it is enough for at least one agent to have a positive initial belief for any hypothesis in order not to get discarded in AA. Moreover, in the case of Gaussian beliefs over continuous hidden variables, repeated application of GA preserves Gaussianity and is related to the covariance-intersection method [36]. In contrast, the properties of AA-combined distributions diverge from the Gaussian distributions.

Algorithm 2.2 Arithmetic social learning

while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta) \tag{2.14}$$

  end for
  agents **exchange** intermediate beliefs with their neighbors
  for each agent $k = 1, 2, \ldots, K$ **do**
    combine received beliefs with **arithmetic averaging** (AA) rule:

$$\boldsymbol{\mu}_{k,i}(\theta) = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \boldsymbol{\psi}_{\ell,i}(\theta) \tag{2.15}$$

  end for
  $i \leftarrow i + 1$
end while

In general, it is more common to use geometric averaging for fusing probability density functions, while arithmetic averaging is typically favored for combining point estimates of random variables [37]. However, this method of application often relies on intuitive reasoning, and theoretical results are notably scarce, particularly in the context of social learning.

Adaptive social learning

The traditional non-Bayesian social learning (NBSL) strategy in (2.7)–(2.9) has the drawback that agents do not prioritize new observations against their old observations. In addition to falling short in modelling human behavior, this strategy can be disadvantageous for engineering applications that require adaptation under non-stationary environments. To tackle this issue, the work [14] proposed the adaptive social learning strategy by changing the adaptation step (2.7) into[1]

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k^{\beta}(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}^{1-\delta}(\theta), \tag{2.16}$$

where $0 < \delta < 1$ and $\beta > 0$ are design parameters. In particular, large values of δ or β place more focus on new observations, whereas small values give importance to past beliefs. The resulting algorithm is listed in Algorithm 2.3.

[1] In fact, [14] only considers the special cases of $\beta = \delta$ and $\beta = 1$. However, their results can be adapted to general $\beta > 0$ straightforwardly.

Algorithm 2.3 Adaptive social learning (ASL)

while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k^{\beta}(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}^{1-\delta}(\theta), \tag{2.17}$$

    where $0 < \delta < 1$ and $\beta > 0$ are design parameters.
  end for
  agents **exchange** intermediate beliefs with their neighbors
  for each agent $k = 1, 2, \ldots, K$ **do**
    combine received beliefs with **geometric averaging** (GA) rule:

$$\boldsymbol{\mu}_{k,i}(\theta) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(\theta)\Big)^{a_{\ell k}} \tag{2.18}$$

  end for
  $i \leftarrow i + 1$
end while

Consensus-based social learning

The algorithms we have discussed so far are motivated by diffusion strategies over graphs [7]. There are also variations based on the consensus strategy [38], where agents combine their intermediate beliefs with the previous beliefs of the neighbors.

For instance, the combination step of the geometric averaging based consensus algorithm replaces (2.9) by

$$\boldsymbol{\mu}_{k,i}(\theta) \propto (\boldsymbol{\psi}_{k,i}(\theta))^{a_{kk}} \prod_{\ell \in \mathcal{N}_k \backslash \{k\}} \Big(\boldsymbol{\psi}_{\ell,i}(\theta)\Big)^{a_{\ell k}} \qquad \text{(GA-Consensus)} \tag{2.19}$$

Similarly, in lieu of (2.13), the combination step in the arithmetic averaging based consensus algorithm [21] is given by

$$\boldsymbol{\mu}_{k,i}(\theta) = a_{kk}\boldsymbol{\psi}_{k,i}(\theta) + \sum_{\ell \in \mathcal{N}_k \backslash \{k\}} a_{\ell k}\boldsymbol{\mu}_{\ell,i}(\theta) \qquad \text{(AA-Consensus)} \tag{2.20}$$

We summarize both algorithms in Algorithm 2.4.

Algorithm 2.4 Consensus social learning

while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta) \tag{2.21}$$

  end for
  for each agent $k = 1, 2, \ldots, K$ **do**
    combine the intermediate beliefs with the previous beliefs of the neighbors

$$\boldsymbol{\mu}_{k,i}(\theta) \propto (\boldsymbol{\psi}_{k,i}(\theta))^{a_{kk}} \prod_{\ell \in \mathcal{N}_k \setminus \{k\}} \Big(\boldsymbol{\psi}_{\ell,i}(\theta)\Big)^{a_{\ell k}} \qquad \text{(GA-Consensus)} \tag{2.22}$$

    by using geometric averaging rule or

$$\boldsymbol{\mu}_{k,i}(\theta) = a_{kk}\boldsymbol{\psi}_{k,i}(\theta) + \sum_{\ell \in \mathcal{N}_k \setminus \{k\}} a_{\ell k}\boldsymbol{\mu}_{\ell,i}(\theta) \qquad \text{(AA-Consensus)} \tag{2.23}$$

    by using arithmetic averaging rule
  end for
  agents **exchange** beliefs with their neighbors
  $i \leftarrow i + 1$
end while

2.4 Network Limiting Behavior

Before presenting the main contributions of this thesis in the following chapters, for the benefit of reader, we briefly review some earlier results on the behavior of agents under the introduced social learning frameworks.

2.4.1 Modeling Conditions

In order to avoid pathological cases, we assume that

$$D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)||L_k(\cdot|\theta)\Big) < \infty \tag{2.24}$$

for each agent k and hypothesis θ. This condition makes sure that the likelihood functions for different hypotheses share the same support; and no observation alone is sufficient to refute any hypothesis.

Moreover, the following condition enables the aggregate of all agents to distinguish the true hypothesis from the wrong ones.

Assumption 2.1 (**Global identifiability**). *For each wrong hypothesis $\theta \neq \theta^\circ$, there exists at least one clear-sighted agent $k^\star$ that can distinguish θ from the true hypothesis θ°, i.e.,*

$$D_{\mathrm{KL}}\Big(L_{k^\star}(\cdot|\theta^\circ)\Big\|L_{k^\star}(\cdot|\theta)\Big) > 0 \tag{2.25}$$

Note that this is a weaker assumption than *local* identifiability, which requires *all* agents (not only one) to have the capability of distinguishing the truth individually.

Furthermore, for geometric social learning, the following condition is necessary in order to avoid trivial cases where agents discard some hypotheses right from the start.

Assumption 2.2 (**Initial beliefs**). *All initial beliefs are strictly positive at all hypotheses, i.e., for each agent k and for all hypotheses $\theta \in \Theta$, $\boldsymbol{\mu}_{k,0}(\theta) > 0$.*

If Assumption 2.2 is violated for a hypothesis θ, due to the nature of the geometric fusion rule, all agents would have zero belief on θ in finite time.

2.4.2 Geometric Social Learning

Next, we present an existing result from the literature [23, 24] that shows agents can discover the true hypothesis θ° with full confidence under the GA-based non-Bayesian social learning (NBSL) strategy (2.7)–(2.9). We also provide the proof of this result which will become useful for our discussions in the sequel.

Theorem 2.1 (**Truth learning**). *Under Assumptions 2.1 and 2.2, the social learning algorithm* (2.7)–(2.9) *with geometric averaging allows each agent k to learn the truth with probability* 1*, i.e.,*

$$\boldsymbol{\mu}_{k,i}(\theta^\circ) \xrightarrow{\text{a.s.}} 1. \tag{2.26}$$

Furthermore, the asymptotic learning of the truth occurs at an exponential rate. That is to say, the asymptotic convergence rate of the beliefs on wrong hypotheses $\theta \neq \theta^\circ$ to zero is given by

$$\frac{1}{i}\log\frac{\boldsymbol{\mu}_{k,i}(\theta^\circ)}{\boldsymbol{\mu}_{k,i}(\theta)} \xrightarrow{\text{a.s.}} \sum_{k=1}^{K} v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\theta)\Big). \tag{2.27}$$

Proof. The algorithm (2.7)–(2.9) gives rise to the following recursion for the log-belief

ratios:

$$\log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} = \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \Big(\log \frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta^\circ)} + \log \frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)}{\boldsymbol{\mu}_{\ell,i-1}(\theta^\circ)} \Big). \tag{2.28}$$

If we iterate this equation over time index i and also divide both sides by i, we arrive at

$$\frac{1}{i} \log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} = \frac{1}{i} \sum_{j=1}^{i} \sum_{\ell=1}^{K} [A^{i-j+1}]_{\ell k} \log \frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} + \frac{1}{i} \sum_{\ell=1}^{K} [A^i]_{\ell k} \log \frac{\boldsymbol{\mu}_{\ell,0}(\theta)}{\boldsymbol{\mu}_{\ell,0}(\theta^\circ)}. \tag{2.29}$$

Since A is a primitive and left-stochastic matrix, it holds by the Perron-Frobenius theorem that $A^i \;\to\; v\mathbb{1}_K^T$ as $i \to \infty$ [39, Chapter 8]. If we incorporate this fact into (2.29), the second term on the right hand side (RHS) vanishes due to Assumption 2.2. Hence, the expression (2.29) transforms into:

$$\frac{1}{i} \log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} \to \frac{1}{i} \sum_{j=1}^{i} \sum_{\ell=1}^{K} v_\ell \log \frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \tag{2.30}$$

as time index i gets larger. Since the observations $\{\boldsymbol{\xi}_{\ell,i}\}$ are i.i.d. over time, and the expectation of the log-likelihood ratios satisfy:

$$\mathbb{E}_{\theta^\circ} \left[\log \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}{L_k(\boldsymbol{\xi}_{k,i}|\theta)} \right] = D_{\mathrm{KL}} \Big(L_k(\cdot|\theta^\circ) || L_k(\cdot|\theta) \Big) < \infty. \tag{2.31}$$

Using (2.31) and applying the strong law of large numbers [40, Chapter 7] to (2.30), we establish the decay rate expression (2.27). Observe that the decay rates are characterized by a weighted average of the KL divergences of the agents — it is known that these entities reflect the inference capacity of an agent for hypothesis testing problems [41].

In addition, under Assumption 2.1, the RHS of (2.27) is strictly positive for each $\theta \neq \theta^\circ$. Consequently,

$$\forall \theta \in \Theta \setminus \{\theta^\circ\}, \quad \boldsymbol{\mu}_{k,i}(\theta) \xrightarrow{\text{a.s.}} 0, \tag{2.32}$$

which in turn implies (2.26). ■

2.4.3 Results on Variations

Arithmetic social learning

The next result [21, 22] states that agents can discover the true hypothesis θ° with full confidence under the arithmetic average fusion-based non-Bayesian social learning (NBSL) strategy as well.

Theorem 2.2 (**Arithmetic truth learning**). *Under Assumptions 2.1 and 2.2, the social learning algorithm* (2.7)–(2.13) *with arithmetic averaging allows each agent* k *to learn the truth with probability* 1*, i.e.,*

$$\boldsymbol{\mu}_{k,i}(\theta^\circ) \xrightarrow{\text{a.s.}} 1. \tag{2.33}$$

However, the proof of Theorem 2.2 relies on martingale arguments [21, 22] which does not provide any information about the speed of learning. It turns out the analysis for arithmetic averaging is more involved compared to the analysis for geometric averaging. In Chapter 3, we tackle this problem and provide results for the learning rate of arithmetic averaging for social learning.

Adaptive social learning

In contrast to Theorems 2.1 and 2.2 where the beliefs converge to the truth almost surely, in the *adaptive social learning* (ASL) strategy defined by steps (2.16) and (2.9), the beliefs will have everlasting random fluctuations that are necessary for keeping adaptation alive. To see the difference between NBSL and ASL, let us introduce the log-belief ratio vector

$$\boldsymbol{\lambda}_i(\theta) \triangleq [\boldsymbol{\lambda}_{1,i}(\theta), \ldots, \boldsymbol{\lambda}_{K,i}(\theta)]^\mathsf{T} \tag{2.34}$$

with individual entries defined as

$$\boldsymbol{\lambda}_{k,i}(\theta) \triangleq \log \frac{\boldsymbol{\mu}_{k,i}(\theta^\circ)}{\boldsymbol{\mu}_{k,i}(\theta)}. \tag{2.35}$$

Equations (2.7)–(2.9) imply that the vector $\boldsymbol{\lambda}_i(\theta)$ evolves according to the stochastic linear recursion

$$\boldsymbol{\lambda}_i(\theta) = A^\mathsf{T}\Big(\boldsymbol{\lambda}_{i-1}(\theta) + \boldsymbol{x}_i(\theta)\Big), \tag{2.36}$$

where $\boldsymbol{x}_i(\theta) \triangleq [\boldsymbol{x}_{1,i}(\theta), \ldots, \boldsymbol{x}_{K,i}(\theta)]^\mathsf{T}$ is the vector of log-likelihood ratios (LLRs) at time instant i:

$$\boldsymbol{x}_{k,i}(\theta) \triangleq \log \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}{L_k(\boldsymbol{\xi}_{k,i}|\theta)}. \tag{2.37}$$

In ASL, the modified adaptation step (2.16) alters the log-belief ratio recursion (2.36) to

$$\boldsymbol{\lambda}_i(\theta) = A^\mathsf{T}\Big((1-\delta)\boldsymbol{\lambda}_{i-1}(\theta) + \beta\boldsymbol{x}_i(\theta)\Big). \tag{2.38}$$

If we iterate this recursion over time, by standard arguments, we arrive at the following result from [14].

Theorem 2.3 (**ASL Convergence in distribution**). *Under the ASL strategy* (2.16) *and* (2.9)*, the log-belief ratios converge in distribution to the following absolutely convergent series:*

$$\boldsymbol{\lambda}_i(\theta) \xrightarrow{\text{dist.}} \beta \sum_{j=1}^{\infty} (1-\delta)^{j-1} (A^{\mathsf{T}})^j \boldsymbol{x}_j(\theta). \tag{2.39}$$

This result implies that the random fluctuations of ASL have a regular behavior in the limit. In particular, the limit of the expectation exists and is given by the following expression.

Corollary 2.1 (**Expected log-belief ratio in ASL**). *Theorem 2.3 implies that the log-belief ratios converge in the mean, i.e.,*

$$\lim_{i\to\infty} \mathbb{E}[\boldsymbol{\lambda}_i(\theta)] = \frac{\beta}{1-\delta} \Big((I - (1-\delta)A^{\mathsf{T}})^{-1} - I \Big) d(\theta) \tag{2.40}$$

where $d(\theta) \triangleq [d_1(\theta), d_2(\theta), \ldots, d_K(\theta)]^{\mathsf{T}}$ *is the vector of network KL divergences.*

Proof. Since the series on the RHS of (2.39) is uniformly integrable, the expectation on $\boldsymbol{\lambda}_i(\theta)$ converges to the expectation of the RHS of (2.39). The result then follows from the fact that $\mathbb{E}[\boldsymbol{x}_j(\theta)] = d(\theta)$ for any time j, and from the closed-form expression for the series of absolutely convergent matrices. ■

Information Exchange Part I

3 Information Fusion

3.1 Introduction[1]

Following the introductory material in the previous two chapters, the main contributions of this dissertation start from this chapter onward. The first problem we address is to compare the arithmetic (AA) and geometric (GA) averaging strategies for information fusion. Since AA and GA have different attributes that can be useful for different applications, as discussed in Chapter 2, it becomes critical to attain a more detailed understanding of their performance. In particular, it is important to examine how much gain or loss in learning rate the distributed system can attain based on its use of the AA or GA fusion rule, and how this gap is influenced by the graph topology. Some earlier studies have been pursued in the literature, however, they are limited in scope in the sense that:

- Only one step of fusion is studied [37, 43, 44];
- The analysis is restricted to well-behaved likelihood functions like Gaussian or Poisson distributions [37, 44–46].

Moreover, in the context of social learning, the asymptotic learning rate of social learning for geometric averaging is already well-established (recall Theorem 2.1). However, results for arithmetic averaging are scarce. An upper bound for the asymptotic learning rate of AA is given in [47]. Even though one can conclude from these earlier studies [24, 47] that GA is faster than AA for learning the truth, the performance difference is still not clearly established in the literature.

To that end, in this chapter, we study the *repeated* application of AA and GA in social learning *without* confining to specific distributions. While doing so, we analyze the

[1] The material in this chapter is based on [42].

performance gap between AA and GA learning rates in detail. In particular, for social learning, we arrive at the following results.

3.1.1 Contributions

- In Theorem 3.2, we prove that the agents learn the truth *exponentially fast* with the AA fusion rule under the standard diffusion algorithm [7]. Furthermore, the decay (i.e., learning) rate of beliefs over a wrong hypothesis is *constant* and does not depend on the agent.

- In Lemma 3.1, we provide upper and lower bounds for the decay rate using superadditive and subadditive functions on matrices. Also, an interesting "inept agent phenomenon" is discovered by using an appropriate superadditive function.

- We provide a variational lower bound on the gap between the decay rates of AA and GA in Theorem 3.3. The bound involves the Dobrushin coefficient [48, Chapter 2.7] as a network connectivity parameter. If the network is geometrically ergodic [48] (for which the Dobrushin coefficient is strictly smaller than 1), then the gap is 0 if, and only if, the agents observe exactly the same data. Otherwise GA performs better in terms of learning rate.

- For the special case of rank-one combination matrices, which includes architectures with fusion center in terms of performance, the exact decay rate of AA in closed form is given in Corollary 3.3.

- For exchangeable networks, where no permutation of data across agents can change the dynamics, we also provide a closed form expression for the gap between the decay rates of AA and GA in Theorem 3.4.

3.2 Existence of Asymptotic Decay Rates

To measure how fast the beliefs converge to the truth (i.e., how fast $\boldsymbol{\mu}_{k,i}(\theta^\circ) \to 1$ and $\boldsymbol{\mu}_{k,i}(\theta \neq \theta^\circ) \to 0$), for each k, it is sufficient to study the asymptotic behavior of agent k's belief on a false hypothesis $\theta \neq \theta^\circ$. More precisely, we will study the exponential decay rates of these beliefs, defined as

$$\boldsymbol{\rho}_k(\theta) \triangleq -\limsup_{i\to\infty} \frac{1}{i} \log \boldsymbol{\mu}_{k,i}(\theta). \tag{3.1}$$

In this section, we show under the diffusion-based arithmetic social learning rule that *exact* limits exist for (3.1) with $\limsup$ replaced by $\lim$ and that these limits are

independent of the agent index k. That is, we will show that

$$\boldsymbol{\rho}^{(\mathrm{AA})}(\theta) \triangleq \lim_{i\to\infty} -\frac{1}{i}\log \boldsymbol{\mu}_{k,i}(\theta) \tag{3.2}$$

exists almost surely and does not depend on k. This means that the beliefs on false hypotheses decay exponentially at the *same rate for all agents*.

3.2.1 Asymptotic Decay Rate of Arithmetic Social Learning

In comparison to GA-diffusion, the analysis for AA-diffusion is more involved. This is because, as shown in Chapter 2, GA-diffusion is amenable to an analysis that studies the evolution of log-belief ratios, and can benefit from an application of the strong law of large numbers. Unfortunately, a similar analysis is not possible for AA-diffusion — if we attempt to study $\log \boldsymbol{\mu}_{k,i}(\theta)$ directly, we end up with

$$\log \boldsymbol{\mu}_{k,i}(\theta) = \log\left(\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\boldsymbol{\psi}_{\ell,i}(\theta)\right), \tag{3.3}$$

which does not provide a simple-to-analyze dynamical system like we had with (2.36). Thus, we need to resort to different methods in the following. The authors of [47] approached the problem of finding $\boldsymbol{\rho}^{(\mathrm{AA})}$ by linearizing the dynamical system (3.3). With this method, they were able to lower bound $\boldsymbol{\rho}^{(\mathrm{AA})}$ by the Lyapunov exponent of the linearized version. We take a different approach by constructing extremal processes that bound $\boldsymbol{\mu}_{k,i}(\theta)$.

Constructing the Extremal Process

We wish to simplify the analysis of the AA algorithm by obtaining an extremal process $\{\boldsymbol{\nu}_{k,i}\}$, which eventually remains above $\{\boldsymbol{\mu}_{k,i}\}$ with probability 1. Studying $\{\boldsymbol{\nu}_{k,i}\}$ will then lead to a bound on the decay rate for AA-diffusion. With this aim, recall that our setting lies in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω represents the space of all data sequence realizations $\{\xi_{k,i}\}$ over time, $\mathcal{F}$ represents the σ-field generated by the sequence of data, and $\mathbb{P}$ represents the probability measure over sample paths $\omega \in \Omega$. In light of Theorem 2.2, we select $\epsilon > 0$ and define the event that all $\boldsymbol{\mu}_{k,i}(\theta^\circ)$ lie above $1-\epsilon$ eventually as

$$\mathcal{G}(\epsilon) \triangleq \{\omega \in \Omega : \exists i_0\ \boldsymbol{\mu}_{k,i}(\theta^\circ) \geq 1-\epsilon,\ \forall i > i_0, \forall k\}. \tag{3.4}$$

We then observe that $\mathbb{P}(\mathcal{G}(\epsilon)) = 1$ as a consequence of Theorem 2.2. Note that $\mathcal{G}(\epsilon)$ can also be interpreted as the event that there exists an $\boldsymbol{i}_0(\omega)$ — which is a random variable since it depends on ω — such that for all agents the true beliefs remain greater than $1-\epsilon$ after $\boldsymbol{i}_0^{\text{th}}$ iteration. We now restrict ourselves to $\mathcal{G}(\epsilon)$ and study the evolution of

$\{\boldsymbol{\mu}_{k,i}\}$ under such restriction.

Consider a false hypothesis $\theta \neq \theta^\circ$. We study the pathwise trajectories of $\{\boldsymbol{\mu}_{k,i}(\theta)\}$ for an outcome in $\mathcal{G}(\epsilon)$, that is, we pick an $\omega \in \mathcal{G}(\epsilon)$. According to the definition of $\mathcal{G}(\epsilon)$, there exists an $\boldsymbol{i}_0(\omega)$ such that for all $i > \boldsymbol{i}_0(\omega)$:

$$\begin{aligned}\boldsymbol{\mu}_{k,i}(\theta) &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta)}{\sum_{\theta'}\boldsymbol{\mu}_{\ell,i-1}(\theta')L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta')}\\ &\leq \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta)}{(1-\epsilon)L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta^\circ)}.\end{aligned} \tag{3.5}$$

Let us define the likelihood ratio of the freshly observed data by agent k at time i between the hypotheses θ and θ°

$$\boldsymbol{r}_{k,i}(\theta) \triangleq \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta)}{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)} \tag{3.6}$$

Then, we define the extremal process $\{\boldsymbol{\nu}_{k,i}\}$ as the process that evolves according to the recursion

$$\boldsymbol{\nu}_{k,i}(\theta) \triangleq (1-\epsilon)^{-1}\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\boldsymbol{\nu}_{\ell,i-1}(\theta)\boldsymbol{r}_{\ell,i}(\theta) \tag{3.7}$$

for all $i \geq \boldsymbol{i}_0(\omega)$ and with $\boldsymbol{\nu}_{k,\boldsymbol{i}_0}(\theta) = \boldsymbol{\mu}_{k,\boldsymbol{i}_0}(\theta)$. Comparing (3.7) with (3.5), we see that $\{\boldsymbol{\mu}_{k,i}(\theta)\}$ is upper bounded by $\{\boldsymbol{\nu}_{k,i}(\theta)\}$ for all $i \geq \boldsymbol{i}_0(\omega)$.

Remark 3.1. *For the rest of the chapter, we fix a false hypothesis $\theta \neq \theta^\circ$ and omit θ as an argument for brevity. This is because the analysis is the same for all $\boldsymbol{\mu}_{k,i}(\theta)$. We write the dependency on θ whenever we feel a need to emphasize them.*

The transition from $\boldsymbol{\nu}_{k,i-1}$ to $\boldsymbol{\nu}_{k,i}$ given by (3.7) is a random linear transform. Let $\boldsymbol{\nu}_i \triangleq [\boldsymbol{\nu}_{1,i},\ldots,\boldsymbol{\nu}_{K,i}]^\mathsf{T}$ and define the diagonal $K\times K$ random diagonal matrices $\boldsymbol{R}_i$ with their k^{th} diagonal element being $\boldsymbol{r}_{k,i}$. Then for all $\omega \in \mathcal{G}(\epsilon)$, expression (3.7) leads to the following vector relation:

$$\boldsymbol{\nu}_i = (1-\epsilon)^{-1}(A^\mathsf{T}\boldsymbol{R}_i)\boldsymbol{\nu}_{i-1}, \quad \forall i > \boldsymbol{i}_0(\omega). \tag{3.8}$$

Equation (3.8) highlights that the asymptotic behavior of the random matrix product

$$\boldsymbol{Y}_i \triangleq \prod_{j=1}^{i}(A^\mathsf{T}\boldsymbol{R}_j) = A^\mathsf{T}\boldsymbol{R}_i A^\mathsf{T}\boldsymbol{R}_{i-1}\ldots A^\mathsf{T}\boldsymbol{R}_1 \tag{3.9}$$

plays an important role. Moreover, observe that since $\{\boldsymbol{R}_i\}$ is a stationary sequence, the asymptotic behavior remains unchanged under any time shift. Hence, starting the

product from $j = 1$ in equation (3.9) is without loss of generality. We now turn our attention to the analysis of the random matrices $\{\boldsymbol{Y}_i\}$.

Asymptotic Behavior of Random Matrix Products

The asymptotic behavior of random matrix products is an important and challenging problem with a long history, which includes the preliminary study [49], and later the seminal work [50]. Although some results exist under certain assumptions on the structure of the random matrices to be multiplied, the problem remains open in general. One difficulty is that the non-commutativity of matrices under multiplication prevents the use of well-known convergence theorems such as the law of large numbers.

Extending the result of [50], reference [51] studied the more general case of "subadditive processes"[2], and derived, under some fairly general conditions, an ergodic theorem known as Kingman's subadditive ergodic theorem [51]. The analysis of random matrix products turns out to be a special case of this result.

Theorem 3.1 (**Kingman's subadditive ergodic theorem [51]**)**.** *Consider a stationary sequence of random matrices* $\{\mathbf{X}_i\}$ *and suppose that the elements of each* $\mathbf{X}_i$ *are positive and that their logarithms have finite expectations. Let* $\boldsymbol{Y}_i \triangleq \prod_{j=1}^{i} \mathbf{X}_j$*. Then, the limit*

$$\boldsymbol{\gamma} = \lim_{i\to\infty} \frac{1}{i} \log [\boldsymbol{Y}_i]_{\ell k} \tag{3.10}$$

exists and is finite almost surely and in the mean, and does not depend on ℓ *or* k*. Furthermore, it holds that*

$$\mathbb{E}\boldsymbol{\gamma} = \lim_{i\to\infty} \frac{1}{i} \mathbb{E} \log [\boldsymbol{Y}_i]_{\ell k} \,. \tag{3.11}$$

We now adapt the above theorem to our setting. First of all, observe that the $\{\boldsymbol{R}_i\}$ is an i.i.d. sequence, and therefore the $\{A^{\mathsf{T}}\boldsymbol{R}_i\}$ is also i.i.d., and hence, stationary. Note that we cannot simply replace the $\mathbf{X}_i$'s with $A^{\mathsf{T}}\boldsymbol{R}_i$'s because the matrix A^{T} need not have all positive entries. However, A is a primitive matrix and thus there must exist some $n \geq 1$ such that every entry of A^n is strictly positive [39, Chapter 8]. Using this observation, we arrive at the following corollary.

[2] A subadditive process is a random sequence $\{\boldsymbol{y}_i\}$ that satisfies the inequality $\boldsymbol{y}_{i+j} \leq \boldsymbol{y}_i + \boldsymbol{y}_j$ for all positive integers i and j.

Corollary 3.1 (Limit of the random matrix product). *Consider* $\boldsymbol{Y}_i = \prod_{j=1}^{i}(A^{\mathsf{T}}\boldsymbol{R}_j)$. *Under Assumption 2.1, the finite limit*

$$\gamma = \lim_{i\to\infty}\frac{1}{i}\log\left[\boldsymbol{Y}_i\right]_{\ell k} = \lim_{i\to\infty}\frac{1}{i}\mathbb{E}\log\left[\boldsymbol{Y}_i\right]_{\ell k} \tag{3.12}$$

exists almost surely, is a constant, and does not depend on ℓ *or* k.

Proof. See Appendix 3.A. ■

Using Corollary 3.1, we are able to characterize the asymptotic decay rate for AA-diffusion and conclude this section.

Theorem 3.2 (Asymptotic decay rate of AA-diffusion). *For any agent* k, *it holds almost surely:*

$$\boldsymbol{\rho}_k^{(\mathrm{AA})}(\theta) = -\lim_{i\to\infty}\frac{1}{i}\log\boldsymbol{\mu}_{k,i}(\theta) = -\gamma(\theta) \tag{3.13}$$

for $\gamma(\theta) < 0$.

Proof. See Appendix 3.B. ■

Theorem 3.2 states that if AA-diffusion (Algorithm 2.2) is executed, the beliefs on a false hypothesis θ decay exponentially almost surely, and the decay rate is constant and is the same among all agents. Note, however, that the decay rate may vary across θ. This is because $\gamma = \gamma(\theta)$ is the limit pertaining to the i.i.d. products of the matrices $A^{\mathsf{T}}\boldsymbol{R}_i(\theta)$, where $\boldsymbol{R}_i(\theta)$ is defined in terms of the $\boldsymbol{r}_{k,i}(\theta)$ — see (3.7).

3.3 Bounds on the Asymptotic Decay Rate

Kingman, in his original work [51] that introduces the subadditive ergodic theorem, writes the following about the constant γ:

"...Pride of place among the unsolved problems of subadditive ergodic theory must go to the calculation of the constant γ*... In none of the applications described here is there an obvious mechanism for obtaining an exact numerical value, and indeed this usually seems to be a problem of some depth..."*

To the best of our knowledge, no standard machinery exists to date for this end. Therefore, we make use of the special structure of the matrix $A^{\mathsf{T}}\boldsymbol{R}_i$ to obtain bounds for γ in Sections 3.3.2 and 3.3.3. But first we provide some simple upper and lower bounds

that hold for the general case.

3.3.1 Bounds Based on Subadditivity

First, it is useful to discuss why products of random matrices are related to subadditive processes. Let $\|X\|$ denote any matrix norm that is submultiplicative, i.e., for any X and Y

$$\|XY\| \leq \|X\|\|Y\|. \tag{3.14}$$

For our problem at hand, replacing X, Y with $\boldsymbol{Y}_i^j$, $\boldsymbol{Y}_j^m$ whose general definition is given as

$$\boldsymbol{Y}_i^m \triangleq \prod_{t=i+1}^{m} (A^\mathsf{T} \boldsymbol{R}_t), \tag{3.15}$$

and taking the logarithms of the both sides yields the subadditive relation

$$\log \|\boldsymbol{Y}_i^m\| \leq \log \|\boldsymbol{Y}_i^j\| + \log \|\boldsymbol{Y}_j^m\|. \tag{3.16}$$

A well-known result pertaining to subadditive functions is Fekete's lemma [52], which, in our context, states that

$$\lim_{i\to\infty} \frac{1}{i} \log \|\boldsymbol{Y}_i\| = \inf_i \frac{1}{i} \log \|\boldsymbol{Y}_i\|, \tag{3.17}$$

where $\boldsymbol{Y}_i$ is defined in (3.9). In fact, this observation is the starting point of the work [50] on random matrix products. Now, consider the norm of matrices with non-negative entries and let $\|X\|_1 \triangleq \max_\ell \sum_k [X]_{k\ell}$, which is submultiplicative. It is then immediate from Corollary 3.1 and (3.17) that for any time instant j:

$$\begin{aligned}
\gamma &= \lim_{i\to\infty} \frac{1}{i} \mathbb{E} \log [\boldsymbol{Y}_i]_{11} \\
&\leq \lim_{i\to\infty} \frac{1}{i} \mathbb{E} \log \|\boldsymbol{Y}_i\|_1 \\
&\overset{(3.17)}{=} \inf_i \frac{1}{i} \mathbb{E} \log \|\boldsymbol{Y}_i\|_1 \\
&\leq \frac{1}{j} \mathbb{E} \log \|\boldsymbol{Y}_j\|_1,
\end{aligned} \tag{3.18}$$

which yields an upper bound for γ.

For the lower bound, we aim to create a supermultiplicative process[3]. Let $\|X\|_- \triangleq \min_\ell \sum_k [X]_{k\ell}$ be the minimum column sum of the matrix X. Note that this is not a

[3] A supermultiplicative process is a sequence $\{\boldsymbol{y}_i\}$ that satisfies the inequality $\boldsymbol{y}_{i+j} \geq \boldsymbol{y}_i \cdot \boldsymbol{y}_j$ for all positive integers i and j. Similarly, a submultiplicative process is a sequence $\{\boldsymbol{y}_i\}$ that satisfies the inequality $\boldsymbol{y}_{i+j} \leq \boldsymbol{y}_i \cdot \boldsymbol{y}_j$ for all i and j.

norm. However, it is supermultiplicative for non-negative matrices, i.e.,

$$\|XY\|_- \geq \|X\|_-\|Y\|_-. \tag{3.19}$$

We then obtain the following result.

Lemma 3.1 (**Bounds based on subadditivity**). *For any* $i, j \geq 1$,

$$\frac{1}{i}\mathbb{E}\Big[\log\|\boldsymbol{Y}_i\|_-\Big] \leq \gamma \leq \frac{1}{j}\mathbb{E}\Big[\log\|\boldsymbol{Y}_j\|_1\Big]. \tag{3.20}$$

Proof. The upper bound follows directly from (3.18). For the lower bound, observe that $-\log\|\boldsymbol{Y}_i\|_-$ is a subadditive process. We then have, once again due to Fekete's lemma,

$$\begin{aligned}\sup_i \frac{1}{i}\mathbb{E}\log\|\boldsymbol{Y}_i\|_- &= \lim_{i\to\infty}\frac{1}{i}\mathbb{E}\log\|\boldsymbol{Y}_i\|_- \\ &= \mathbb{E}\left[\lim_{i\to\infty}\frac{1}{i}\log\|\boldsymbol{Y}_i\|_-\right] \\ &= \gamma. \end{aligned}\tag{3.21}$$

where the interchange of the limit and expectation operations is due to the general form of the subadditive ergodic theorem — see Theorem 3.5 in Appendix 3.C. ■

We point out that $\|\cdot\|_-$ and $\|\cdot\|_1$ can be replaced with any suitable functions on positive matrices that are super/submultiplicative, respectively; and satisfy the conditions of the subadditive ergodic theorem in Appendix 3.C. This way, one can obtain other lower and upper bounds for the rate of AA-diffusion using a method similar to the approach in Lemma 3.1.

In light of the above observation, we remark that $\|\cdot\|_-$ can also be replaced with an element of $\boldsymbol{Y}_i$, e.g., with any mapping $\boldsymbol{Y}_i \mapsto [\boldsymbol{Y}_i]_{kk}$, as long as it remains strictly positive for all i — this is to ensure that the logarithm remains finite. It can be verified that $[\boldsymbol{Y}_i]_{kk}$ is supermultiplicative as long as $\boldsymbol{Y}_i$'s are non-negative. Recall from Sec. 2.1.1 that at least one agent k° has a self-loop, that is, $a_{k^\circ k^\circ} > 0$. The self-loop assumption ensures that $[\boldsymbol{Y}_i]_{k^\circ k^\circ}$ remains strictly positive. Then, according to Lemma 3.1, we have

$$\mathbb{E}\log[\boldsymbol{Y}_1]_{k^\circ k^\circ} = \log a_{k^\circ k^\circ} + \mathbb{E}\log \boldsymbol{r}_{k^\circ,1} \leq \gamma \tag{3.22}$$

which, by Theorem 3.2, implies that

$$\begin{aligned}\rho^{(\mathrm{AA})} &\leq -\log a_{k^\circ k^\circ} - \mathbb{E}\log \boldsymbol{r}_{k^\circ,1} \\ &= -\log a_{k^\circ k^\circ} + D_{\mathrm{KL}}(L_{k^\circ}(\cdot|\theta^\circ)||L_{k^\circ}(\cdot|\theta)). \end{aligned}\tag{3.23}$$

The inequality (3.23) gives rise to an interesting observation of the learning model under Algorithm 2.2.

For a small $\delta > 0$, suppose that (i) $a_{k^\circ k^\circ} = 1 - \delta$, and (ii) $D_{\mathrm{KL}}(L_{k^\circ}(\cdot|\theta^\circ)||L_{k^\circ}(\cdot|\theta)) \leq \delta$. Here, (i) suggests that agent k° is highly self-confident with a self-loop close to 1, whereas (ii) suggests that agent k° has limited learning abilities with small informativeness δ. For this special case, observe that

$$\rho^{(\mathrm{AA})} \leq -\log(1-\delta) + \delta \leq \delta/(1-\delta) + \delta \tag{3.24}$$

is also small. Since all agents learn the truth at the same rate $\rho^{(\mathrm{AA})}$ by Theorem 3.2, we conclude that even a single agent k° can drastically decrease the learning ability of the *whole* network.

This phenomenon can be avoided under GA-diffusion as follows. Recall from Theorem 2.1 that asymptotic learning rate under Algorithm 2.1 is given by

$$\rho^{(\mathrm{GA})} = \sum_{k=1}^{K} v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big|\Big|L_k(\cdot|\theta)\Big). \tag{3.25}$$

If the remaining agents $k \neq k^\circ$ do not trust agent k°, i.e., if $a_{k^\circ k}$'s are small, then the Perron entry v_{k° can be kept small as well. It is evident from (3.25) that such isolation of agent k° from the network will preserve (and might even boost) the learning rate of the remaining agents. This observation is consistent with the numerical results in Section 3.4.

3.3.2 Bounds for Distributed Inference

In this section, we take advantage of the special structure of $\{\boldsymbol{Y}_i\}$ in our distributed setting and derive bounds for γ. To that end, observe first from (3.9) that

$$[\boldsymbol{Y}_i]_{1k} = \boldsymbol{r}_{k,i} \sum_{\ell=1}^{K} [\boldsymbol{Y}_{i-1}]_{1\ell} a_{k\ell}. \tag{3.26}$$

This implies that

$$\log[\boldsymbol{Y}_i]_{1k} = \log \boldsymbol{r}_{k,i} + \log\left(\sum_{\ell=1}^{K} [\boldsymbol{Y}_{i-1}]_{1\ell} a_{k\ell}\right), \tag{3.27}$$

and averaging over the network with weights v_k yields

$$\sum_{k=1}^{K} v_k \log[\boldsymbol{Y}_i]_{1k} = \sum_{k=1}^{K} v_k \log \boldsymbol{r}_{k,i} + \sum_{k=1}^{K} v_k \log \sum_{\ell} [\boldsymbol{Y}_{i-1}]_{1\ell} a_{k\ell}. \tag{3.28}$$

We subtract $\sum_{k=1}^{K} v_k \log[\boldsymbol{Y}_{i-1}]_{1k}$ from both sides of this equation to obtain

$$\sum_{k=1}^{K} v_k \log \frac{[\boldsymbol{Y}_i]_{1k}}{[\boldsymbol{Y}_{i-1}]_{1k}} = \sum_{k=1}^{K} v_k \log \boldsymbol{r}_{k,i} + \sum_{k=1}^{K} v_k \log \frac{\sum_\ell [\boldsymbol{Y}_{i-1}]_{1\ell} a_{k\ell}}{[\boldsymbol{Y}_{i-1}]_{1k}}. \tag{3.29}$$

If we take the time average of (3.29) from 2 to i, the left-hand side becomes a telescoping sum where the intermediate terms cancel each other, and we arrive at the following relation:

$$\frac{1}{i}\sum_{k=1}^{K} v_k \log \frac{[\boldsymbol{Y}_i]_{1k}}{[\boldsymbol{Y}_1]_{1k}} = \frac{1}{i}\sum_{j=2}^{i}\Big(\sum_{k=1}^{K} v_k \log \boldsymbol{r}_{k,j}\Big) + \frac{1}{i}\sum_{j=2}^{i}\sum_{k=1}^{K} v_k \log \frac{\sum_\ell [\boldsymbol{Y}_{j-1}]_{1\ell} a_{k\ell}}{[\boldsymbol{Y}_{j-1}]_{1k}}. \tag{3.30}$$

Assume $[\boldsymbol{Y}_1]_{1k} > 0$ for simplicity. Note that there is no loss of generality here since the strong connectivity assumption on the network ensures the primitiveness of the combination matrix A which in turn ensures $[\boldsymbol{Y}_1]_{1k}$ to be strictly positive eventually—see Appendix 3.A. The left-hand side tends to γ by Corollary 3.1 and Theorem 3.2, and since the $\boldsymbol{r}_{k,i}$ are i.i.d., the first term on the right-hand side tends to its mean by the law of large numbers and is equal to the negative of the decay rate of GA-diffusion from Theorem 2.1. In other words,

$$\gamma = \sum_{k=1}^{K} v_k \mathbb{E}[\log \boldsymbol{r}_{k,j}] + \lim_{i\to\infty} \frac{1}{i}\sum_{j=2}^{i}\sum_{k=1}^{K} v_k \log \frac{\sum_\ell [\boldsymbol{Y}_{j-1}]_{1\ell} a_{k\ell}}{[\boldsymbol{Y}_{j-1}]_{1k}} \tag{3.31}$$

and accordingly,

$$\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} = \lim_{i\to\infty} \frac{1}{i}\sum_{j=2}^{i}\sum_{k=1}^{K} v_k \log \frac{\sum_{\ell=1}^{K} [\boldsymbol{Y}_{j-1}]_{1\ell} a_{k\ell}}{[\boldsymbol{Y}_{j-1}]_{1k}} \tag{3.32}$$

The above equation quantifies the gap between the decay rates. The gap turns out to be the difference of two KL divergences averaged over time. To see this, let $\boldsymbol{u}$ denote the probability vector obtained by normalization of the first row of $\boldsymbol{Y}$, i.e.,

$$[\boldsymbol{u}_i]_k \triangleq \frac{[\boldsymbol{Y}_i]_{1k}}{\sum_\ell [\boldsymbol{Y}_i]_{1\ell}}. \tag{3.33}$$

Then, it follows that

$$\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} = \lim_{i\to\infty} \frac{1}{i}\sum_{j=2}^{i}\Big[D_{\mathrm{KL}}(v||\boldsymbol{u}_{j-1}) - D_{\mathrm{KL}}(v||A\boldsymbol{u}_{j-1})\Big] \tag{3.34}$$

Equation (3.34) has an interesting interpretation. Note that $A\boldsymbol{u}_{j-1}$ is another probability vector that is obtained by passing $\boldsymbol{u}_{j-1}$ through the Markov matrix A. Furthermore, $Av = v$ is the unique invariant distribution of the kernel A and thus the difference

term in (3.34) is equal to

$$\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} = D_{\mathrm{KL}}(v||\boldsymbol{u}_{j-1}) - D_{\mathrm{KL}}(Av||A\boldsymbol{u}_{j-1}). \tag{3.35}$$

The well-known data processing theorem [53] ensures that this difference is non-negative. Hence,

Corollary 3.2. *Asymptotic learning rate under geometric averaging based social learning (Algorithm 2.1) is always larger than the rate under arithmetic averaging based social learning (Algorithm 2.2). Namely,*

$$\rho^{(\mathrm{GA})} \geq \rho^{(\mathrm{AA})} \tag{3.36}$$

holds almost surely.

The only case where the limit in (3.34) — the performance gap — tends to zero is when $D_{\mathrm{KL}}(v||\boldsymbol{u}_i) \to 0$, or simply when $\boldsymbol{u}_i \to v$. We will show that if the network has sufficient connectivity, this can only happen when all agents receive the same data, i.e., $\boldsymbol{r}_{k,i} = \boldsymbol{r}_{\ell,i}$ for all agents k and ℓ. It is obvious that in this case most fusion methods, and in particular AA and GA, are equivalent. In fact, there is no need for an agent to communicate with its neighbors — every neighbor is equivalent to the agent itself.

If the network is connected enough, and when the agents do not observe the same data, $\rho^{(\mathrm{GA})}$ is *strictly* greater than $\rho^{(\mathrm{AA})}$, and the gap is quantified as the limit term in (3.34). To study this term, we verify that the $\{\boldsymbol{u}_i\}$ can be modeled in terms of a Markov chain. We denote the K-dimensional probability simplex as $\mathbb{S}_K$, and define the time-homogeneous transition map $T : \mathbb{S}_K \times \mathbb{R}^{K\times K} \to \mathbb{S}_K$ as

$$T(u, \boldsymbol{R}) \triangleq \frac{\boldsymbol{R}Au}{\sum_\ell [\boldsymbol{R}Au]_\ell} \tag{3.37}$$

where $\boldsymbol{R}$ is the diagonal matrix with $[\boldsymbol{R}]_{kk} = \boldsymbol{r}_k \triangleq \boldsymbol{r}_{k,1}$ due to time-homogeneity. Note that the $\boldsymbol{u}_i$'s in (3.34) obey this map with $\boldsymbol{u}_i = T(\boldsymbol{u}_{i-1}, \boldsymbol{R}_i)$, and with $\boldsymbol{R}_i$ independent of $\boldsymbol{u}_{i-1}$. It can be verified that the Markov chain governed by the mapping T has at least one invariant distribution $\mathbb{Q}$ on $\mathbb{S}_K$, i.e., if $\boldsymbol{u}$ has distribution $\mathbb{Q}$, so has $T(\boldsymbol{u}, \boldsymbol{R})$ — this comes from the observation that T is Feller continuous and Krylov-Bogolyubov theorem for compact spaces [54]. However, it may be a futile attempt to find such invariant distributions. Furthermore, it is not certain if there is a unique invariant distribution although the limit in (3.34) exists. Hence, we resort to a different method and study a lower bound for the gap.

Since the state space of the Markov chain $\mathbb{S}_K$ is compact, there must exist a u_0 and a small neighborhood $\mathcal{V}_0$ around it that is visited infinitely often. So, the limit in (3.34) is

lower bounded by

$$\lim_{i\to\infty}\frac{1}{i}\sum_{j=2}^{i}\Big[D_{\mathrm{KL}}(v||\boldsymbol{u}_{j-1})-D_{\mathrm{KL}}(v||A\boldsymbol{u}_{j-1})\Big]\geq\inf_{u\in\mathcal{V}_0}\mathbb{E}[D_{\mathrm{KL}}(v||\boldsymbol{u}_1)-D_{\mathrm{KL}}(v||A\boldsymbol{u}_1)]$$
$$\geq\inf_{u\in\mathbb{S}_K}\mathbb{E}[D_{\mathrm{KL}}(v||\boldsymbol{u}_1)-D_{\mathrm{KL}}(v||A\boldsymbol{u}_1)] \tag{3.38}$$

with $\boldsymbol{u}_1=T(u,\boldsymbol{R})$. Moreover, we can use the strong data processing inequality [55] to obtain a further lower bound. It is useful to give the following definition.

Definition 3.1 (**Contraction coefficient [55]**). *Let $A_{K\times K}$ be a probability transition matrix. Then the contraction coefficient associated with A is given by*

$$\eta_A\triangleq\sup_{\substack{u,\pi\in\mathbb{S}_K\\u\neq\pi}}\left\{\frac{D_{\mathrm{KL}}(Au||A\pi)}{D_{\mathrm{KL}}(u||\pi)}\right\}\leq 1. \tag{3.39}$$

The coefficient η_A is lower bounded by the second largest absolute eigenvalue of A [55]. Hence, dense graphs usually lead to lower η_A that is close to 0 in value, while sparse graphs usually lead to η_A that is close to 1 in value. The strong data processing inequality implies that

$$D_{\mathrm{KL}}(Au||A\pi)\leq\eta_A D_{\mathrm{KL}}(u||\pi) \tag{3.40}$$

for all K-dimensional probability vectors u, π. Applying this inequality to (3.38), we obtain

$$\rho^{(\mathrm{GA})}-\rho^{(\mathrm{AA})}\geq(1-\eta_A)\inf_{u\in\mathbb{S}_K}\mathbb{E}D_{\mathrm{KL}}(v||\boldsymbol{u}_1). \tag{3.41}$$

Remark 3.2. *Applying the strong data processing inequality directly to* (3.34)*, it is evident that*

$$\lim_{i\to\infty}\frac{1}{i}\sum_{j=1}^{i}D_{\mathrm{KL}}(v||\boldsymbol{u}_j)\geq\rho^{(\mathrm{GA})}-\rho^{(\mathrm{AA})}\geq(1-\eta_A)\lim_{i\to\infty}\frac{1}{i}\sum_{j=1}^{i}D_{\mathrm{KL}}(v||\boldsymbol{u}_j) \tag{3.42}$$

Therefore, when $\eta_A=0$, we have equality. This corresponds to the case when A is rank one, with every column of A being v. For this special case, we will show in Sec. 3.3.3 that γ has a simple form.

We now focus on the infimization problem that shows up in (3.41). Writing explicitly, it

is equivalent to

$$\inf_{\substack{u \\ \pi=uA^{\mathsf{T}}}} \mathbb{E}\left[\sum_{k=1}^{K} v_k \log \frac{v_k}{\pi_k \boldsymbol{r}_k} + \log(\sum_{k=1}^{K} \pi_k \boldsymbol{r}_k)\right] = \inf_{\substack{u \\ \pi=uA^{\mathsf{T}}}} \mathbb{E}\left[\sum_{k=1}^{K} v_k \log \frac{v_k}{\pi_k} + \log\Big(\sum_{k=1}^{K} \pi_k \boldsymbol{r}_k\Big)\right] + \rho^{(\mathrm{GA})}$$
$$\geq \inf_{\pi} \sum_{k=1}^{K} v_k \log \frac{v_k}{\pi_k} + \mathbb{E}\left[\log\Big(\sum_{k=1}^{K} \pi_k \boldsymbol{r}_k\Big)\right] + \rho^{(\mathrm{GA})} \tag{3.43}$$

We give the Karush-Kuhn-Tucker (KKT) conditions for this problem in the next result.

Lemma 3.2 (Optimality conditions). *The KKT conditions for the optimization problem in* (3.43)*, namely the infimization of*

$$F(\pi) \triangleq \sum_{k=1}^{K} v_k \log \frac{v_k}{\pi_k} + \mathbb{E}\left[\log\Big(\sum_{k=1}^{K} \pi_k \boldsymbol{r}_k\Big)\right] + \rho^{(\mathrm{GA})}, \tag{3.44}$$

are given by

$$\frac{v_k}{\pi_k} = \mathbb{E}\left[\frac{\boldsymbol{r}_k}{\sum_{\ell=1}^{K} \boldsymbol{r}_\ell \pi_\ell}\right] \quad \forall k. \tag{3.45}$$

Proof. See Appendix 3.D. ■

Note that $F(\pi)$ cannot approach the infimum over the boundary of the K-dimensional simplex as $F(\pi)$ tends to infinity close to the boundary. Using the KKT conditions above, and replacing $\frac{v_k}{\pi_k}$ with $\mathbb{E}\big[\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\big]$ in the first summation term in $F(\pi)$, we consider the infimization of

$$G(\pi) \triangleq \sum_{k=1}^{K} v_k \log \mathbb{E}\left[\frac{\boldsymbol{r}_k}{\sum_{\ell=1}^{K} \pi_\ell \boldsymbol{r}_\ell}\right] + \mathbb{E}\left[\log\Big(\sum_{k=1}^{K} \pi_k \boldsymbol{r}_k\Big)\right] + \rho^{(\mathrm{GA})}. \tag{3.46}$$

Observe that $\inf_\pi F(\pi) \geq \inf_\pi G(\pi)$. We summarize the above results in the theorem below.

Theorem 3.3 (Variational lower bound for the gap). *The performance gap can be lower bounded as*

$$\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} \geq (1-\eta_A) \inf_{\pi} F(\pi) \geq (1-\eta_A) \inf_{\pi} G(\pi) \tag{3.47}$$

with $F(\pi)$ defined in (3.44) *and $G(\pi)$ in* (3.46)*. Furthermore if $\eta_A < 1$, then $\rho^{(\mathrm{GA})} = \rho^{(\mathrm{AA})}$ if and only if $\boldsymbol{r}_k = \boldsymbol{r}_\ell$ for all k, ℓ.*

Proof. See Appendix 3.E. ■

Remark 3.3. *The bound we have provided in this section partially captures the effect of network structure through the quantity η_A. It is known that $\eta_A \leq \mathsf{Dob}_A$, the Dobrushin coefficient of A [55]:*

$$\mathsf{Dob}_A \triangleq \max_{k,\ell} \frac{1}{2}\|a_k - a_\ell\|_1 \tag{3.48}$$

where a_k denotes the kth column of A. Furthermore, $\eta_A < 1$ if and only if $\mathsf{Dob}_A < 1$ [55, 56]. Hence, for the broad class of combination matrices where $\mathsf{Dob}_A < 1$, i.e., geometrically ergodic ones [48, Chapter 2.7], the bound (3.47) *is non-trivial.*

Theorem 3.3 points out that the decay rate of AA-diffusion is dependent on network connectivity via η_A as opposed to the decay rate of GA-diffusion, whose decay rate only depends on the network centrality, i.e., Perron vector v. In [47], the authors study the effect of network regularity under AA-consensus on the upper bound $\rho^{(\mathrm{GA})}$. Different from their work, the bound on the performance gap in Theorem 3.3 captures the effect of network connectivity.

3.3.3 Special Networks

The bounds given in the previous sections are in variational form. Although we have studied certain characteristics of these bounds, i.e., found the KKT conditions, still, the bounds are dependent on the joint distribution of the data across the users — recall that the decay rate of GA-diffusion, $\rho^{(\mathrm{GA})}$, has a closed form expression and only depends on the marginals. Moreover, the extremal points satisfying the KKT conditions are difficult to find in general. Hence, we study two special cases in this section.

Rank-one Combination Matrices

When A is a rank one matrix, it turns out that $\rho^{(\mathrm{AA})}$ has a closed form. In this case, A can be written as $A = v\mathbb{1}_K^\mathsf{T}$. Then,

$$\begin{aligned} A^\mathsf{T}\boldsymbol{R}A^\mathsf{T} &= \mathbb{1}_K v^\mathsf{T}\boldsymbol{R}\mathbb{1}_K v^\mathsf{T} \\ &= \Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_k\Big)\mathbb{1}_K v^\mathsf{T} \\ &= \Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_k\Big)A^\mathsf{T} \end{aligned} \tag{3.49}$$

and it follows that

$$\boldsymbol{Y}_i = (A^{\mathsf{T}}\boldsymbol{R}_i)\prod_{j=1}^{i-1}(A^{\mathsf{T}}\boldsymbol{R}_j) = (A^{\mathsf{T}}\boldsymbol{R}_i)\prod_{j=1}^{i-1}\Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_{k,j}\Big). \tag{3.50}$$

If we choose k, ℓ such that $a_{\ell k} > 0$, it holds that

$$\frac{1}{i}\log[\boldsymbol{Y}_i]_{k\ell} = \frac{1}{i}\sum_{j=1}^{i-1}\log\Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_{k,j}\Big) + \frac{1}{i}\log(a_{\ell k}\boldsymbol{r}_{\ell,i}) \tag{3.51}$$

and from the law of large numbers

$$\lim_{i\to\infty}\frac{1}{i}\log[\boldsymbol{Y}_i]_{k\ell} = \mathbb{E}\Big[\log\Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_k\Big)\Big]. \tag{3.52}$$

Corollary 3.1 and Theorem 3.2 then lead to the following statement.

Corollary 3.3 (**Exact rate under rank-one topologies**). *Under network architectures with rank-one combination matrices, including fully-connected networks and federated architectures, the learning rate of arithmetic social learning is given by*

$$\rho^{(\mathrm{AA})} = -\gamma = -\mathbb{E}\Big[\log\Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_k\Big)\Big]. \tag{3.53}$$

Notice that Corollary 3.3 implies the performance gap between geometric and arithmetic social learning algorithms is equal to the expectation of a Jensen's inequality gap

$$\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} = \mathbb{E}\Big[\log\Big(\sum_{k=1}^{K} v_k \boldsymbol{r}_k\Big) - \sum_{k=1}^{K} v_k \log \boldsymbol{r}_k\Big] \tag{3.54}$$

for rank-one combination matrices. Here, recall from Chapter 2 that fully-connected networks with rank-one combination matrices are equivalent to architectures with fusion center in terms of performance, because each agent computes the same weighted average of all beliefs across the network. In this sense, each agent acts like a fusion center. Therefore, Corollary 3.3 also gives the learning rate of inference with federated architectures.

Exchangeable Networks

In this special case, we assume that the data is exchangeable across the users. More precisely, $\{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K\}$ constitutes a set of exchangeable random variables.

Definition 3.2 (**Exchangeable random variables**). *A set of random variables is called exchangeable if the distribution of $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_K$ remains unchanged under any permutation over the index set, i.e., for all $\xi_1, \ldots, \xi_k$*

$$\mathbb{P}\left(\boldsymbol{\xi}_1 \leq \xi_1, \ldots, \boldsymbol{\xi}_K \leq \xi_k\right) = \mathbb{P}\left(\boldsymbol{\xi}_{\sigma_1} \leq \xi_1, \ldots, \boldsymbol{\xi}_{\sigma_K} \leq \xi_k\right) \tag{3.55}$$

with σ being any permutation of $\{1, \ldots, K\}$.

Exchangeability is weaker than the i.i.d. assumption, as the data need not be independent across the agents. Observe that exchangeable networks could be of particular interest as they model fair networks — since exchangeability requires identical distributions of data across the agents, no agent learns better than another if there was no cooperation. For this particular example, we also assume that A is doubly stochastic, i.e., $\sum_k a_{\ell k} = 1$ as well. Then, it is easily seen that each element of the Perron vector of A is $v_k = 1/K$. Under this assumption, we are able to solve the KKT conditions of problem (3.43).

Theorem 3.4 (**Lower bound under exchangeable networks**). *If the data across the agents is exchangeable and A is doubly stochastic, the Perron eigenvector v is the unique solution of* (3.43). *Hence,*

$$\rho^{(\mathrm{AA})} \leq \eta_A \rho^{(\mathrm{GA})} - (1-\eta_A)\mathbb{E}\Big[\log\Big(\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}_k\Big)\Big] \tag{3.56}$$

and, moreover,

$$\begin{aligned}\rho^{(\mathrm{GA})} - \rho^{(\mathrm{AA})} &\geq (1-\eta_A)\left(\rho^{(\mathrm{GA})} + \mathbb{E}\Big[\log\Big(\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}_k\Big)\Big]\right) \\ &= (1-\eta_A)\mathbb{E}\Big[\log\Big(\sum_{k=1}^{K}\frac{1}{K}\boldsymbol{r}_k\Big) - \sum_{k=1}^{K}\frac{1}{K}\log\boldsymbol{r}_k\Big]\end{aligned} \tag{3.57}$$

Proof. It immediately follows that $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_K$ is also exchangeable. Also, the KKT conditions in (3.45) imply

$$\mathbb{E}\left[\frac{\boldsymbol{r}_1\pi_1}{\sum_\ell \boldsymbol{r}_\ell\pi_\ell}\right] = \mathbb{E}\left[\frac{\boldsymbol{r}_2\pi_2}{\sum_\ell \boldsymbol{r}_\ell\pi_\ell}\right] \tag{3.58}$$

and because of exchangeability,

$$\mathbb{E}\left[\frac{\boldsymbol{r}_1\pi_1}{\boldsymbol{r}_1\pi_1 + \boldsymbol{r}_2\pi_2 + \boldsymbol{s}}\right] = \mathbb{E}\left[\frac{\boldsymbol{r}_1\pi_2}{\boldsymbol{r}_2\pi_1 + \boldsymbol{r}_1\pi_2 + \boldsymbol{s}}\right] \tag{3.59}$$

where $\boldsymbol{s} \triangleq \sum_{\ell>2} \boldsymbol{r}_\ell \pi_\ell$. Suppose that $\pi_1 > \pi_2$. The difference of the two terms in (3.59) then becomes

$$\mathbb{E}\left[\frac{\boldsymbol{r}_1\pi_1}{\boldsymbol{r}_1\pi_1+\boldsymbol{r}_2\pi_2+\boldsymbol{s}} - \frac{\boldsymbol{r}_1\pi_2}{\boldsymbol{r}_2\pi_1+\boldsymbol{r}_1\pi_2+\boldsymbol{s}}\right] = (\pi_1-\pi_2)\mathbb{E}\left[\frac{\boldsymbol{r}_1\boldsymbol{r}_2(\pi_1+\pi_2)+\boldsymbol{r}_1\boldsymbol{s}}{(\boldsymbol{r}_1\pi_1+\boldsymbol{r}_2\pi_2+\boldsymbol{s})(\boldsymbol{r}_2\pi_1+\boldsymbol{r}_1\pi_2+\boldsymbol{s})}\right] \tag{3.60}$$

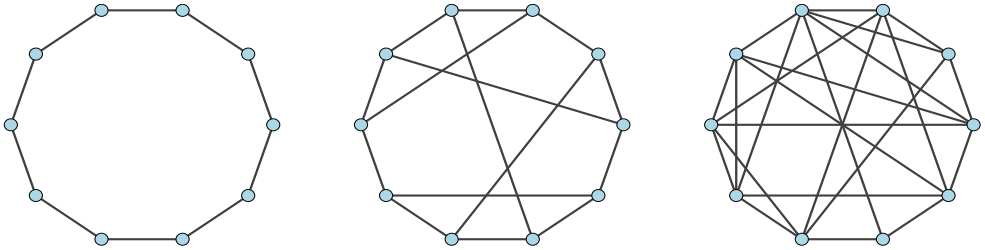
Since we assumed $\pi_1 > \pi_2$, the difference in (3.59) must be strictly greater than zero. A similar argument for $\pi_2 > \pi_1$ yields the same result. This implies $\pi_1 = \pi_2$, and by symmetry, $\pi_k = \pi_\ell$ for all k, ℓ. Therefore, π must be equal to v^T. ■

It is seen from Theorem 3.4 that the lower bound on the performance gap increases as the network becomes more connected, i.e., small η_A. Moreover, the Jensen's inequality gap in (3.54) and Theorem 3.4 increases when the observations become more diverse across the agents. The performance gap drifts away from zero. This results in an interesting observation: If the individual agents, or any subgroup of agents, have the same learning abilities (implied by exchangeability), the overall learning speed of the network decreases with respect to the increased amount of collaboration under AA-diffusion. However, the learning abilities under GA-diffusion is not affected — every agent in the network learns at the same speed regardless of the amount of collaboration.

3.4 Numerical Simulations

In this section, we provide numerical results to study the gap between GA-based Alg. 2.1 and AA-based Alg. 2.2; and to study the effect of network connectivity on the decay rate of AA-diffusion. We simulated the networks given in Figure 3.1a, 3.1b, 3.1c. All networks consist of $K = 10$ nodes. The network 3.1a is 2-regular, 3.1b is 3-regular and 3.1c is not a regular graph. Recall that a graph is called D-regular when each vertex has D neighbors.

Our first simulation compares the 2-regular and 3-regular networks. The combination matrices are denoted by A_2 and A_3, and are set as follows. For an $\alpha \in [0,1]$, if nodes ℓ and k are connected, then $[A_2]_{\ell k} = \frac{1-\alpha}{2}$ and $[A_3]_{\ell k} = \frac{1-\alpha}{3}$ respectively; and $[A_2]_{\ell\ell} = [A_3]_{\ell\ell} = \alpha$. The other elements are necessarily set to zero. Observe that we ensure the diagonal elements of A_2 and A_3 are the same and equal to α — this setting enables us to compare the decay rates with the bound given in [47]. We have chosen $\alpha = 0.05$. Note that A_2 and A_3 are doubly stochastic and thus their Perron vectors v have entries $v_k = 1/K$. Moreover, we assume $H = 2$. Under the true hypothesis θ°, the agents observe exponential random variables with parameter 1, and under the alternative hypothesis θ, each agent k observes exponential random variables with parameter



(a) (b) (c)

Figure 3.1: The three networks consist of $K = 10$ nodes. (a) is a 2-regular network, (b) is a 3-regular network and (c) is non-regular with 24 edges. Self-loops are omitted for visual simplicity in all figures.

β_k, with $\beta_k \in [0.500, 0.300, 0.025, 0.750, 1.200, 2.250, 0.900, 1, 0.250, 0.025]$. For simplicity, we assume that the data is independent across the agents.

The results of the first experiment are given in Figure 3.2i. We have plotted agent 1's belief decay rates on hypothesis θ, with AA and GA-diffusion algorithms (denoted with the superscript AA and GA, respectively) and on networks 3.1a and 3.1b (denoted with the subscript added to the algorithm description). For instance, the AA_3 in the superscript refers to the decay rate with AA-diffusion and on the 3-regular network 3.1b. First, observe the significant performance gap between GA and AA-diffusion learning rates. As expected, the decay rate of GA-diffusion is not affected by the network regularity — we know it only depends on the Perron vector v. However, the AA-diffusion decay rate is visibly affected, which is expected according to Theorem 3.3.

We point out that in [47], the authors upper bounded the decay rate of AA with consensus algorithm (not AA-diffusion) with $\alpha\rho^{(\mathrm{GA})}$. This is not true for AA-diffusion as for our setting $\alpha = 0.05$, $\rho^{(\mathrm{GA})} = 0.7261$, and $\rho^{(\mathrm{AA}_2)}$, $\rho^{(\mathrm{AA}_4)}$ both seem to be above $\alpha\rho^{(\mathrm{GA})} = 0.0363$. This also shows that since the AA-diffusion decay rate is above the upper bound given for AA-consensus, the agents learn faster with the diffusion algorithm than with consensus. This behavior is consistent with the results in [57], which showed that diffusion strategies are superior to consensus strategies in terms of performance in distributed estimation. Figure 3.2i complements this result in the sense that it shows diffusion outperforms consensus in the AA-social learning setting as well. Such distinction was not present for GA-social learning at least asymptotically. The diffusion and consensus based geometric social learning algorithms have the same asymptotic learning rate.

The second experiment involves an exchangeable network. To this end, we set $\beta_k = 3$ for all k. Referring to the result in Theorem 3.4, we define

$$B_A \triangleq \mathsf{Dob}_A\, \rho^{(\mathrm{GA})} - (1 - \mathsf{Dob}_A)\mathbb{E}\Big[\log\Big(\frac{1}{K}\sum_k \boldsymbol{r}_k\Big)\Big]. \tag{3.61}$$

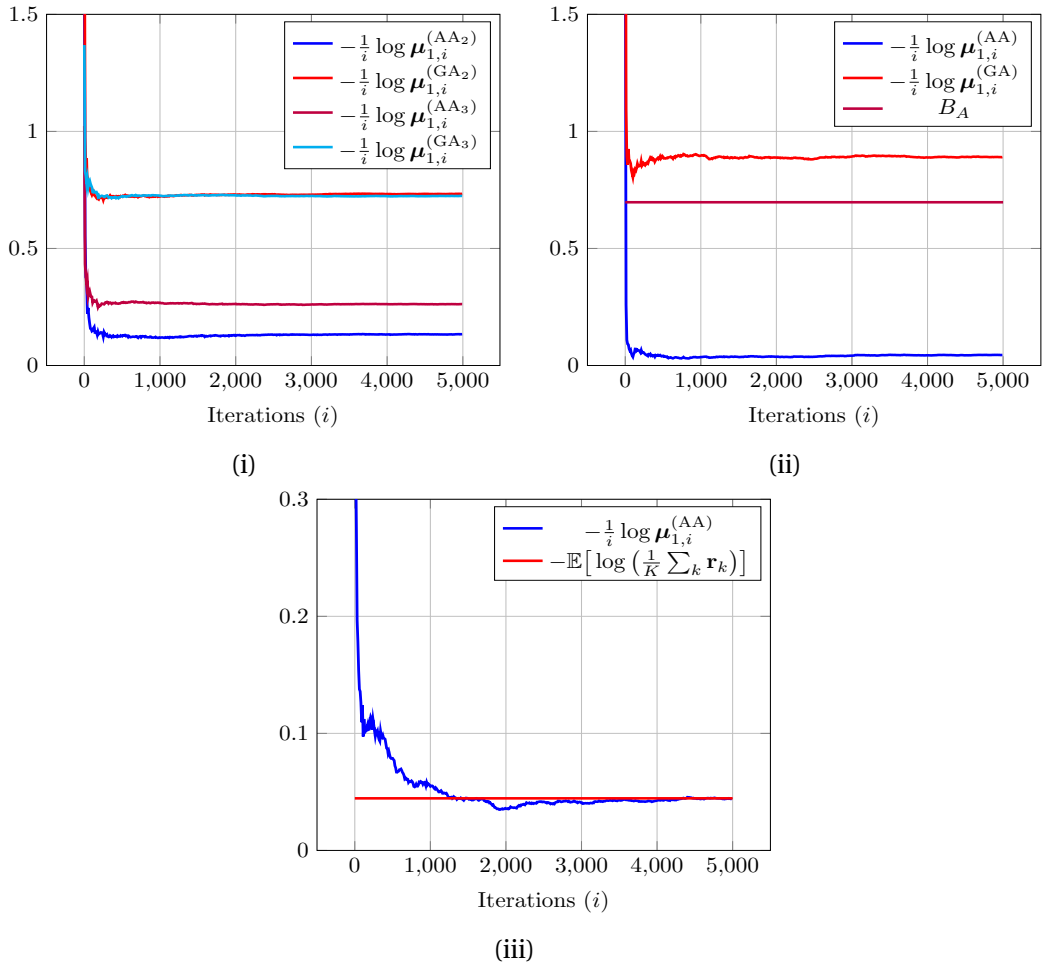


Figure 3.2: (i) Comparison of the decay rates of different network connectivities. The superscripts on the decay rates indicate the algorithm, and the network where we execute the algorithm, e.g., AA_2 means AA-diffusion is executed on the 2-regular network 3.1a with combination matrix A_2. The decay rates of GA-diffusion are the same for networks 3.1a and 3.1b whereas they differ when AA-diffusion is run. (ii) The decay rates of the non-regular network 3.1c are plotted under AA and GA-diffusion in the exchangeable setting; and the closed form bound B_A is shown. (iii) Time-scaled minus log-belief of agent 1 in the fully connected network with a rank-one combination matrix. This entity approaches the closed form expression we provided in Corollary 3.3.

Recall that Dob_A was defined in (3.48). We simulated AA-diffusion on the non-regular network $3.1c$ with its combination matrix chosen according to a lazy Metropolis rule [58]. More precisely, we take $B \triangleq [b_{\ell k}]$ with $b_{\ell k} = \max\{\deg(\ell), \deg(k)\}^{-1}$ for $\ell \neq k$ and $b_{\ell\ell} = 1 - \sum_{k\neq\ell} b_{\ell k}$. Then, we set $A_{\mathrm{non}} = \alpha I + (1-\alpha)B$. The matrix A_{non} is also doubly stochastic, hence, its Perron vector is the same as networks 3.1a and 3.1b. We have plotted the decay rates of agent 1 with the combination matrix A_{non} and also indicated the bound B_A in Figure 3.2ii. Since the Dobrushin coefficient $\mathsf{Dob}_A = 0.81$, B_A gives a non-trivial bound.

The third experiment includes a fully-connected network with combination matrix $a_{\ell k} = 1/K$. Observe A is rank-one and from Corollary 3.3, we know

$$\rho^{(\mathrm{AA})} = -\mathbb{E}\Big[\log\Big(\frac{1}{K}\sum_k \boldsymbol{r}_k\Big)\Big] \approx 0.0457. \tag{3.62}$$

The AA-diffusion decay rate corresponding to the final experiment is plotted in Figure 3.2iii and it is readily seen that the minus log-belief with time scaling approaches 0.0457.

The next part of this section consists of numerical examples where the "inept agent" phenomenon is observed — recall the end of Section 3.3.1. We set the inept agent to agent 1, and select $\beta_1 = 1$, i.e., $D_{\mathrm{KL}}(L_1(\cdot|\theta^\circ)||L_1(\cdot|\theta)) = 0$. The other β_k's remain the same as in the simulations in Figure 3.2i. The simulations take place over the non-regular network 3.1c, and four different connecivity matrices A_{α_1}, A_{α_2}, A_{α_3}, A_{α_4} are set with the same lazy Metropolis rule in 3.2ii, with four different α values $\alpha_1 = 0.01$, $\alpha_1 = 0.5$, $\alpha_1 = 0.8$, $\alpha_4 = 0.95$. Note that the Perron vector remains unchanged when we change α, hence the decay rate of GA should also remain unchanged. This is observed in Figure 3.3i. However, as α increases, the inept agent 1 becomes more self-confident and the learning rate for AA decreases drastically, which is evident from Figure 3.3i.

For the remaining simulations, we assume that, under θ°, agents observe standard Gaussian random variables and under θ they observe unit-variance Gaussian random variables with mean 10. The $K \times K$ covariance matrix of the data under both hypotheses is given by

$$\Sigma(c) \triangleq c\mathbb{1}_K\mathbb{1}_K^\mathsf{T} + (1-c)I_K \tag{3.63}$$

for a $c \in [0,1]$. Observe that $c = 1$ corresponds to the case where all agents observe the same data and $c = 0$ corresponds to the i.i.d. case. This setup ensures that the data is exchangeable.

The following simulation aims to investigate the effect of network connectivity on the decay rates of AA-diffusion under exchangeable networks. We choose $c = 0.5$, and the network is simulated on $D =$ 2,3,4,6-regular networks. The D regular networks are constructed as follows: The neighbor of agent k is set as

$$\mathcal{N}_k = \{k - \lfloor D/2 \rfloor, \ldots k-1, k+1, \ldots, k + \lceil D/2 \rceil\} \mod K. \tag{3.64}$$

Note that the 2-regular network constructed with this method is equivalent to that in Figure 3.1a whereas the 3-regular network is different from the one in Figure 3.1b. The connectivity matrices are set as in Figure 3.2i with the same $\alpha = 0.05$. In line with Theorem 3.4, Figure 3.3ii illustrates that the performance gap increases with connectivity. Finally, we compare the decay rates under various joint distributions under exchangeable setting. We simulate AA-diffusion under the same Gaussian

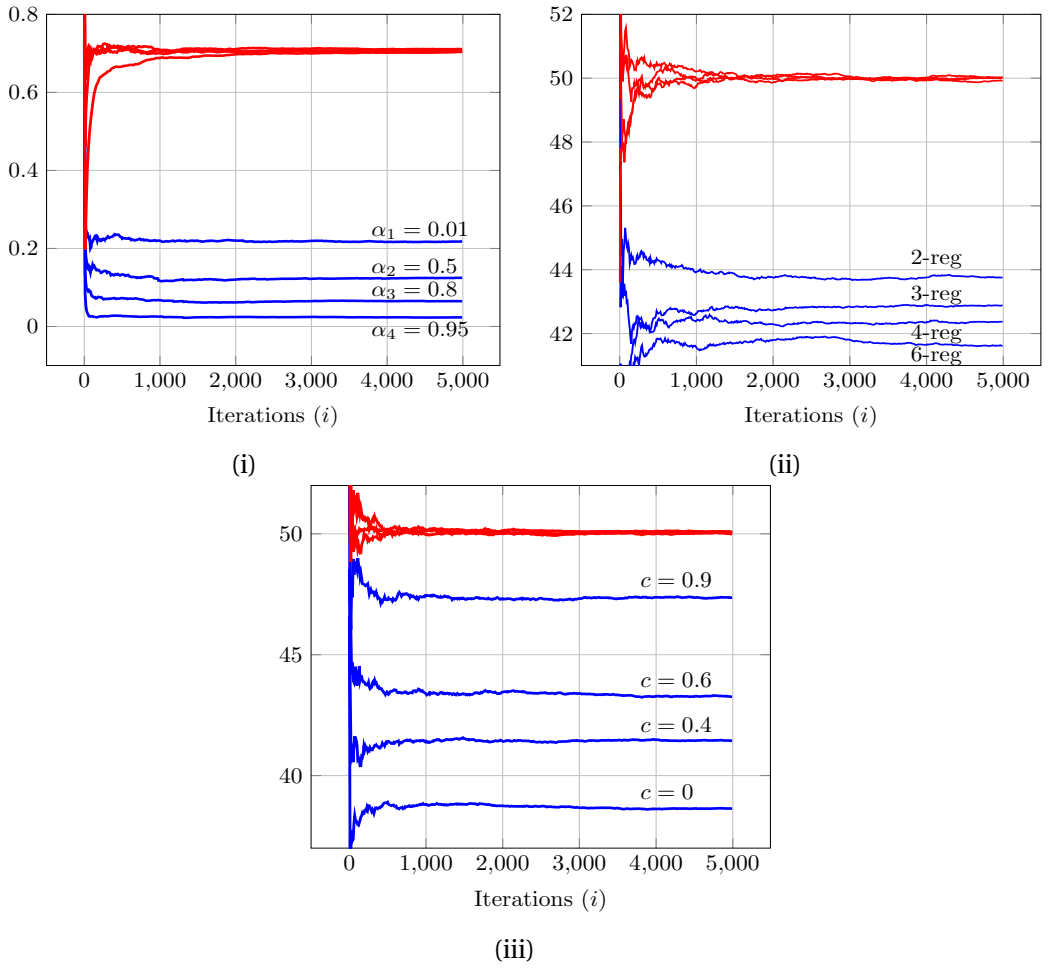


Figure 3.3: (i) Numerical example for the "inept agent" phenomenon mentioned at the end of Section 3.3.1. We have drawn the time-scaled minus log-beliefs, i.e., $-\frac{1}{i}\log \boldsymbol{\mu}_{1,i}$ of agent 1 over the network 3.1c and with connectivity matrices A_{α_1}, A_{α_2}, A_{α_3}, A_{α_4}. The red curves correspond to the decay rates of GA-diffusion and the blue curves correspond to the decay rates of AA-diffusion. The decrease in the AA-diffusion decay rate with respect to the increase in α is evident. (ii) Comparison of different network connectivities in an exchangeable setting. The red curves correspond to the decay rates of GA-diffusion and the blue curves correspond to the decay rates of AA-diffusion simulated on 2,3,4,6-regular networks. It is clearly visible that the decay rate decreases with network connectivity. (iii) Comparison of decay rates under different joint distributions for the Gaussian setting. The red curves correspond to the decay rates of GA-diffusion and the blue curves correspond to the decay rates of AA-diffusion simulated for $\Sigma(c)$, $c = 0, 0.4, 0.6, 0.9$.

setting for $c = 0, 0.4, 0.6, 0.9$. Recall that the covariance matrix was given by $\Sigma(c)$ from (3.63). We also observe that the decay rate increases when c increases, and in particular, is equal to the decay rate under GA-diffusion for $c = 1$ — every agent observes same data if $c = 1$.

3.5 Concluding Remarks

In this chapter, we compared the decay rates of beliefs under arithmetic and geometric averaging for social learning over networks. For arithmetic averaging strategies, we established that the beliefs on wrong hypotheses decay exponentially almost surely, and the decay rates are the same among all agents. We provided upper and lower bounds for the decay rates, one being the decay rate corresponding to geometric averaging (Corollary 3.2), while the other one revealing the network learning rate's sensitivity to individual agents.

We studied the performance gap between arithmetic and geometric averaging with a formulation that permitted the use of the strong data processing inequality. We have proved that for a broad class of networks, there is no performance gap if, and only if, all agents observe the same data. We also obtained closed form bounds and expressions for the performance gap for some special instances.

An interesting future direction is to examine how our results relate to the distributed estimation and filtering. Such methods primarily focus on the inference of continuous variables, and they typically rely on the fusion of point estimates. Nevertheless, they can be interpreted in terms of fusion of probability density functions [35]. Therefore, a natural question is the applicability of our results for social learning to that line of work. Furthermore, all results in this work are asymptotic. In the subsequent work [59], we established the asymptotically normal behavior of learning rates for both arithmetic and geometric averaging under federated architectures. Employing the Berry-Esseen theorem [60] and related tools on top of this work to understand finite sample behaviors of social learning can be another interesting direction to study.

3.A Proof of Corollary 3.1

To make use of Theorem 3.1, we first need to ensure that the random matrices have all positive entries. Recall that there must exist an $n \geq 1$ such that every entry of A^n is strictly positive. We choose the smallest such n. Moreover, the $\{\boldsymbol{r}_{k,i}\}$ are strictly positive with probability 1 as well — otherwise some of the KL divergences would be infinite. We therefore replace $\mathbf{X}_i$ with the expression

$$\widetilde{\mathbf{X}}_i \triangleq A^{\mathsf{T}} \boldsymbol{R}_{n(i-1)+1} \ldots A^{\mathsf{T}} \boldsymbol{R}_{ni}, \tag{3.65}$$

which turn out to have all positive entries. To see this, observe that

$$[\widetilde{\mathbf{X}}_i]_{11} = \sum_{\ell_1,\ldots,\ell_{n-1}} a_{\ell_1 1} a_{\ell_2 \ell_1} \ldots a_{1\ell_{n-1}} \boldsymbol{r}_{\ell_1,1} \ldots \boldsymbol{r}_{1,n}. \tag{3.66}$$

Since A is primitive, there must exist $\ell_1, \ldots, \ell_{n-1}$ such that $a_{\ell_1 1} \ldots a_{1\ell_{n-1}} > 0$. If we choose such a path, we can observe that

$$[\widetilde{\mathbf{X}}_i]_{11} \geq a_{\ell_1 1} a_{\ell_2 \ell_1} \ldots a_{1\ell_{n-1}} \boldsymbol{r}_{\ell_1,1} \ldots \boldsymbol{r}_{1,n}. \tag{3.67}$$

As mentioned, the $\boldsymbol{r}_{k,i}$ are strictly positive and thus $[\widetilde{\mathbf{X}}_i]_{11}$ is strictly positive as well. Similar arguments hold for other entries $[\widetilde{\mathbf{X}}_i]_{\ell k}$ as well, which proves that all entries of $\widetilde{\mathbf{X}}_i$ are positive. Note that the $\widetilde{\mathbf{X}}_i$'s are also i.i.d.. The next step is to check if the logarithms of the entries of $\widetilde{\mathbf{X}}_i$ have finite expectations. First of all, since $\mathbb{E}[\boldsymbol{R}_i] = I$ and the $\boldsymbol{R}_i$ are i.i.d., it holds that

$$\mathbb{E}[\widetilde{\mathbf{X}}_i] = \mathbb{E}[A^\mathsf{T} \boldsymbol{R}_{n(i-1)+1} \ldots A^\mathsf{T} \boldsymbol{R}_{ni}] = (A^\mathsf{T})^n, \tag{3.68}$$

which further implies from Jensen's inequality that

$$\mathbb{E}[\log[\widetilde{\mathbf{X}}]_{\ell k}] \leq \log \mathbb{E}[[\widetilde{\mathbf{X}}]_{\ell k}] = \log[A^n]_{k\ell} \leq 0. \tag{3.69}$$

Therefore, the expectations of the logarithms are bounded from above. To check if they are also bounded from below, if we take the logarithms of both sides in (3.67), we obtain

$$\mathbb{E}[\log[\widetilde{\mathbf{X}}_i]_{11}] \geq \log(a_{\ell_1 1} a_{\ell_2 \ell_1} \ldots a_{1\ell_{n-1}}) + \mathbb{E}[\log \boldsymbol{r}_{\ell_1,1}] + \cdots + \mathbb{E}[\log \boldsymbol{r}_{1,n}]. \tag{3.70}$$

By finiteness of KL divergences, it holds that

$$\mathbb{E} \log \boldsymbol{r}_{k,1} = -D_{\mathrm{KL}}(L_k(.|\theta^\circ)||L_k(.|\theta)) > -\infty. \tag{3.71}$$

Consequently, $\mathbb{E}[\log[\widetilde{\mathbf{X}}_i]_{11}]$ is bounded from below. A similar argument holds for the other terms $\mathbb{E}[\log[\widetilde{\mathbf{X}}_i]_{\ell k}]$ as well.

Since the $[\widetilde{\mathbf{X}}_i]_{\ell k}$ are strictly positive and their logarithms have finite expectations, we can invoke Theorem 3.1 to obtain

$$\tilde{\boldsymbol{\gamma}} = \lim_{i \to \infty} \frac{1}{i} \log[\widetilde{\boldsymbol{Y}}_i]_{\ell k} \tag{3.72}$$

with $\widetilde{\boldsymbol{Y}}_i \triangleq \prod_{j=1}^{i} \widetilde{\mathbf{X}}_j$. Since $\widetilde{\boldsymbol{Y}}_i = \boldsymbol{Y}_{ni}$, the above equation implies

$$\boldsymbol{\gamma} = \frac{\tilde{\boldsymbol{\gamma}}}{n} = \lim_{i \to \infty} \frac{1}{ni} \log[\boldsymbol{Y}_{ni}]_{\ell k} = \lim_{j \to \infty} \frac{1}{j} \log[\boldsymbol{Y}_j]_{\ell k}. \tag{3.73}$$

Moreover, since $\{A^\mathsf{T} \boldsymbol{R}_i\}$ is an i.i.d. sequence, Kolmogorov's zero-one law [40] implies that the finite limit $\boldsymbol{\gamma}$ is almost surely a constant. Hence,

$$\gamma \triangleq \mathbb{E}[\boldsymbol{\gamma}] = \lim_{i \to \infty} \frac{1}{i} \mathbb{E}\Big[\log[\boldsymbol{Y}_i]_{\ell k} \Big], \tag{3.74}$$

and the proof is complete.

3.B Proof of Theorem 3.2

It is sufficient to establish the result for one agent k; a similar argument applies to the other agents. We establish the proof in two parts by showing:

- (i) $\limsup_{i\to\infty} \frac{1}{i}\log \boldsymbol{\mu}_{k,i} \leq \gamma$,
- (ii) $\liminf_{i\to\infty} \frac{1}{i}\log \boldsymbol{\mu}_{k,i} \geq \gamma$.

Part (i) of the proof makes use of the extremal process $\{\boldsymbol{\nu}_i\}$. Selecting a $\delta > 0$, we define the events

$$\mathcal{H}_{i_0}^{+}(\delta) \triangleq \left\{\omega \in \Omega : \exists i_1 \geq i_0, \forall i \geq i_1, \frac{1}{i-i_0}\max_{\ell,k} \log[\boldsymbol{Y}_{i_0}^{i}]_{\ell k} \leq \gamma + \delta\right\} \tag{3.75}$$

for every $i_0 \geq 1$ with

$$\boldsymbol{Y}_{i_0}^{i} \triangleq \prod_{j=i_0+1}^{i} (A^{\mathsf{T}}\boldsymbol{R}_j). \tag{3.76}$$

In words, $\mathcal{H}_{i_0}^{+}(\delta)$ is the event that the logarithms of all entries of $\boldsymbol{Y}_{i_0}^{i}$ eventually become smaller than $(\gamma+\delta)(i-i_0)$. Since $\{A^{\mathsf{T}}\boldsymbol{R}_i\}$ is i.i.d., $\boldsymbol{Y}_i$ is stationary; and $\mathbb{P}\left(\mathcal{H}_{i_0}^{+}(\delta)\right)$ does not depend on i_0. Corollary 3.1 states that $\mathbb{P}\left(\mathcal{H}_{0}^{+}(\delta)\right) = 1$, and we deduce that $\mathbb{P}\left(\mathcal{H}_{i_0}^{+}(\delta)\right) = 1$. As any countable intersection of unit-probability events is also unit-probability, we have $\mathbb{P}\left(\mathcal{H}^{+}(\delta)\right) = 1$ where

$$\mathcal{H}^{+}(\delta) \triangleq \bigcap_{i_0} \mathcal{H}_{i_0}^{+}(\delta). \tag{3.77}$$

Consider an $\omega \in \mathcal{G}(\epsilon) \cap \mathcal{H}^{+}(\delta)$, with $\mathcal{G}(\epsilon)$ defined in (3.4). Repeated application of (3.8) yields

$$\boldsymbol{\nu}_i = (1-\epsilon)^{\boldsymbol{i}_0 - i}\boldsymbol{Y}_{\boldsymbol{i}_0}^{i}\boldsymbol{\nu}_{\boldsymbol{i}_0}, \quad \forall i \geq \boldsymbol{i}_0(\omega). \tag{3.78}$$

Furthermore, since $\omega \in \mathcal{H}^{+}(\delta)$ implies

$$[\boldsymbol{Y}_{\boldsymbol{i}_0}^{i}]_{\ell k} \leq e^{(i-\boldsymbol{i}_0)(\gamma+\delta)} \tag{3.79}$$

for all $i \geq \boldsymbol{i}_1(\omega)$,

$$\boldsymbol{\nu}_{k,i} \leq e^{(i-\boldsymbol{i}_0)(\gamma+\delta+\epsilon')}, \quad \forall i \geq \boldsymbol{i}_1(\omega) \tag{3.80}$$

with some $\epsilon' \triangleq -\log(1-\epsilon)$. Thus,

$$\limsup_{i\to\infty} \frac{1}{i}\log \boldsymbol{\mu}_{k,i} \le \limsup_{i\to\infty} \frac{1}{i}\log \boldsymbol{\nu}_{k,i} \le \gamma + \delta + \epsilon' \tag{3.81}$$

for all $\omega \in \mathcal{G}(\epsilon) \cap \mathcal{H}^+(\delta)$. Since $\mathcal{G}(\epsilon)$ and $\mathcal{H}^+(\delta)$ are both probability one events, so is their intersection. This completes part (i) of the proof.

Part (ii) of the proof requires the construction of another extremal process $\{\boldsymbol{z}_{k,i}\}$, which lower bounds $\{\boldsymbol{\mu}_{k,i}\}$. Similar to part (i), we define the events

$$\mathcal{H}^-_{i_0}(\delta) \triangleq \left\{\omega \in \Omega : \exists i_1 \ge i_0, \forall i \ge i_1, \frac{1}{i-i_0}\min_{\ell,k} \log[\boldsymbol{Y}^i_{i_0}]_{\ell k} \ge \gamma - \delta\right\} \tag{3.82}$$

and

$$\mathcal{H}^-(\delta) \triangleq \bigcap_{i_0} \mathcal{H}^-_{i_0}(\delta). \tag{3.83}$$

Then, for any $\omega \in \mathcal{G}(\epsilon) \cap \mathcal{H}^-(\delta)$, we have $\boldsymbol{\mu}_{k,i}(\theta) \le \epsilon$ for $i \ge \boldsymbol{i}_0(\omega)$ and

$$\begin{aligned}\boldsymbol{\mu}_{k,i}(\theta) &\ge \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)\boldsymbol{r}_{\ell,i}(\theta)}{1+\epsilon\sum\limits_{\theta'\neq\theta^\circ}\boldsymbol{r}_{\ell,i}(\theta)} \\ &\ge \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)\boldsymbol{r}_{\ell,i}(\theta)}{1+\epsilon\sum\limits_{\theta'\neq\theta^\circ}\sum_\ell \boldsymbol{r}_{\ell,i}(\theta')}.\end{aligned} \tag{3.84}$$

We introduce the vector $\boldsymbol{z}_i = [\boldsymbol{z}_{1,i}, \boldsymbol{z}_{2,i}, \ldots, \boldsymbol{z}_{K,i}]^\mathsf{T}$ and define its evolution for $i \ge \boldsymbol{i}_0$ as

$$\boldsymbol{z}_i = (A^\mathsf{T}\boldsymbol{R}_i)\boldsymbol{z}_{i-1}, \quad \boldsymbol{z}_{\boldsymbol{i}_0} = \boldsymbol{\mu}_{\boldsymbol{i}_0}. \tag{3.85}$$

This implies that

$$\begin{aligned}\frac{1}{i}\log\boldsymbol{\mu}_{k,i} &\ge \frac{1}{i}\log\boldsymbol{z}_{k,i} - \frac{1}{i}\sum_{j=i_0}^{i}\log\left(1+\epsilon\sum_{\theta'\neq\theta^\circ}\sum_\ell \boldsymbol{r}_{\ell,j}(\theta')\right) \\ &\overset{(a)}{\ge} \frac{1}{i}\log\boldsymbol{z}_{k,i} - \epsilon\frac{1}{i}\sum_{j=i_0}^{i}\sum_{\theta'\neq\theta^\circ}\sum_\ell \boldsymbol{r}_{\ell,j}(\theta')\end{aligned} \tag{3.86}$$

where (a) follows from using $\log(1+x) \le x$. Observe that by the strong law of large numbers

$$\begin{aligned}\lim_{i\to\infty}\epsilon\frac{1}{i}\sum_{j=i_0}^{i}\sum_{\theta'\neq\theta^\circ}\sum_\ell \boldsymbol{r}_{\ell,j}(\theta') &= \epsilon K(H-1)\mathbb{E}[\boldsymbol{r}_{\ell,j}] \\ &= \epsilon K(H-1).\end{aligned} \tag{3.87}$$

Hence, proceeding similarly to the previous part we obtain that almost surely

$$\begin{aligned}\liminf_{i\to\infty}\frac{1}{i}\log\boldsymbol{\mu}_{k,i} &\geq \liminf_{i\to\infty}\frac{1}{i}\log\boldsymbol{z}_{k,i} - \epsilon K(H-1)\\ &\geq \gamma-\delta-\epsilon K(H-1).\end{aligned} \tag{3.88}$$

Since δ and ϵ are arbitrary, the proof is complete.

3.C Conditions for the Subadditive Ergodic Theorem

We first restate the subadditive ergodic theorem.

Theorem 3.5 (**Subadditive ergodic theorem [51, Theorem 1]**). *Let $\{\boldsymbol{x}_{ij}\}_{i\leq j}$ be a doubly indexed random sequence that satisfies the following conditions:*

(i) The distribution of $\boldsymbol{x}_{ij}$ depends only on $j-i$.

(ii) $\boldsymbol{x}_{ij}\leq \boldsymbol{x}_{ik}+\boldsymbol{x}_{kj}$, for all $i\leq k\leq j$.

(iii) $\frac{1}{j}\mathbb{E}[\boldsymbol{x}_{1j}]\geq -\kappa$ for some constant κ and for all $j\geq 1$.

Then, the finite limit

$$\boldsymbol{\gamma}=\lim_{i\to\infty}\frac{1}{i}\boldsymbol{x}_{1i} \tag{3.89}$$

exists almost surely and in the mean, and furthermore,

$$\mathbb{E}[\boldsymbol{\gamma}]=\lim_{i\to\infty}\frac{1}{i}\mathbb{E}[\boldsymbol{x}_{1i}]. \tag{3.90}$$

To use this result, we replace $\boldsymbol{x}_{ij}$ with $-\log\|\boldsymbol{Y}_i^j\|_-$ (defined in (3.76)). Since $\{A^{\mathsf{T}}\boldsymbol{R}_i\}$ is an i.i.d. sequence, $\boldsymbol{Y}_i^j$ is stationary and (i) is satisfied. In the relation

$$\|\boldsymbol{Y}_i^k\|_-\|\boldsymbol{Y}_k^j\|_-\leq\|\boldsymbol{Y}_i^j\|_-, \tag{3.91}$$

taking the logarithm and negating both sides shows immediately that (ii) is satisfied. The only remaining condition to verify is (iii), i.e., whether

$$\mathbb{E}[-\log\|\boldsymbol{Y}_i\|_-]\geq -\kappa i \tag{3.92}$$

for some constant κ. We achieve this result by following similar steps to the proof

of [51, Theorem 5]:

$$\begin{aligned}\mathbb{E}[\log \|\boldsymbol{Y}_i\|_-] &\leq \mathbb{E}[\log \|\boldsymbol{Y}_i\|_1] \\ &\overset{(a)}{\leq} i\mathbb{E}[\log \|\boldsymbol{Y}_1\|_1] \\ &\leq i\mathbb{E}[\log(K \max_{k,\ell}[\boldsymbol{Y}_1]_{k\ell})] \\ &\leq i\left(\mathbb{E}\left[\sum_{k,\ell}(\log[\boldsymbol{Y}_1]_{k\ell})^+\right] + \log K\right) \\ &\overset{(b)}{\leq} i\kappa \end{aligned} \tag{3.93}$$

where (a) follows from subadditivity, and (b) follows from Assumption 1. Hence, $-\log \|\boldsymbol{Y}_i\|_-$ satisfies the conditions of the subadditive ergodic theorem.

3.D Proof of Lemma 3.2

For $\pi \in \mathbb{S}_K$, it is known that [53] for some finite constant c, the KKT conditions are given by

$$\frac{\partial F}{\partial \pi_k} = c, \quad \pi_k > 0 \tag{3.94}$$

$$\frac{\partial F}{\partial \pi_k} \leq c, \quad \pi_k = 0. \tag{3.95}$$

We however have to justify the interchange of differentiation and expectation. Observe that

$$\begin{aligned}\frac{\partial}{\partial \pi_k} \log\Big(\sum_\ell \pi_\ell \boldsymbol{r}_\ell\Big) &= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\left(\log\Big(\sum_\ell \pi_\ell \boldsymbol{r}_\ell + \epsilon \boldsymbol{r}_k\Big) - \log\Big(\sum_\ell \pi_\ell \boldsymbol{r}_\ell\Big)\right) \\ &\leq \frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell} \end{aligned} \tag{3.96}$$

where we used the inequality $\log(1+x) \leq x$. Since all the random variables in (3.96) are uniformly bounded by $\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}$, and $\mathbb{E}\Big[\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\Big] < \infty$, the dominated convergence theorem [40] justifies interchanging differentiation and expectation. Therefore,

$$\frac{\partial F}{\partial \pi_k} = -\frac{v_k}{\pi_k} + \mathbb{E}\left[\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\right]. \tag{3.97}$$

Also, v, being the Perron vector of a primitive matrix A, has all strictly positive entries. Hence, setting any $v_k = 0$ will make $F(\pi)$ infinite. Therefore, all π_k must have positive

values. As a result, from (3.94),

$$-\frac{v_k}{\pi_k}+\mathbb{E}\left[\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\right]=\mu. \tag{3.98}$$

In this equation, multiplying both sides with π_k and summing over each agent k, we get $\mu=0$, which concludes the proof.

3.E Proof of Theorem 3.3

We first show that $\rho^{(\mathrm{GA})}=\rho^{(\mathrm{AA})}$ implies $\boldsymbol{r}_k=\boldsymbol{r}_\ell$ for all k,ℓ. Applying Jensen's inequality to exchange the logarithm and expectation, we observe that

$$G(\pi)\geq \sum_k v_k \mathbb{E}\left[\log\Big(\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\Big)\right]+\mathbb{E}\left[\log\Big(\sum_k \pi_k \boldsymbol{r}_k\Big)\right]+\rho^{(\mathrm{GA})}=0. \tag{3.99}$$

Since the logarithm is a strictly convex function, $G(\pi)=0$ only when

$$\log \mathbb{E}\left[\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\right]=\mathbb{E}\left[\log \frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\right]. \tag{3.100}$$

In other words, $G(\pi)=0$ only when $\dfrac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}$ is constant with probability one for all k. This means that for any k,ℓ,

$$\frac{\boldsymbol{r}_k}{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}\frac{\sum_\ell \pi_\ell \boldsymbol{r}_\ell}{\boldsymbol{r}_m}=\frac{\boldsymbol{r}_k}{\boldsymbol{r}_m} \tag{3.101}$$

is also constant. Also, since by definition

$$\mathbb{E}[\boldsymbol{r}_k]=\mathbb{E}[\boldsymbol{r}_m]=1, \tag{3.102}$$

$\boldsymbol{r}_k$ must be equal to $\boldsymbol{r}_m$ with probability one. This shows that if $\boldsymbol{r}_k\neq\boldsymbol{r}_\ell$ for some k, ℓ with non-zero probability, then $G(\pi)$ is strictly positive. Also, due to the fact that

$$\inf_{u\in\mathbb{S}_K}\mathbb{E}[D_{\mathrm{KL}}(v||\boldsymbol{u}_1)]\geq \inf_\pi F(\pi)\geq \inf_\pi G(\pi)>0 \tag{3.103}$$

and $\eta_A<1$, Eq. (3.41) implies $\rho^{(\mathrm{GA})}>\rho^{(\mathrm{AA})}$.

The other direction is more straightforward to establish. If for all k, ℓ, we have the relation $\boldsymbol{r}_k=\boldsymbol{r}_\ell=\boldsymbol{r}$, it holds that

$$[\boldsymbol{Y}_i]_{k\ell}=\prod_{j=1}^{i}\boldsymbol{r}_j[A^i]_{\ell k}. \tag{3.104}$$

This, in turn, implies

$$\frac{1}{i}\log[\boldsymbol{Y}_i]_{k\ell} \to \mathbb{E}[\log \boldsymbol{r}] = -\rho^{(\mathrm{GA})} \tag{3.105}$$

since every column of A^i tends to its Perron vector v.

4 Information Sharing

4.1 Introduction[1]

An implicit assumption common to our discussion so far is that agents are willing to share with neighbors their *full* belief vector. That is, they share their beliefs about all possible hypotheses. In the context of social networks, this can be an unrealistic assumption.

For instance, oftentimes, Twitter users concentrate on particular topics that constitute Twitter Trends. If candidate A gave a recent press release, Twitter users will likely focus on A when exchanging opinions and ignore other candidates. Another example can be users trying to determine the best mobile operator — see Fig. 4.1, which displays the evolution of discussions on Twitter over time about two operators in Switzerland. The communication trends in this case can change based on the campaigns and advertisements by the mobile service providers. Furthermore, in the context of engineering systems, transmitting partial beliefs rather than full beliefs enables the design of communication-efficient systems under limited resources.

Motivated by these examples, in this chapter, we are interested in the case where social agents share information on a random hypothesis of interest at each iteration.

4.1.1 Contributions.

- We propose a social learning algorithm, where agents share their beliefs on *only one* randomly chosen hypothesis at each time instant — see Section 4.3. Compared to the prior work [62], the hypothesis being exchanged between agents is allowed to change over time. Moreover, under this partial information sharing scheme, agents complete the missing components of the received beliefs by using their own beliefs.

[1]The material in this chapter is based on [61].

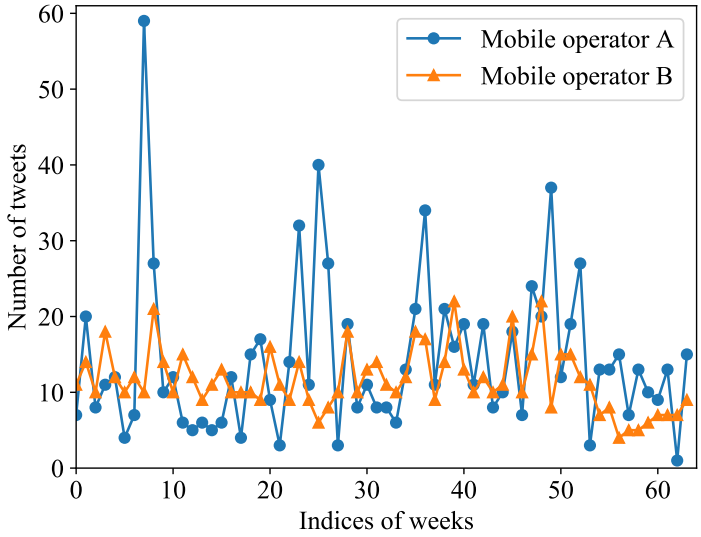


Figure 4.1: Number of tweets about two mobile operators for each week. (Mobile operator A- Swisscom, B- SunRise). Total of 64 weeks between 01/01/2021 and 30/04/2022, tweets in Switzerland.

- When a wrong hypothesis is exchanged with positive probability, we show that beliefs evaluated at that hypothesis decay exponentially. Theorem 4.3 establishes that the decay rate is the same as the asymptotic learning rate of traditional geometric social learning algorithm in Algorithm 2.1. As a result, if each wrong hypothesis is exchanged with positive probability, then learning occurs with probability 1 (Corollary 4.1).

- We develop proof techniques to tackle the randomness in the combination policy stemming from the shared hypothesis. The constraint that the beliefs belong to the probability simplex couples the processes for different hypotheses, making the standard application of the strong-law-of-large-numbers nonfeasible. Thus, we utilize martingale arguments to handle the non-linearity.

- We provide a counter-example in Section 4.5 to show that sharing information about the true hypothesis is not sufficient for truth learning with full confidence when agents use their own beliefs as priors for other agents' beliefs.

- On the other hand, Theorem 4.4 states that agents will never discard the truth completely. Namely, their beliefs on the true hypothesis will never be zero, which also means that they will never be fully confident on a wrong hypothesis being the true hypothesis. This contrasts with the findings in [62], where truth sharing is shown to lead to truth learning when agents adopt uniform priors for the missing components. However, with such an approach, agents might also become fully confident in an incorrect hypothesis.

The randomness in the shared hypothesis and the incomplete information received from the neighbors introduce new challenges into the analysis, in comparison to the proof of full information sharing setting described in Theorem 2.1. This aspect is treated in later sections.

4.2 Partial Information Sharing

The canonical social learning setting described in Chapter 2 assumes that agents share their *entire* beliefs with each other at every iteration. The work [62] considers an alternative setting where agents share their beliefs on only one hypothesis of interest, denoted by $\tau \in \Theta$, which is *fixed* over time. In particular, after agents perform the local adaptation step (2.7), each agent k receives $\boldsymbol{\psi}_{\ell,i}(\tau)$ from its neighbors $\ell \in \mathcal{N}_k$. Then, each agent k completes the missing entries of the received belief vectors by using the construction:

$$\widehat{\boldsymbol{\psi}}_{\ell,i}(\theta) = \begin{cases} \boldsymbol{\psi}_{\ell,i}(\tau), & \theta = \tau \\ \dfrac{1-\boldsymbol{\psi}_{\ell,i}(\tau)}{H-1}, & \theta \neq \tau \end{cases} \tag{4.1}$$

for $\ell \in \mathcal{N}_k$. Observe that the received component of the intermediate belief is used as is, while the remaining components are assigned uniform weight. Then, these modified beliefs are used in the combination step:

$$\boldsymbol{\mu}_{k,i}(\theta) = \frac{\prod\limits_{\ell \in \mathcal{N}_k} (\widehat{\boldsymbol{\psi}}_{\ell,i}(\theta))^{a_{\ell k}}}{\sum\limits_{\theta' \in \Theta} \prod\limits_{\ell \in \mathcal{N}_k} (\widehat{\boldsymbol{\psi}}_{\ell,i}(\theta'))^{a_{\ell k}}}. \tag{4.2}$$

The work [62] provides a detailed analysis of this algorithm and proposes a self-aware variant of it where each agent k uses $\boldsymbol{\psi}_{k,i}(\theta)$ instead of the approximation $\widehat{\boldsymbol{\psi}}_{k,i}(\theta)$ in (4.2). For the purposes of this chapter, it is enough to state the following results, which we will refer to in the sequel. Let us introduce the following notation for the average likelihood function. The average is over the non-transmitted components (denoted by the symbol $\overline{\tau}$):

$$L_k(\xi|\overline{\tau}) \triangleq \frac{1}{H-1} \sum_{\theta \in \Theta \setminus \{\tau\}} L_k(\xi|\theta). \tag{4.3}$$

Then, the following two results are from [62].

Theorem 4.1 (**Learning with truth sharing [62, Theorem 1]**). *If $\tau = \theta^\circ$, under Assumption 2.2 and a modified global identifiability assumption, i.e., if there exists an agent such that*

$$D_{\mathrm{KL}}\Big(L_k(\cdot|\tau)\Big\|L_k(\cdot|\overline{\tau})\Big) > 0, \tag{4.4}$$

then, truth learning occurs with full confidence:

$$\boldsymbol{\mu}_{k,i}(\tau) \xrightarrow{\text{a.s.}} 1 \tag{4.5}$$

for each agent k under (4.1)–(4.2).

Theorem 4.1 shows that if the hypothesis τ that agents exchange information about happens to be the truth θ°, then this is sufficient for truth learning with full confidence. However, if they are not discussing the truth, this can lead to mislearning a wrong hypothesis as the next result illustrates.

Theorem 4.2 (**Total mislearning [62, Theorem 3]**). *If $\tau \neq \theta^\circ$, under Assumption 2.2 and the condition that*

$$\sum_{k=1}^{K} v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\tau)\Big) < \sum_{k=1}^{K} v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\overline{\tau})\Big) \tag{4.6}$$

all agents learn a wrong hypothesis with full confidence:

$$\boldsymbol{\mu}_{k,i}(\tau) \xrightarrow{\text{a.s.}} 1 \tag{4.7}$$

for each agent k under (4.1)–(4.2).

In the algorithm we propose in this chapter (Algorithm 4.1), agents will fill the missing beliefs using their own information, instead of assigning uniform values to them. This practice will result in outcomes that are opposite to Theorems 4.1 and 4.2. Namely, we show in the following that exclusive truth sharing does not suffice for truth learning (Sec. 4.5), while total mislearning can never occur under Algorithm 4.1.

4.3 Social Learning under Trending Topics

As opposed to existing works that require transmission of the entire beliefs at each iteration, or exchanging a fixed component of the beliefs, in this work we allow agents to share a random component at each iteration. Similar to [62], each agent will continue to compute its intermediate belief $\boldsymbol{\psi}_{k,i}(\theta)$ for every possible $\theta \in \Theta$ according to (2.7), and the agents will continue to share with their neighbors information about their

belief, albeit for *only one* of the hypotheses. However, the agents will now be sharing information about the same trending hypothesis $\boldsymbol{\tau}_i \in \Theta$, where, $\boldsymbol{\tau}_i$ is randomly selected and allowed to change over time.

The motivation for this setting is two-fold. First, people tend to concentrate on particular topics of discussion over social networks, as indicated earlier in the introduction. Second, if agents (e.g., sensors, robots) have the same random key, they can randomly select the same hypothesis at each iteration without needing a central controller, and then exchange beliefs on that hypothesis alone, as opposed to the more costly approach of exchanging entire beliefs. Utilizing randomness this way can prove useful compared to a periodic scheduling scheme for security reasons, since deciphering a random key is harder for eavesdroppers than inferring a fixed schedule.

We denote the distribution of $\boldsymbol{\tau}_i$ by π, where π is a probability vector over the set of hypotheses Θ, and write $\mathbb{P}(\boldsymbol{\tau}_i = \theta) = \pi_\theta$. We assume that $\boldsymbol{\tau}_i$ is i.i.d. over time and also independent of all observations $\{\xi_{k,j}\}$ over space and time.

Again, since the agents receive incomplete belief vectors from their neighbors (actually, they receive only one entry from these vectors), the agents will need to complete the missing entries. In the proposed strategy, the agents use their own intermediate local beliefs to fill in for the missing beliefs from their neighbors by using the following construction. Namely, agent k completes the belief vector received from its neighbor ℓ by using

$$\boldsymbol{\phi}_{\ell,i}^{(k)}(\theta) = \begin{cases} \boldsymbol{\psi}_{\ell,i}(\theta), & \theta = \boldsymbol{\tau}_i \\ \boldsymbol{\psi}_{k,i}(\theta), & \theta \neq \boldsymbol{\tau}_i \end{cases}. \tag{4.8}$$

By doing so, the entries of $\boldsymbol{\phi}_{\ell,i}^{(k)}$ need not add up to 1. For this reason, agent k normalizes (4.8) according to:

$$\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta) = \frac{\boldsymbol{\phi}_{\ell,i}^{(k)}(\theta)}{\sum_{\theta' \in \Theta} \boldsymbol{\phi}_{\ell,i}^{(k)}(\theta')} \tag{4.9}$$

whose denominator can be written as

$$\sum_{\theta' \in \Theta} \boldsymbol{\phi}_{\ell,i}^{(k)}(\theta') = 1 - \boldsymbol{\psi}_{k,i}(\boldsymbol{\tau}_i) + \boldsymbol{\psi}_{\ell,i}(\boldsymbol{\tau}_i). \tag{4.10}$$

We refer to (4.8)–(4.9) as a bootstrapping step. Subsequently, the agents combine the *approximate* intermediate beliefs from (4.9) to update their beliefs as in (4.2):

$$\boldsymbol{\mu}_{k,i}(\theta) = \frac{\prod_{\ell \in \mathcal{N}_k} (\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta))^{a_{\ell k}}}{\sum_{\theta' \in \Theta} \prod_{\ell \in \mathcal{N}_k} (\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta'))^{a_{\ell k}}}. \tag{4.11}$$

The procedure is summarized in Algorithm 4.1. Observe that the algorithm is *self-aware* because for each agent k, it follows from (4.13) that $\widehat{\boldsymbol{\psi}}_{k,i}^{(k)}(\theta) = \boldsymbol{\psi}_{k,i}(\theta)$. In other words, agents use their own intermediate beliefs as is.

Algorithm 4.1 Social learning with trending hypothesis

set initial priors $\boldsymbol{\mu}_{k,0}(\theta) > 0$, $\forall \theta \in \Theta$ and $\forall k \in \mathcal{N}$
while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{k,i-1}(\theta) \tag{4.12}$$

  end for
  agents **exchange** $\{\boldsymbol{\psi}_{k,i}(\boldsymbol{\tau}_i)\}$ on the current hypothesis of interest $\boldsymbol{\tau}_i \sim \pi$
  for each agent $k = 1, 2, \ldots, K$ **do**
    approximate the intermediate beliefs for $\ell \in \mathcal{N}_k$ by **bootstrapping**:

$$\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta) = \begin{cases} \dfrac{\boldsymbol{\psi}_{\ell,i}(\theta)}{1 - \boldsymbol{\psi}_{k,i}(\boldsymbol{\tau}_i) + \boldsymbol{\psi}_{\ell,i}(\boldsymbol{\tau}_i)}, & \theta = \boldsymbol{\tau}_i \\ \dfrac{\boldsymbol{\psi}_{k,i}(\theta)}{1 - \boldsymbol{\psi}_{k,i}(\boldsymbol{\tau}_i) + \boldsymbol{\psi}_{\ell,i}(\boldsymbol{\tau}_i)}, & \theta \neq \boldsymbol{\tau}_i \end{cases} \tag{4.13}$$

    combine approximate beliefs:

$$\boldsymbol{\mu}_{k,i}(\theta) \propto \prod_{\ell \in \mathcal{N}_k} (\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta))^{a_{\ell k}} \tag{4.14}$$

  end for
  $i \leftarrow i + 1$
end while

Recall that in geometric non-Bayesian social learning Algorithm 2.1, the entire belief vectors are exchanged and hence there is no approximation of the intermediate beliefs, that is,

$$\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta) = \boldsymbol{\psi}_{\ell,i}(\theta). \tag{4.15}$$

In addition, in the fixed hypothesis sharing case (4.1), there is a fixed transmitted hypothesis $\boldsymbol{\tau}_i = \tau$ and non-transmitted hypotheses are assumed to be uniformly likely. In contrast, in (4.13) agents exploit their own beliefs as prior information, and the random hypothesis $\boldsymbol{\tau}_i$ changes over time.

4.4 Truth Learning

In this section, we present results that characterize truth learning under certain conditions. First, we define the following loss function in order to assess the disagreement between the truth and the belief of each agent k at time i:

$$Q(\boldsymbol{\mu}_{k,i}) \triangleq D_{\mathrm{KL}}(e_{\theta^\circ}||\boldsymbol{\mu}_{k,i}), \tag{4.16}$$

where we denote the true probability mass function by the basis vector with unit entry at location θ°:

$$e_{\theta^\circ}(\theta) = \begin{cases} 1, & \theta = \theta^\circ \\ 0, & \theta \neq \theta^\circ \end{cases}. \tag{4.17}$$

Observe that

$$Q(\boldsymbol{\mu}_{k,i}) = \sum_{\theta\in\Theta} e_{\theta^\circ}(\theta) \log \frac{e_{\theta^\circ}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta)} = -\log \boldsymbol{\mu}_{k,i}(\theta^\circ), \tag{4.18}$$

where we use the convention that $0 \log 0 = 0$. The network loss is defined as the weighted average of the individual loss functions, where the weighting is given by the Perron entries:

$$Q(\boldsymbol{\mu}_i) \triangleq \sum_{k=1}^{K} v_k Q(\boldsymbol{\mu}_{k,i}) = -\sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i}(\theta^\circ). \tag{4.19}$$

We denote the history of observations and transmitted hypotheses up to time i by $\boldsymbol{\mathcal{F}}_i \triangleq \{\boldsymbol{\tau}_i, \boldsymbol{\xi}_i, \boldsymbol{\tau}_{i-1}, \boldsymbol{\xi}_{i-1}, \dots\}$, where we introduced the aggregate vector of observations $\boldsymbol{\xi}_i \triangleq \{\boldsymbol{\xi}_{k,i}\}_{k=1}^K$. The following result shows that the conditional expectation of the network loss does not increase, given the history $\boldsymbol{\mathcal{F}}_{i-1}$.

Lemma 4.1 (**Network average loss**). *The network loss $Q(\boldsymbol{\mu}_i)$ is a super-martingale, namely,*

$$\mathbb{E}\Big[Q(\boldsymbol{\mu}_i)\Big|\boldsymbol{\mathcal{F}}_{i-1}\Big] \leq Q(\boldsymbol{\mu}_{i-1}) \tag{4.20}$$

Proof. See Appendix 4.A. ■

In view of this lemma, Algorithm 4.1 leads to a robust design in the sense that the loss does not increase in expectation. Note that this result holds for any possible transmission distribution π. Similar to Eq. (2.27) in Theorem 2.1 for geometric social learning, we can also consider the decay rate of the belief ratio for some wrong hypothesis $\theta \in \Theta \setminus \{\theta^\circ\}$ under the proposed strategy (4.12)–(4.14).

Theorem 4.3 (**Asymptotic learning rate**). *For each wrong hypothesis $\theta \in \Theta \setminus \{\theta^\circ\}$, if the transmission probability is strictly positive, i.e., $\pi_\theta > 0$, then the belief on that wrong hypothesis will converge to 0 at an asymptotically exponential rate under Assumption 2.2. Namely, for each agent k:*

$$\frac{1}{i} \log \frac{\boldsymbol{\mu}_{k,i}(\theta^\circ)}{\boldsymbol{\mu}_{k,i}(\theta)} \xrightarrow{\text{a.s.}} \sum_{k=1}^{K} v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ) \Big\| L_k(\cdot|\theta) \Big). \tag{4.21}$$

Proof. See Appendix 4.B. ■

The convergence rate expression in (4.21) is an average of the local KL divergences of agents, which measures their individual informativeness, weighted by their network centralities (i.e., Perron entries). Observe that this rate matches the rate (2.27) for canonical geometric social learning algorithms, which require transmission of full beliefs. Therefore, a positive probability of transmitting the wrong hypothesis suffices for achieving the same asymptotic performance with probability 1, regardless of the transmission probabilities for the remaining hypotheses. Combining Theorem 4.3 and Assumption 2.1 yields the following sufficient conditions for truth learning.

Corollary 4.1 (**Truth learning**). *Under Assumptions 2.1 and 2.2, if $\pi_\theta > 0$ for all wrong hypotheses $\theta \in \Theta \setminus \{\theta^\circ\}$, then each agent k learns the truth with probability 1, i.e.,*

$$\boldsymbol{\mu}_{k,i}(\theta^\circ) \xrightarrow{\text{a.s.}} 1. \tag{4.22}$$

Notice that any asymmetry between the entries of π does not affect the learning. In particular, more or less frequent communication of a hypothesis does not change the asymptotic rate of convergence. In [62], truth learning (in the sense of Definition 2.1) occurs if, and only if, the fixed transmitted hypothesis is the true hypothesis. Corollary 4.1 shows that if agents are bootstrapping as opposed to using uniform weights [62], then learning can occur as long as $\pi_\theta > 0$ for all wrong hypotheses θ. This implies that they can learn the truth even if they do not discuss the true hypothesis, i.e. even if $\pi_{\theta^\circ} = 0$.

4.5 Truth Sharing

In the previous section, we drew the following two conclusions: In Theorem 4.3, we established that if there is a positive probability $\pi_\theta > 0$ of transmitting a wrong hypothesis $\theta \neq \theta^\circ$, this is sufficient to reject θ, i.e., beliefs on that hypothesis will go to 0 exponentially fast. Building upon this, in Corollary 4.1 we showed that exchanging

all wrong hypotheses with positive probability (i.e., $\pi_\theta > 0,\ \forall \theta \neq \theta^\circ$) is sufficient for agents to infer the truth.

Given these findings, the question arises: Is the *exclusive* exchange ($\pi_{\theta^\circ} = 1$) of the true hypothesis alone enough for learning? We give a negative answer to this question by providing a toy counter-example where agents do not learn even when $\pi_{\theta^\circ} = 1$.

Consider a fully-connected network of 3 agents (see Fig. 4.2). The hypotheses set is $\Theta = \{1, 2, 3, 4\}$ where incidentally $\theta^\circ = 4$. Assume that agent k cannot distinguish between the true hypothesis $\theta^\circ = 4$ and the hypothesis $\theta_k = k$, i.e., $D_{\mathrm{KL}}(L_k(\cdot|\theta^\circ)||L_k(\cdot|\theta_k)) = 0$. Assume further that each agent k is capable of distinguishing hypotheses other than θ_k and θ° (e.g., agent 1 can distinguish hypotheses 2 and 3, but cannot distinguish 1 and 4). Since each wrong hypothesis $\theta \neq \theta^\circ$ can be distinguished by at least one agent, the problem is globally identifiable, satisfying Assumption 2.1.

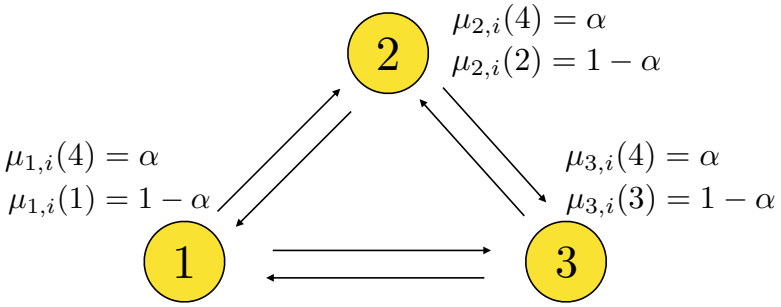


Figure 4.2: A network of 3 agents. Each agent k can distinguish all but the hypotheses $\theta^\circ = 4$ and $\theta_k = k$. In the assumed scenario, belief values on $\theta^\circ = 4$ are assumed to be $\alpha \in (0, 1)$ for all agents. Belief values at locally distinguishable hypotheses are assumed to be 0.

Imagine that at time i, $\boldsymbol{\mu}_{k,i}(\theta^\circ) = \alpha$ and $\boldsymbol{\mu}_{k,i}(\theta_k) = 1 - \alpha$ for each agent k. In other words, for each agent k, the beliefs on hypotheses that are locally distinguishable are equal to 0 [2]. Using (2.7) yields the intermediate beliefs:

$$\boldsymbol{\psi}_{k,i+1}(\theta^\circ) = \frac{L_k(\boldsymbol{\xi}_{k,i+1}|\theta^\circ)\boldsymbol{\mu}_{k,i}(\theta^\circ)}{L_k(\boldsymbol{\xi}_{k,i+1}|\theta^\circ)\boldsymbol{\mu}_{k,i}(\theta^\circ) + L_k(\boldsymbol{\xi}_{k,i+1}|\theta_k)\boldsymbol{\mu}_{k,i}(\theta_k)}. \tag{4.23}$$

Since $L_k(\boldsymbol{\xi}_{k,i+1}|\theta^\circ) = L_k(\boldsymbol{\xi}_{k,i+1}|\theta_k)$ (their KL divergence is assumed to be 0):

$$\boldsymbol{\psi}_{k,i+1}(\theta^\circ) = \frac{\boldsymbol{\mu}_{k,i}(\theta^\circ)}{\boldsymbol{\mu}_{k,i}(\theta_k) + \boldsymbol{\mu}_{k,i}(\theta^\circ)} = \boldsymbol{\mu}_{k,i}(\theta^\circ) = \alpha \tag{4.24}$$

Similarly,

$$\boldsymbol{\psi}_{k,i+1}(\theta_k) = \boldsymbol{\mu}_{k,i}(\theta_k) = 1 - \alpha. \tag{4.25}$$

[2] Under Assumption 2.2 and finite KL divergence condition, the belief values remain nonzero for all finite time instants. For ease of presentation, we assume that sufficient time has elapsed so that the beliefs on locally distinguishable hypotheses are sufficiently close to 0.

Agents fill the received intermediate beliefs by using their own beliefs (i.e, they bootstrap) according to (4.13). Here, since $\pi_{\theta^\circ} = 1$ (i.e., $\boldsymbol{\tau}_{i+1} = \theta^\circ$) we get

$$\widehat{\boldsymbol{\psi}}^{(k)}_{\ell,i+1}(\theta^\circ) = \boldsymbol{\psi}_{\ell,i+1}(\theta^\circ) = \alpha, \tag{4.26}$$

and

$$\widehat{\boldsymbol{\psi}}^{(k)}_{\ell,i+1}(\theta_k) = \boldsymbol{\psi}_{k,i+1}(\theta_k) = 1 - \alpha. \tag{4.27}$$

Finally, after the combination of the approximate intermediate beliefs using (4.14), we arrive at

$$\boldsymbol{\mu}_{k,i+1}(\theta^\circ) = \boldsymbol{\mu}_{k,i}(\theta^\circ) = \alpha, \tag{4.28}$$

and

$$\boldsymbol{\mu}_{k,i+1}(\theta_k) = \boldsymbol{\mu}_{k,i}(\theta_k) = 1 - \alpha. \tag{4.29}$$

This is an equilibrium (fixed point) for the algorithm. Regardless of the observations, the beliefs of agents will not change over time. Consequently, if α is small, agents can get stuck in beliefs where their confidence levels on the wrong hypotheses are higher than their confidence on the true hypothesis.

The results in Sec. 4.4 and the counter-example in this section reveal that it is the exchange of the wrong hypotheses that promotes truth learning, and not truth sharing. The intuition behind this rather surprising outcome is the following. In the current strategy, agents fill the missing belief components with their own beliefs (using bootstrapping). If their beliefs happen to align closely on the true hypothesis, it is difficult to change them. This undesirable equilibrium is bypassed when agents exchange beliefs on wrong hypotheses. This is due to Assumption 2.1, which states that there exists at least one agent that is able to drive the beliefs on a wrong hypothesis to 0.

In [62] (see Eq. (4.1)), when the fixed transmitted hypothesis corresponds to the truth, truth learning in the sense of Definition 2.1 occurs almost surely. The example described in this section suggests that using one's own belief for estimating non-transmitted components of neighbors, i.e., bootstrapping, as opposed to using uniform priors, may lead regular agents to become *conservative* about their own opinions. It can prevent learning under partial information sharing. In addition to not learning the truth, the network can also fail to reach consensus and opinion clusters might emerge. In Fig. 4.2, agents having positive beliefs for different hypotheses can have a major effect especially when α is small. It leads to a strong network disagreement. Network disagreement phenomena were observed in the works [63–66] when there are special agents who never change their opinions, i.e., stubborn agents. Our result, on the other hand, indicates that even when the network is only composed of regular

agents, limited communication can hinder concurrence and truth learning.

4.6 Impossibility of Mislearning

The previous section demonstrated that bootstrapping might induce network disagreement and poor equilibrium when the belief about the true state is shared. In this section, we provide a positive result in the opposite direction: Agents will never be fully confident on a wrong hypothesis. Total mislearning cannot occur.

Theorem 4.4 (**Impossibility of mislearning**). *Under Assumption 2.2, agents will always have positive confidence on the true hypothesis. Namely, for each agent k, $\boldsymbol{\mu}_{k,i}(\theta^\circ) > 0$ for any finite time i, and*

$$\mathbb{P}\Big(\liminf_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta^\circ) = 0\Big) = 0, \tag{4.30}$$

or alternatively, for $\theta \in \Theta \setminus \{\theta^\circ\}$

$$\mathbb{P}\Big(\limsup_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta) = 1\Big) = 0. \tag{4.31}$$

Proof. See Appendix 4.C. ■

Notice that there is no assumption on the transmission probabilities in Theorem 4.4. With bootstrapping, agents never learn a wrong hypothesis. In [62], it was shown that agents might mislearn a wrong hypothesis if the fixed transmitted hypothesis is not the true hypothesis — recall Theorem 4.2. As a matter of fact, bootstrapping leads to a more robust design in the face of partial communication.

4.7 Numerical Simulations

Consider a $10-$agent strongly-connected network (see the topology in Fig. 4.3i). The combination matrix is designed using the Metropolis rule [7], yielding a doubly-stochastic matrix. Agents are trying to detect the true state θ° among a set of five hypotheses, namely $\Theta \triangleq \{1,2,3,4,5\}$. We assume that $\theta^\circ = 1$. To accomplish this task, agents use the protocol described in (4.12)-(4.14), where the random shared hypothesis $\boldsymbol{\tau}_i$ is distributed according to the following probability mass function:

$$\mathbb{P}(\boldsymbol{\tau}_i = \theta) = \pi_\theta = \begin{cases} 0, & \text{if } \theta = \theta^\circ \\ 0.25, & \text{otherwise.} \end{cases} \tag{4.32}$$

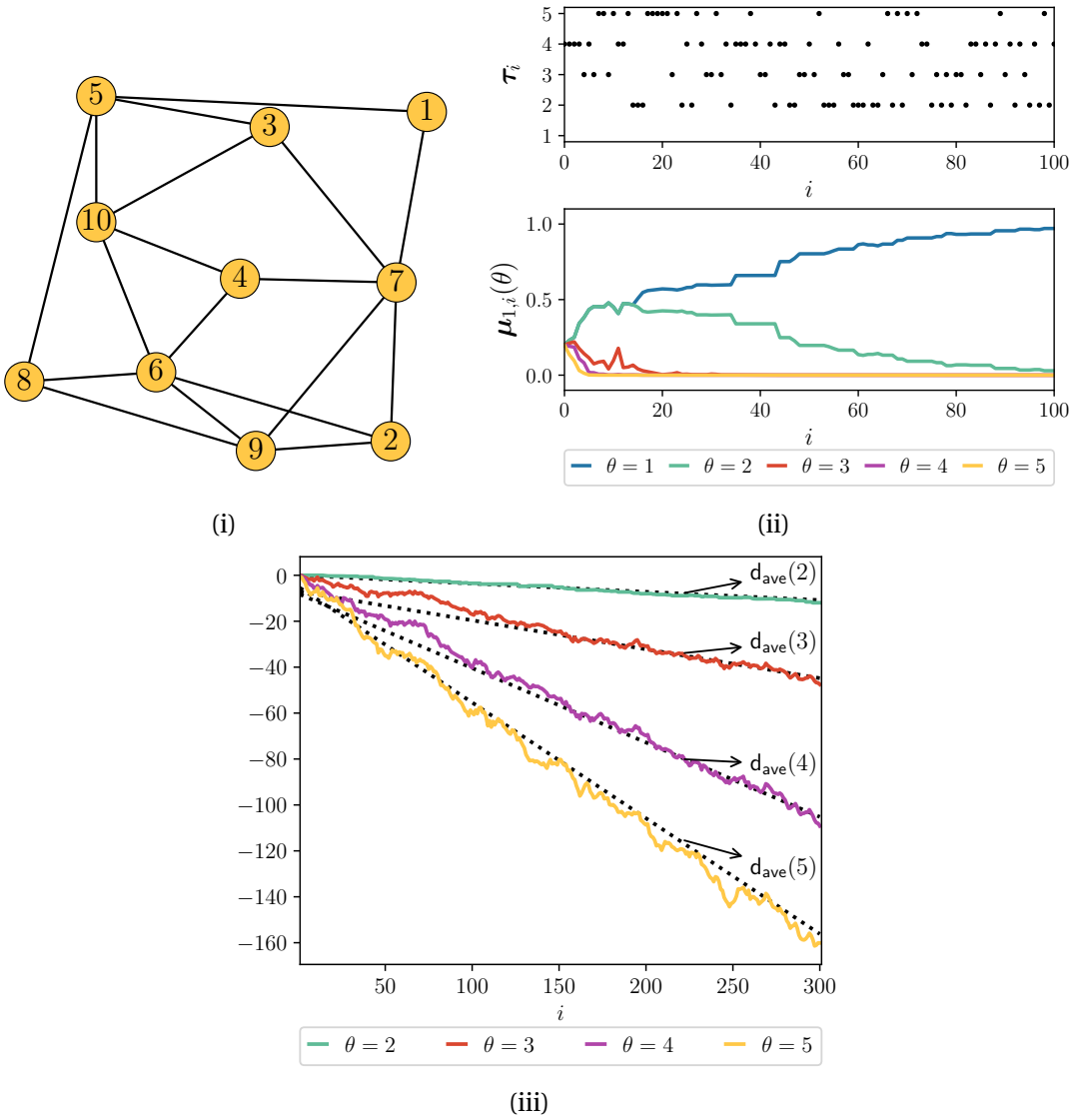


Figure 4.3: (i) Network topology. (ii) Evolution of the shared hypothesis τ_i over time in the upper panel, and belief evolution for agent 1 showing *truth learning* in the bottom panel. (iii) Experimental rates of convergence for agent 1 for different hypotheses (in colored lines), compared with the theoretical asymptotic rate of convergence, i.e., $\mathsf{d}_{\mathsf{ave}}$ defined in (4.35) (in black dotted lines).

The observations at the agents are generated by a family of unit-variance Gaussian densities:

$$f_n(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(\xi - 0.3n)^2}{2}\right\} \tag{4.33}$$

for $n = 1, 2, 3, 4, 5$. The likelihoods of agents are chosen from among these Gaussian densities according to the identifiability setup in Table 4.1. For example, note that agents 8–10 cannot distinguish hypotheses 1 and 5. Observe that the global identifiability condition in Assumption 2.1 is satisfied.

Agent k	**Likelihood Function:** $L_k(\xi\|\theta)$				
	$\theta=1$	$\theta=2$	$\theta=3$	$\theta=4$	$\theta=5$
$1-2$	f_1	f_1	f_3	f_4	f_5
$3-5$	f_1	f_2	f_1	f_4	f_5
$6-7$	f_1	f_2	f_3	f_1	f_5
$8-10$	f_1	f_2	f_3	f_4	f_1

Table 4.1: Identifiability Setup for Network in Fig. 4.3

In Fig. 4.3ii, we see the evolution of the belief for agent 1, which shows that, although the agents never share information about the true hypothesis, i.e., $\pi_{\theta^\circ}=0$, the agent asymptotically learns the truth, as suggested by Corollary 4.1. A similar behavior happens for the remaining agents. Fig. 4.3iii shows that the experimental convergence rates for agent 1, i.e.,

$$\log \frac{\boldsymbol{\mu}_{1,i}(\theta)}{\boldsymbol{\mu}_{1,i}(\theta^\circ)} \tag{4.34}$$

which are shown in colored lines, approach the asymptotic convergence rates of traditional social learning (black dotted lines):

$$\mathsf{d}_{\mathsf{ave}}(\theta) \triangleq \sum_{k=1}^{K} -v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\theta)\Big) \tag{4.35}$$

as predicted by Theorem 4.3. This means that regarding the asymptotic convergence rate, there is no performance loss when only one hypothesis is exchanged at each iteration as long as all wrong hypotheses have positive probability of being transmitted.

In the next simulation, we illustrate that truth sharing is not sufficient for truth learning. For that purpose, we fix the transmitted hypothesis at the true hypothesis $\boldsymbol{\tau}_i=\theta^\circ=1$ for all $i=1,2,\ldots$. The result can be seen in Fig. 4.4, where we show the evolution of the belief of agent 1 over time. Despite sharing the true hypothesis, the *conservative* behavior described in Section 4.5 hinders the ability of the agent to learn the truth. We note that agent 1 cannot *decidedly* distinguish between hypothesis 1 and 2, which are indistinguishable from its local point of view (see Table 4.1). This was suggested by the example in Section 4.5, where agents are caught in an equilibrium where they have non-zero belief values for locally indistinguishable hypotheses. Notice that although truth learning is not observed, there is no total mislearning phenomena as well. As suggested by Theorem 4.4, the confidence in the truth is not going to 0.

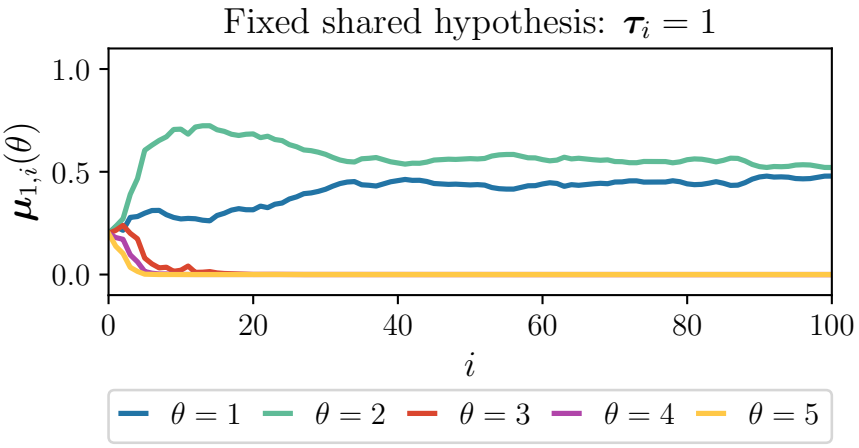


Figure 4.4: Belief evolution of agent 1 when the shared hypothesis is *fixed over time* to be the true state of nature. This demonstrates that while there is no truth learning with full confidence, there is no complete rejection of the true hypothesis as well.

4.8 Concluding Remarks

In this chapter, we studied social learning under partial information sharing where the transmitted hypothesis is changing at random at every iteration. In Sec. 4.3, we proposed an algorithm for this setting in which agents fill the latent belief components with their own beliefs. In Sec. 4.4, we derived the rate of convergence under the proposed algorithm and provided sufficient conditions for truth learning. Then, in Secs. 4.5 and 4.6, we demonstrated that by exchanging beliefs exclusively on the true hypotheses, agents will neither learn the truth with full confidence nor mislearn, i.e., learn a false hypothesis with full confidence. Instead they will be unsure about the truth among their indistinguishable hypotheses.

There are many possible extensions to the setting considered in this work. For instance, each agent k can choose a possibly different hypothesis $\tau_{k,i}$ to transmit at each iteration. Alternatively, the transmitted hypothesis τ_i can evolve according to some Markovian model, instead of independently over time. These extensions introduce additional complexities to the technical analysis, especially when managing the random matrix products. Another interesting extension would be to see if the current work on finite hypothesis sets can extend to continuous hypothesis sets, such as compact sets as in [67] or non-compact sets as in the distributed estimation literature (see e.g., [9, 68, 69]).

4.A Proof of Lemma 4.1

From the combination step (4.14),

$$\begin{aligned}
\log \boldsymbol{\mu}_{k,i}(\theta^\circ) &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) - \log \sum_{\theta'\in\Theta} \exp\Big\{ \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta') \Big\} \\
&\overset{(a)}{\geq} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) - \log \Big(\sum_{\theta'\in\Theta} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta') \Big) \\
&= \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) - \log \Big(\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \sum_{\theta'\in\Theta} \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta') \Big) \\
&\overset{(b)}{=} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ)
\end{aligned} \tag{4.36}$$

where (a) follows from Jensen's inequality, and (b) follows from the fact that $\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}$ is a pmf and A is left-stochastic. Therefore, the conditional expectations satisfy

$$\begin{aligned}
\mathbb{E}\Big[\log \boldsymbol{\mu}_{k,i}(\theta^\circ) \Big| \boldsymbol{\mathcal{F}}_{i-1} \Big] &\geq \mathbb{E}\Big[\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) \Big| \boldsymbol{\mathcal{F}}_{i-1} \Big] \\
&\overset{(a)}{=} \mathbb{E}_{\boldsymbol{\xi}_i,\boldsymbol{\tau}_i}\Big[\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) \Big] \\
&\overset{(b)}{=} \mathbb{E}_{\boldsymbol{\xi}_i} \mathbb{E}_{\boldsymbol{\tau}_i|\boldsymbol{\xi}_i}\Big[\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) \Big] \\
&\overset{(c)}{=} \mathbb{E}_{\boldsymbol{\xi}_i} \mathbb{E}_{\boldsymbol{\tau}_i}\Big[\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) \Big]
\end{aligned} \tag{4.37}$$

where (a) follows from the fact that arguments inside the expectation are functions of $\{\boldsymbol{\xi}_i, \boldsymbol{\tau}_i, \boldsymbol{\mathcal{F}}_{i-1}\}$, (b) follows from the tower rule of expectation, and (c) follows from the fact that $\boldsymbol{\xi}_i$ and $\boldsymbol{\tau}_i$ are independent. The inner expectation can be written as

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\tau}_i}\Big[\sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ) \Big] \\
&= \pi_{\theta^\circ} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{\ell,i}(\theta^\circ)}{1 - \boldsymbol{\psi}_{k,i}(\theta^\circ) + \boldsymbol{\psi}_{\ell,i}(\theta^\circ)} + \sum_{\tau\neq\theta^\circ} \pi_\tau \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{k,i}(\theta^\circ)}{1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{\ell,i}(\tau)} \\
&= \pi_{\theta^\circ} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \boldsymbol{\psi}_{\ell,i}(\theta^\circ) + \sum_{\tau\neq\theta^\circ} \pi_\tau \log \boldsymbol{\psi}_{k,i}(\theta^\circ) - \sum_{\tau\in\Theta} \pi_\tau \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \Big(1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{\ell,i}(\tau) \Big)
\end{aligned} \tag{4.38}$$

Using the Perron vector defined by (2.2), taking the expectation of (4.36) with respect to $\boldsymbol{\tau}_i$, and using (4.38) we get

$$\mathbb{E}_{\boldsymbol{\tau}_i}\Big[\sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i}(\theta^\circ) \Big]$$

$$
\begin{aligned}
&\geq \pi_{\theta^\circ} \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) + \sum_{\tau \neq \theta^\circ} \pi_\tau \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) \\
&\qquad\qquad - \sum_{\tau \in \Theta} \pi_\tau \sum_{k=1}^{K} v_k \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \Big(1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{\ell,i}(\tau)\Big) \\
&= \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) - \sum_{\tau \in \Theta} \pi_\tau \sum_{k=1}^{K} v_k \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \Big(1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{\ell,i}(\tau)\Big) \\
&\overset{(a)}{\geq} \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) - \log \left(\sum_{\tau \in \Theta} \pi_\tau \sum_{k=1}^{K} v_k \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \Big(1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{\ell,i}(\tau)\Big) \right) \\
&= \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) - \log \left(\sum_{\tau \in \Theta} \pi_\tau \sum_{k=1}^{K} v_k \Big(1 - \boldsymbol{\psi}_{k,i}(\tau) + \boldsymbol{\psi}_{k,i}(\tau)\Big) \right) \\
&= \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) - \log(1) \\
&= \sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ), \qquad (4.39)
\end{aligned}
$$

where (a) follows from Jensen's inequality. Applying the expectation with respect to $\boldsymbol{\xi}_i$ to both sides of (4.39), we arrive at

$$
\begin{aligned}
&\mathbb{E}_{\xi_i} \mathbb{E}_{\tau_i} \Big[\sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i}(\theta^\circ) \Big] \\
&\qquad \overset{(4.39)}{\geq} \mathbb{E}_{\xi_i} \Big[\sum_{k=1}^{K} v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) \Big] \\
&\qquad \overset{(a)}{=} \mathbb{E}_{\xi_i} \Big[\sum_{k=1}^{K} v_k \log \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}{\sum_{\theta' \in \Theta} L_k(\boldsymbol{\xi}_{k,i}|\theta') \boldsymbol{\mu}_{k,i-1}(\theta')} \Big] + \sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i-1}(\theta^\circ) \\
&\qquad = \sum_{k=1}^{K} v_k D_{\mathrm{KL}} \Big(L_k(\cdot|\theta^\circ) \Big\| \sum_{\theta' \in \Theta} L_k(\cdot|\theta') \boldsymbol{\mu}_{k,i-1}(\theta') \Big) + \sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i-1}(\theta^\circ) \\
&\qquad \overset{(b)}{\geq} \sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i-1}(\theta^\circ) \qquad (4.40)
\end{aligned}
$$

where (a) follows from the fact that belief vectors at time $i-1$ are independent of the new observations at time i, and (b) follows from the fact that KL-divergences are non-negative. It is worth noting that this proof holds even in cases where the transmission distribution π is time-dependent, or when it depends on the observations. Therefore, the result is more general than what is explicitly stated. However, to maintain consistency with the other parts of the paper, we used the fixed distribution notation π.

4.B Proof of Theorem 4.3

The log-belief ratio can be written as

$$\begin{aligned}\log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta)}{\widehat{\boldsymbol{\psi}}_{\ell,i}^{(k)}(\theta^\circ)} \\ &= \mathbb{I}\{\boldsymbol{\tau}_i=\theta\} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{\ell,i}(\theta)}{\boldsymbol{\psi}_{k,i}(\theta^\circ)} + \mathbb{I}\{\boldsymbol{\tau}_i=\theta^\circ\} \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{k,i}(\theta)}{\boldsymbol{\psi}_{\ell,i}(\theta^\circ)} \\ &\quad + \Big(1-\mathbb{I}\{\boldsymbol{\tau}_i=\theta\}-\mathbb{I}\{\boldsymbol{\tau}_i=\theta^\circ\}\Big) \log \frac{\boldsymbol{\psi}_{k,i}(\theta)}{\boldsymbol{\psi}_{k,i}(\theta^\circ)}. \end{aligned} \tag{4.41}$$

Observe from (4.41) that the log-belief ratio is a random variable given the intermediate beliefs, because of the randomness of the trending topic $\boldsymbol{\tau}_i$. Next, we fix a wrong hypothesis $\theta\neq\theta^\circ$ and define the effective combination matrix at time i as

$$\widetilde{\boldsymbol{A}}_i \triangleq \begin{cases} A, & \theta=\boldsymbol{\tau}_i \\ I, & \theta\neq\boldsymbol{\tau}_i \end{cases}. \tag{4.42}$$

This is a binary random variable taking the value of the original combination matrix A if the hypothesis is exchanged at iteration i, and the identity matrix I otherwise. More compactly,

$$\widetilde{\boldsymbol{A}}_i = \mathbb{I}\{\boldsymbol{\tau}_i=\theta\}A + (1-\mathbb{I}\{\boldsymbol{\tau}_i=\theta\})I. \tag{4.43}$$

Using this definition, the relation in (4.41) can be rewritten for $\theta\neq\theta^\circ$ as:

$$\begin{aligned}\log \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} &= \sum_{\ell\in\mathcal{N}_k} [\widetilde{\boldsymbol{A}}_i]_{\ell k} \log \frac{\boldsymbol{\psi}_{\ell,i}(\theta)}{\boldsymbol{\psi}_{\ell,i}(\theta^\circ)} + (\mathbb{I}\{\boldsymbol{\tau}_i=\theta^\circ\}-\mathbb{I}\{\boldsymbol{\tau}_i=\theta\}) \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{k,i}(\theta^\circ)}{\boldsymbol{\psi}_{\ell,i}(\theta^\circ)} \\ &\overset{(4.12)}{=} \sum_{\ell\in\mathcal{N}_k} [\widetilde{\boldsymbol{A}}_i]_{\ell k} \log \frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta^\circ)} + \sum_{\ell\in\mathcal{N}_k} [\widetilde{\boldsymbol{A}}_i]_{\ell k} \log \frac{\boldsymbol{\mu}_{\ell,i-1}(\theta)}{\boldsymbol{\mu}_{\ell,i-1}(\theta^\circ)} \\ &\quad + \Big(\mathbb{I}\{\boldsymbol{\tau}_i=\theta^\circ\}-\mathbb{I}\{\boldsymbol{\tau}_i=\theta\}\Big) \sum_{\ell\in\mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{\psi}_{k,i}(\theta^\circ)}{\boldsymbol{\psi}_{\ell,i}(\theta^\circ)}. \end{aligned} \tag{4.44}$$

The first two terms on the RHS of (4.44) are analogous to the terms that arise in the standard log-linear social learning analysis (see Eq. (2.28)), albeit with the random matrix $\widetilde{\boldsymbol{A}}_i$ in place of the original combination matrix A. The last term in (4.44) is a residue term due to the network disagreement. In order to expand the recursion over time, we introduce the following notation for the product of the effective combination matrices for $j\leq i$:[3]

$$\widetilde{\boldsymbol{A}}^{j\to i} \triangleq \widetilde{\boldsymbol{A}}_j \widetilde{\boldsymbol{A}}_{j+1} \ldots \widetilde{\boldsymbol{A}}_{i-1} \widetilde{\boldsymbol{A}}_i, \tag{4.45}$$

[3] If $j>i$, we set $\widetilde{\boldsymbol{A}}^{j\to i}=I$.

and also for the residue terms for $j \leq i$ as

$$\boldsymbol{R}_j^k \triangleq \Big(\mathbb{I}\{\boldsymbol{\tau}_j = \theta^\circ\} - \mathbb{I}\{\boldsymbol{\tau}_j = \theta\}\Big) \times \sum_{\ell=1}^{K} [\widetilde{\boldsymbol{A}}^{j+1\to i}]_{\ell k} \sum_{m=1}^{K} a_{m\ell} \log \frac{\boldsymbol{\psi}_{\ell,j}(\theta^\circ)}{\boldsymbol{\psi}_{m,j}(\theta^\circ)}. \tag{4.46}$$

In other words, $\boldsymbol{R}_j^k$ denotes the residual term at time i caused by the network disagreement at time j. Expanding (4.44) with these definitions and dividing both sides by i, we arrive at the following expression for the convergence rate.

$$\frac{1}{i}\log\frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} = \frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{j\to i}]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} + \frac{1}{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{1\to i}]_{\ell k}\log\frac{\boldsymbol{\mu}_{\ell,0}(\theta)}{\boldsymbol{\mu}_{\ell,0}(\theta^\circ)} + \frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j^k. \tag{4.47}$$

In Lemma 4.3 further ahead, we show that the summation of the residue terms in (4.47) stays finite with probability 1 as i grows, so that

$$\frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j^k \xrightarrow{\text{a.s.}} 0. \tag{4.48}$$

Therefore, the residue terms do not affect the convergence rate in (4.47) as $i \to \infty$. We proceed to study the remaining terms in (4.47). Observe that the finiteness of the KL divergence of likelihood functions can be expressed as

$$\left|\mathbb{E}\left[\log\frac{L_k(\boldsymbol{\xi}_{k,i}|\theta)}{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}\right]\right| < \infty \tag{4.49}$$

which in turn implies

$$\left|\log\frac{L_k(\boldsymbol{\xi}_{k,i}|\theta)}{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}\right| \overset{\text{a.s.}}{<} \infty. \tag{4.50}$$

The set of events such that

$$\lim_{n\to\infty}[\widetilde{\boldsymbol{A}}^{j\to j+n} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \neq 0 \tag{4.51}$$

is a subset of the union of the event sets:

$$\lim_{n\to\infty}[\widetilde{\boldsymbol{A}}^{j\to j+n} - v\mathbb{1}_K^\top]_{\ell k} \neq 0 \tag{4.52}$$

and

$$\left|\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)}\right| = \infty. \tag{4.53}$$

But we know that these two sets are null sets because of the auxiliary Lemma 4.4 and (4.50), respectively. Therefore, for any time instant j, it holds that

$$\lim_{n\to\infty}[\widetilde{\boldsymbol{A}}^{j\to j+n} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \overset{\text{a.s.}}{=} 0. \tag{4.54}$$

This implies for the convergence of the Cesàro mean [70] that

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=0}^{t-1}[\widetilde{\boldsymbol{A}}^{j\to j+n} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \overset{\text{a.s.}}{=} 0. \tag{4.55}$$

By using $j = i - n$, this can alternatively be written as

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=0}^{t-1}[\widetilde{\boldsymbol{A}}^{i-n\to i} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,i-n}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,i-n}|\theta^\circ)} \overset{\text{a.s.}}{=} 0. \tag{4.56}$$

Since this holds for any time instant $i \geq t$ (which ensures $j \geq 1$), we can set $i = t$. By changing the summation index n to j, we arrive at

$$\lim_{i\to\infty}\frac{1}{i}\sum_{j=1}^{i}[\widetilde{\boldsymbol{A}}^{j\to i} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \overset{\text{a.s.}}{=} 0. \tag{4.57}$$

As a result, the first term in (4.47) can be written as:

$$\begin{aligned}
&\frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{j\to i}]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \\
&= \frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{j\to i} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} + \frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} \\
&= \frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{j\to i} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} + \frac{1}{i}\sum_{j=1}^{i}\sum_{k=1}^{K} v_k\log\frac{L_k(\boldsymbol{\xi}_{k,j}|\theta)}{L_k(\boldsymbol{\xi}_{k,j}|\theta^\circ)} \\
&= \frac{1}{i}\sum_{j=1}^{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{j\to i} - v\mathbb{1}_K^\top]_{\ell k}\log\frac{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta)}{L_\ell(\boldsymbol{\xi}_{\ell,j}|\theta^\circ)} + \sum_{k=1}^{K} v_k\frac{1}{i}\sum_{j=1}^{i}\log\frac{L_k(\boldsymbol{\xi}_{k,j}|\theta)}{L_k(\boldsymbol{\xi}_{k,j}|\theta^\circ)} \\
&\xrightarrow{\text{a.s.}} \sum_{k=1}^{K} -v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\theta)\Big)
\end{aligned} \tag{4.58}$$

where the last step follows from (4.57) and the strong law of large numbers [40, Chapter 7]. Also, since by Lemma 4.4 $\widetilde{\boldsymbol{A}}^{1\to i} \xrightarrow{\text{a.s.}} v\mathbb{1}_K^\top$ and by Assumption 2.2 the initial beliefs are nonzero, it follows that

$$\frac{1}{i}\sum_{\ell=1}^{K}[\widetilde{\boldsymbol{A}}^{1\to i}]_{\ell k}\log\frac{\boldsymbol{\mu}_{\ell,0}(\theta)}{\boldsymbol{\mu}_{\ell,0}(\theta^\circ)} \xrightarrow{\text{a.s.}} 0. \tag{4.59}$$

The asymptotic convergence rate then becomes

$$\frac{1}{i}\log\frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} \xrightarrow{\text{a.s.}} \sum_{k=1}^{K} -v_k D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)\Big\|L_k(\cdot|\theta)\Big). \tag{4.60}$$

■

4.C Proof of Theorem 4.4

Since $Q(\boldsymbol{\mu}_i)$ is a super-martingale (Lemma 4.1) and also non-negative (i.e., uniformly bounded from below), by Doob's forward martingale convergence theorem [40, Chapter 11.5], there exists a finite random variable $\boldsymbol{Q}_\infty$ such that, as $i \to \infty$,

$$Q(\boldsymbol{\mu}_i) \xrightarrow{\text{a.s.}} \boldsymbol{Q}_\infty. \tag{4.61}$$

Since $\boldsymbol{Q}_\infty$ is finite, it holds that:

$$\begin{aligned}
&\lim_{i\to\infty}\sum_{k=1}^{K} v_k \log \boldsymbol{\mu}_{k,i}(\theta^\circ) > -\infty \\
&\Longrightarrow \liminf_{i\to\infty} \log \boldsymbol{\mu}_{k,i}(\theta^\circ) > -\infty, \quad \forall k \in \mathcal{N} \\
&\Longrightarrow \liminf_{i\to\infty} \boldsymbol{\mu}_{k,i}(\theta^\circ) > 0
\end{aligned} \tag{4.62}$$

with probability 1.

4.D Auxiliary Results

4.D.1 Vanishing Matrix Norm

Lemma 4.2. *If the wrong hypothesis θ is transmitted with positive probability, i.e., $\pi_\theta > 0$, then, for any induced matrix norm,*

$$\mathbb{E}\Big[\,\big\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top (I - A^\top)\big\|\,\Big] = \mathcal{O}(\widetilde{\lambda}^{i-j}) \tag{4.63}$$

for a constant $\widetilde{\lambda}$ that satisfies $0 \le \widetilde{\lambda} < 1$.

Proof. Define the time difference $n \triangleq i - j$. Since the matrices $\widetilde{\boldsymbol{A}}_i$ are i.i.d. binary random variables over time, as defined by (4.42), for time differences $0 \le m \le n$, we

get

$$\mathbb{P}\Big(\widetilde{\boldsymbol{A}}^{j+1\to i}=A^m\Big)=\binom{n}{m}(\pi_\theta)^m(1-\pi_\theta)^{n-m}. \tag{4.64}$$

Moreover, since A is a primitive stochastic matrix, for consecutive time instants, there exists a non-negative constant $\lambda<1$ such that [39, Eq. (8.2.10)]:

$$\|A^m-A^{m+1}\|\le C\lambda^m(1-\lambda) \tag{4.65}$$

where C is a constant independent of m. Then, it follows that

$$\begin{aligned}
\mathbb{E}\Big[\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I-A^\top)\|\Big] &= \sum_{m=0}^{n}\mathbb{P}(\widetilde{\boldsymbol{A}}^{j+1\to i}=A^m)\|A^m-A^{m+1}\| \\
&\overset{(4.65)}{\le} \sum_{m=0}^{n}\mathbb{P}(\widetilde{\boldsymbol{A}}^{j+1\to i}=A^m)C\lambda^m(1-\lambda) \\
&\overset{(4.64)}{=} C(1-\lambda)\sum_{m=0}^{n}\binom{n}{m}(\pi_\theta)^m(1-\pi_\theta)^{n-m}\lambda^m \\
&= C(1-\lambda)\Big(\lambda\pi_\theta+(1-\pi_\theta)\Big)^n \\
&= C(1-\lambda)\Big(1-(1-\lambda)\pi_\theta\Big)^n \\
&= \mathcal{O}(\widetilde{\lambda}^n)
\end{aligned} \tag{4.66}$$

where $\widetilde{\lambda}\triangleq 1-(1-\lambda)\pi_\theta$, which is a constant strictly smaller than 1 as long as $\pi_\theta>0$. ■

4.D.2 Finiteness of the Residual Sum

Lemma 4.3. *As $i\to\infty$, if $\pi_\theta>0$, then, under Assumption 2.2*

$$\frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j^k\xrightarrow{\text{a.s.}}0. \tag{4.67}$$

Proof. First, we aggregate the residue terms and log-intermediate beliefs from across the network into the following vectors:

$$\boldsymbol{R}_j\triangleq\text{col}\Big\{\boldsymbol{R}_j^k\Big\}_{k=1}^{K},\quad \boldsymbol{\Psi}_i\triangleq\text{col}\Big\{\log\boldsymbol{\psi}_{k,i}(\theta^\circ)\Big\}_{k=1}^{K}. \tag{4.68}$$

By using these definitions, expression (4.46) can be transformed into the following

vector equation form:

$$\boldsymbol{R}_j = \Big(\mathbb{I}\{\boldsymbol{\tau}_j = \theta^\circ\} - \mathbb{I}\{\boldsymbol{\tau}_j = \theta\}\Big)(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(\boldsymbol{\Psi}_j - A^\top\boldsymbol{\Psi}_j). \tag{4.69}$$

To bound the terms in this expression, we start by using Lemma 4.2 and Markov's inequality to obtain:

$$\mathbb{P}\Big(\|(\widetilde{\boldsymbol{A}}^{j+1\to j+n})^\top(I - A^\top)\| \geq \epsilon\Big) \leq \frac{C(1-\lambda)\widetilde{\lambda}^n}{\epsilon}, \tag{4.70}$$

where recall that $n = i - j$. Since $\widetilde{\lambda} < 1$, it holds that

$$\sum_{n=0}^{\infty} \frac{C(1-\lambda)\widetilde{\lambda}^n}{\epsilon} < \infty. \tag{4.71}$$

Then, the first Borel-Cantelli Lemma [40, Chapter 2.7] implies

$$\lim_{n\to\infty} \Big\|(\widetilde{\boldsymbol{A}}^{j+1\to j+n})^\top(I - A^\top)\Big\| \overset{\text{a.s.}}{=} 0, \tag{4.72}$$

for any value of j. Moreover, if we bound the norm of (4.69), it holds almost surely that

$$\begin{aligned}
\|\boldsymbol{R}_j\| &= \Big\|\Big(\mathbb{I}\{\boldsymbol{\tau}_j = \theta^\circ\} - \mathbb{I}\{\boldsymbol{\tau}_j = \theta\}\Big)(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(\boldsymbol{\Psi}_j - A^\top\boldsymbol{\Psi}_j)\Big\| \\
&\overset{(a)}{\leq} \Big\|\mathbb{I}\{\boldsymbol{\tau}_j = \theta^\circ\} - \mathbb{I}\{\boldsymbol{\tau}_j = \theta\}\Big\| \times \Big\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I - A^\top)\Big\| \times \Big\|\boldsymbol{\Psi}_j\Big\| \\
&\leq \Big\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I - A^\top)\Big\| \times \Big\|\boldsymbol{\Psi}_j\Big\| \\
&\overset{(b)}{\leq} \Big\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I - A^\top)\Big\|\boldsymbol{\Psi},
\end{aligned} \tag{4.73}$$

where (a) follows from the submultiplicative property of the norm, and (b) follows from the definition

$$\boldsymbol{\Psi} \triangleq \sup_{j\geq 1} \|\boldsymbol{\Psi}_j\|, \tag{4.74}$$

which is shown to be finite under Assumption 2.2 in Appendix 4.D.4. Subsequently, the norm of the Cesàro mean satisfies

$$\Big\|\frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j\Big\| \leq \frac{1}{i}\sum_{j=1}^{i}\Big\|\boldsymbol{R}_j\Big\| \overset{(4.73)}{\leq} \boldsymbol{\Psi}\frac{1}{i}\sum_{j=1}^{i}\Big\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I - A^\top)\Big\|. \tag{4.75}$$

Observe that (4.72) can alternatively be written as (by using the definition $n = i - j$)

$$\lim_{n\to\infty} \Big\|(\widetilde{\boldsymbol{A}}^{i-n+1\to i})^\top(I - A^\top)\Big\| \overset{\text{a.s.}}{=} 0. \tag{4.76}$$

As a result, the Cesàro mean satisfies

$$\lim_{t\to\infty}\frac{1}{t}\sum_{n=0}^{t-1}\left\|(\widetilde{\boldsymbol{A}}^{i-n+1\to i})^\top(I-A^\top)\right\|\overset{\text{a.s.}}{=}0, \tag{4.77}$$

for any $i\geq t$ (so that $j=i-n\geq 1$). If we set $i=t$, and change the indices from n to $j=i-n$, we get

$$\lim_{i\to\infty}\frac{1}{i}\sum_{j=1}^{i}\left\|(\widetilde{\boldsymbol{A}}^{j+1\to i})^\top(I-A^\top)\right\|\overset{\text{a.s.}}{=}0. \tag{4.78}$$

Incorporating this into (4.75), we conclude that, as $i\to\infty$,

$$\left\|\frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j\right\|\xrightarrow{\text{a.s.}}0,\quad\Longrightarrow\frac{1}{i}\sum_{j=1}^{i}\boldsymbol{R}_j^k\xrightarrow{\text{a.s.}}0. \tag{4.79}$$

■

4.D.3 Convergence of the Matrix Product

Lemma 4.4. *For a fixed time instant j, if $\pi_\theta>0$, then*

$$\widetilde{\boldsymbol{A}}^{j\to i}\xrightarrow{\text{a.s.}}v\mathbb{1}_K^\top \tag{4.80}$$

as $i\to\infty$.

Proof. Recall from (4.42) that $\widetilde{\boldsymbol{A}}_i$ is a binary random variable and let E_i denote the event that $\widetilde{\boldsymbol{A}}_i=A$. Since $\{E_i\}_{i=1}^\infty$ are independent events across time and their probability of occurrence satisfies $\pi_\theta>0$, it holds that

$$\sum_{i=1}^{\infty}\mathbb{P}(E_i)=\sum_{i=1}^{\infty}\pi_\theta=\infty. \tag{4.81}$$

Subsequently, by the second Borel-Cantelli Lemma [40, Chapter 4.3], we conclude that

$$\mathbb{P}(E_i\text{ occurs for infinitely many }i)=1. \tag{4.82}$$

Notice that in the product of random matrices

$$\widetilde{\boldsymbol{A}}^{j\to i}=\widetilde{\boldsymbol{A}}_j\widetilde{\boldsymbol{A}}_{j+1}\dots\widetilde{\boldsymbol{A}}_{i-1}\widetilde{\boldsymbol{A}}_i, \tag{4.83}$$

the realization of random matrices will either be equal to A or the identity matrix I. Multiplication with identity matrices has no effect on the product and by (4.82), there will be infinitely many A's in the product with probability 1, as $n=i-j\to\infty$. Thus, if

we define $A^\infty \triangleq \lim_{i\to\infty} A^i$,

$$\mathbb{P}(\lim_{n\to\infty} \widetilde{\boldsymbol{A}}^{j\to j+n} = A^\infty) = 1. \tag{4.84}$$

Finally, since $A^\infty = v\mathbb{1}_K^\top$ due to A being a primitive matrix, we conclude that

$$\lim_{n\to\infty} \widetilde{\boldsymbol{A}}^{j\to j+n} \overset{\text{a.s.}}{=} v\mathbb{1}_K^\top \tag{4.85}$$

for any time instant $j \geq 1$. ■

4.D.4 Uniform Boundedness

In this section, we show that, under Assumption 2.2

$$\boldsymbol{\Psi} = \sup_{j\geq 1} \|\boldsymbol{\Psi}_j\| \overset{\text{a.s.}}{<} \infty. \tag{4.86}$$

For that purpose, first, we show that Lemma 4.1 implies that $\sum_{k=1}^K v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ)$ is a sub-martingale. To see this, observe that by (4.40),

$$\mathbb{E}_{\xi_i}\Big[\sum_{k=1}^K v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ)\Big] \geq \sum_{k=1}^K v_k \log \boldsymbol{\mu}_{k,i-1}(\theta^\circ), \tag{4.87}$$

and in turn, by (4.39),

$$\mathbb{E}_{\tau_{i-1}}\Big[\sum_{k=1}^K v_k \log \boldsymbol{\mu}_{k,i-1}(\theta^\circ)\Big] \geq \sum_{k=1}^K v_k \log \boldsymbol{\psi}_{k,i-1}(\theta^\circ). \tag{4.88}$$

Since it is also a non-positive sub-martingale, it converges to a finite limit almost surely [40, Chapter 11.5], which means that as $i \to \infty$,

$$\begin{aligned}
&\sum_{k=1}^K v_k \log \boldsymbol{\psi}_{k,i}(\theta^\circ) \overset{\text{a.s.}}{>} -\infty \\
&\quad\Longrightarrow \log \boldsymbol{\psi}_{k,i}(\theta^\circ) \overset{\text{a.s.}}{>} -\infty, \quad \forall k \\
\Longrightarrow &\sum_{k=1}^K -\log \boldsymbol{\psi}_{k,i}(\theta^\circ) \overset{\text{a.s.}}{<} \infty \\
&\quad\Longrightarrow \|\boldsymbol{\Psi}_i\| \overset{\text{a.s.}}{<} \infty.
\end{aligned} \tag{4.89}$$

In addition, for any finite time instant j, it is true that $\boldsymbol{\psi}_{k,i}(\theta^\circ) > 0$ for each agent k for the following reason. Due to Assumption 2.2, the initial beliefs are positive. Moreover, the likelihood at the true hypothesis by definition cannot be 0 for the received obser-

vations. Furthermore, the geometric combination rule results in the intersection of the supports of its arguments [35]. Consequently, for any time instant j, it is true that $\|\mathbf{\Psi}_j\| < \infty$. Combining this with (4.89) establishes (4.86).

Part II Dynamical Conditions

5 Markovian States

5.1 Introduction[1]

In many applications, the hidden state or hypothesis that the agents are interested in tracking is time-varying, such as the position of a moving object, the concentration of air pollutants, and the product quality of a brand. In all situations, the agents will attempt to cooperatively track the dynamic state by using observations emitted by the underlying physical systems.

This setting is general enough and can be used in many engineering applications, including target tracking, environmental monitoring, and opinion formation over networks. For example, consider an economic network where the individual agents are trying to decide which currency (e.g., USD, EUR, CHF) is the best option to buy now. The optimal choice (true hypothesis) can be changing rapidly. Most of the literature on social learning ignores the dynamic nature of the truth, or assumes slow transition models.

In this chapter, we propose a networked filtering algorithm to track the state of a general hidden Markov model (HMM). We also analyze the performance and steady-state behavior of the resulting distributed strategy. In this process, we clarify questions about the benefit of cooperation and the nature of equilibria in social networks under dynamic environments.

5.1.1 Contributions

- In Sec. 5.3.1, we propose an HMM filtering algorithm for multi-agent networks. The algorithm requires only one round of communication between agents per state change. Moreover, it utilizes the knowledge of the transition model, which allows it to track highly dynamic states.

[1] The material in this chapter is based on [71, 72].

- In Sec. 5.4, we study the deviation of the proposed algorithm from the optimal centralized strategy defined in Sec. 5.2.1. Geometric ergodicity is the only assumption on the transition model. The specialization of the results to the single-agent case is a contribution to the Bayesian filtering literature.

- In Sec. 5.5, we provide recursive expressions for the probability of error across the network for the binary hypothesis testing case. Furthermore, under Gaussian data distributions, we obtain an asymptotic convergence result in *distribution*. The result implies that the agents attain steady-state probability of errors, which can vary across the agents depending on their centrality.

5.2 Problem Setting

In this chapter, we allow the state of nature to be dynamic. It is now a time-dependent random variable, and we denote it by $\boldsymbol{\theta}_i^\circ$.

The value $\mu_{k,i}(\theta)$ represents the confidence level that agent k has at time i about θ being the true hypothesis $\boldsymbol{\theta}_i^\circ$ (which is assumed to belong to a set Θ). The true hypothesis is a random variable and it will be assumed to evolve according to some Markov chain, known to all agents. We use the following notation for the transition model:

$$\mathbb{T}(\theta_i|\theta_{i-1}) \triangleq \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta_i|\boldsymbol{\theta}_{i-1}^\circ = \theta_{i-1}). \tag{5.1}$$

Conditioned on $\boldsymbol{\theta}_i^\circ$, the observation is distributed according to some likelihood function known to agent k, and denoted by $L_k(\boldsymbol{\xi}_{k,i}|\boldsymbol{\theta}_i^\circ)$. These agent-specific likelihoods can be probability density or mass functions depending on whether the observations are continuous or discrete. In this chapter, for ease of notation, we assume the observations are continuous. Nevertheless, our analysis is also valid for discrete observations with proper adjustments, e.g., by changing integrals to summations. Moreover, in this chapter, we have the following additional conditions.

Assumption 5.1 (**Independent observations).** *Conditioned on the true state, the observations are independent over space. More specifically, let* $\boldsymbol{\xi}_i \triangleq \{\boldsymbol{\xi}_{k,i}\}_{k=1}^K$, *collect all observations from across the agents at time* i. *Then, the joint likelihood is given by,*

$$L(\boldsymbol{\xi}_i|\boldsymbol{\theta}_i^\circ) = \prod_{k=1}^{K} L_k(\boldsymbol{\xi}_{k,i}|\boldsymbol{\theta}_i^\circ). \tag{5.2}$$

In this work, agents will be required to communicate only once with their neighbors per iteration. The underlying communication topology is assumed to satisfy the following condition.

Assumption 5.2 (Doubly-stochastic matrix). *We assume that A is a doubly-stochastic and symmetric matrix, namely,*

$$A\mathbb{1}_K = \mathbb{1}_K, \quad A = A^{\mathsf{T}}. \tag{5.3}$$

We have already seen that strong connectivity causes the information to disperse throughout the entire network given sufficient iterations. When the true hypothesis is fixed (i.e., $\boldsymbol{\theta}_i^\circ = \theta^\circ$), this allows the agents to reach agreement and learn θ° almost surely, as discussed in Chapter 2. However, strong connectivity is not sufficient for network agreement if the true hypothesis is changing rapidly before local information reaches other agents, as in the current chapter.

5.2.1 Optimal Centralized Belief Recursion

Let us denote the observation history of all agents across the network up to time i by $\boldsymbol{\mathcal{F}}_i \triangleq \{\boldsymbol{\xi}_j\}_{j=1}^i$. Likewise, let us denote the posterior distribution (or belief), which is a probability mass function (pmf) due to the assumed finite state-space model, by the notation:

$$\boldsymbol{\mu}_i^\star(\theta_i) \triangleq \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta_i | \boldsymbol{\mathcal{F}}_i). \tag{5.4}$$

It is known that the above distribution satisfies the optimal Bayesian filtering recursion [48]:

$$\boldsymbol{\mu}_i^\star(\theta_i) \propto L(\boldsymbol{\xi}_i|\theta_i)\boldsymbol{\eta}_i^\star(\theta_i), \tag{5.5}$$

where $\boldsymbol{\eta}_i^\star(\theta_i)$ is the time-adjusted prior defined by

$$\boldsymbol{\eta}_i^\star(\theta_i) \triangleq \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta_i | \boldsymbol{\mathcal{F}}_{i-1}) = \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\boldsymbol{\mu}_{i-1}^\star(\theta_{i-1}). \tag{5.6}$$

Once the posterior is updated by (5.4), the state estimator at time i is obtained from the maximum a-posteriori construction:

$$\widehat{\boldsymbol{\theta}}_i^\star \triangleq \arg\max_{\theta_i\in\Theta} \boldsymbol{\mu}_i^\star(\theta_i) \tag{5.7}$$

The main challenge with this solution method is that it requires a fusion center to gather all data from across time and agents. In Section 5.4, we will examine how close the beliefs generated by the proposed decentralized algorithm will get to the above centralized posterior given by (5.5)–(5.6).

Remark 5.1 (**Sequence estimation).** *In this chapter, the focus is on estimating the* current *state from past observations. If a sequence of hidden states, including both future or past states, is to be estimated, then single state-estimators can be combined with dynamic programming principles, such as in the Viterbi algorithm [48, 73].*

5.3 Decentralized Bayesian Filtering

The centralized solution (5.5)–(5.6) can be disadvantageous for various reasons: (i) collecting all data at a single fusion location makes the system vulnerable with a single point of failure; (ii) the agents may be reluctant to share their raw data with a remote central processor for privacy or security reasons; and (iii) communications back and forth with a remote fusion center is costly. For these reasons, we pursue instead a decentralized solution that is able to approach the performance of the centralized solution. In the decentralized approach, agents will only share data with their immediate neighbors; actually, the agents will not be required to share their raw data but only their updated belief vectors. The resulting solution will be more robust to node or link failure and more communication efficient, and will lead to an effective solution method.

5.3.1 Diffusion HMM Filtering

The streaming observations arriving at each agent are generally only partially informative about the true state of nature, $\boldsymbol{\theta}_i^\circ$. For this reason, agents will need to cooperate with their neighbors, thus leading to a learning mechanism that allows information to diffuse through the network for enhanced performance. To do so, we propose a *social* Bayesian filtering algorithm where cooperation among agents takes advantage of the notion of diffusion learning (see, e.g., [7, 9]).

Specifically, at each time instant i, every agent k first time-adjusts or evolves its belief from $i-1$, denoted by $\boldsymbol{\mu}_{k,i-1}(\theta_{i-1})$, via the Chapman-Kolmogorov equation [48] and generates an updated prior denoted by $\boldsymbol{\eta}_{k,i}(\theta_i)$:

$$\boldsymbol{\eta}_{k,i}(\theta_i) = \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\boldsymbol{\mu}_{k,i-1}(\theta_{i-1}) \qquad (\textbf{Evolve}) \tag{5.8}$$

This relation is motivated by the optimal update (5.6), except that optimal beliefs are replaced by their local versions at agent k. In the next step, agents seek to incorporate the information from their newly arrived private observations. This can be achieved

by considering the following regularized optimization problem:

$$\min_{\psi\in\Delta_H}\left\{D_{\mathrm{KL}}(\psi||\boldsymbol{\eta}_{k,i})-\gamma\,\mathbb{E}_\psi\log L_k(\boldsymbol{\xi}_{k,i}|\theta_i)\right\} \tag{5.9}$$

where Δ_H is the probability simplex of dimension H, and $\mathbb{E}_\psi$ is the expectation computed with respect to ψ, i.e.,

$$\mathbb{E}_\psi\log L_k(\boldsymbol{\xi}_{k,i}|\theta_i)\triangleq\sum_{\theta_i\in\Theta}\psi(\theta_i)\log L_k(\boldsymbol{\xi}_{k,i}|\theta_i). \tag{5.10}$$

The objective function in (5.9) consists of two terms. The first term is the KL-divergence term that penalizes the disagreement with the time-adjusted prior $\{\boldsymbol{\eta}_{k,i}\}$. The second term corresponds to the log-likelihood of the observation $\boldsymbol{\xi}_{k,i}$ averaged over the hypotheses with respect to ψ. The cost in (5.9) then seeks to minimize disagreement with the prior while maximizing the likelihood of the observation; the two terms are coupled by a regularization parameter $\gamma>0$. As we show in the sequel, different special cases of the problem setting might necessitate different γ values, e.g., $\gamma=K,\gamma=1$. Therefore, we continue with a general parameter $\gamma>0$. The objective function in (5.9) can be expanded as

$$\sum_{\theta_i\in\Theta}\psi(\theta_i)\bigg(\log\frac{\psi(\theta_i)}{\boldsymbol{\eta}_{k,i}(\theta_i)}-\gamma\log L_k(\boldsymbol{\xi}_{k,i}|\theta_i)\bigg)=\sum_{\theta_i\in\Theta}\psi(\theta_i)\log\frac{\psi(\theta_i)}{\boldsymbol{\eta}_{k,i}(\theta_i)(L_k(\boldsymbol{\xi}_{k,i}|\theta_i))^\gamma} \tag{5.11}$$

The RHS of (5.11) is a KL-divergence under a proper normalization (to make sure the term in the denominator is a pmf). Minimizing (5.11) over ψ results in the following *local* γ-scaled Bayesian adaptation step for each agent:

$$\boldsymbol{\psi}_{k,i}(\theta_i)\propto(L_k(\boldsymbol{\xi}_{k,i}|\theta_i))^\gamma\boldsymbol{\eta}_{k,i}(\theta_i)\qquad(\textbf{Adapt}) \tag{5.12}$$

where $\gamma>0$ scales the likelihood of the new observation against prior information. After agents independently obtain their intermediate beliefs $\boldsymbol{\psi}_{k,i}$ according to (5.8) and (5.12), they exchange these beliefs with their neighbors. Each agent k will then need to fuse the beliefs received from the neighbors, and one way to do so is to seek the belief vector μ that solves [35, 74]:

$$\min_{\mu\in\Delta_H}\left\{\sum_{\ell\in\mathcal{N}_k}a_{\ell k}D_{\mathrm{KL}}(\mu||\boldsymbol{\psi}_{\ell,i})\right\}. \tag{5.13}$$

This objective function penalizes the average disagreement with the neighbors' intermediate beliefs and it can be expanded as

$$\sum_{\theta_i\in\Theta}\mu(\theta_i)\sum_{\ell\in\mathcal{N}_k}a_{\ell k}\log\frac{\mu(\theta_i)}{\boldsymbol{\psi}_{\ell,i}(\theta_i)}=\sum_{\theta_i\in\Theta}\mu(\theta_i)\log\frac{\mu(\theta_i)}{\prod_{\ell\in\mathcal{N}_k}[\boldsymbol{\psi}_{\ell,i}(\theta_i)]^{a_{\ell k}}} \tag{5.14}$$

where the term on the RHS of (5.14) can be seen as a KL divergence under proper normalization, whose minimizer is given by the following geometric-average combination:

$$\boldsymbol{\mu}_{k,i}(\theta_i) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(\theta_i)\Big)^{a_{\ell k}} \quad (\textbf{Combine}). \tag{5.15}$$

Exchanging and combining the beliefs repeatedly allows the local information to diffuse through the network. The complete procedure leads to the *diffusion HMM strategy* (DHS), which is listed in (5.16)–(5.18).

Algorithm 5.1 Diffusion HMM strategy (DHS)

set initial priors $\boldsymbol{\mu}_{k,0}(\theta) > 0$, $\forall \theta \in \Theta$ and $\forall k \in \mathcal{N}$
choose $\gamma > 0$
while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    time-adjust the belief from previous iteration:

$$\boldsymbol{\eta}_{k,i}(\theta_i) = \sum_{\theta_{i-1} \in \Theta} \mathbb{T}(\theta_i|\theta_{i-1}) \boldsymbol{\mu}_{k,i-1}(\theta_{i-1}) \tag{5.16}$$

    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt locally to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta_i) \propto (L_k(\boldsymbol{\xi}_{k,i}|\theta_i))^{\gamma} \boldsymbol{\eta}_{k,i}(\theta_i) \tag{5.17}$$

    combine received beliefs from the neighbors:

$$\boldsymbol{\mu}_{k,i}(\theta_i) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(\theta_i)\Big)^{a_{\ell k}} \tag{5.18}$$

  end for
  $i \leftarrow i + 1$
end while

The proposed DHS algorithm is a generalization of the following special cases:

- When the network consists of a single-agent, i.e., $K = 1$, the strategy is equivalent to the traditional optimal Bayesian filtering algorithm [48, Chapter 3] when the local updates are Bayesian, i.e., when $\gamma = 1$.
- If $\gamma = 1$, and the true hypothesis is *fixed*, i.e.,

$$\mathbb{T}(\theta_i|\theta_{i-1}) = \begin{cases} 1, & \theta_i = \theta_{i-1} \\ 0, & \theta_i \neq \theta_{i-1} \end{cases}, \tag{5.19}$$

  then, the algorithm reduces to the canonical geometric social learning listed in

Algorithm 2.1.

- The beliefs of agents will match the optimal centralized belief (5.5) *exactly*, if the network is fully-connected with $a_{\ell k} = 1/K \; \forall \ell, k \in \mathcal{N}$, all initial priors are equal ($\mu_0^\star = \mu_{k,0}, \forall k \in \mathcal{N}$), and $\gamma = K$. This conclusion follows by induction. First, assume that for each agent k, $\boldsymbol{\mu}_{k,i-1} = \boldsymbol{\mu}_{i-1}^\star$. This would imply that $\boldsymbol{\eta}_{k,i} = \boldsymbol{\eta}_i^\star$ by the equivalence of the time-adjustment steps (5.6) and (5.16). Combining the adapt (5.17) and combine (5.18) steps, the belief of agent k at time i then becomes

$$\begin{aligned} \boldsymbol{\mu}_{k,i}(\theta_i) &\propto \prod_{\ell \in \mathcal{N}_k} \Big(L_k(\boldsymbol{\xi}_{k,i}|\theta_i) \Big)^{\gamma a_{\ell k}} \Big(\boldsymbol{\eta}_{k,i}(\theta_i) \Big)^{a_{\ell k}} \\ &\propto \boldsymbol{\eta}_i^\star(\theta_i) \prod_{\ell=1}^{K} L_k(\boldsymbol{\xi}_{k,i}|\theta_i) \end{aligned} \tag{5.20}$$

  which is equivalent to the centralized update (5.5). Since the base case $\mu_0^\star = \mu_{k,0}$ also holds, we conclude by induction that, the beliefs at all iterations will match the centralized belief.

We continue with the general strategy (5.16)–(5.18). Since different γ values correspond to different special cases, we continue to use a general step-size $\gamma > 0$.

5.4 Optimality Gap

In this section, we analyze the disagreement between the diffusion HMM strategy (5.16)–(5.18) and the centralized solution (5.5)–(5.6). For this section alone, we assume a regularity condition on the likelihood functions for technical reasons (similar to what was done in [75]).

Assumption 5.3 (Regularity condition). *The absolute log-likelihood functions are uniformly bounded over their support for all agents:*

$$\Big| \log L_k(\xi|\theta) \Big| \leq C_L, \qquad \forall k \in \mathcal{N}, \theta \in \Theta. \tag{5.21}$$

Assumption 5.3 implies that the likelihood functions do not get arbitrarily close to zero or arbitrarily large in their support. This ensures that each private signal ξ has bounded informativeness. For example, discrete signal space models or truncated Gaussian likelihoods satisfy this assumption.

Next, we introduce conditions on the transition model, which will play a crucial part in the analysis.

5.4.1 Transition Model

We assume that the transition Markov chain is *irreducible* and *aperiodic* [10, Chapter 2]. This means that there exists a constant integer $n > 0$ such that for any two hypotheses $\theta, \theta' \in \Theta$:

$$\mathbb{T}^n(\theta|\theta') > 0, \tag{5.22}$$

where $\mathbb{T}^n$ is the n-fold application of the transition kernel. This condition also implies that the Markov chain is *ergodic* because the number of hypotheses H is finite [10, Chapter 2]. In other words, repeated application of the transition kernel $\mathbb{T}$ will converge to a limiting distribution regardless of the initial input distribution. More formally, for any input distribution $\mu \in \Delta_H$,

$$\lim_{n\to\infty} \sum_{\theta'\in\Theta} \mathbb{T}^n(\theta|\theta')\mu(\theta') = \pi(\theta), \tag{5.23}$$

where π is the Perron vector of the $H \times H$ transition matrix $T \triangleq [\mathbb{T}(\theta|\theta')]$. In this work, we consider the *geometrically ergodic* [48, Chapter 2] subclass of transition models. In the following, we define these models using the strong-data processing inequality (SDPI) [55].

Definition 5.1 (**Strong-data processing inequality (SDPI) [55]**). *Consider any two discrete distributions over* Θ, μ^a *and* μ^b, *satisfying* $0 < D_{\mathrm{KL}}(\mu^a||\mu^b) < \infty$, *and introduce their time-adjusted versions according to the Chapman-Kolmogorov equation as in* (5.16)*:*

$$\eta^a(\theta_i) = \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\mu^a(\theta_{i-1}) \tag{5.24}$$

and similarly for η^b. *Then, the SDPI states that:*

$$D_{\mathrm{KL}}(\eta^a||\eta^b) \leq \kappa_{KL}(\mathbb{T}) D_{\mathrm{KL}}(\mu^a||\mu^b) \tag{5.25}$$

where $\kappa_{KL}(\mathbb{T}) \in [0,1]$ *is a contraction coefficient defined as*

$$\kappa_{KL}(\mathbb{T}) \triangleq \sup_{\mu^a,\mu^b} \frac{D_{\mathrm{KL}}(\eta^a||\eta^b)}{D_{\mathrm{KL}}(\mu^a||\mu^b)}. \tag{5.26}$$

Observe that the coefficient is only dependent on the transition model, and is not a function of the input distributions. An upper bound on $\kappa_{\mathrm{KL}}(\mathbb{T})$ is given by the Do-

brushin's contraction coefficient [48, 55, 76], which is defined by

$$\kappa(\mathbb{T}) \triangleq \sup_{\theta',\theta''\in\Theta} \frac{1}{2}\sum_{\theta\in\Theta}\left|\mathbb{T}(\theta|\theta') - \mathbb{T}(\theta|\theta'')\right| \quad \in [0,1], \tag{5.27}$$

It is known that [55]:

$$\kappa_{\mathrm{KL}}(\mathbb{T}) \leq \kappa(\mathbb{T}), \qquad \kappa_{\mathrm{KL}}(\mathbb{T}) = 1 \iff \kappa(\mathbb{T}) = 1. \tag{5.28}$$

For example,

- If the transition model is a binary symmetric channel:

$$\mathbb{T}(\theta_i|\theta_{i-1}) = \begin{cases} 1-\alpha, & \theta_i = \theta_{i-1} \\ \alpha, & \theta_i \neq \theta_{i-1} \end{cases}, \tag{5.29}$$

  then $\kappa(\mathbb{T}) = |1-2\alpha|$ [55]. Notice that this is a symmetric function around the transition probability $\alpha = 0.5$, e.g., $\alpha = 0.2$ and $\alpha = 0.8$ yield the same coefficient.

- The coefficient $\kappa(\mathbb{T}) = 0$ if, and only if [48, Chapter 2],

$$\mathbb{T}(\theta_i|\theta_{i-1}) = \pi(\theta_i). \tag{5.30}$$

  Note that this implies, for any $\mu \in \Delta_H$,

$$\sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\mu(\theta_{i-1}) = \sum_{\theta_{i-1}\in\Theta} \pi(\theta_i)\mu(\theta_{i-1}) = \pi(\theta_i) \tag{5.31}$$

  In other words, the transition kernel will output the same distribution $\pi(\theta)$ regardless of the input distribution $\mu(\theta)$. Observe that the rapidly mixing binary symmetric channel with transition probability $\alpha = 0.5$ is an example of this case.

In general, for two input distributions μ^a and μ^b, the output distributions resulting from an n-fold application of the transition kernel, i.e.,

$$\eta_n^a(\theta_i) = \sum_{\theta_{i-n}\in\Theta} \mathbb{T}^n(\theta_i|\theta_{i-n})\mu^a(\theta_{i-n}), \tag{5.32}$$

and similarly for η_n^b, satisfy the following SDPI:

$$D_{\mathrm{KL}}(\eta_n^a||\eta_n^b) \leq (\kappa(\mathbb{T}))^n D_{\mathrm{KL}}(\mu^a||\mu^b) \tag{5.33}$$

It is clear that if $\kappa(\mathbb{T}) < 1$, then the disagreement between any two input distributions will approach 0 exponentially fast. Transition models for which $\kappa(\mathbb{T}) < 1$ are said to be geometrically ergodic [48, Chapter 2]. It is seen from (5.33) that the coefficient $\kappa(\mathbb{T})$ is a

measure of how rapidly the initial conditions are forgotten. In particular, as $\kappa(\mathbb{T}) \to 0$, forgetting is faster.

Assumption 5.4 (Transition model). *The transition model $\mathbb{T}$ is assumed to be geometrically ergodic, i.e., $\kappa(\mathbb{T}) < 1$.*

The class of geometric ergodic transition models comprises a large group of models. For instance, non-deterministic binary symmetric channels, i.e., with a transition probability $\alpha \in (0,1)$, are geometrically ergodic. Moreover, transition matrices with all positive elements, and in general, those that satisfy a minorization condition [48, Theorem 2.7.4] are examples of geometrically ergodic transition models. However, the geometrically ergodic class excludes some models such as the fixed hypothesis case, where $\kappa(\mathbb{T}) = 1$. We elaborate more on this issue in the sequel.

5.4.2 Disagreement with the Centralized Strategy

We introduce the following time-varying risks to compare the performance of the diffusion HMM strategy (5.16)–(5.18) with the centralized solution (5.5)–(5.6):

$$J_{k,i} \triangleq \mathbb{E}_{\mathcal{F}_i} D_{\mathrm{KL}}(\boldsymbol{\mu}_i^\star || \boldsymbol{\mu}_{k,i}) \tag{5.34}$$

and

$$\widetilde{J}_{k,i} \triangleq \mathbb{E}_{\mathcal{F}_{i-1}} D_{\mathrm{KL}}(\boldsymbol{\eta}_i^\star || \boldsymbol{\eta}_{k,i}) \tag{5.35}$$

where $\mathbb{E}_{\mathcal{F}_i}$ represents expectation over the distribution of $\mathcal{F}_i$, which collects all observations from across the network until time i. Notice that the risks in (5.34)–(5.35) are not random variables, since the corresponding KL-divergences are averaged over all possible realizations of observations. The *posterior* risk $J_{k,i}$ in (5.34) is the disagreement between the belief of agent k and the centralized belief at time i, *after* the observations $\boldsymbol{\xi}_i$ are received. In comparison, the risk $\widetilde{J}_{k,i}$ in (5.35) is the divergence of time-adjusted *priors*, which measures the disagreement *before* the observations have been emitted.

Our first result establishes that the disagreement between the centralized and distributed solutions is asymptotically bounded for all agents in the network.

Theorem 5.1 (**Asymptotic bounds).** *For each agent k, under Assumptions 5.1–5.4, the risks* (5.34) *and* (5.35) *are asymptotically bounded, namely,*

$$\limsup_{i\to\infty} J_{k,i} \leq \frac{2\sqrt{K}\gamma\lambda C_L}{1-\kappa(\mathbb{T})} \tag{5.36}$$

and

$$\limsup_{i\to\infty} \widetilde{J}_{k,i} \leq \frac{2\kappa(\mathbb{T})\sqrt{K}\gamma\lambda C_L}{1-\kappa(\mathbb{T})} \tag{5.37}$$

where $\lambda \triangleq \max\{|1-\frac{K}{\gamma}|, \rho_2\}$, and ρ_2 is the second largest modulus eigenvalue of A.

Proof. See Appendix 5.A. ■

Swift and random changes in the environment can prevent a strongly-connected network from approaching the performance of the centralized solution close enough—especially when the network is sparse and it takes more time for the information to diffuse to all agents than the rate at which the state is changing. Therefore, the bounds in Theorem 5.1 are not generally close to 0. The following remarks are now in place. Simulations that support these observations appear in Section 5.6.

- **Matching the centralized strategy**: The bounds (5.36) and (5.37) will be tight when the distributed solution reduces to the centralized implementation, which happens with a uniformly weighted fully-connected network ($a_{\ell k} = 1/K\ \forall \ell, k \in \mathcal{N}$, $\rho_2 = 0$), same initial priors ($\mu_0^\star = \mu_{k,0}, \forall k \in \mathcal{N}$), and $\gamma = K$ as shown in (5.20). In this case, the upper bounds will become zero as expected.

- **Stability**: The bounds (5.36) and (5.37) are independent of the initial beliefs as long as Assumption 2.2 is satisfied. Indeed, geometric ergodicity is sufficient to asymptotically forget the initial conditions. In this way, the filter is robust to incorrect initializations.

- **Network connectivity**: Note that as $\gamma \to K$, the bounds become proportional to ρ_2, the mixing rate of the graph. This aspect emphasizes the benefit of cooperation. Highly-connected graphs, with small ρ_2, will be closer to the centralized solution while sparse networks or non-cooperative agents will have higher deviation.

- **Network size**: The bounds are also proportional to the number of agents. The disagreement between the centralized solution and the individual agents increases with the square-root of the network size. Note that this does not mean agents would perform worse if new agents join the network. It is the *relative*

performance compared to the optimal solution, which has access to data from all agents, that could decrease.

- **Ergodicity**: Notice that when $\kappa(\mathbb{T}) \to 0$, from the bound (5.37), it is obvious that $\widetilde{J}_{k,i} \to 0$. This is anticipated because if the coefficient $\kappa(\mathbb{T}) = 0$, the time-adjustment steps of the centralized and decentralized strategies (5.6) and (5.16) become

$$\begin{aligned}\eta_i^\star(\theta_i) &= \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\boldsymbol{\mu}_{i-1}^\star(\theta_{i-1}) \overset{(5.31)}{=} \pi(\theta_i),\\ \eta_{k,i}(\theta_i) &= \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\boldsymbol{\mu}_{k,i-1}(\theta_{i-1}) \overset{(5.31)}{=} \pi(\theta_i),\end{aligned} \tag{5.38}$$

  regardless of the input distributions $\boldsymbol{\mu}_{i-1}^\star(\theta_{i-1})$ and $\boldsymbol{\mu}_{k,i-1}(\theta_{i-1})$. As a result,

$$\widetilde{J}_{k,i} \overset{(5.35)}{=} \mathbb{E}_{\mathcal{F}_{i-1}} D_{\mathrm{KL}}(\boldsymbol{\eta}_i^\star||\boldsymbol{\eta}_{k,i}) \overset{(5.38)}{=} 0. \tag{5.39}$$

  Therefore, the bound (5.37) captures the effect of the ergodicity of the transition model via the ergodicity coefficient $\kappa(\mathbb{T})$ accurately. In particular, the results imply that the bound is tight for rapidly changing binary symmetric transition models where $\kappa(\mathbb{T}) \to 0$.

- **Informativeness of observations**: Nonetheless, the bounds fall short in capturing the effect of the observations. For example, if the true hypothesis is fixed, then the transition model $\mathbb{T}$ satisfies (5.19). In this case, for two distributions $\mu^a, \mu^b \in \Delta_H$,

$$\eta^a(\theta_i) = \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\mu^a(\theta_{i-1}) \overset{(5.19)}{=} \mu^a(\theta_i) \tag{5.40}$$

  and similarly for η^b. Then,

$$D_{\mathrm{KL}}(\eta^a||\eta^b) = D_{\mathrm{KL}}(\mu^a||\mu^b), \tag{5.41}$$

  and consequently, $\kappa(\mathbb{T}) = 1$. Since this transition model is not geometrically ergodic, the bounds do not cover this case. However, it is known from the standard social learning literature (recall Theorems 2.1 and 2.2) that when the observations of the agents are informative enough, that is to say, the observations provide sufficient information about the underlying state (Assumption 2.1), all agents will learn the true hypothesis eventually. In other words, beliefs of agents, as well as the belief of the centralized strategy, on wrong hypotheses go to 0. This means that log-beliefs on wrong hypotheses become degenerate, and hence, it is not clear how the KL-divergences in the risks (5.34)–(5.35) would behave in this situation, or whether the KL-divergence is a meaningful metric here. We leave examining this interesting regime of $\kappa(\mathbb{T}) \to 1$, where the dominant factor is the informativeness of observations rather than the ergodicity, to future work.

- **Single-agent case**: In fact, the distinction between the ergodicity of the transition model and the informativeness of the observations arises in the analysis of single-agent strategies as well. The stability of the algorithm would refer to the conditions under which a wrongly initialized belief converges to the true posterior distribution in HMMs. In the notation of this work, this is the special case of $\gamma = K = 1$ and $\rho_2 = 0$. Remember that in this special case the proposed diffusion HMM strategy is equivalent to a traditional optimal Bayes filter [48, Chapter 3]. One criterion that the literature on the single-agent case uses is whether the total variation distance between the true posterior and the agent's belief vanishes in the mean asymptotically, i.e., whether (setting the single-agent index $k = 1$),

$$\lim_{i\to\infty} \mathbb{E}_{\mathcal{F}_i}\Big[\sum_{\theta\in\Theta}\Big|\boldsymbol{\mu}_i^\star(\theta) - \boldsymbol{\mu}_{1,i}(\theta)\Big|\Big] \stackrel{?}{=} 0, \tag{5.42}$$

when the initial belief of the agent is not accurate, i.e., when

$$\mathbb{P}(\boldsymbol{\theta}_0^\circ = \theta) = \mu_0^\star(\theta) \neq \mu_{1,0}(\theta). \tag{5.43}$$

The total variation distance in (5.42) can vanish because of two mechanisms: either the observations are sufficiently informative about the true state, or the transition model is sufficiently ergodic. Even though there are some works that focus on the informativeness of observations [77], most of the results in the literature rely on ergodicity to establish stability [78, 79]. Similarly, Theorem 5.1 also depends on the ergodicity of the transition model. In particular, reference [79] studies the stability via the Dobrushin coefficient $\kappa(\mathbb{T})$ and concludes that as long as $\kappa(\mathbb{T}) < 1/2$, regardless of the observation model, (5.42) is satisfied. In comparison, our results are in terms of KL-divergences, but they can be expressed in terms of total variation distances if we use Pinsker's inequality [13, Chapter 3]. For the single-agent case, we have

$$\sum_{\theta\in\Theta}\Big|\boldsymbol{\mu}_i^\star(\theta) - \boldsymbol{\mu}_{1,i}(\theta)\Big| \leq \Big(2D_{\mathrm{KL}}(\boldsymbol{\mu}_i^\star||\boldsymbol{\mu}_{1,i})\Big)^{\frac{1}{2}}. \tag{5.44}$$

If we take expectations, this relation implies

$$\mathbb{E}_{\mathcal{F}_i}\Big[\sum_{\theta\in\Theta}\Big|\boldsymbol{\mu}_i^\star(\theta) - \boldsymbol{\mu}_{1,i}(\theta)\Big|\Big] \leq \mathbb{E}_{\mathcal{F}_i}\Big(2D_{\mathrm{KL}}(\boldsymbol{\mu}_i^\star||\boldsymbol{\mu}_{1,i})\Big)^{\frac{1}{2}} \stackrel{(a)}{\leq} \Big(2J_{1,i}\Big)^{\frac{1}{2}}, \tag{5.45}$$

where (a) follows from Jensen's inequality. Finally, taking the limit of both sides, we arrive at

$$\lim_{i\to\infty} \mathbb{E}_{\mathcal{F}_i}\Big[\sum_{\theta\in\Theta}\Big|\boldsymbol{\mu}_i^\star(\theta) - \boldsymbol{\mu}_{1,i}(\theta)\Big|\Big] \leq \sqrt{2}\lim_{i\to\infty}\Big(J_{1,i}\Big)^{\frac{1}{2}} \stackrel{(b)}{=} 0, \tag{5.46}$$

where (b) follows from the fact that the bound (5.36) is equal to zero if $\gamma = K = 1$

and $\rho_2 = 0$. This is a more general result, namely that the single-agent Bayesian filter is stable whenever $\kappa(\mathbb{T}) < 1$, as long as the informativeness of observations are bounded (Assumption 5.3).

5.5 Probability of Error and Convergence

The previous section analyzed the closeness of the diffusion HMM strategy to the optimal centralized solution. In this section, we study the error probability for each agent. There are a few results for the probability of error of HMM filtering, even in the single agent case. One notable result is [80] where error recursions are obtained for the single-agent binary hypothesis setting case. In a similar spirit, we will obtain recursive equations for the probability density functions (pdfs) that capture the stochastic behavior of the underlying system. This will not only provide recursive formulas for the error probabilities, but will also allow us to deduce asymptotic properties for the diffusion HMM strategy. Similar to [80], for this section, we shift our focus to the binary hypothesis setting, i.e., throughout this section, we set $H = 2$ and $\Theta = \{0, 1\}$. We do not need to restrict ourselves to bounded log-likelihood signal models in this section, as was the case in Assumption 5.3. In this setting, the MAP-classifier at agent k at time instant i becomes

$$\widehat{\boldsymbol{\theta}}_{k,i} = \begin{cases} 1, & \text{if} \quad \boldsymbol{\mu}_{k,i}(1) > \boldsymbol{\mu}_{k,i}(0) \\ 0, & \text{if} \quad \boldsymbol{\mu}_{k,i}(0) \geq \boldsymbol{\mu}_{k,i}(1) \end{cases} . \tag{5.47}$$

This estimator is equivalent to

$$\widehat{\boldsymbol{\theta}}_{k,i} = \begin{cases} 1, & \text{if} \quad \boldsymbol{w}_{k,i} > 0 \\ 0, & \text{if} \quad \boldsymbol{w}_{k,i} \leq 0 \end{cases} \tag{5.48}$$

in terms of the log-belief ratio $\boldsymbol{w}_{k,i}$ defined as

$$\boldsymbol{w}_{k,i} \triangleq \log \frac{\boldsymbol{\mu}_{k,i}(1)}{\boldsymbol{\mu}_{k,i}(0)}. \tag{5.49}$$

As such, the probability of error for agent k at time instant i is given by:

$$p_{k,i} \triangleq \mathbb{P}(\boldsymbol{\theta}_i^\circ = 1, \boldsymbol{w}_{k,i} \leq 0) + \mathbb{P}(\boldsymbol{\theta}_i^\circ = 0, \boldsymbol{w}_{k,i} > 0). \tag{5.50}$$

Let $f_{k,i}(\theta, w_k)$ denote the probability density function of the joint variables $\{\boldsymbol{\theta}_i^\circ, \boldsymbol{w}_{k,i}\}$ for agent k at time i. Note that the joint variables $\{\boldsymbol{\theta}_i^\circ, \boldsymbol{w}_{k,i}\}$ mix a discrete and a continuous random variable. Here, we use the general definition of density, i.e., density is the Radon-Nikodym derivative with respect to a measure. The corresponding measure for $\{\boldsymbol{\theta}_i^\circ, \boldsymbol{w}_{k,i}\}$ is the product measure of the counting measure for $\boldsymbol{\theta}_i^\circ$ and the Lebesgue measure for $\boldsymbol{w}_{k,i}$ [81]. For example, we can evaluate the probability that $\boldsymbol{\theta}_i^\circ = \theta$ and

$\boldsymbol{w}_{k,i}$ lies within the infinitesimal interval $(w_k, w_k + dw_k)$ by computing

$$f_{k,i}(\theta, w_k)dw_k = \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta, \boldsymbol{w}_{k,i} \in (w_k, w_k + dw_k)) \tag{5.51}$$

In this way, the probability of error (5.50) is given by

$$p_{k,i} = \int_{w_k=-\infty}^{0} f_{k,i}(1, w_k)dw_k + \int_{w_k=0}^{\infty} f_{k,i}(0, w_k)dw_k. \tag{5.52}$$

We further consider the probability density function involving the log-beliefs across all agents, namely,

$$f_i(\theta, w)dw_1 dw_2 \cdots dw_K \triangleq \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta, \boldsymbol{w}_i \in (w, w + dw)), \tag{5.53}$$

in terms of the aggregate variables:

$$\boldsymbol{w}_i \triangleq \text{col}\{\boldsymbol{w}_{\ell,i}\}_{\ell=1}^K, \quad w \triangleq \text{col}\{w_\ell\}_{\ell=1}^K. \tag{5.54}$$

If we integrate (5.53) over all agents with the exception of agent k, we can determine the marginal density for agent k, namely,

$$f_{k,i}(\theta, w_k) = \int\!\!\ldots\!\!\int f_i(\theta, w)dw_1 \cdots dw_{k-1}dw_{k+1} \cdots dw_K. \tag{5.55}$$

In what follows, we will derive a temporal recursion for the joint density given by (5.53), from which agent-specific densities can then be deduced. To this end, first observe that the diffusion equations (5.16)-(5.18) can be written compactly in terms of the log-belief ratio:

$$\begin{aligned}\boldsymbol{w}_{k,i} &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\gamma \log\frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|1)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|0)} + \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\log\frac{\boldsymbol{\eta}_{\ell,i}(1)}{\boldsymbol{\eta}_{\ell,i}(0)} \\ &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\gamma \log\frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|1)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|0)} + \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\log\frac{\mathbb{T}(1|0) + \mathbb{T}(1|1)\text{exp}\{\boldsymbol{w}_{\ell,i-1}\}}{\mathbb{T}(0|0) + \mathbb{T}(0|1)\text{exp}\{\boldsymbol{w}_{\ell,i-1}\}}.\end{aligned} \tag{5.56}$$

In vector form, equation (5.56) leads to

$$\boldsymbol{w}_i = A^\mathsf{T}\boldsymbol{\nu}_i + A^\mathsf{T}\boldsymbol{\chi}_i \quad \text{(diffusion HMM)}, \tag{5.57}$$

where we are introducing the vector of γ-scaled log-likelihood ratios (LLR) across the network:

$$\boldsymbol{\nu}_i \triangleq \text{col}\left\{\gamma\frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|1)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|0)}\right\}_{\ell=1}^K, \tag{5.58}$$

and the vector of time-adjusted prior belief log-ratios across the network:

$$\boldsymbol{\chi}_i \triangleq \text{col}\left\{ \log \frac{\mathbb{T}(1|0) + \mathbb{T}(1|1)\exp\{\boldsymbol{w}_{\ell,i-1}\}}{\mathbb{T}(0|0) + \mathbb{T}(0|1)\exp\{\boldsymbol{w}_{\ell,i-1}\}} \right\}_{\ell=1}^{K}. \tag{5.59}$$

If the underlying distributed strategy is instead the consensus-based (recall Algorithm 2.4 from Chapter 2) in lieu of the diffusion strategy (5.16)–(5.18), then (5.57) would be replaced by

$$\boldsymbol{w}_i = \boldsymbol{\nu}_i + A^{\mathsf{T}}\boldsymbol{\chi}_i \quad \text{(consensus HMM)}. \tag{5.60}$$

Observe that the joint density over two consecutive time instants of $\{\boldsymbol{\theta}_{i-1}^{\circ}, \boldsymbol{w}_{i-1}, \boldsymbol{\theta}_i^{\circ}, \boldsymbol{w}_i\}$ satisfies

$$\begin{aligned}\mathbb{P}\Big(\boldsymbol{\theta}_i^{\circ} = \theta, &\boldsymbol{w}_i \in (w, w+dw), \boldsymbol{\theta}_{i-1}^{\circ} = \theta', \boldsymbol{w}_{i-1} \in (w', w'+dw')\Big) \\ &= S_i^{(\theta)}(w, w')\mathbb{T}(\theta|\theta')f_{i-1}(\theta', w')dW\,dW'\end{aligned} \tag{5.61}$$

where we are using $dW \triangleq dw_1 dw_2 \cdots dw_K$ and $dW' \triangleq dw_1' dw_2' \cdots dw_K'$ for notational brevity, and where we are introducing the conditional probability

$$\begin{aligned}S_i^{(\theta)}(w, w')dW &\triangleq \mathbb{P}(\boldsymbol{w}_i \in (w, w+dw)|\boldsymbol{\theta}_i^{\circ} = \theta, \boldsymbol{\theta}_{i-1}^{\circ} = \theta', \boldsymbol{w}_{i-1} = w') \\ &\overset{(a)}{=} \mathbb{P}(\boldsymbol{w}_i \in (w, w+dw)|\boldsymbol{\theta}_i^{\circ} = \theta, \boldsymbol{w}_{i-1} = w')\end{aligned} \tag{5.62}$$

where (a) follows from the fact that $\boldsymbol{w}_i$ is a function of $\boldsymbol{\xi}_i$ only, once $\boldsymbol{w}_{i-1}$ and $\boldsymbol{\theta}_i^{\circ}$ are given. Therefore, the log-belief ratio $\boldsymbol{w}_i$ is conditionally independent of $\boldsymbol{\theta}_{i-1}^{\circ}$ — see (5.57) for diffusion and (5.60) for consensus. Note that, for diffusion algorithms, in general, even under the independence Assumption 5.1:

$$S_i^{(\theta)}(w, w')dW \neq \prod_{\ell=1}^{K} \mathbb{P}(\boldsymbol{w}_{\ell,i} \in (w_\ell, w_\ell + dw_\ell)|\boldsymbol{\theta}_i^{\circ} = \theta, \boldsymbol{w}_{i-1} = w') \tag{5.63}$$

because the newly arrived data $\boldsymbol{\xi}_{k,i}$ is utilized by agent k as well as by its neighbors in the same iteration. On the other hand, for consensus, under Assumption 5.1,

$$S_i^{(\theta)}(w, w')dW = \prod_{\ell=1}^{K} \mathbb{P}(\boldsymbol{w}_{\ell,i} \in (w_\ell, w_\ell + dw_\ell)|\boldsymbol{\theta}_i^{\circ} = \theta, \boldsymbol{w}_{i-1} = w') \tag{5.64}$$

since the fresh data is used by the observing agent only. In other words, in the consensus implementation, each log-belief $\boldsymbol{w}_{k,i}$ is a function of that agent's observation $\boldsymbol{\xi}_{k,i}$ only (given $\boldsymbol{\theta}_i^{\circ}$ and $\boldsymbol{w}_{i-1}$) and does not depend on the observations at the other agents at that time. This distinction between diffusion and consensus can be seen from (5.57) and (5.60).

Now, marginalizing (5.61) with respect to the state θ' and the log-belief ratio w' at the earlier time instant yields the following temporal recursion for the joint density in (5.53):

$$f_i(\theta,w)dW = \sum_{\theta'} \mathbb{T}(\theta|\theta') \left[\int_{w'} \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta, \boldsymbol{w}_i \in (w, w+dw), \boldsymbol{\theta}_{i-1}^\circ = \theta', \boldsymbol{w}_{i-1} \in (w', w'+dw')) \right] \tag{5.65}$$

which implies that

$$f_i(\theta,w) = \sum_{\theta'} \mathbb{T}(\theta|\theta') \left[\int_{w'} S_i^{(\theta)}(w,w') f_{i-1}(\theta',w') dW' \right] \tag{5.66}$$

To summarize, the probability of error at each time instant i can be computed by (i) using (5.66) to find the joint density at time i, (ii) marginalizing the joint density to find the agent-specific density by (5.55), and finally (iii) integrating the agent density as in (5.52).

Remark 5.2 (**Evaluation of integrals).** *Finding closed-form expressions to the integral expressions (e.g.,* (5.55)*,* (5.66)*) might not be feasible. One can apply numerical integration methods such as Monte Carlo techniques [82] to compute the desired integrals.*

The analysis until here holds for general transition and likelihood models. However, the kernel $S_i^{(\theta)}(w,w')$ might not be obtained in closed form in general. Therefore, for stronger results, we focus on Gaussian likelihood models in the next section.

5.5.1 Gaussian Likelihoods

Thus, let us consider now Gaussian models of the form:

$$\begin{aligned} L_k(\xi_{k,i}|\boldsymbol{\theta}_i^\circ = 1) &= \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{ -\frac{(\xi_{k,i} - \zeta(1))^2}{2\sigma_k^2} \right\} \\ L_k(\xi_{k,i}|\boldsymbol{\theta}_i^\circ = 0) &= \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left\{ -\frac{(\xi_{k,i} - \zeta(0))^2}{2\sigma_k^2} \right\} \end{aligned} \tag{5.67}$$

where the means are assumed to satisfy $\zeta(0) = -\zeta(1) = \zeta$ for a constant value $\zeta \neq 0$, and where the agent-specific variances satisfy $\sigma_k^2 > 0$. The corresponding log-

likelihood ratio appearing in (5.56) is then given by

$$\log \frac{L_k(\xi_{k,i}|\boldsymbol{\theta}_i^\circ = 1)}{L_k(\xi_{k,i}|\boldsymbol{\theta}_i^\circ = 0)} = \frac{-2\zeta\xi_{k,i}}{\sigma_k^2}, \tag{5.68}$$

and accordingly, the vector $\boldsymbol{\nu}_i$ of γ-scaled LLRs across agents is given by

$$\boldsymbol{\nu}_i = \text{col}\Big\{\frac{-2\gamma\zeta}{\sigma_\ell^2}\boldsymbol{\xi}_{\ell,i}\Big\}_{\ell=1}^K. \tag{5.69}$$

By Assumption 5.1, the $\{\boldsymbol{\xi}_{\ell,i}\}_{\ell=1}^K$ are independent random variables conditioned on $\boldsymbol{\theta}_i^\circ$. This implies that $\boldsymbol{\nu}_i$ is a multivariate Gaussian random variable conditioned on the true hypothesis $\boldsymbol{\theta}_i^\circ$ at time instant i,

$$\boldsymbol{\nu}_i\Big|_{\boldsymbol{\theta}_i^\circ} \sim \mathcal{G}\Big(\boldsymbol{\beta}^{(\theta_i^\circ)}, \Sigma\Big) \tag{5.70}$$

with mean

$$\boldsymbol{\beta}^{(\theta_i^\circ)} \triangleq \text{col}\Big\{\frac{-2(-1)^{\boldsymbol{\theta}_i^\circ}\gamma\zeta^2}{\sigma_\ell^2}\Big\}_{\ell=1}^K, \tag{5.71}$$

and covariance matrix

$$\Sigma \triangleq \text{diag}\Big\{\frac{4\gamma^2\zeta^2}{\sigma_\ell^2}\Big\}_{\ell=1}^K. \tag{5.72}$$

Next, we treat the consensus and diffusion cases separately. The consensus case is straightforward and useful to understand the diffusion case.

Consensus

Using (5.60) and the distribution (5.70) for $\boldsymbol{\nu}_i$, we conclude that the conditional pdf of $\boldsymbol{w}_i$ given the current state and the prior log-belief ratio vector $\boldsymbol{w}_{i-1}$ is also Gaussian and equal to

$$S_i^{(\theta)}(w, w') = \frac{\exp\Big\{-\frac{1}{2}(w - \rho^{(\theta)}(w'))^\mathsf{T}\Sigma^{-1}(w - \rho^{(\theta)}(w'))\Big\}}{\sqrt{(2\pi)^K \det(\Sigma)}} \tag{5.73}$$

where the mean is defined by

$$\rho^{(\theta)}(w') \triangleq \boldsymbol{\beta}^{(\theta_i^\circ)} + A^\mathsf{T}\boldsymbol{\chi}_i\Big|_{\boldsymbol{\theta}_i^\circ=\theta, \boldsymbol{w}_{i-1}=w'}. \tag{5.74}$$

Observe that since Σ is a diagonal matrix, Eq. (5.73) can also be written as the multiplication of individual conditional densities, as already suggested by (5.64).

Diffusion

The covariance matrix Σ in (5.72) is non-singular since it is diagonal with positive diagonal entries. Consequently, $\boldsymbol{\nu}_i$ in (5.70) is a non-degenerate random variable. In the consensus implementation (5.60), the variable $\boldsymbol{w}_i$ is an additive shift of $\boldsymbol{\nu}_i$ conditioned on $\boldsymbol{w}_{i-1}$. Therefore, $\boldsymbol{w}_i$ is also a non-degenerate random variable and it admits the conditional density (5.73).

In diffusion, however, $\boldsymbol{w}_i$ is an affine transformation of $\boldsymbol{\nu}_i$—see (5.57). The combination matrix A need not be invertible and hence, $\boldsymbol{w}_i$ might not admit a density in $\mathbb{R}^K$ in general. In Appendix 5.C, we show that by representing $\boldsymbol{w}_i$ in an r-dimensional subspace, where r is the rank of A, no information is lost and the analysis and conclusions can be adjusted accordingly.

Remark 5.3 (Difference from [80]). *In [80], the probability of error recursions are studied for the* single-agent *case only. Moreover, the recursions are based on belief differences instead of log-belief ratios. In that case, transition kernels are not Gaussian even under Gaussian observation models, as opposed to* (5.73) *and* (5.131).

5.5.2 Asymptotic Convergence

In addition to providing a way for calculating the error probabilities, the density evolution recursion (5.66) also allows us to show that agents exhibit a regular behavior in the limit. In particular, the distributions of the beliefs $\boldsymbol{\mu}_{k,i}$ and log-belief ratios $\boldsymbol{w}_{k,i}$ will converge to the distribution of some time-independent random variables. That is to say, they will converge *in distribution* [40, Chapter 17]. A sequence (over time index i) of random variables $\boldsymbol{x}_i$ converges to a limiting random variable $\boldsymbol{x}$ in distribution if it holds that

$$\lim_{i\to\infty} \mathbb{P}(\boldsymbol{x}_i \in \mathcal{X}) = \mathbb{P}(\boldsymbol{x} \in \mathcal{X}) \tag{5.75}$$

for a set $\mathcal{X}$ of $\boldsymbol{x}$, whose boundary has zero probability under the limiting distribution. We denote this by writing

$$\boldsymbol{x}_i \overset{d}{\rightsquigarrow} \boldsymbol{x}. \tag{5.76}$$

Although beliefs can demonstrate random behavior with fluctuations in the limit, convergence in distribution implies the existence of limiting statistics such as steady-state probability of errors, as shown in the following result.

Theorem 5.2 (**Asymptotic probability of error).** *The diffusion and consensus HMM strategies are asymptotically stable (in the sense of [83]) under binary classification, Gaussian likelihood models* (5.67)*, and non-deterministic transition models (i.e.,* $\mathbb{T}(\theta|\theta') < 1\ \forall \theta, \theta' \in \Theta$*). That is, the density function* f_i *satisfies*

$$\lim_{i\to\infty} \|f_i - f_\infty\|_{\text{TV}} = 0, \tag{5.77}$$

where f_∞ *is a unique stationary density, and* $\|\cdot\|_{\text{TV}}$ *is the total variation norm defined with respect to the product measure, i.e., for any two densities* f *and* g

$$\|f - g\|_{\text{TV}} \triangleq \frac{1}{2} \sum_{\theta\in\Theta} \int \Big| f(\theta, w) - g(\theta, w) \Big| dW. \tag{5.78}$$

This result implies the convergence of the distribution of the log-belief ratios, and as a special case, the agent-specific probability of errors converge as well:

$$\boldsymbol{w}_{k,i} \overset{d}{\rightsquigarrow} \boldsymbol{w}_{k,\infty}, \qquad \lim_{i\to\infty} p_{k,i} = p_{k,\infty}. \tag{5.79}$$

Proof. See Appendix 5.D. ■

In order to establish this result, we employ in Appendix 5.D a known result from ergodic theory [83, Theorem 5.7.4]. As is common in ergodic theory, even though we can affirm that there exist limiting distributions; we do not know exactly what these distributions are. This task might be of formidable complexity in general.

Note that this result is in contrast to fixed-hypothesis social learning, where log-belief ratios do not converge, i.e., $\boldsymbol{w}_{k,i} = \log \dfrac{\boldsymbol{\mu}_{k,i}(1)}{\boldsymbol{\mu}_{k,i}(0)} \to -\infty$ (assuming w.l.o.g. that $\theta = 0$ is the fixed true hypothesis).

Theorem 5.2 states that the agents' probability of error will approach a steady-state value despite the fact that the belief vectors can fluctuate randomly in the limit.

Corollary 5.1 (**Asymptotic beliefs).** *Theorem 5.2 implies that beliefs of the agents converge in distribution. More formally,*

$$\boldsymbol{\mu}_{k,i} \overset{d}{\rightsquigarrow} \boldsymbol{\mu}_{k,\infty}, \tag{5.80}$$

for a time-independent random variable $\boldsymbol{\mu}_{k,\infty}$.

Proof. By definition (5.49),

$$\boldsymbol{w}_{k,i} = \log \frac{1-\boldsymbol{\mu}_{k,i}(0)}{\boldsymbol{\mu}_{k,i}(0)} \Longleftrightarrow \boldsymbol{\mu}_{k,i}(0) = \frac{1}{1+\exp\{\boldsymbol{w}_{k,i}\}}. \tag{5.81}$$

Since this is a continuous and non-degenerate transformation, by Theorem 5.2 and the continuous mapping theorem [84] it holds that $\boldsymbol{\mu}_{k,i} \overset{d}{\rightsquigarrow} \boldsymbol{\mu}_{k,\infty}$, for some time-independent random variable $\boldsymbol{\mu}_{k,\infty}$. This implies that the statistics of the belief distribution also converge. ■

Corollary 5.1 suggests that in general, the beliefs of agents will have random characteristics and will fluctuate in the long-run. This is in contrast to conventional social learning models where beliefs on the true *fixed* hypothesis converge to 1 almost surely (recall Theorems 2.1 and 2.2). In other words, all agents come to an agreement on the truth eventually. In comparison, in the current *dynamic* hypothesis scenario, agents do not even come to an agreement as shown in the next result.

Lemma 5.1 (**Network disagreement**). *In general, the agents' beliefs do not converge to the same random variable in distribution. Namely, for any agent pair (ℓ,k), the limiting variables $\boldsymbol{\mu}_{k,\infty}$ and $\boldsymbol{\mu}_{\ell,\infty}$ are not necessarily identically distributed. Moreover, agents will have different performance in the long run, namely:*

$$p_{k,\infty} \neq p_{\ell,\infty}, \qquad J_{k,\infty} \neq J_{\ell,\infty}. \tag{5.82}$$

Proof. We prove this by a counter-example in Appendix 5.E. ■

Lemma 5.1 implies that rapidly changing states prevent learning the truth with full confidence, as well as eventual network agreement, even under strongly-connected networks where information can flow thoroughly in all directions. Moreover, when the true state of nature is changing, agents can have different and non-vanishing asymptotic error probabilities. In traditional social learning, agents can have different and non-zero error probabilities in *finite-time*. But as time elapses, all probabilities of error vanish, i.e., they all become 0. So, unlike the traditional setting, the dynamic truth model gives rise to an equilibrium of *wise* and *unwise* agents in asymptotics. That is to say, some agents will be more successful in predicting the truth than others in steady-state. The agents' error probabilities will be dependent on their observations' informativeness and their relative location in the network. Indeed, this "wise agent phenomenon" is more in line with what we observe in real-world, as against to eventual agreement of agents on the correct hypothesis that traditional social learning literature concludes. This observation shows the importance of incorporating the changing

behavior of the state of nature into social learning models.

As discussed before, most of the literature on learning over strongly-connected social networks conclude consensus across agents, although there are exceptions. The works [63] and [64] show that when there are *stubborn* agents in the network that never change their opinion, the beliefs in the long-run can fluctuate and vary, as in the current work. Moreover, if agents tend to communicate with other agents that think alike [85, 86]; or if they tend to use their own beliefs as substitutes for others' beliefs in the case of limited communication [61], then opinion clusters can emerge. The current work gives evidence for another reason of disagreement, namely, the rapidly changing truth.

5.6 Numerical Simulations

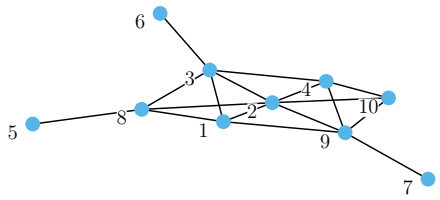


Figure 5.1: The network topology.

We consider the $10-$agent network displayed in Fig. 5.1. The combination weights are given by the Metropolis rule [7, 58], which results in a doubly-stochastic and symmetric matrix with the mixing rate $\rho_2 = 0.86$, satisfying Assumption 5.2.

The agents over the network aim to track the true state of nature from a set of two hypotheses, $\Theta = \{0, 1\}$. For the initial simulations, all agents possess the same family of truncated Gaussian likelihoods, satisfying Assumption 5.3:

$$L_k(\xi|\theta) = \begin{cases} \dfrac{1}{Z_\theta}\dfrac{1}{\sqrt{2\pi}}\exp\left\{-\dfrac{1}{2}(\xi-(1.5\times\theta)^2\right\}, -1 \le \xi \le 2 \\ 0, \qquad\qquad \text{otherwise} \end{cases}$$

for each agent $k \in \mathcal{N}$, where Z_θ is a normalization constant:

$$Z_\theta \triangleq \int_{-1}^{2} \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}(\xi-(1.5\times\theta))^2\right\} d\xi \tag{5.83}$$

The observations are independent conditioned on the true state, satisying Assumption 5.1. The hidden state is changing with respect to the following transition model:

$$\mathbb{T}(\theta_i|\theta_{i-1}) = \begin{cases} 0.9, & \theta_i = \theta_{i-1} \\ 0.1, & \theta_i \neq \theta_{i-1} \end{cases} \tag{5.84}$$

for which the Dobrushin coefficient is given by $\kappa(\mathbb{T}) = 0.8$.

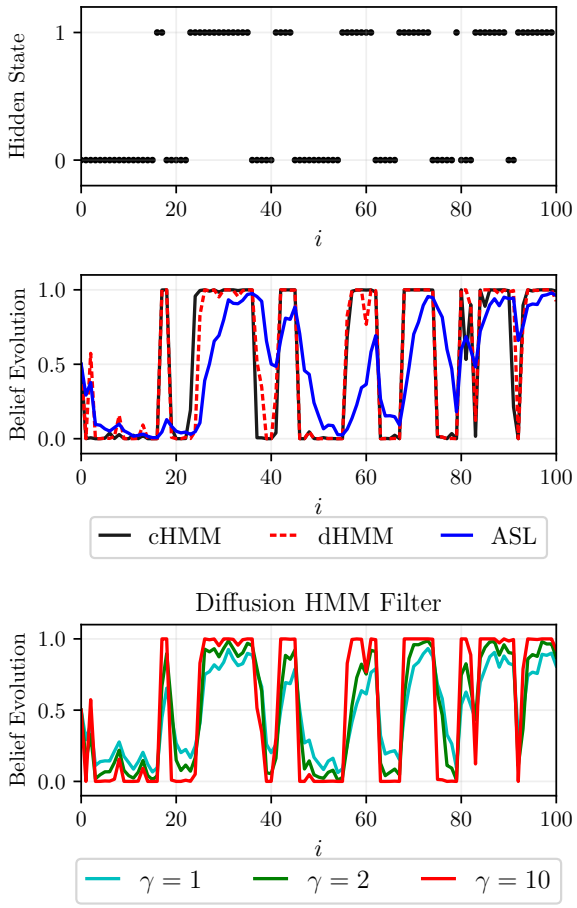


Figure 5.2: *Top panel*: A realization of the true hidden state. *Middle panel*: Belief evolution over time for different algorithms (cHMM, dHMM, and ASL). *Bottom panel*: Belief evolution over time for different γ.

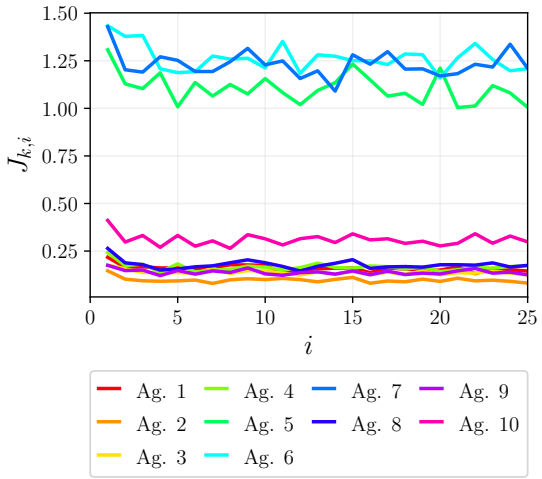


Figure 5.3: Risk functions over time belonging to different agents.

K	dHMM	DBF [87]	ρ_2
10	0.49	2.59	0.86
20	0.53	5.64	0.82
30	0.67	8.63	0.81
40	0.98	11.88	0.80
70	1.23	21.29	0.77

Table 5.1: Number of agents and average asymptotic risks across agents $\frac{1}{K}\sum_{k=1}^{K} J_{k,\infty}$

The top panel of Fig. 5.2 demonstrates a particular realization of hidden states $\boldsymbol{\theta}_i^\circ$. In the middle panel, the belief evolution under this realization is shown for the following algorithms: the proposed diffusion HMM filter (dHMM) with the choice $\gamma = K$, the centralized HMM filter (cHMM), and ASL (Algorithm 2.3) with the choice of $\delta = 0.1$. Notice that dHMM and cHMM behave similarly with a remarkable performance for tracking the abrupt changes in the true state. They are faster in responding to state changes compared to ASL, which does not utilize knowledge of the transition model.

The bottom panel of Fig. 5.2 provides the belief evolution over time for different choices of γ in the diffusion HMM filter. As γ gets closer to $K = 10$, we can see that the tracking capacity of the algorithm increases, approaching the centralized algorithm.

The evolution of different agents' risk functions $J_{k,i}$ over time is provided in Fig. 5.3. Although they all exhibit a regular and bounded behavior as suggested by Theorem 5.1, they are different with respect to different agents. More central agents have less divergence from the optimal centralized solution as expected, whereas marginal agents, such as agents 5, 6 and 7 present higher divergence.

Fig. 5.4a illustrates the network average J_i of asymptotic agent-specific risks $J_{k,i}$ over different network topologies. The sparse networks are associated with higher ρ_2 values, whereas smaller values correspond to dense networks. The risks were approximated by averaging 2000 Monte Carlo simulations with $\gamma = K$. It can be seen that the average risk is increasing with increasing ρ_2. In other words, the average deviation from the centralized solution decreases with increasing network connectivity. This observation is consistent with Theorem 5.1. Specifically, when the network is fully-connected, the risk vanishes as expected since the filter is stable (i.e., corrects wrong initialization) as argued in Section 5.4.

In Table 5.1, we compare the average asymptotic risks of networks with different sizes. From the bottom panel of Fig. 5.2 we know that choosing $\gamma \to K$ boosts performance, so we set $\gamma = K$ for all cases. From Fig. 5.4a we observe that increasing the network connectivity, i.e., $\rho_2 \to 0$, boosts performance. Hence, for a fair comparison, we choose

smaller ρ_2 for larger networks—it is a challenging task to get different-sized graphs with exactly the same ρ_2. Despite this advantage, larger networks have higher risk values, in other words, higher disagreement with the optimal centralized solution, supporting Theorem 5.1. We also provide the average risk values for the DBF strategy from [87]. The risk values are significantly higher compared to the proposed dHMM algorithm. Moreover, dHMM is more scalable in the sense that the growth of the risk values with network size is worse in the DBF case.

The effect of $\kappa(\mathbb{T})$ on the average time-adjusted prior divergence $\widetilde{J}_\infty \triangleq \frac{1}{K}\sum_{k=1}^{K}\widetilde{J}_{k,\infty}$ can be examined from Fig. 5.4b. Remember that $\kappa(\mathbb{T})$ is closer to 0 for rapidly mixing transition models. Theorem 5.1 suggests that the risks should increase with increasing $\kappa(\mathbb{T})$. It is visible that this is the case for $\kappa(\mathbb{T}) \leq 0.8$, and even more, the risk is equal to 0 for $\kappa(\mathbb{T}) = 0$ as revealed by Theorem 5.1. However, when the informativeness of the observations starts to dominate the ergodicity of the transition model, i.e., $\kappa(\mathbb{T}) \to 1$, the setting gets closer to traditional social learning setup and the risk vanishes, which is unfortunately not explainable with the analysis of the present work. Also note that binary symmetric channels (BSCs) with the same $\kappa(\mathbb{T})$ result in the same divergence. For example, $\kappa(\mathbb{T}) = 0.8$ represents both BSC with change probability 0.1 and change probability 0.9.

In Fig. 5.4c, we compare the average risk values of the analyzed diffusion with geometric averaging (GA) to (i) consensus with GA and (ii) diffusion with arithmetic averaging (AA). The age of the utilized information is critical for highly dynamic state transitions. Since diffusion-based strategies use the neighbors' updated information, they outperform the consensus strategy, which can be seen from Fig. 5.4c. Also, diffusion-AA has smaller deviation from the optimal solution compared to the GA-based strategy. However, this observation is not directly transferable to probability of error comparison, as we discuss in the sequel.

For simulations on probability of error, we consider Gaussian likelihoods, as in Section 5.5.1:

$$L_k(\xi|\theta) = \begin{cases} \dfrac{1}{\sqrt{2\pi}}\exp\left\{-\dfrac{(\xi+1)^2}{2}\right\}, & \theta = 0 \\ \dfrac{1}{\sqrt{2\pi}}\exp\left\{-\dfrac{(\xi-1)^2}{2}\right\}, & \theta = 1 \end{cases}$$

The plots for error probability are based on 10000 Monte Carlo simulations. We first see in Fig. 5.5 that the error probabilities of agents rapidly converge, supporting Theorem 5.2. Moreover, more central agents are better, i.e., wise, in tracking the state of nature compared to less central agents.

The network average error probability of diffusion-GA is compared to the (i) centralized and ASL strategies in Fig. 5.6a and to (ii) diffusion-AA and consensus-GA in

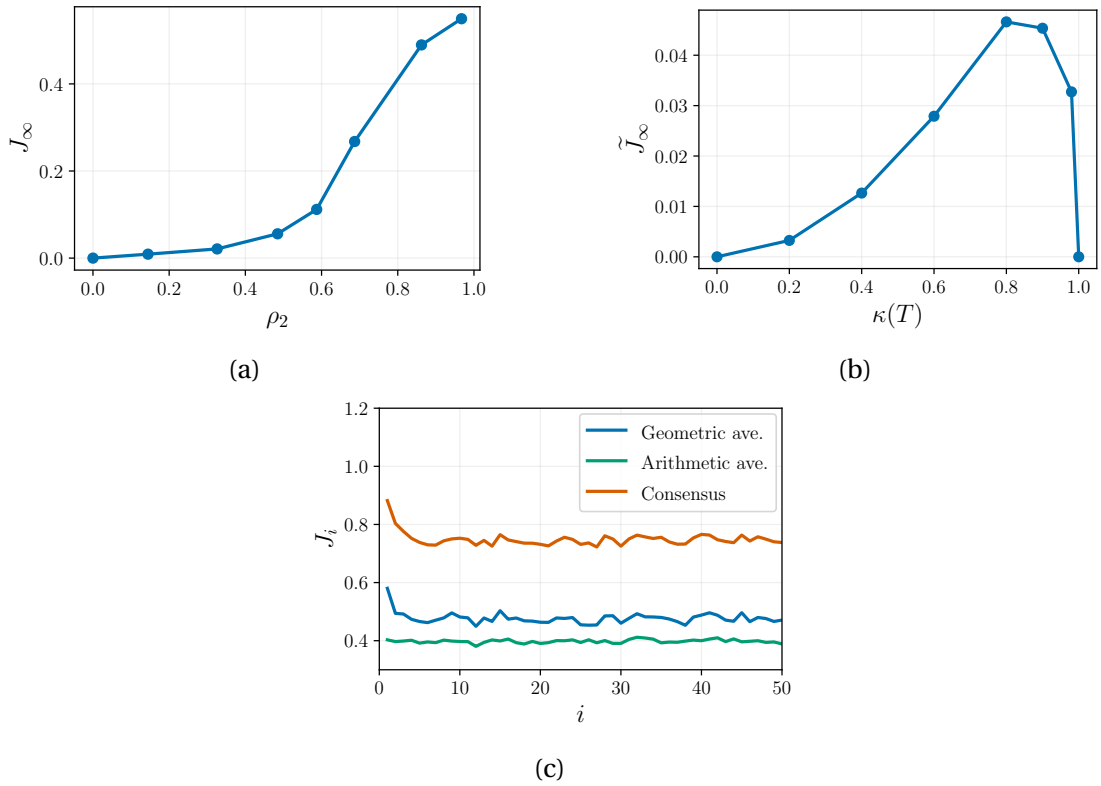


Figure 5.4: (a) Average asymptotic risk function and the mixing rate of the network, (b) Average asymptotic time-adjusted prior risk and the Dobrushin coefficient of the transition model, (c) Average risk over time for Diffusion-GA, Diffusion-AA and Consensus-GA.

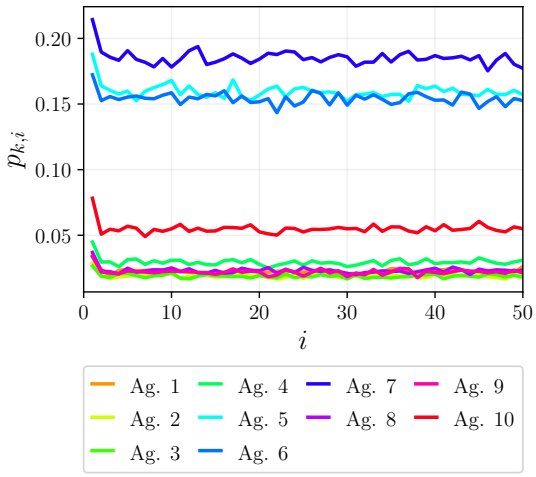


Figure 5.5: Probability of error across different agents over time.

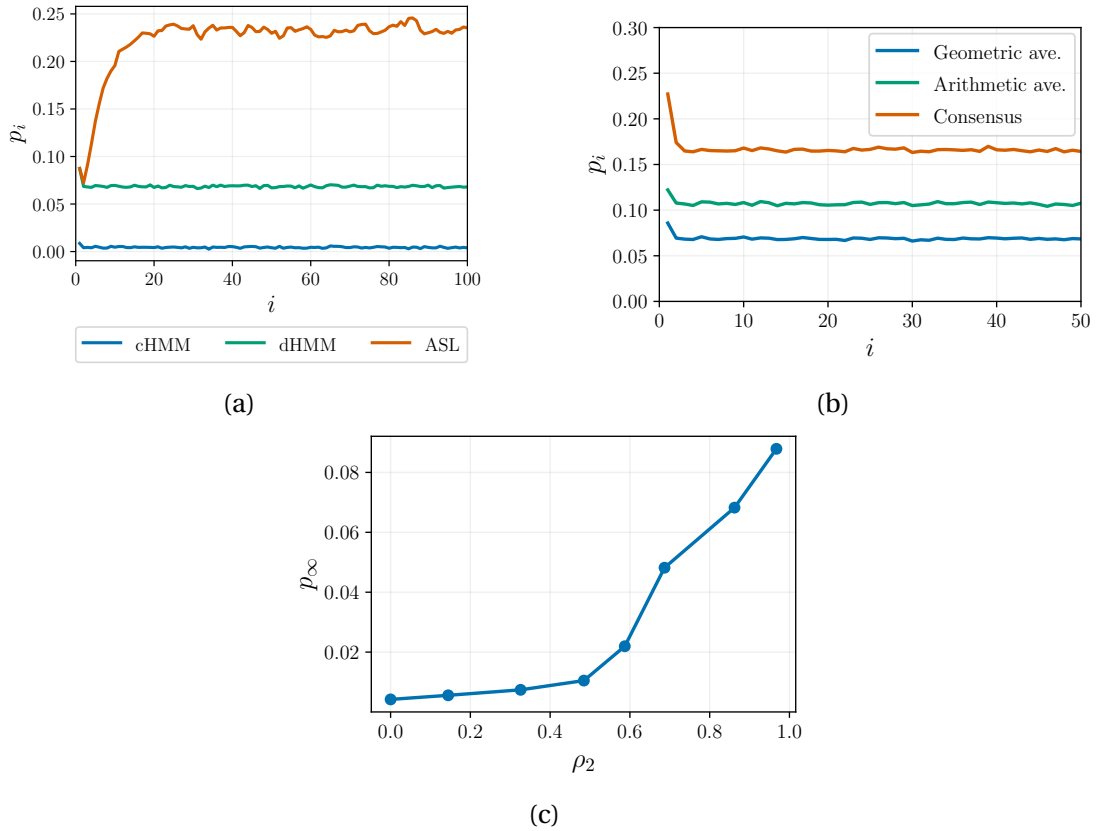


Figure 5.6: Average probability of error over network (a) for cHMM, dHMM and ASL, (b) for diffusion-GA, diffusion-AA, and consensus-GA, over time, (c) in steady-state with respect to different network connectivity.

Fig. 5.6b. It is seen that the diffusion-GA strategy (5.16)–(5.18) outperforms other distributed solutions, but there is still a gap to the centralized solution which can be removed completely only with fully-connected networks. In particular, diffusion-GA has smaller error probability than diffusion-AA, as opposed to the risk function case. A detailed comparison between these two algorithms can be an interesting future work. Finally, Fig. 5.6c shows that the error probability decreases with increasing network connectivity which highlights the benefit of cooperation.

5.7 Concluding Remarks

In this chapter, we proposed a distributed and online state estimation algorithm for finite-state HMMs. Based on ergodicity of the underlying transition model, we provided asymptotic bounds on the disagreement between the distributed strategy and the optimal centralized strategy. We also examined the error probability in steady-state and established convergence in distribution under Gaussian observation models.

In addition to state estimation, the proposed algorithm can be used for the prediction of incoming data, by averaging the data-state likelihood functions with respect to the belief over states. More formally, for well-defined cases, agent k can estimate the incoming data at time i by using

$$\widehat{\boldsymbol{\xi}}_{k,i} \triangleq \arg\max_{\xi} \sum_{\theta_i} L_k(\xi|\theta_i)\boldsymbol{\eta}_{k,i}(\theta_i). \tag{5.85}$$

Furthermore, the algorithm can also be used for continuous state estimation. This would require changing the summations over the states to integrals. These can be numerically tractable under some conditions. For instance, the exponential family of observation models can lead to closed-form formulas for the integral expressions, as in [69].

Finally, the theoretical analysis in the current work utilizes the ergodicity of the transition model to establish performance bounds. A future challenge is to incorporate the informativeness of the observations as well.

5.A Proof of Theorem 5.1

The risk function can be written as

$$\begin{aligned}
J_{k,i} &= \mathbb{E}_{\mathcal{F}_i} D_{\mathrm{KL}}(\boldsymbol{\mu}_i^\star || \boldsymbol{\mu}_{k,i}) \\
&= \mathbb{E}_{\mathcal{F}_i}\Big[\sum_{\theta_i\in\Theta} \boldsymbol{\mu}_i^\star(\theta_i)\log\frac{\boldsymbol{\mu}_i^\star(\theta_i)}{\boldsymbol{\mu}_{k,i}(\theta_i)}\Big] \\
&\overset{(a)}{=} \mathbb{E}_{\mathcal{F}_i}\Big[\sum_{\theta_i\in\Theta} \mathbb{P}(\boldsymbol{\theta}_i^\circ=\theta_i|\boldsymbol{\mathcal{F}}_i)\log\frac{\boldsymbol{\mu}_i^\star(\theta_i)}{\boldsymbol{\mu}_{k,i}(\theta_i)}\Big] \\
&\overset{(b)}{=} \mathbb{E}_{\mathcal{F}_i}\Big[\mathbb{E}_{\theta_i^\circ|\mathcal{F}_i}\Big(\log\frac{\boldsymbol{\mu}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\mu}_{k,i}(\boldsymbol{\theta}_i^\circ)}\Big)\Big] \\
&= \mathbb{E}_{\mathcal{F}_i,\theta_i^\circ}\Big[\log\frac{\boldsymbol{\mu}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\mu}_{k,i}(\boldsymbol{\theta}_i^\circ)}\Big] \\
&\overset{(c)}{=} \mathbb{E}_{\mathcal{F}_i,\theta_i^\circ}\Big[\log\boldsymbol{\mu}_i^\star(\boldsymbol{\theta}_i^\circ) - \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\log\boldsymbol{\psi}_{\ell,i}(\boldsymbol{\theta}_i^\circ)\Big] + \mathbb{E}_{\mathcal{F}_i}\Big[\log\sum_{\theta_i'\in\Theta}\exp\Big\{\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\log\boldsymbol{\psi}_{\ell,i}(\theta_i')\Big\}\Big] \\
&= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_i,\theta_i^\circ}\Big[\log\frac{\boldsymbol{\mu}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\psi}_{\ell,i}(\boldsymbol{\theta}_i^\circ)}\Big]\mathbb{E}_{\mathcal{F}_i}\Big[\log\sum_{\theta_i'\in\Theta}\exp\Big\{\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\log\boldsymbol{\psi}_{\ell,i}(\theta_i')\Big\}\Big]
\end{aligned} \tag{5.86}$$

where (a) follows from definition (5.4), (b) follows from the definition of conditional expectation with respect to θ_i° given $\mathcal{F}_i$, and (c) follows from the combination step (5.18). From the centralized update (5.5) and the adaptation step (5.17), we have:

$$\log\frac{\boldsymbol{\mu}_i^\star(\theta_i)}{\boldsymbol{\psi}_{\ell,i}(\theta_i)} = \log\frac{L(\boldsymbol{\xi}_i|\theta_i)}{(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i))^\gamma} + \log\frac{\boldsymbol{\eta}_i^\star(\theta_i)}{\boldsymbol{\eta}_{\ell,i}(\theta_i)} - \log\frac{\boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)}{\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i})}. \tag{5.87}$$

where we are introducing the following marginal distribution for the new data given the past data:

$$\begin{aligned}\boldsymbol{m}_i^\star(\xi_i) \triangleq \mathbb{P}(\boldsymbol{\xi}_i = \xi_i|\boldsymbol{\mathcal{F}}_{i-1}) &= \sum_{\theta_i'\in\Theta} \mathbb{P}(\boldsymbol{\xi}_i = \xi_i, \boldsymbol{\theta}_i^\circ = \theta_i'|\boldsymbol{\mathcal{F}}_{i-1}) \\ &= \sum_{\theta_i'\in\Theta} L(\xi_i|\theta_i')\mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta_i'|\boldsymbol{\mathcal{F}}_{i-1}) \\ &= \sum_{\theta_i'\in\Theta} L(\xi_i|\theta_i')\boldsymbol{\eta}_i^\star(\theta_i'), \end{aligned} \tag{5.88}$$

as well as the agent-specific pseudo-marginal distribution:

$$\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i}) \triangleq \sum_{\theta_i'\in\Theta} (L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))^\gamma \boldsymbol{\eta}_{\ell,i}(\theta_i'). \tag{5.89}$$

Note that $\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i})$ is not a real distribution, i.e., it is not summing up to 1 because $\gamma \neq 1$, in general. To rewrite (5.86) using (5.87), we first observe that

$$\begin{aligned}&\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\xi_i,\theta_i^\circ}\left[\log \frac{L(\boldsymbol{\xi}_i|\boldsymbol{\theta}_i^\circ)}{(L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ))^\gamma}\right] \\ &\qquad \overset{(a)}{=} \mathbb{E}_{\xi_i,\theta_i^\circ}\left[\sum_{\ell=1}^K \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\right] - \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\xi_{\ell,i},\theta_i^\circ}\left[\gamma \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\right] \\ &\qquad = \mathbb{E}_{\xi_i,\theta_i^\circ}\left[\sum_{\ell=1}^K (1-\gamma a_{\ell k}) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\right] \end{aligned} \tag{5.90}$$

where in (a) we used the independence from Assumption 5.1. Moreover, the divergence of time-adjusted priors can be bounded as:

$$\begin{aligned}\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_i,\theta_i^\circ}\left[\log \frac{\boldsymbol{\eta}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{\theta}_i^\circ)}\right] &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1},\theta_i^\circ}\left[\mathbb{E}_{\xi_i|\mathcal{F}_{i-1},\theta_i^\circ}\left(\log \frac{\boldsymbol{\eta}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{\theta}_i^\circ)}\right)\right] \\ &\overset{(a)}{=} \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1},\theta_i^\circ}\left[\log \frac{\boldsymbol{\eta}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{\theta}_i^\circ)}\right] \\ &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}}\left[\mathbb{E}_{\theta_i^\circ|\mathcal{F}_{i-1}}\left(\log \frac{\boldsymbol{\eta}_i^\star(\boldsymbol{\theta}_i^\circ)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{\theta}_i^\circ)}\right)\right] \\ &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}}\left[\sum_{\theta_i\in\Theta} \mathbb{P}(\boldsymbol{\theta}_i^\circ = \theta_i|\boldsymbol{\mathcal{F}}_{i-1}) \log \frac{\boldsymbol{\eta}_i^\star(\theta_i)}{\boldsymbol{\eta}_{\ell,i}(\theta_i)}\right] \\ &\overset{(b)}{=} \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}}\left[\sum_{\theta_i\in\Theta} \boldsymbol{\eta}_i^\star(\theta_i) \log \frac{\boldsymbol{\eta}_i^\star(\theta_i)}{\boldsymbol{\eta}_{\ell,i}(\theta_i)}\right] \\ &= \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}}\left[D_{\mathrm{KL}}(\boldsymbol{\eta}_i^\star||\boldsymbol{\eta}_{\ell,i})\right] \end{aligned}$$

$$\overset{(c)}{\leq} \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \kappa(\mathbb{T}) \mathbb{E}_{\mathcal{F}_{i-1}} \Big[D_{\mathrm{KL}}(\boldsymbol{\mu}^{\star}_{i-1} || \boldsymbol{\mu}_{\ell,i-1}) \Big] \tag{5.91}$$

where (a) follows from the fact that the time-adjusted priors evaluated at the true hypothesis are deterministic given the old history and the true hypothesis, (b) follows from definition (5.6), and (c) follows from the strong data processing inequality.

Combining (5.86), (5.87), (5.90), and (5.91) yields:

$$\begin{aligned} J_{k,i} \leq \mathbb{E}_{\xi_i, \theta_i^{\circ}} \Big[\sum_{\ell=1}^{K} (1 - \gamma a_{\ell k}) \log L_{\ell}(\boldsymbol{\xi}_{\ell,i} | \boldsymbol{\theta}_i^{\circ}) \Big] + \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \kappa(\mathbb{T}) \mathbb{E}_{\mathcal{F}_{i-1}} \Big[D_{\mathrm{KL}}(\boldsymbol{\mu}^{\star}_{i-1} || \boldsymbol{\mu}_{\ell,i-1}) \Big] \\ + \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp \Big\{ \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \boldsymbol{\psi}_{\ell,i}(\theta_i') \Big\} \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{m}^{\star}_i(\boldsymbol{\xi}_i)}{\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i})} \Big]. \end{aligned} \tag{5.92}$$

Furthermore, the normalization term satisfies:

$$\begin{aligned} &\mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp \Big\{ \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \boldsymbol{\psi}_{\ell,i}(\theta_i') \Big\} \Big] \\ &\quad = \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \prod_{\ell=1}^{K} \exp \Big\{ a_{\ell k} \log \boldsymbol{\psi}_{\ell,i}(\theta_i') \Big\} \Big] \\ &\quad = \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \prod_{\ell=1}^{K} (\boldsymbol{\psi}_{\ell,i}(\theta_i'))^{a_{\ell k}} \Big] \\ &\quad \overset{(a)}{=} \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \Big(\prod_{\ell=1}^{K} (L_{\ell}(\boldsymbol{\xi}_{\ell,i} | \theta_i'))^{\gamma a_{\ell k}} \prod_{\ell=1}^{K} (\boldsymbol{\eta}_{\ell,i}(\theta_i'))^{a_{\ell k}} \Big) \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i}) \Big] \end{aligned} \tag{5.93}$$

where (a) follows from (5.17) and (5.89). Therefore, the last two terms in (5.92) can be bounded as:

$$\begin{aligned} &\mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp \Big\{ \sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \boldsymbol{\psi}_{\ell,i}(\theta_i') \Big\} \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{m}^{\star}_i(\boldsymbol{\xi}_i)}{\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i})} \Big] \\ &\quad \overset{(a)}{=} \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \Big(\prod_{\ell=1}^{K} (L_{\ell}(\boldsymbol{\xi}_{\ell,i} | \theta_i'))^{\gamma a_{\ell k}} \prod_{\ell=1}^{K} (\boldsymbol{\eta}_{\ell,i}(\theta_i'))^{a_{\ell k}} \Big) \Big] \\ &\qquad - \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i}) \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{\ell \in \mathcal{N}_k} a_{\ell k} \log \frac{\boldsymbol{m}^{\star}_i(\boldsymbol{\xi}_i)}{\boldsymbol{m}_{\ell,i}(\boldsymbol{\xi}_{\ell,i})} \Big] \\ &\quad = \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \Big(\prod_{\ell=1}^{K} (L_{\ell}(\boldsymbol{\xi}_{\ell,i} | \theta_i'))^{\gamma a_{\ell k}} \prod_{\ell=1}^{K} (\boldsymbol{\eta}_{\ell,i}(\theta_i'))^{a_{\ell k}} \Big) \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\log \boldsymbol{m}^{\star}_i(\boldsymbol{\xi}_i) \Big] \end{aligned}$$

$$
\begin{aligned}
&\overset{(b)}{\leq} \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))^{\gamma a_{\ell k}}\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] - \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)\Bigg] \\
&= \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))^{\gamma a_{\ell k}}\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] \\
&\quad - \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i')\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] \\
&\quad + \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i')\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] - \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)\Bigg] \\
&\overset{(c)}{\leq} \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))^{\gamma a_{\ell k}}\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] \\
&\qquad - \mathbb{E}_{\mathcal{F}_i}\Bigg[\log \sum_{\theta_i'\in\Theta}\Bigg(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i')\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Bigg)\Bigg] \qquad (5.94)
\end{aligned}
$$

where (a) follows from inserting (5.93), (b) follows from the weighted arithmetic-geometric mean inequality, (c) follows from the fact that:

$$
\begin{aligned}
-\mathbb{E}_{\mathcal{F}_i}\Bigg[\log \frac{\boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)}{\sum_{\theta_i'\in\Theta}\Big[\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Big]}\Bigg] &= -\mathbb{E}_{\mathcal{F}_{i-1}}\mathbb{E}_{\xi_i|\mathcal{F}_{i-1}}\Bigg[\log \frac{\boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)}{\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i)}\Bigg] \\
&= -\mathbb{E}_{\mathcal{F}_{i-1}}D_{\mathrm{KL}}(\boldsymbol{m}_i^\star(\boldsymbol{\xi}_i)||\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i)) \\
&\leq 0 \qquad (5.95)
\end{aligned}
$$

where we defined the probability density function:

$$
\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i) \triangleq \sum_{\theta_i'\in\Theta}\Big[\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Big] \qquad (5.96)
$$

which can be verified to be a density as follows:

$$
\begin{aligned}
\int_{\xi_i}\boldsymbol{m}_i^\dagger(\xi_i)d\xi_i &= \int_{\xi_i}\sum_{\theta_i'\in\Theta}\Big[\prod_{\ell=1}^{K}(L_\ell(\xi_{\ell,i}|\theta_i'))\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Big]d\xi_i \\
&= \sum_{\theta_i'\in\Theta}\Big[\underbrace{\int_{\xi_i}\prod_{\ell=1}^{K}(L_\ell(\xi_{\ell,i}|\theta_i'))d\xi_i}_{1}\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Big] \\
&= \sum_{\theta_i'\in\Theta}\Big[\sum_{\ell=1}^{K}a_{\ell k}\boldsymbol{\eta}_{\ell,i}(\theta_i')\Big]
\end{aligned}
$$

$$= \sum_{\ell=1}^{K} a_{\ell k} \Big[\underbrace{\sum_{\theta_i' \in \Theta} \boldsymbol{\eta}_{\ell,i}(\theta_i')}_{=1} \Big]$$

$$= 1 \tag{5.97}$$

Let us introduce the following vectors of dimension H over all hypotheses for notational convenience:

$$\boldsymbol{\vartheta}_i^{+} \triangleq \mathrm{col}\Big\{ \log\Big(\prod_{\ell=1}^{K} (L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)^{\gamma a_{\ell k}} \sum_{\ell=1}^{K} a_{\ell k} \boldsymbol{\eta}_{\ell,i}(\theta_i) \Big) \Big\}_{\theta_i=0}^{H-1} \tag{5.98}$$

and

$$\boldsymbol{\vartheta}_i^{-} \triangleq \mathrm{col}\Big\{ \log\Big(\prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i) \sum_{\ell=1}^{K} a_{\ell k} \boldsymbol{\eta}_{\ell,i}(\theta_i) \Big) \Big\}_{\theta_i=0}^{H-1}. \tag{5.99}$$

Then, the bound in (5.94) can be expressed as

$$\mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \Big[\prod_{\ell=1}^{K} (L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i'))^{\gamma a_{\ell k}} \sum_{\ell=1}^{K} a_{\ell k} \boldsymbol{\eta}_{\ell,i}(\theta_i') \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \Big[\prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i') \sum_{\ell=1}^{K} a_{\ell k} \boldsymbol{\eta}_{\ell,i}(\theta_i') \Big]$$

$$= \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp\{\boldsymbol{\vartheta}_i^{+}(\theta_i')\} \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp\{\boldsymbol{\vartheta}_i^{-}(\theta_i')\} \Big] \tag{5.100}$$

Note that the difference of the vectors satisfy

$$\boldsymbol{\vartheta}_i^{+} - \boldsymbol{\vartheta}_i^{-} = \mathrm{col}\Big\{ \sum_{\ell=1}^{K} (\gamma a_{\ell k} - 1) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i) \Big\}_{\theta_i=0}^{H-1}. \tag{5.101}$$

It is useful to introduce the LogSumExp function f:

$$f(\upsilon) \triangleq \log \sum_{\theta \in \Theta} \exp\{\upsilon(\theta)\}, \tag{5.102}$$

whose gradient is given by

$$\nabla_\upsilon f(\upsilon) \triangleq \mathrm{col}\Big\{ \frac{\partial f(\upsilon)}{\partial \upsilon(\theta)} \Big\}_{\theta \in \Theta} = \mathrm{col}\Big\{ \frac{\exp\{\upsilon(\theta)\}}{\sum_{\theta'} \exp\{\upsilon(\theta')\}} \Big\}_{\theta \in \Theta}. \tag{5.103}$$

By applying the mean-value theorem (MVT) to function f and taking the expectation we get

$$\mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp\{\boldsymbol{\vartheta}_i^{+}(\theta_i')\} \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[\log \sum_{\theta_i' \in \Theta} \exp\{\boldsymbol{\vartheta}_i^{-}(\theta_i')\} \Big]$$

$$
\begin{aligned}
&\overset{(5.102)}{=} \mathbb{E}_{\mathcal{F}_i}\Big[f(\boldsymbol{\vartheta}_i^+)\Big] - \mathbb{E}_{\mathcal{F}_i}\Big[f(\boldsymbol{\vartheta}_i^-)\Big] \\
&\overset{(\text{MVT})}{=} \mathbb{E}_{\mathcal{F}_i}\Big[(\nabla_v f(\boldsymbol{\vartheta}_i))^{\mathsf{T}} \cdot (\boldsymbol{\vartheta}_i^+ - \boldsymbol{\vartheta}_i^-)\Big] \\
&\overset{(5.101),(5.103)}{=} \mathbb{E}_{\mathcal{F}_i}\left[\operatorname{col}\Big\{\frac{\exp\{\boldsymbol{\vartheta}_i(\theta_i)\}}{\sum_{\theta_i'}\exp\{\boldsymbol{\vartheta}_i(\theta_i')\}}\Big\}^{\mathsf{T}} \cdot \operatorname{col}\Big\{\sum_{\ell=1}^{K}(\gamma a_{\ell k}-1)\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)\Big\}_{\theta_i=0}^{H-1}\right]
\end{aligned}
\tag{5.104}
$$

for some $\boldsymbol{\vartheta}_i$ lying on the line segment between $\boldsymbol{\vartheta}_i^-$ and $\boldsymbol{\vartheta}_i^+$. The absolute value of (5.104) can be bounded for any time instant i:

$$
\begin{aligned}
&\left|\mathbb{E}_{\mathcal{F}_i}\left[\operatorname{col}\Big\{\frac{\exp\{\boldsymbol{\vartheta}_i(\theta_i)\}}{\sum_{\theta_i'}\exp\{\boldsymbol{\vartheta}_i(\theta_i')\}}\Big\}^{\mathsf{T}} \cdot \operatorname{col}\Big\{\sum_{\ell=1}^{K}(\gamma a_{\ell k}-1)\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)\Big\}_{\theta_i=0}^{H-1}\right]\right| \\
&\quad\overset{(a)}{\leq} \mathbb{E}_{\mathcal{F}_i}\left|\operatorname{col}\Big\{\frac{\exp\{\boldsymbol{\vartheta}_i(\theta_i)\}}{\sum_{\theta_i'}\exp\{\boldsymbol{\vartheta}_i(\theta_i')\}}\Big\}^{\mathsf{T}} \cdot \operatorname{col}\Big\{\sum_{\ell=1}^{K}(\gamma a_{\ell k}-1)\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)\Big\}_{\theta_i=0}^{H-1}\right| \\
&\quad\overset{(b)}{\leq} \mathbb{E}_{\xi_i}\left\|\operatorname{col}\Big\{\sum_{\ell=1}^{K}(\gamma a_{\ell k}-1)\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)\Big\}_{\theta_i=0}^{H-1}\right\|_\infty \\
&\quad\overset{(c)}{\leq} \sqrt{K}\gamma\lambda\mathbb{E}_{\xi_i}\left\|\operatorname{col}\Big\{\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\theta_i)\Big\}_{\theta_i=0}^{H-1}\right\|_\infty \\
&\quad\overset{(d)}{\leq} \sqrt{K}\gamma\lambda C_L
\end{aligned}
\tag{5.105}
$$

where (a) follows from Jensen's inequality, (b) follows from Hölder's inequality and the fact that

$$
\left\|\operatorname{col}\Big\{\frac{\exp\{\boldsymbol{\vartheta}_i(\theta_i)\}}{\sum_{\theta_i'}\exp\{\boldsymbol{\vartheta}_i(\theta_i')\}}\Big\}\right\|_1 = 1, \tag{5.106}
$$

and the last step (d) follows from Assumption 5.3. Step (c) follows from

$$
\begin{aligned}
\sum_{\ell=1}^{K}|\gamma a_{\ell k}-1| &\leq \gamma\left\|A-\frac{1}{\gamma}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}}\right\|_1 \\
&\leq \gamma\sqrt{K}\left\|A-\frac{1}{\gamma}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}}\right\|_2 \\
&\overset{(e)}{=} \gamma\sqrt{K}\lambda.
\end{aligned}
\tag{5.107}
$$

Step (e) follows from the fact that for symmetric matrices, their ℓ_2-induced norm is equal to the spectral radius. Here, we also use the fact that since A is primitive and doubly-stochastic, it has a unique eigenvalue at 1, and all other eigenvalues lie inside the unit circle. The spectral radius of $A-\frac{1}{\gamma}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}}$ becomes the maximum absolute

difference between the eigenvalues of A and $\frac{1}{\gamma}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}}$, i.e., λ. Next, we combine (5.92), (5.94), and (5.105) to obtain the bound on the risk function:

$$J_{k,i} \leq \sum_{\ell=1}^{K}(1-\gamma a_{\ell k})\mathbb{E}_{\xi_{\ell,i},\theta_i^\circ}\Big[\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\Big] + \kappa(\mathbb{T})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\underbrace{\mathbb{E}_{\mathcal{F}_{i-1}}\Big[D_{\mathrm{KL}}(\boldsymbol{\mu}_{i-1}^\star||\boldsymbol{\mu}_{\ell,i-1})\Big]}_{J_{\ell,i-1}} + \sqrt{K}\gamma\lambda C_L. \tag{5.108}$$

Iterating this bound over time for all agents results in

$$\begin{aligned}
J_{k,i} &\leq \sum_{\ell=1}^{K}(1-\gamma a_{\ell k})\mathbb{E}_{\xi_{\ell,i},\theta_i^\circ}\Big[\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\Big] \\
&\quad + \kappa(\mathbb{T})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\sum_{m\in\mathcal{N}_\ell}\Big((1-\gamma a_{m\ell})\times\mathbb{E}_{\xi_{m,i-1},\theta_{i-1}^\circ}\Big[\log L_m(\boldsymbol{\xi}_{m,i-1}|\boldsymbol{\theta}_{i-1}^\circ)\Big]\Big) \\
&\quad + \kappa(\mathbb{T})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\kappa(\mathbb{T})\sum_{m\in\mathcal{N}_\ell} a_{m\ell}J_{m,i-2} + \kappa(\mathbb{T})\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\sqrt{K}\gamma\lambda C_L + \sqrt{K}\gamma\lambda C_L \\
&\overset{(a)}{\leq} \sum_{\ell=1}^{K}(1-\gamma a_{\ell k})\mathbb{E}_{\xi_{\ell,i},\theta_i^\circ}\Big[\log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{\theta}_i^\circ)\Big] \\
&\quad + \kappa(\mathbb{T})\sum_{m=1}^{K}\Big((1-\gamma[A^2]_{mk})\times\mathbb{E}_{\xi_{m,i-1},\theta_{i-1}^\circ}\Big[\log L_m(\boldsymbol{\xi}_{m,i-1}|\boldsymbol{\theta}_{i-1}^\circ)\Big]\Big) \\
&\quad + \kappa(\mathbb{T})^2\sum_{m=1}^{K}[A^2]_{mk}J_{m,i-2} + (1+\kappa(\mathbb{T}))\sqrt{K}\gamma\lambda C_L \\
&\leq \sum_{j=0}^{i-1}(\kappa(\mathbb{T}))^j\sum_{\ell=1}^{K}(1-\gamma[A^{j+1}]_{\ell k})\times\mathbb{E}_{\xi_{\ell,i-j},\theta_{i-j}^\circ}\Big[\log L_\ell(\boldsymbol{\xi}_{\ell,i-j}|\boldsymbol{\theta}_{i-j}^\circ)\Big] \\
&\quad + \sum_{j=0}^{i-1}(\kappa(\mathbb{T}))^j\sqrt{K}\gamma\lambda C_L + (\kappa(\mathbb{T}))^i\sum_{\ell=1}^{K}[A^i]_{\ell k}J_{\ell,0},
\end{aligned} \tag{5.109}$$

where (a) follows from the fact that

$$\begin{aligned}
&\sum_{\ell\in\mathcal{N}_k} a_{\ell k}\sum_{m\in\mathcal{N}_\ell}\Big((1-\gamma a_{m\ell})\mathbb{E}_{\xi_{m,i-1},\theta_{i-1}^\circ}\Big[\log L_m(\boldsymbol{\xi}_{m,i-1}|\boldsymbol{\theta}_{i-1}^\circ)\Big]\Big) \\
&\qquad = \sum_{m=1}^{K}\mathbb{E}_{\xi_{m,i-1},\theta_{i-1}^\circ}\Big[\log L_m(\boldsymbol{\xi}_{m,i-1}|\boldsymbol{\theta}_{i-1}^\circ)\Big]\sum_{\ell=1}^{K}a_{\ell k}(1-\gamma a_{m\ell}) \\
&\qquad = \sum_{m=1}^{K}(1-\gamma[A^2]_{mk})\mathbb{E}_{\xi_{m,i-1},\theta_{i-1}^\circ}\Big[\log L_m(\boldsymbol{\xi}_{m,i-1}|\boldsymbol{\theta}_{i-1}^\circ)\Big].
\end{aligned} \tag{5.110}$$

The first summation in the bound (5.109) can be bounded by the inequality

$$\begin{aligned}\left|\sum_{\ell=1}^{K}(1-\gamma[A^{j+1}]_{\ell k})\mathbb{E}_{\xi_{\ell,i-j},\theta^{\circ}_{i-j}}\left[\log L_{\ell}(\boldsymbol{\xi}_{\ell,i-j}|\boldsymbol{\theta}^{\circ}_{i-j})\right]\right| & \\ \overset{(a)}{\leq}\sum_{\ell=1}^{K}\left|1-\gamma[A^{j+1}]_{\ell k}\right|\times\left|\mathbb{E}_{\xi_{\ell,i-j},\theta^{\circ}_{i-j}}\left[\log L_{\ell}(\boldsymbol{\xi}_{\ell,i-j}|\boldsymbol{\theta}^{\circ}_{i-j})\right]\right| & \\ \overset{(b)}{\leq}\sum_{\ell=1}^{K}\left|1-\gamma[A^{j+1}]_{\ell k}\right|C_L & \\ \overset{(c)}{\leq}\sqrt{K}\gamma\lambda_j C_L & \qquad (5.111)\end{aligned}$$

where $\lambda_j > \max\{|1-\frac{K}{\gamma}|, \rho_2^{j+1}\}$ is a positive constant, (a) follows from the triangle inequality, (b) follows from Assumption 5.3, and (c) follows from (5.107) applied to A^{j+1} instead of A. Inserting (5.111) into (5.109) we can bound the risk function as:

$$\begin{aligned}J_{k,i} &\leq \sum_{j=0}^{i-1}(\kappa(\mathbb{T}))^j\sqrt{K}\gamma\lambda_j C_L + \sum_{j=0}^{i-1}(\kappa(\mathbb{T}))^j\sqrt{K}\gamma\lambda C_L + (\kappa(\mathbb{T}))^i\sum_{\ell=1}^{K}[A^i]_{\ell k}J_{\ell,0} \\ &\overset{(a)}{\leq} 2\sum_{j=0}^{i-1}(\kappa(\mathbb{T}))^j\sqrt{K}\gamma\lambda C_L + (\kappa(\mathbb{T}))^i\sum_{\ell=1}^{K}[A^i]_{\ell k}J_{\ell,0} \\ &= 2\frac{1-(\kappa(\mathbb{T}))^i}{1-\kappa(\mathbb{T})}\sqrt{K}\gamma\lambda C_L + (\kappa(\mathbb{T}))^i\sum_{\ell=1}^{K}[A^i]_{\ell k}J_{\ell,0} \qquad (5.112)\end{aligned}$$

where (a) follows from $\lambda_j \leq \lambda$ for all j. Using $\kappa(\mathbb{T}) < 1$, the risk function is asymptotically bounded by

$$\limsup_{i\to\infty} J_{k,i} \leq \frac{2\sqrt{K}\gamma\lambda C_L}{1-\kappa(\mathbb{T})}. \qquad (5.113)$$

This also means that

$$\begin{aligned}\widetilde{J}_{k,i} &\overset{(a)}{\leq} \kappa(\mathbb{T})J_{k,i-1} \\ \Longrightarrow \limsup_{i\to\infty}\widetilde{J}_{k,i} &\leq \limsup_{i\to\infty}\kappa(\mathbb{T})J_{k,i-1} \leq \kappa(\mathbb{T})\frac{2\sqrt{K}\gamma\lambda C_L}{1-\kappa(\mathbb{T})} \qquad (5.114)\end{aligned}$$

where (a) follows from the strong-data processing inequality, for any time instant i.

5.B An Auxiliary Lemma

We present a general matrix result in the following lemma.

Lemma 5.2 (Lower dimensional representation). *Consider the $K \times K$ doubly stochastic and symmetric combination matrix A. Let $r = \mathrm{rank}(A)$. Then, for any positive-definite diagonal covariance matrix Σ, there exists an $r \times K$ matrix Q such that:*

- $A^{\mathsf{T}}\Sigma A = Q^{\mathsf{T}}Q$.
- *For any vector $v \in \mathbb{R}^K$, there exists a unique vector $v_Q \in \mathbb{R}^r$ that satisfies:*

$$A^{\mathsf{T}}v = Q^{\mathsf{T}}v_Q. \tag{5.115}$$

In other words, Q has full row rank and

$$v_Q = (Q^{\mathsf{T}})^{\dagger}A^{\mathsf{T}}v, \tag{5.116}$$

where $(Q^{\mathsf{T}})^{\dagger}$ is the pseudo-inverse matrix

$$(Q^{\mathsf{T}})^{\dagger} \triangleq (QQ^{\mathsf{T}})^{-1}Q. \tag{5.117}$$

Proof. Observe that

$$\begin{aligned}\mathrm{rank}(A) &\overset{(a)}{=} \mathrm{rank}(\Sigma^{1/2}A)\\ &= \mathrm{rank}\left((\Sigma^{1/2}A)^{\mathsf{T}}\Sigma^{1/2}A\right)\\ &= \mathrm{rank}(A^{\mathsf{T}}\Sigma A) = r\end{aligned} \tag{5.118}$$

where (a) follows from the fact that Σ is positive-definite and $\Sigma^{1/2}$ is its square-root. Moreover, since $A^{\mathsf{T}}\Sigma A$ is a real symmetric matrix, it can be decomposed as

$$A^{\mathsf{T}}\Sigma A = U\Lambda U^{\mathsf{T}} \tag{5.119}$$

where U is $K \times r$ with orthonormal columns and Λ is $r \times r$ with the positive eigenvalues of $A^{\mathsf{T}}\Sigma A$. Let $Q = \Lambda^{1/2}U^{\mathsf{T}}$, which has full row rank. Then, $A^{\mathsf{T}}\Sigma A = Q^{\mathsf{T}}Q$. Note that Q is not unique since we can modify it by any orthonormal transformation. It is also obvious that, in terms of null (NULL) and range (RAN) spaces,

$$\begin{aligned}\mathrm{NULL}(A) &= \mathrm{NULL}(\Sigma^{1/2}A)\\ &= \mathrm{NULL}(A^{\mathsf{T}}\Sigma A)\\ &= \mathrm{NULL}(Q^{\mathsf{T}}Q)\\ &= \mathrm{NULL}(Q)\end{aligned} \tag{5.120}$$

and, hence, $\mathrm{RAN}(A^{\mathsf{T}}) = \mathrm{RAN}(Q^{\mathsf{T}})$. It follows that for any vector $v \in \mathbb{R}^K$, there exists a

vector $v_Q \in \mathbb{R}^r$ such that

$$A^{\mathsf{T}} v = Q^{\mathsf{T}} v_Q. \tag{5.121}$$

Multiplying both sides of (5.121) from the left by the pseudo-inverse of Q^{T} gives

$$v_Q = (Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} v. \tag{5.122}$$

■

5.C Error Recursion for Diffusion

In light of Lemma 5.2 from Appendix 5.B, there exist vectors in $\mathbb{R}^r$ such that:

$$\begin{aligned} \boldsymbol{w}_i &= A^{\mathsf{T}}(\boldsymbol{\nu}_i + \boldsymbol{\chi}_i) \\ &= Q^{\mathsf{T}}(\boldsymbol{\nu}_{Q,i} + \boldsymbol{\chi}_{Q,i}) \\ &= Q^{\mathsf{T}} \boldsymbol{w}_{Q,i} \end{aligned} \tag{5.123}$$

where

$$\boldsymbol{w}_{Q,i} \triangleq (Q^{\mathsf{T}})^{\dagger} \boldsymbol{w}_i \tag{5.124}$$

$$\boldsymbol{\chi}_{Q,i} \triangleq (Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} \boldsymbol{\chi}_i \tag{5.125}$$

$$\boldsymbol{\nu}_{Q,i} \triangleq (Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} \boldsymbol{\nu}_i. \tag{5.126}$$

Then, it follows from (5.70) and (5.126) that

$$\boldsymbol{\nu}_{Q,i} \sim \mathcal{G}\Big((Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} \boldsymbol{\beta}^{(\theta_i^{\circ})}, I_{r\times r} \Big). \tag{5.127}$$

where the covariance term follows from Lemma 5.2:

$$(Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} \Sigma A Q^{\dagger} = (Q^{\mathsf{T}})^{\dagger} (Q^{\mathsf{T}} Q) Q^{\dagger} = I_{r\times r}. \tag{5.128}$$

Moreover, from the definition (5.59) of $\boldsymbol{\chi}_i$ in terms of $\boldsymbol{w}_{i-1}$ and (5.123) we can alternatively write

$$\boldsymbol{\chi}_i = \mathrm{col}\left\{ \log \frac{\mathbb{T}(1|0) + \mathbb{T}(1|1)\exp\{[Q^{\mathsf{T}} \boldsymbol{w}_{Q,i-1}]_\ell\}}{\mathbb{T}(0|0) + \mathbb{T}(0|1)\exp\{[Q^{\mathsf{T}} \boldsymbol{w}_{Q,i-1}]_\ell\}} \right\}_{\ell=1}^{K}. \tag{5.129}$$

Now note from (5.124)–(5.126) that

$$\boldsymbol{w}_{Q,i} = \boldsymbol{\nu}_{Q,i} + \boldsymbol{\chi}_{Q,i} \overset{(5.125)}{=} \boldsymbol{\nu}_{Q,i} + (Q^{\mathsf{T}})^{\dagger} A^{\mathsf{T}} \boldsymbol{\chi}_i \tag{5.130}$$

which represents a transformation from $\boldsymbol{w}_{Q,i-1}$ to $\boldsymbol{w}_{Q,i}$ directly in light of (5.129). This indicates that we can work with $\boldsymbol{w}_{Q,i}$ over time and transform it into the original vector $\boldsymbol{w}_i$ via (5.123). Repeating arguments similar to (5.61) and (5.62) and replacing $\boldsymbol{w}_i$ by $\boldsymbol{w}_{Q,i}$ we obtain:

$$
\begin{aligned}
S_{Q,i}^{(\theta)}(w_Q, w'_Q)dW_Q &\triangleq \mathbb{P}(\boldsymbol{w}_{Q,i} \in (w_Q, w_Q + dw_Q)|\boldsymbol{\theta}_i^\circ = \theta, \boldsymbol{w}_{Q,i-1} = w'_Q) \\
S_{Q,i}^{(\theta)}(w_Q, w'_Q) &\stackrel{(a)}{=} \frac{1}{(2\pi)^{r/2}} \exp\left\{ -\frac{1}{2}\|w_Q - \rho_Q^{(\theta)}(w'_Q)\|^2 \right\}
\end{aligned}
\tag{5.131}
$$

where (a) follows from (5.127) and

$$
\rho_Q^{(\theta)}(w'_Q) \triangleq (Q^\mathsf{T})^\dagger A^\mathsf{T} \boldsymbol{\beta}^{(\theta_i^\circ)} + \boldsymbol{\chi}_{Q,i}\Big|_{\boldsymbol{\theta}_i^\circ=\theta, \boldsymbol{w}_{Q,i-1}=w'_Q}. \tag{5.132}
$$

Observe that the density (5.131) for $\boldsymbol{w}_{Q,i}$ exists in $\mathbb{R}^r$, even if $\boldsymbol{w}_i$ does not admit a density in $\mathbb{R}^K$. Furthermore, the effective temporal recursion becomes

$$
f_{Q,i}(\theta, w_Q) = \sum_{\theta'} \mathbb{T}(\theta|\theta') \left[\int_{w'_Q} S_{Q,i}^{(\theta)}(w_Q, w'_Q) f_{Q,i-1}(\theta', w'_Q) dW'_Q \right] \tag{5.133}
$$

Using this information along with (5.123), which allows us to recover the agent-specific log-belief ratio $\boldsymbol{w}_{k,i}$ from the low-dimensional representation $\boldsymbol{w}_{Q,i}$, namely,

$$
\boldsymbol{w}_{k,i} = q_k^\mathsf{T} \boldsymbol{w}_{Q,i}, \tag{5.134}
$$

where q_k is the kth column of Q, we arrive at the following probability of error calculation for the diffusion HMM strategy:

$$
p_{k,i} = \int_{q_k^\mathsf{T} w_Q \le 0} \dots \int f_{Q,i}(1, w_Q) dW_Q + \int_{q_k^\mathsf{T} w_Q > 0} \dots \int f_{Q,i}(0, w_Q) dW_Q. \tag{5.135}
$$

5.D Proof of Theorem 5.2

First, observe from (5.59) that for a given transition model $0 < \mathbb{T}(\theta|\theta') < 1\, \forall \theta, \theta' \in \Theta$, $\boldsymbol{\chi}_i$ is bounded in norm. Moreover, the Gaussian mean $\boldsymbol{\beta}^{(\theta_i^\circ)}$ is also bounded—see (5.71). These in turn imply

$$
\|\rho^{(\theta)}(w')\|_{\Sigma^{-1}} \le \widetilde{\rho} \tag{5.136}
$$

for some constant $\widetilde{\rho} > 0$. Let us define the spherical region $\mathcal{R} \triangleq \{w : \|w\|_{\Sigma^{-1}} \le \widetilde{\rho}\}$. For any vector w that satisfies $\|w\|_{\Sigma^{-1}} \ge \widetilde{\rho}$, the projection to this region is given by $\widetilde{\rho}\frac{w}{\|w\|_{\Sigma^{-1}}}$, which can be verified by following the same steps for finding a vector's projection on

the ℓ_2-ball. This implies that for any w outside the region $\mathcal{R}$, and for any w':

$$\min_{w_p\in\mathcal{R}} \|w-w_p\|_{\Sigma^{-1}} = \left\|w-\widetilde{\rho}\frac{w}{\|w\|_{\Sigma^{-1}}}\right\|_{\Sigma^{-1}} \le \|w-\rho^{(\theta)}(w')\|, \tag{5.137}$$

since $\rho^{(\theta)}(w')\in\mathcal{R}$ for any w'. We incorporate this result into (5.73) and observe that for consensus we have

$$0 < S_i^{(\theta)}(w,w') \le \widetilde{S}(w) \tag{5.138}$$

for a Lebesgue integrable function $\widetilde{S}(w)$:

$$\widetilde{S}(w) \triangleq \begin{cases} \dfrac{1}{\sqrt{(2\pi)^K \det(\Sigma)}}, & \|w\|_{\Sigma^{-1}} < \widetilde{\rho} \\ \dfrac{\exp\left\{-\frac{1}{2}\left\|w-\widetilde{\rho}\frac{w}{\|w\|_{\Sigma^{-1}}}\right\|^2_{\Sigma^{-1}}\right\}}{\sqrt{(2\pi)^K\det(\Sigma)}}, & \text{elsewhere} \end{cases}. \tag{5.139}$$

Therefore, the kernel $\mathbb{T}(\theta|\theta')S_i^{(\theta)}(w,w')$ of the recursion (5.66) satisfies the conditions required by [83, Theorem 5.7.4] and we conclude that:

$$\lim_{i\to\infty} \left\|f_i - f_\infty\right\|_{\text{TV}} = 0. \tag{5.140}$$

Consider time independent random variables $\{\boldsymbol{\theta}_\infty^\circ, \boldsymbol{w}_\infty\}$ whose joint pdf is given by f_∞. The convergence in total variation (5.140) implies convergence in distribution (defined in (5.75)), which means that $\{\boldsymbol{\theta}_i^\circ, \boldsymbol{w}_i\}$ converge to limiting random variables $\{\boldsymbol{\theta}_\infty^\circ, \boldsymbol{w}_\infty\}$ in distribution, i.e., $\{\boldsymbol{\theta}_i^\circ, \boldsymbol{w}_i\} \overset{d}{\rightsquigarrow} \{\boldsymbol{\theta}_\infty^\circ, \boldsymbol{w}_\infty\}$. As a result, if we define for consensus probability of error

$$p_{k,\infty} = \int_{w_k=-\infty}^{0} f_{k,\infty}(1,w_k)dw_k + \int_{w_k=0}^{\infty} f_{k,\infty}(0,w_k)dw_k, \tag{5.141}$$

where

$$f_{k,\infty}(\theta,w_k) = \int\cdots\int f_\infty(\theta,w)dw_1\cdots dw_{k-1}dw_{k+1}\cdots dw_K, \tag{5.142}$$

we obtain convergence to the steady-state error probability

$$\lim_{i\to\infty} p_{k,i} = p_{k,\infty}. \tag{5.143}$$

Similarly for diffusion.

5.E Proof of Lemma 5.1

We verify that in general there is no network agreement by providing a counter-example. Consider the following special case of the binary hypothesis testing problem described in Sec. 5.5:

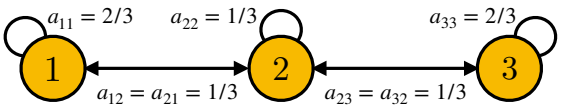


Figure 5.7: Network of $K = 3$ agents.

The model: Consider the network of $K = 3$ agents in Fig. 5.7. Observe that the network is strongly-connected. Moreover, the combination weights satisfy the doubly-stochastic and symmetric matrix Assumption 5.2. Assume for simplicity of notation that all agents have the same Gaussian observation models with $\sigma_k^2 = \sigma^2$. The transition Markov chain is a binary symmetric channel with transition probability $\alpha = 0.5$.

System equilibrium: It follows that for each agent k, regardless of the belief $\boldsymbol{\mu}_{k,i-1}(\theta_{i-1})$,

$$\begin{aligned}\boldsymbol{\eta}_{k,i}(\theta_i) &= \sum_{\theta_{i-1}\in\Theta} \mathbb{T}(\theta_i|\theta_{i-1})\boldsymbol{\mu}_{k,i-1}(\theta_{i-1}) \\ &= \alpha\boldsymbol{\mu}_{k,i-1}(0) + \alpha\boldsymbol{\mu}_{k,i-1}(1) = 0.5. \end{aligned} \tag{5.144}$$

This also means that, for each agent k,

$$\log \frac{\boldsymbol{\eta}_{k,i}(1)}{\boldsymbol{\eta}_{k,i}(0)} = 0, \tag{5.145}$$

and according to (5.56),

$$\boldsymbol{w}_{k,i} = \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\gamma \log \frac{L_\ell(\boldsymbol{\xi}_{\ell,i}|1)}{L_\ell(\boldsymbol{\xi}_{\ell,i}|0)}. \tag{5.146}$$

Moreover, since the likelihood functions of the agents are identical, the entries

$$\boldsymbol{\nu}_{k,i} \triangleq \gamma \log \frac{L_k(\boldsymbol{\xi}_{k,i}|1)}{L_k(\boldsymbol{\xi}_{k,i}|0)} \tag{5.147}$$

of $\boldsymbol{\nu}_i \in \mathbb{R}^3$ are i.i.d. Gaussian random variables (with mean in (5.71) and variance in (5.72)) given the true hypothesis. For the peripheral agent 1, it holds that

$$\boldsymbol{w}_{1,i} = \sum_{\ell\in\mathcal{N}_k} a_{\ell k}\boldsymbol{\nu}_{\ell,i} = \frac{2}{3}\boldsymbol{\nu}_{1,i} + \frac{1}{3}\boldsymbol{\nu}_{2,i}, \tag{5.148}$$

and for the central agent 2,

$$\boldsymbol{w}_{2,i} = \frac{1}{3}\boldsymbol{\nu}_{1,i} + \frac{1}{3}\boldsymbol{\nu}_{2,i} + \frac{1}{3}\boldsymbol{\nu}_{3,i}. \tag{5.149}$$

Observe that the log-belief ratio $\boldsymbol{w}_{k,i}$ for each agent k is a Gaussian random variable, with mean and variance parameters that do not depend on time. However, even though $\boldsymbol{w}_{1,i}$ and $\boldsymbol{w}_{2,i}$ have the same mean, their variances and hence their distributions are not the same. This proves that agents do not converge to the same random variable and do not have the same steady-state error probability. In this particular example, the central agent has less error probability due to smaller variance. For the consensus case, a similar counter-example can be formed by allowing agents to have different likelihoods.

6 Policy Evaluation in Dec-POMDPs

6.1 Introduction[1]

In this chapter, we expand the dynamic state of nature studied in Chapter 5 to the case in which agents not only update their beliefs but also take actions based on these beliefs. Specifically, we consider the multi-agent reinforcement learning (MARL) [89, 90] paradigm, which is useful for searching optimal policies in sequential decision making tasks involving a group of agents.

Most works on MARL focus on the case where agents can directly observe the global state of the environment. In many scenarios of interest, however, agents can only have access to partial information about the state. The decentralized partially observable Markov decision process (Dec-POMDP) framework [91] is applicable to these types of situations. A majority of MARL works assume that Dec-POMDPs observe data that are deterministic and known functions of the underlying state, which is not the case in general. Consider, for example, robots that receive noisy observations from their sensors. The underlying observation model is stochastic in this case.

Given the difficulty of forming beliefs in a decentralized manner, most MARL algorithms [92–94] resort to model-free and end-to-end approaches where agents try to simultaneously learn a policy and an embedding of the history that can replace the beliefs (using, e.g., recurrent neural networks (RNNs)). Nevertheless, some empirical results suggest that this model-free approach can be sub-optimal when the underlying signals of the environment are too weak to train a model such as RNNs [95, 96]. Add to this the fact that RNNs (or other similar machine learning models) correspond to black boxes that provide little insight into the inner workings of the methodologies. In other words, these algorithms lack interpretability, which is critical for the design of trustworthy systems (see [8]). In addition, although end-to-end approaches have shown remarkable performance empirically, they are still based on heuristics and lack

[1] The material in this chapter is based on [88].

theoretical guarantees on their performance. Compared to modular approaches, they are inefficient in terms of adaptability and generalization to similar tasks.

As an alternative, in this chapter, we examine belief-based strategies for MARL, and in particular, study the multi-agent policy evaluation problem. We focus on solving the problem of approximating the *global* Bayesian posterior in a *distributed* manner in a Dec-POMDP context.

6.1.1 Contributions

- We consider a setting where agents only get partial observations from the underlying state of nature, as opposed to prior works on MARL over networks [97–105] that assume agents have full state information. Moreover, as opposed to the literature on decentralized stochastic control [106–109], in our setting, agents need to learn their value functions from data. More specifically, in our framework, agents only know their local observations, actions, and rewards and are allowed to communicate with their immediate neighbors over a graph. In the proposed strategy (Algorithm 6.2), agents exchange both their belief and value function estimates.

- We show in Theorem 6.1 that by exchanging beliefs, agents keep a bounded disagreement with the global posterior distribution, which requires fusing all observations and actions. Also, exchanging value function parameters enables agents to cluster around the network centroid for sufficiently small learning rates (Theorem 6.2). Furthermore, we prove that the network centroid attains a bounded difference from centralized training (Theorem 6.3).

- By means of simulations, we illustrate that agents attain a small mean-square distance from the network centroid. Moreover, the squared Bellman error (SBE) averaged over the network is shown to be comparable to the SBE of the centralized strategy.

6.2 Problem Setting

In this chapter, we are interested in multi-agent policy evaluation under partially observable stochastic environments. For clarity of exposition and to motivate the notation, we briefly review the procedure of single-agent policy evaluation under both fully and partially observable states. Note that there are small differences of notation in comparison to the other chapters of the thesis in order to be consistent with the canonical notation of the reinforcement learning literature.

6.2.1 Fully-Observable Case

For modeling a learning agent under fully observable and dynamic environments, the traditional setting is a finite Markov Decision Process (MDP). An MDP is defined by the quintuple $(\mathcal{S}, \mathcal{A}, \mathbb{T}, \boldsymbol{r}, \gamma)$, where $\mathcal{S}$ is a set of states with cardinality $|\mathcal{S}| = S$, $\mathcal{A}$ is a set of actions, $\mathbb{T}$ is a transition model where $\mathbb{T}(s'|a,s)$ denotes the probability of transitioning from $s \in \mathcal{S}$ to $s' \in \mathcal{S}$ when the agent executes action $a \in \mathcal{A}$, $\boldsymbol{r}(s,a,s')$ denotes the reward the agent receives when it executes action a and the environment transitions from state s to s', and $\gamma \in [0,1)$ is a discount factor that determines the importance given to immediate rewards ($\gamma \to 0$) or the total reward ($\gamma \to 1$).

The goal of policy evaluation is to learn the value function $V^{\pi}(s)$ of a target policy $\pi(a|s)$, where the value function is the expected return if the agent starts from state s and follows policy π, namely,

$$V^{\pi}(s) = \mathbb{E}\Big[\sum_{i=0}^{\infty} \gamma^i \boldsymbol{r}(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1}) | \boldsymbol{s}_0 = s\Big], \tag{6.1}$$

where $\boldsymbol{s}_i$ is the state at time i and $\boldsymbol{a}_i$ is the action chosen by the agent according to the policy, $\boldsymbol{a}_i \sim \pi(a|\boldsymbol{s}_i)$. In many applications, the state space is too large (or infinite), which makes it impractical to keep track of the value function for all states. Therefore, function approximations are used to reduce the dimension of the problem. For instance, linear approximations, which are the focus of the theoretical analysis of this work, correspond to using a parameter $w^{\circ} \in \mathbb{R}^M$ to approximate $V^{\pi}(s) \approx \phi(s)^{\mathsf{T}} w^{\circ}$, where $\phi : \mathcal{S} \to \mathbb{R}^M$ is a *pre-defined* feature mapping for representing state s.

A standard stochastic approximation algorithm to learn w° from data is TD-learning [8, 110] such as the TD(0) strategy [111] and variations thereof. If we denote the value function estimate at $w \in \mathbb{R}^M$ by $\widehat{V}(s,w) \triangleq \phi(s)^{\mathsf{T}} w$, then, under this strategy, the agent first computes the TD-error $\boldsymbol{\delta}_i$ at time i by using the observed transition tuple $(\boldsymbol{s}_i, \boldsymbol{r}_i, \boldsymbol{s}_{i+1})$:

$$\boldsymbol{\delta}_i = \boldsymbol{r}_i + \gamma \widehat{V}(\boldsymbol{s}_{i+1}, \boldsymbol{w}_i) - \widehat{V}(\boldsymbol{s}_i, \boldsymbol{w}_i), \tag{6.2}$$

where $\boldsymbol{r}_i \triangleq \boldsymbol{r}(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1})$ is the instantaneous reward at time i. Subsequently, the agent uses this error to update the current parameter estimate $\boldsymbol{w}_i$ to

$$\boldsymbol{w}_{i+1} = \boldsymbol{w}_i + \alpha \boldsymbol{\delta}_i \nabla_w \widehat{V}(\boldsymbol{s}_i, \boldsymbol{w}_i), \tag{6.3}$$

where $\alpha > 0$ is the learning rate, and

$$\nabla_w \widehat{V}(\boldsymbol{s}_i, \boldsymbol{w}_i) = \phi(\boldsymbol{s}_i) \tag{6.4}$$

for the linear function approximation case. This algorithm can be viewed as a "stochastic gradient algorithm" where the effective stochastic gradient is $\boldsymbol{g}_i \triangleq -\boldsymbol{\delta}_i \phi(\boldsymbol{s}_i)$. In this work, we consider an ℓ_2-regularized version of the algorithm, which changes the

update step (6.3) to

$$\boldsymbol{w}_{i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_i + \alpha\boldsymbol{\delta}_i\nabla_w\widehat{V}(\boldsymbol{s}_i, \boldsymbol{w}_i), \tag{6.5}$$

where $\rho > 0$ is a constant hyper-parameter. As opposed to supervised learning, regularization is rather under-explored in reinforcement learning, with notable exceptions in [112, 113]. However, recent work [114, 115] suggests that regularization can increase generalization and sample-efficiency in function approximation with over-parameterized models.

6.2.2 Partially-Observable Case

In many applications, the agent does not directly observe the state $\boldsymbol{s}_i$. For instance, a robot may receive noisy and partially informative observations from its sensors about the environment. The observation $\boldsymbol{\xi}_i$ that the agent receives at time i is generally assumed to be distributed according to some likelihood function linking it to the unobservable state, say, $\boldsymbol{\xi}_i \sim L(\xi|\boldsymbol{s}_i)$, which is conditioned on $\boldsymbol{s}_i$. In these scenarios, the agent will need to estimate the latent state first from the observations. To do so, the agent will need to learn a probability vector $\boldsymbol{\mu}_i \in \mathcal{M}(S)$ over the set of states $\mathcal{S}$, which is called the *belief* vector [8, 116]. Here, $\mathcal{M}(S)$ denotes the S-dimensional probability simplex, and the entry $\boldsymbol{\mu}_i(s) \in [0, 1]$ of the belief vector quantifies the confidence the agent has about state s being the true state at time i. The value of $\boldsymbol{\mu}_i(s)$ corresponds to the posterior probability of s conditioned on the action-observation history (a.k.a. trajectory):

$$\boldsymbol{\mathcal{F}}_i \triangleq \{\boldsymbol{\xi}_i, \boldsymbol{a}_{i-1}, \boldsymbol{\xi}_{i-1}, \dots\}, \tag{6.6}$$

which means

$$\boldsymbol{\mu}_i(s) \triangleq \mathbb{P}(\boldsymbol{s}_i = s|\boldsymbol{\mathcal{F}}_i). \tag{6.7}$$

As already explained in the previous chapter in Sec. 5.2.1, this posterior satisfies the following temporal recursion [8, 48, 116]:

$$\boldsymbol{\mu}_i(s) \propto L(\boldsymbol{\xi}_i|s)\boldsymbol{\eta}_i(s), \tag{6.8}$$

where $\boldsymbol{\eta}_i(s)$ is the time-adjusted prior defined by

$$\boldsymbol{\eta}_i(s) \triangleq \mathbb{P}(\boldsymbol{s}_i = s|\boldsymbol{\mathcal{F}}^a_{i-1}) = \sum_{s' \in \mathcal{S}} \mathbb{T}(s|s', \boldsymbol{a}_{i-1})\boldsymbol{\mu}_{i-1}(s'). \tag{6.9}$$

Here, $\boldsymbol{\mathcal{F}}^a_{i-1}$ is the collection of past observations and actions (i.e., the current observation $\boldsymbol{\xi}_i$ is excluded):

$$\boldsymbol{\mathcal{F}}^a_{i-1} \triangleq \{\boldsymbol{a}_{i-1}, \boldsymbol{\xi}_{i-1}, \boldsymbol{a}_{i-2}, \dots\}. \tag{6.10}$$

Note that $\boldsymbol{\mathcal{F}}_i = \{\boldsymbol{\xi}_i\} \cup \boldsymbol{\mathcal{F}}^a_{i-1}$. If beliefs are used as substitutes for hidden states, then partially-observable MDPs (POMDPs) can be treated as *continuous* MDPs, since beliefs

are continuous even if the number of states is finite. In this way, the policy evaluation problem would correspond to evaluating $V^{\pi}(\mu)$ where the value function is now defined as the expected return when the agent starts from the belief state μ and follows the policy $\pi(a|\mu)$, namely [8, 116]:

$$V^{\pi}(\mu) = \mathbb{E}\Big[\sum_{i=0}^{\infty} \gamma^i \boldsymbol{r}_i | \boldsymbol{\mu}_0 = \mu\Big]. \tag{6.11}$$

Observe that, in contrast to the fully-observable case, the agent now chooses action $\boldsymbol{a}_i$ according to the policy $\boldsymbol{a}_i \sim \pi(a|\boldsymbol{\mu}_i)$, which is conditioned on the belief vector.

Algorithm (6.2)–(6.5) can be adjusted for POMDPs by using the belief vectors $(\boldsymbol{\mu}_i, \boldsymbol{\eta}_{i+1})$ instead of the states $(\boldsymbol{s}_i, \boldsymbol{s}_{i+1})$. Thus, we let

$$\boldsymbol{\delta}_i = \boldsymbol{r}_i + \gamma \widehat{V}(\boldsymbol{\eta}_{i+1}, \boldsymbol{w}_i) - \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i), \tag{6.12}$$

and

$$\boldsymbol{w}_{i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_i + \alpha \boldsymbol{\delta}_i \nabla_w \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i), \tag{6.13}$$

where the approximation $\widehat{V}(\mu, w)$ is computed by using the feature vectors $\phi(\mu)$, now dependent on μ, to evaluate $\widehat{V}(\mu, w) \triangleq \phi(\mu)^{\mathsf{T}} w$. Note that from now on $\phi : \mathcal{M}(S) \to \mathbb{R}^M$ is a different feature mapping that represents μ instead of s, and agents' goal is to learn w° that satisfies $V^{\pi}(\mu) \approx \phi(\mu)^{\mathsf{T}} w^{\circ}$.

Observe from (6.8)–(6.9) that in order for the agent to update the belief vectors $(\boldsymbol{\mu}_i, \boldsymbol{\eta}_{i+1})$, it needs to know the transition model $\mathbb{T}$ and the likelihood functions $L(\boldsymbol{\xi}_i|s)$ for each state. However, the agent does not need to know the underlying reward model $\boldsymbol{r}$. It can use instantaneous reward samples $\boldsymbol{r}_i$ to run the algorithm. In this sense, the algorithm is a mixture of model-based and model-free reinforcement learning. Motivation for this approach is at least two-fold.

First, in some applications, learning the transition and observation models from data is inherently easier than learning the reward function. This is because the reward function can depend on some latent characteristics of the environment or some human expert, which may be challenging to estimate. One example where this scenario can arise is autonomous cars [117]. In this case, the observations from environmental sensors and cameras are processed with a learned likelihood model such as a convolutional neural network. The transition dynamics of the car depends on various parameters such as speed, acceleration, position, and incline, and can be modeled based on physics laws and a mapping of the surroundings. However, learning a reward function for this application is notoriously difficult, as it is challenging to cover all possible situations [118].

Second, the agent can still run (6.12)–(6.13) even if beliefs are not formed through

(6.8)–(6.9), but estimated by some other approach, as in [119–121].

6.3 Multi-agent Policy Evaluation

We now consider a set $\mathcal{K}$ of K cooperative agents that aim to evaluate the average value function under a joint policy $\pi = \{\pi_k\}_{k=1}^K$ that consists of individual policies π_k. The framework we consider is a *decentralized* POMDP (Dec-POMDP) [91], which is defined by the sextuple $(\mathcal{S}, \mathcal{A}_k, \mathcal{O}_k, \mathbb{T}, \boldsymbol{r}_k, \gamma)$. Here, the set of states $\mathcal{S}$ and the transition model $\mathbb{T}$ are common to all agents, where the notation $\mathbb{T}(s|s',a)$ now specifies the probability that the environment transitions from s' to s when the agents execute the joint action $a = \{a_k\}_{k=1}^K$. The individual action a_k of each agent k takes values from the set $\mathcal{A}_k$, and $\boldsymbol{r}_k(s,a,s')$ is the *local* reward k gets when the agents execute the collection of actions a and the environment transitions from s to s'. Note that this setting covers general teamwork scenarios where the local reward of an individual agent can be dependent on all actions, and not only on its own actions. Specifically, it covers the scenario where all agents observe the same reward, i.e., $\boldsymbol{r}_k(s,a,s') = \boldsymbol{r}(s,a,s'), \forall k \in \mathcal{K}$. Remember that agents receive instantaneous rewards as they progress through the POMDP, and they are not required to know the joint action a from all agents. Moreover, $\mathcal{O}_k$ is a set of *private* observations. At each time instant i, agent k receives observation $\boldsymbol{\xi}_{k,i} \in \mathcal{O}_k$ emitted by state $\boldsymbol{s}_i$, and assumed to be distributed according to the local *marginal* likelihood $L_k(\xi_k|\boldsymbol{s}_i)$.

Similar to the single-agent case, Dec-POMDPs can in principal be treated as multi-agent belief MDPs by replacing the hidden states with joint centralized beliefs defined by [91, Chapter 2]

$$\boldsymbol{\mu}_i(s) \triangleq \mathbb{P}(\boldsymbol{s}_i = s|\boldsymbol{\mathcal{F}}_i) \propto L(\boldsymbol{\xi}_i|s)\boldsymbol{\eta}_i(s). \tag{6.14}$$

Here, $\boldsymbol{\mathcal{F}}_i$ denotes the history of all observations and past actions from across all agents until time i, $\boldsymbol{\xi}_i \triangleq \{\boldsymbol{\xi}_{k,i}\}_{k=1}^K$ is now the aggregate of the observations from across the network, and $\boldsymbol{a}_{i-1}$ is a tuple aggregating actions from all agents at time $i-1$. Moreover, under spatial independence, the joint likelihood $L(\boldsymbol{\xi}_i|s)$ appearing in (6.14) is given by

$$L(\boldsymbol{\xi}_i|s) = \prod_{k=1}^{K} L_k(\boldsymbol{\xi}_{k,i}|s). \tag{6.15}$$

In a manner similar to the single-agent case, the belief $\boldsymbol{\eta}_i(s)$ is the time-adjusted prior conditioned on $\boldsymbol{\mathcal{F}}_{i-1}^a$ (6.10):

$$\boldsymbol{\eta}_i(s) \triangleq \mathbb{P}(\boldsymbol{s}_i = s|\boldsymbol{\mathcal{F}}_{i-1}^a) = \sum_{s' \in \mathcal{S}} \mathbb{T}(s|s', \boldsymbol{a}_{i-1})\boldsymbol{\mu}_{i-1}(s'). \tag{6.16}$$

The goal of policy evaluation is to learn the *team* value function, which is the expected

average reward of all agents starting from some belief state μ, i.e.,

$$V^{\pi}(\mu) = \mathbb{E}\Big[\sum_{i=0}^{\infty}\gamma^{i}\Big(\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}_{k,i}\Big)\Big|\boldsymbol{\mu}_0 = \mu\Big], \tag{6.17}$$

where $\boldsymbol{r}_{k,i}$ denotes the instantaneous local reward agent k gets at time i.

There is one major inconvenience with this approach. In order to compute the joint belief (6.14), it is necessary to fuse all observations and actions from across the agents in a central location. This is possible in settings where there exists a fusion center. However, many applications rely solely on localized processing. In the following, we discuss and compare two strategies for multi-agent reinforcement learning under partial observations: (i) a centralized strategy, (ii) and a fully decentralized strategy.

6.3.1 Centralized Strategy

In the fully centralized strategy, the state estimation and policy evaluation phases are centralized and, hence, the setting is equivalent to a single-agent POMDP, already discussed in Sec. 6.2.2, using the joint likelihood $L(\boldsymbol{\xi}_i|s)$ and the average reward $\boldsymbol{r}_i \triangleq K^{-1}\sum_{k=1}^{K}\boldsymbol{r}_{k,i}$. The fusion center computes the joint belief (6.14), and agents take actions based on this joint belief, i.e., $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_i)$. The fusion center then computes the centralized TD-error:

$$\boldsymbol{\delta}_i = \boldsymbol{r}_i + \gamma\widehat{V}(\boldsymbol{\eta}_{i+1}, \boldsymbol{w}_i) - \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i), \tag{6.18}$$

and updates the estimate to

$$\boldsymbol{w}_{i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_i + \alpha\boldsymbol{\delta}_i\nabla_w\widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i). \tag{6.19}$$

This construction is listed under Algorithm 6.1.

6.3.2 Decentralized Strategy

The centralized strategy is disadvantageous in the sense that (i) failure of the fusion center results in failure of the system; (ii) there can be communication bottlenecks at the fusion center; (iii) and agents can be spatially distributed to begin with. Therefore, in this section, we propose a fully decentralized strategy for policy evaluation where agents communicate with their immediate neighbors only. In this chapter, we denote the combination matrix with $C = [c_{\ell k}]$ and assume that C is symmetric and doubly-stochastic.

Algorithm 6.1 Centralized policy evaluation under POMDPs

set initial prior $\eta_0(s) > 0, \forall s \in \mathcal{S}$
initialize w_0
while $i \geq 1$ **do**
  each agent k observes $\boldsymbol{\xi}_{k,i}$
  collect all observations $\boldsymbol{\xi}_i \triangleq \{\boldsymbol{\xi}_{k,i}\}_{k=1}^K$ and **evaluate**

$$\boldsymbol{\mu}_i(s) \propto L(\boldsymbol{\xi}_i|s)\boldsymbol{\eta}_i(s) \tag{6.20}$$

  for each agent $k = 1, 2, \ldots, K$ **do**
    take action $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_i)$
    get reward $\boldsymbol{r}_{k,i} = \boldsymbol{r}_k(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1})$
  end for
  then, **evolve** according to:

$$\boldsymbol{\eta}_{i+1}(s) = \sum_{s' \in \mathcal{S}} \mathbb{T}(s|s', \boldsymbol{a}_i)\boldsymbol{\mu}_i(s') \tag{6.21}$$

  average the rewards $\boldsymbol{r}_i = \frac{1}{K}\sum_{k=1}^K \boldsymbol{r}_{k,i}$
  update the model:

$$\boldsymbol{\delta}_i = \boldsymbol{r}_i + \gamma \widehat{V}(\boldsymbol{\eta}_{i+1}, \boldsymbol{w}_i) - \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i) \tag{6.22}$$

$$\boldsymbol{w}_{i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_i + \alpha\boldsymbol{\delta}_i \nabla_w \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i) \tag{6.23}$$

  $i \leftarrow i + 1$
end while

Local Belief Formation

In the fully decentralized strategy, the agents cannot form the joint belief (6.14) since they do not have access to the observations and actions of all other agents. They, however, can construct local beliefs. To do so, we will extend the diffusion HMM strategy (DHS) from previous chapter (Algorithm 5.1), originally designed for hidden Markov models, to the current POMDP setting.

Recall from previous chapter that In DHS (Algorithm 5.1), the global belief vectors $\{\boldsymbol{\mu}_i, \boldsymbol{\eta}_i\}$ are replaced by local belief vectors $\{\boldsymbol{\mu}_{k,i}, \boldsymbol{\eta}_{k,i}\}$, and the latter are updated by using local observations and by relying solely on interactions with the immediate neighbors. The original DHS algorithm is designed for actionless partially observable Markov chains, and each agent can use the same global transition model. However, in POMDPs, transition of the global state depends on the joint action, and the agents cannot perform a *centralized* time-adjustment step as in (6.21) since they do not know the actions of all agents in the network.

Therefore, one strategy is to use a transition model that is obtained by marginalizing

over actions that are unknown to agent k. More specifically, let $a_{\mathcal{N}_k} \in \mathcal{A}_{\mathcal{N}_k}$ denote a tuple of actions taken by the set of neighbors of agent k (which we are denoting by $\mathcal{N}_k$). These actions can be assumed to be known by agent k if, for instance, agents share their actions with their neighbors. Let $a^c_{\mathcal{N}_k} \in \mathcal{A}^c_{\mathcal{N}_k}$ denote the remaining actions by all other agents in the network, so that $a = a_{\mathcal{N}_k} \cup a^c_{\mathcal{N}_k}$. Then, each agent can use the following *local* transition model approximation:

$$\mathbb{T}^{\pi}_k(s|s', a_{\mathcal{N}_k}) \propto \sum_{a^c_{\mathcal{N}_k} \in \mathcal{A}^c_{\mathcal{N}_k}} \mathbb{T}(s|s', a_{\mathcal{N}_k}, a^c_{\mathcal{N}_k}) \pi(a_{\mathcal{N}_k}, a^c_{\mathcal{N}_k}|s') \tag{6.24}$$

in lieu of $\mathbb{T}(s|s', a)$, to time-adjust its local belief:

$$\boldsymbol{\eta}_{k,i}(s) = \sum_{s' \in \mathcal{S}} \mathbb{T}^{\pi}_k(s|s', \boldsymbol{a}_{\mathcal{N}_k, i-1}) \boldsymbol{\mu}_{k,i-1}(s'), \tag{6.25}$$

Here, $\boldsymbol{a}_{\mathcal{N}_k,i-1}$ is the tuple of actions taken by the neighbors of agent k at time instant $i-1$. Moreover, in (6.24), the notation $\pi(a_{\mathcal{N}_k}, a^c_{\mathcal{N}_k}|s')$ represents the joint action probability:

$$\pi(a_{\mathcal{N}_k}, a^c_{\mathcal{N}_k}|s') = \prod_{\ell=1}^{K} \pi_\ell(a_\ell|s'), \tag{6.26}$$

where the notation $\pi(a|s)$ is now a shorthand for $\pi(a|\mu)$ when

$$\mu = [0 \ldots 1 \ldots 0]^{\mathsf{T}}, \tag{6.27}$$

i.e., when the belief attains value 1 for state s and is 0 otherwise. Note that this construction leads to a richer scenario compared to the previous chapter, with transition models that are different across the agents.

The rest of the algorithm is the same as the DHS strategy. Following (6.25), and based on the personal observation $\boldsymbol{\xi}_{k,i}$, each agent k forms an *intermediate* belief using a β-scaled Bayesian update of the form:

$$\boldsymbol{\psi}_{k,i}(s) \propto (L_k(\boldsymbol{\xi}_{k,i}|s))^{\beta} \boldsymbol{\eta}_{k,i}(s), \tag{6.28}$$

where $\beta > 0$. Next, agents in the neighborhood of k share their intermediate beliefs, which allows agent k to update its belief using the weighted geometric average expression:

$$\boldsymbol{\mu}_{k,i}(s) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(s)\Big)^{c_{\ell k}}. \tag{6.29}$$

This procedure of repeated updating and exchanging of beliefs allows information to

diffuse over the network.

Diffusion Policy Evaluation

In the fully decentralized strategy, the local belief formation strategy is used during both training and execution phases. Namely, the target value function in (6.17) represents the average return agents get when they execute the policy π with their local beliefs formed via the DHS strategy. Moreover, since the policy evaluation is also decentralized, during the training phase, they again need to use DHS to approximate the global belief state μ on top of the function approximation. More specifically, using its local belief vectors, each agent k computes a local TD error:

$$\boldsymbol{\delta}_{k,i} = \boldsymbol{r}_{k,i} + \gamma \widehat{V}(\boldsymbol{\eta}_{k,i+1}, \boldsymbol{w}_{k,i}) - \widehat{V}(\boldsymbol{\mu}_{k,i}, \boldsymbol{w}_{k,i}), \tag{6.30}$$

where $\boldsymbol{r}_{k,i} = \boldsymbol{r}_k(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1})$ is also a function of the local beliefs since each agent k now executes the action $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_{k,i})$. Subsequently, each agent k forms an intermediate parameter estimate denoted by

$$\boldsymbol{z}_{k,i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_{k,i} + \alpha\boldsymbol{\delta}_{k,i}\nabla_w \widehat{V}(\boldsymbol{\mu}_{k,i}, \boldsymbol{w}_{k,i}). \tag{6.31}$$

After receiving the intermediate estimates from its neighbors, agent k updates $\boldsymbol{w}_{k,i}$ to

$$\boldsymbol{w}_{k,i+1} = \sum_{\ell \in \mathcal{N}_k} c_{\ell k}\boldsymbol{z}_{\ell,i+1}. \tag{6.32}$$

The local adaptation step (6.31) followed by the combination step (6.32) are reminiscent of diffusion strategies for distributed learning [7,8]. Observe that there are actually two combination steps involved in diffusion policy evaluation: the belief combination (6.29) with geometric averaging (GA), and the parameter combination (6.32) with arithmetic averaging (AA). The listing of the proposed diffusion policy evaluation strategy for POMDPs appears in Algorithm 6.2.

Algorithm 6.2 has the following advantages:

- **Decentralized information structure:** The algorithm is designed to be fully decentralized, with each agent only having access to its own private data, such as observations and rewards, without the need to share this information with other agents. Importantly, agents do not require knowledge of the joint distribution of observations or the network topology. They only know their own marginal likelihood function, and their actions are only known by (or transmitted to) their immediate neighbors. If agents happen to know their own marginal transition models, they do not need to know the policies of other agents or the global transition model. However, if the application requires them to approximate it

Algorithm 6.2 Diffusion policy evaluation under POMDPs

set initial priors $\eta_{k,0}(s) > 0$, $\forall s \in \mathcal{S}$ and $\forall k \in \mathcal{K}$
choose $\beta > 0$
initialize $w_{k,0}$ for $\forall k \in \mathcal{K}$
while $i \geq 0$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive personal **observation** $\boldsymbol{\xi}_{k,i}$ and **adapt** for each state $s \in \mathcal{S}$:

$$\boldsymbol{\psi}_{k,i}(s) \propto (L_k(\boldsymbol{\xi}_{k,i}|s))^{\beta} \boldsymbol{\eta}_{k,i}(s) \tag{6.33}$$

$$\boldsymbol{\mu}_{k,i}(s) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(s)\Big)^{c_{\ell k}} \tag{6.34}$$

    take **action** $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_{k,i})$
    get **reward** $\boldsymbol{r}_{k,i} = \boldsymbol{r}_k(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1})$
  end for
  for each agent $k = 1, 2, \ldots, K$ **do**
    compute $\mathbb{T}_k^{\pi}(s|s', \boldsymbol{a}_{\mathcal{N}_k,i})$ using (6.24), and

$$\boldsymbol{\eta}_{k,i+1}(s) = \sum_{s' \in \mathcal{S}} \mathbb{T}_k^{\pi}(s|s', \boldsymbol{a}_{\mathcal{N}_k,i}) \boldsymbol{\mu}_{k,i}(s') \tag{6.35}$$

  end for
  for each agent $k = 1, 2, \ldots, K$ **do**

$$\boldsymbol{\delta}_{k,i} = \boldsymbol{r}_{k,i} + \gamma \widehat{V}(\boldsymbol{\eta}_{k,i+1}, \boldsymbol{w}_{k,i}) - \widehat{V}(\boldsymbol{\mu}_{k,i}, \boldsymbol{w}_{k,i}) \tag{6.36}$$

$$\boldsymbol{z}_{k,i+1} = (1 - 2\rho\alpha)\boldsymbol{w}_{k,i} + \alpha \boldsymbol{\delta}_{k,i} \nabla_w \widehat{V}(\boldsymbol{\mu}_{k,i}, \boldsymbol{w}_{k,i}) \tag{6.37}$$

  end for
  for each agent $k = 1, 2, \ldots, K$ **combine**

$$\boldsymbol{w}_{k,i+1} = \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \boldsymbol{z}_{\ell,i+1} \tag{6.38}$$

  end for
  $i \leftarrow i + 1$
end while

themselves, they require knowledge of the other policies and the global transition model.

- **Privacy**: The algorithm is also advantageous in terms of privacy since (i) communicating beliefs allows information diffusion without explicitly sharing raw observational data, and (ii) exchanging value parameters allows agents to learn the cumulative reward across network without explicitly sharing local rewards.

- **Complexity**: (i) The memory requirement is constant over time, with each agent only needing to store its value function parameter estimate (M-dimensional) and local belief (S-dimensional), as well as the necessary model functions. (ii) The communication requirement is also manageable, with each agent communicating only with its immediate neighbors through belief and parameter sharing. The communication load is not affected by the network size, making our algorithm scalable and avoiding communication bottlenecks. (iii) The computational complexity depends on whether the application at hand allows agents to have access to the local transition model. If this is the case, then the computational complexity is equivalent to the single-agent Bayesian filtering case, which is $O(S^2)$. The combination steps add only linear additional complexity $O(S)$ with fixed neighborhood size. However, if agents need to approximate the transition model themselves, the computational complexity increases with the network size, and becomes $O(KS^2)$. This is due to the need to average over non-neighbors' actions in (6.24), whose size grows with the network size in general. Compared to alternative approaches such as relaying raw data, incremental approaches [122], or Bayesian belief forming [20], our algorithm is much lighter in terms of complexity. Relaying raw data, for example, would result in an exponential increase of memory and communication overload at each hop, making it highly impractical. The incremental approach of relaying over a cyclic path (which is NP-hard to find [123]) that visits each agent once would reduce the overload. However, it is not robust against failures and not scalable, making it impractical for a decentralized setting. The Bayesian belief forming strategy requires knowledge of the network topology and other agents' functions, and known to be NP-hard, even in the much simpler case of fixed state and no action setting [124].

6.4 Theoretical Results

In this section, we analyze the performance of the decentralized strategy in Algorithm 6.2. In particular, we first show in Sec. 6.4.2 that the value function parameters $\{\boldsymbol{w}_{k,i}\}$ of the agents cluster around the network centroid. Then, in Sec. 6.4.3, we show that this network centroid has a bounded difference from the parameter of a baseline strategy (which will be presented in Algorithm 6.3). Our analysis relies on bounding the disagreement between the joint centralized belief $\boldsymbol{\mu}_i$ and the local estimate $\boldsymbol{\mu}_{k,i}$,

which is presented next.

6.4.1 Belief Disagreement

In a manner similar to previous chapter, we introduce the following risk functions in order to assess the disagreement between the local beliefs formed via (6.33)–(6.35) with the joint centralized beliefs formed via (6.20)–(6.21):

$$J_{k,i} \triangleq \mathbb{E}_{\mathcal{F}_i} D_{\mathrm{KL}}(\boldsymbol{\mu}_i || \boldsymbol{\mu}_{k,i}), \tag{6.39}$$

and

$$\widetilde{J}_{k,i} \triangleq \mathbb{E}_{\mathcal{F}^a_{i-1}} D_{\mathrm{KL}}(\boldsymbol{\eta}_i || \boldsymbol{\eta}_{k,i}). \tag{6.40}$$

The risks in (6.39) and (6.40) measure the disagreement after and before the joint observation $\boldsymbol{\xi}_i$, respectively. In comparison to the previous chapter, in the current chapter, each agent uses a local approximation for the global transition model based on (6.24). Therefore, we need to make some non-trivial adjustments to the belief disagreement analysis. We begin with adjusting the assumptions from previous chapter to our model.

Modeling Conditions

- **Likelihood functions**: We assume that

$$\Big| \log L_k(\xi|s) \Big| \leq B \tag{6.41}$$

  over its support for each state $s \in \mathcal{S}$ and agent $k \in \mathcal{K}$.

- **Transition model**: The Markov chain induced by any joint action $a \in \mathcal{A}$ is *irreducible* and *aperiodic*. Since the number of states is finite, this assumption implies that the transition model $\mathbb{T}(s|s',a)$ is ergodic [10, Chapter 2]. As explained in previous chapter, we focus on the important class of *geometrically* ergodic models, which additionally satisfy the relation $\kappa(\mathbb{T}^a) \leq \kappa(\mathbb{T})$ for some constant $\kappa(\mathbb{T}) < 1$. Here, $\kappa(\mathbb{T}^a)$ is the Dobrushin coefficient [48, Chapter 2] defined by:

$$\kappa(\mathbb{T}^a) \triangleq \sup_{s',s'' \in \mathcal{S}} \frac{1}{2} \sum_{s \in \mathcal{S}} \Big| T^a_{ss'} - T^a_{ss''} \Big|, \tag{6.42}$$

  where $T^a_{ss'} \triangleq \mathbb{T}(s|s',a)$ is a generic entry of the $S \times S$ transition matrix T^a. Due to space limitations, we refer the reader to [48, Chapter 2] for a comprehensive discussion on the Dobrushin coefficient $\kappa(\mathbb{T}^a)$. In short, $\kappa(\mathbb{T}^a)$ quantifies how fast the transition model forgets its initial conditions. Namely, as $\kappa(\mathbb{T}^a) \to 0$,

past conditions are forgotten faster. Instances of geometrically ergodic transition models include transition matrices with all positive elements, or that satisfy the minorization condition in [48, Theorem 2.7.4]. In addition to this condition from Chapter 5, we have an additional assumption on the transition model to regulate the disagreement stemming from the local transition model estimates:

Assumption 6.1 (**Transition model disagreement**)**.** *For each agent k, consider the n-hop neighbors set $\mathcal{N}_{k^n}$ and its complement $\mathcal{N}^c_{k^n}$. Here, $\mathcal{N}_{k^n}$ is the set of agents that have at most n-hop distance to agent k. We define the transition model approximation that uses n-hop neighbors' actions as follows:*

$$\mathbb{T}^\pi_k(s|s', a_{\mathcal{N}_{k^n}}) \propto \sum_{a^c_{\mathcal{N}_{k^n}} \in \mathcal{A}^c_{\mathcal{N}_{k^n}}} \mathbb{T}(s|s', a_{\mathcal{N}_{k^n}}, a^c_{\mathcal{N}_{k^n}})\pi(a_{\mathcal{N}_{k^n}}, a^c_{\mathcal{N}_{k^n}}|s'). \tag{6.43}$$

We assume that

$$D_{\mathrm{KL}}\left(\mathbb{T}^\pi_k\left(s\middle|s', a_{\mathcal{N}_{k^n}}\right)\middle|\middle|\mathbb{T}^\pi_k\left(s\middle|s', a_{\mathcal{N}_{k^{n+1}}}\right)\right) < \infty, \tag{6.44}$$

which ensures that transition model approximations induced from n-hop and $(n+1)$-hop neighbors' actions share the same support. Moreover, we assume that over the shared support,

$$\left|\log \frac{\mathbb{T}^\pi_k\left(s|s', a_{\mathcal{N}_{k^n}}\right)}{\mathbb{T}^\pi_k\left(s|s', a_{\mathcal{N}_{k^{n+1}}}\right)}\right| \leq \tau. \tag{6.45}$$

for $n \geq 1$.

This assumption basically makes sure that the increase in the error of the transition model approximation of agents due to lack of information about actions is bounded at each geodesic distance increase to that agent.

Difference with Centralized Strategy

The following result provides upper bounds on the disagreement measures in (6.39)–(6.40).

Theorem 6.1 (**Bounds on belief disagreement).** *For each agent k, the belief disagreement risks* (6.39) *and* (6.40) *get bounded with a linear rate of $\kappa(\mathbb{T})$, namely, as $i \to \infty$,*

$$J_{k,i} \leq \frac{2\sqrt{K}\beta\lambda B}{1-\kappa(\mathbb{T})} + \frac{(K - d_{\min})\,\tau}{1-\kappa(\mathbb{T})} \tag{6.46}$$

and

$$\widetilde{J}_{k,i} \leq \frac{2\kappa(\mathbb{T})\sqrt{K}\beta\lambda B}{1-\kappa(\mathbb{T})} + \frac{(K - d_{\min})\,\tau}{1-\kappa(\mathbb{T})} \tag{6.47}$$

where $d_{\min}$ is the minimum degree over the graph, i.e., the minimum number of neighbors over the network, and $\lambda \triangleq \max\{|1-\frac{K}{\beta}|, \lambda_2\}$ where $\lambda_2 < 1$ is the mixing rate (second largest modulus eigenvalue) of C .

Proof. See Appendix 6.A. ■

In Theorem 6.1, the first terms in both bounds are equivalent to the bounds obtained in Theorem 5.1. However, the terms proportional to $(K - d_{\min})\tau$ are new, and they arise from the fact that agents do not observe the joint actions and hence only have a local estimate of the transition model. Nevertheless, the bounds get smaller with increasing network connectivity, i.e., as $\lambda_2 \to 0$ and $d_{\min} \to K$, which shows the benefit of cooperation. In particular, if $\beta = K$ and the network is fully connected $(\lambda_2 = 0, d_{\min} = K)$, then the bounds are equal to 0. In other words, local beliefs match the centralized belief in this situation. It is important to note that the linear term $(K - d_{\min})$ represents a worst-case bound that holds true for any strongly connected network topology.

For instance, in a scenario where each agent has $N > 1$ neighbors, it is straightforward to modify the proof and show that these linear terms will instead be logarithmic, i.e., proportional to $\log K / \log N$.

We use Theorem 6.1 in the performance analysis of the diffusion policy evaluation. To that regard, we first present the following consequence of Theorem 6.1, which provides a bound in terms of disagreement norms.

Corollary 6.1 (Bounds on disagreement norms). *Theorem 6.1 implies that, as* $i \to \infty$*,*

$$\mathbb{E}\Big\|\boldsymbol{\mu}_i - \boldsymbol{\mu}_{k,i}\Big\| \leq B_{\text{TV}} \tag{6.48}$$

and

$$\mathbb{E}\Big\|\boldsymbol{\eta}_i - \boldsymbol{\eta}_{k,i}\Big\| \leq \widetilde{B}_{\text{TV}}, \tag{6.49}$$

where we introduce the constants

$$B_{\text{TV}} \triangleq 2\left(1 - \exp\left\{-\frac{2\sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau}{1-\kappa(\mathbb{T})}\right\}\right)^{1/2} \tag{6.50}$$

and

$$\widetilde{B}_{\text{TV}} \triangleq 2\left(1 - \exp\left\{-\frac{2\kappa(\mathbb{T})\sqrt{K}\beta\lambda B + (K - d_{\min})\tau}{1-\kappa(\mathbb{T})}\right\}\right)^{1/2} \tag{6.51}$$

Proof. See Appendix 6.B. ■

6.4.2 Network Disagreement

In this section, we study the variation of agent parameters from the network centroid. To that end, let us incorporate the linear approximation $\widehat{V}(\mu, w) = \phi(\mu)^{\mathsf{T}} w$ into the TD-error expression (6.36) to obtain the following relation:

$$\boldsymbol{\delta}_{k,i} = \boldsymbol{r}_{k,i} + \gamma\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}}\boldsymbol{w}_{k,i} - \phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}}\boldsymbol{w}_{k,i}. \tag{6.52}$$

Since $\nabla_w \widehat{V}(\mu, w) = \phi(\mu)$ for the linear case, it follows that

$$\boldsymbol{z}_{k,i+1} = \Big((1 - 2\rho\alpha)I - \alpha\boldsymbol{H}_{k,i}\Big)\boldsymbol{w}_{k,i} + \alpha\boldsymbol{d}_{k,i}, \tag{6.53}$$

where

$$\boldsymbol{H}_{k,i} \triangleq \phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \gamma\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}}, \tag{6.54}$$

and

$$\boldsymbol{d}_{k,i} \triangleq \boldsymbol{r}_{k,i}\phi(\boldsymbol{\mu}_{k,i}). \tag{6.55}$$

To proceed, we introduce the following regularity assumption on the feature vector.

Assumption 6.2 (**Feature vector**). *The feature mapping $\phi(\mu)$ is bounded and Lipschitz continuous in the domain of the S-dimensional probability simplex. Namely, for any vectors $\mu_1, \mu_2 \in \mathcal{M}(S)$,*

$$\|\phi(\mu_1) - \phi(\mu_2)\| \leq L_\phi \|\mu_1 - \mu_2\|, \quad \|\phi(\mu_1)\| \leq B_\phi. \tag{6.56}$$

Lemma 6.1 (**Belief feature difference**). *For each agent $k \in \mathcal{K}$, the belief feature matrix $\boldsymbol{H}_{k,i}$ in* (6.54) *has bounded expected difference in relation to the centralized belief feature matrix $\boldsymbol{H}_i^\star$, defined below, i.e.,*

$$\mathbb{E}\|\boldsymbol{H}_{k,i} - \boldsymbol{H}_i^\star\| \leq 2B_\phi L_\phi B_{\text{TV}}(1+\gamma), \tag{6.57}$$

where

$$\boldsymbol{H}_i^\star \triangleq \phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\mu}_i)^{\mathsf{T}} - \gamma\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}. \tag{6.58}$$

Proof. See Appendix 6.C. ■

We further assume that all rewards are non-negative and uniformly bounded, i.e., $0 \leq \boldsymbol{r}_{k,i} \leq R_{\max}$ for each agent $k \in \mathcal{K}$, and all time instants i. To study the network disagreement, we define the network centroid as

$$\boldsymbol{w}_{c,i} \triangleq \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{w}_{k,i}, \tag{6.59}$$

which is an average of the parameters of all agents. The following result shows that the agents cluster around this network centroid after sufficient iterations.

Theorem 6.2 (**Network agreement**). *The average distance to the network centroid is bounded for $\rho > \gamma B_\phi L_\phi/\sqrt{2}$ after sufficient number of iterations. In particular, if $\rho \geq 0.75\gamma B_\phi L_\phi$, then*

$$\frac{1}{K}\sum_{k=1}^{K} \mathbb{E}\|\boldsymbol{w}_{k,i} - \boldsymbol{w}_{c,i}\| \leq \frac{\alpha\lambda_2\epsilon}{(1-\lambda_2)} + O(\alpha^2) \tag{6.60}$$

where

$$\epsilon \triangleq R_{\max} B_\phi \Big(\frac{2B_{\text{TV}}(1+\gamma)}{0.08\gamma} + 1 \Big). \tag{6.61}$$

Proof. See Appendix 6.D. ■

Theorem 6.2 states that the parameter estimates by the agents cluster around the network centroid within mean ℓ_2-distance on the order of $O(\alpha\lambda_2)$ in the limit as $i \to \infty$. This result confirms that agents can get arbitrarily close to each other by setting the learning rate α sufficiently small. Besides, dense networks have in general small λ_2, which results in a small disagreement within the network.

6.4.3 Performance of Diffusion Policy Evaluation

We can therefore use the network centroid as a proxy for all agents to show that the disagreement between the fully decentralized strategy of Alg. 6.2 and a baseline strategy that requires a central processor during training is bounded. We start by describing this baseline strategy and explain why it is a more suitable baseline compared to using the fully centralized strategy Alg. 6.1.

In some applications, even though agents are supposed to work in a decentralized fashion once implemented in the field, they can nevertheless rely on central processing during the training phase in order to learn the best policy. In the literature, this paradigm is referred to as *centralized training for decentralized execution* [94, 125]. For our problem, the crucial point is that during training the centralized processor can form beliefs based on all observations, but it should keep in mind that agents will execute their actions based on *local* beliefs once implemented. Therefore, in the baseline strategy, actions and rewards are based on local beliefs as in (6.33)–(6.35), whereas parameter updates are based on the centralized posterior as in (6.20)–(6.21). Algorithm 6.3 lists this baseline procedure. Notice that the algorithm consists of both local belief construction (see (6.62), (6.63), and (6.65)) and centralized belief construction (see (6.64) and (6.66)). The former is used for action execution $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_{k,i})$, while the latter is used for value function parameter updates in (6.67)–(6.68).

In the fully centralized strategy of Alg. 6.1, the actions by the agents and the subsequent rewards are based on the centralized belief. Therefore, the target value function that Alg. 6.1 aims to learn corresponds to the average cumulative reward obtained under centralized execution. In comparison, the target value functions that Algs. 6.2 and 6.3 try to learn are the same and they correspond to the average cumulative reward under decentralized execution. While trying to learn the same parameter w°, the baseline strategy can utilize centralized processing, but the diffusion strategy is fully decentralized. Nonetheless, the following result illustrates that the expected disagreement between the baseline strategy and the fully decentralized strategy remains bounded.

Algorithm 6.3 Centralized evaluation for decentralized execution

set initial priors $\eta_{k,0}(s) > 0, \eta_0(s) > 0$, for $\forall s \in \mathcal{S}$ and $\forall k \in \mathcal{K}$
choose $\beta > 0$
initialize $w_0^\star$
while $i \geq 0$ **do**
  each agent k observes $\boldsymbol{\xi}_{k,i}$
  for each agent $k = 1, 2, \ldots, K$ and $s \in \mathcal{S}$ **adapt** and **combine**

$$\boldsymbol{\psi}_{k,i}(s) \propto (L_k(\boldsymbol{\xi}_{k,i}|s))^\beta \boldsymbol{\eta}_{k,i}(s) \tag{6.62}$$

$$\boldsymbol{\mu}_{k,i}(s) \propto \prod_{\ell \in \mathcal{N}_k} \Big(\boldsymbol{\psi}_{\ell,i}(s)\Big)^{a_{\ell k}} \tag{6.63}$$

  end for
  to form centralized belief with joint observation $\boldsymbol{\xi}_i \triangleq \{\boldsymbol{\xi}_{k,i}\}_{k=1}^K$, **adapt**

$$\boldsymbol{\mu}_i(s) \propto L(\boldsymbol{\xi}_i|s)\boldsymbol{\eta}_i(s) \tag{6.64}$$

  for each agent $k = 1, 2, \ldots, K$ **do**
    take action $\boldsymbol{a}_{k,i} \sim \pi_k(a_k|\boldsymbol{\mu}_{k,i})$
    get reward $\boldsymbol{r}_{k,i} = \boldsymbol{r}_k(\boldsymbol{s}_i, \boldsymbol{a}_i, \boldsymbol{s}_{i+1})$
  end for
  average the rewards $\boldsymbol{r}_i^\star = \frac{1}{K}\sum_{k=1}^K \boldsymbol{r}_{k,i}$
  for each agent $k = 1, 2, \ldots, K$ **evolve**
    Compute $\mathbb{T}_k^\pi(s|s', \boldsymbol{a}_{\mathcal{N}_k,i})$ using (6.24), and

$$\boldsymbol{\eta}_{k,i+1}(s) = \sum_{s' \in \mathcal{S}} \mathbb{T}_k^\pi(s|s', \boldsymbol{a}_{\mathcal{N}_k,i})\boldsymbol{\mu}_{k,i}(s') \tag{6.65}$$

  end for
  evolve the centralized belief

$$\boldsymbol{\eta}_{i+1}(s) = \sum_{s' \in \mathcal{S}} \mathbb{T}(s|s', \boldsymbol{a}_i)\boldsymbol{\mu}_i(s') \tag{6.66}$$

  update value function parameter

$$\boldsymbol{\delta}_i^\star = \boldsymbol{r}_i^\star + \gamma \widehat{V}(\boldsymbol{\eta}_{i+1}, \boldsymbol{w}_i^\star) - \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i^\star) \tag{6.67}$$

$$\boldsymbol{w}_{i+1}^\star = (1 - 2\rho\alpha)\boldsymbol{w}_i^\star + \alpha\boldsymbol{\delta}_i^\star \nabla_w \widehat{V}(\boldsymbol{\mu}_i, \boldsymbol{w}_i^\star) \tag{6.68}$$

  $i \leftarrow i + 1$
end while

Theorem 6.3 (Disagreement with the baseline solution). *The expected distance between the baseline strategy and the network centroid is bounded after sufficient iterations for $\rho > \gamma B_\phi L_\phi/\sqrt{2}$. In particular, if $\rho \geq 0.75\gamma B_\phi L_\phi$, then*

$$\mathbb{E}\|\boldsymbol{w}_i^\star - \boldsymbol{w}_{c,i}\| \leq \frac{B_{\text{TV}} R_{\max} \epsilon'}{0.08\gamma B_\phi L_\phi} \tag{6.69}$$

after $i \geq i_0 = o\left(1/(\alpha\gamma B_\phi L_\phi)\right)$ iterations, where $\epsilon' > 0$ is a constant defined by

$$\epsilon' \triangleq \frac{2B_\phi(1+\gamma)}{0.08\gamma} + L_\phi. \tag{6.70}$$

Proof. See Appendix 6.E. ■

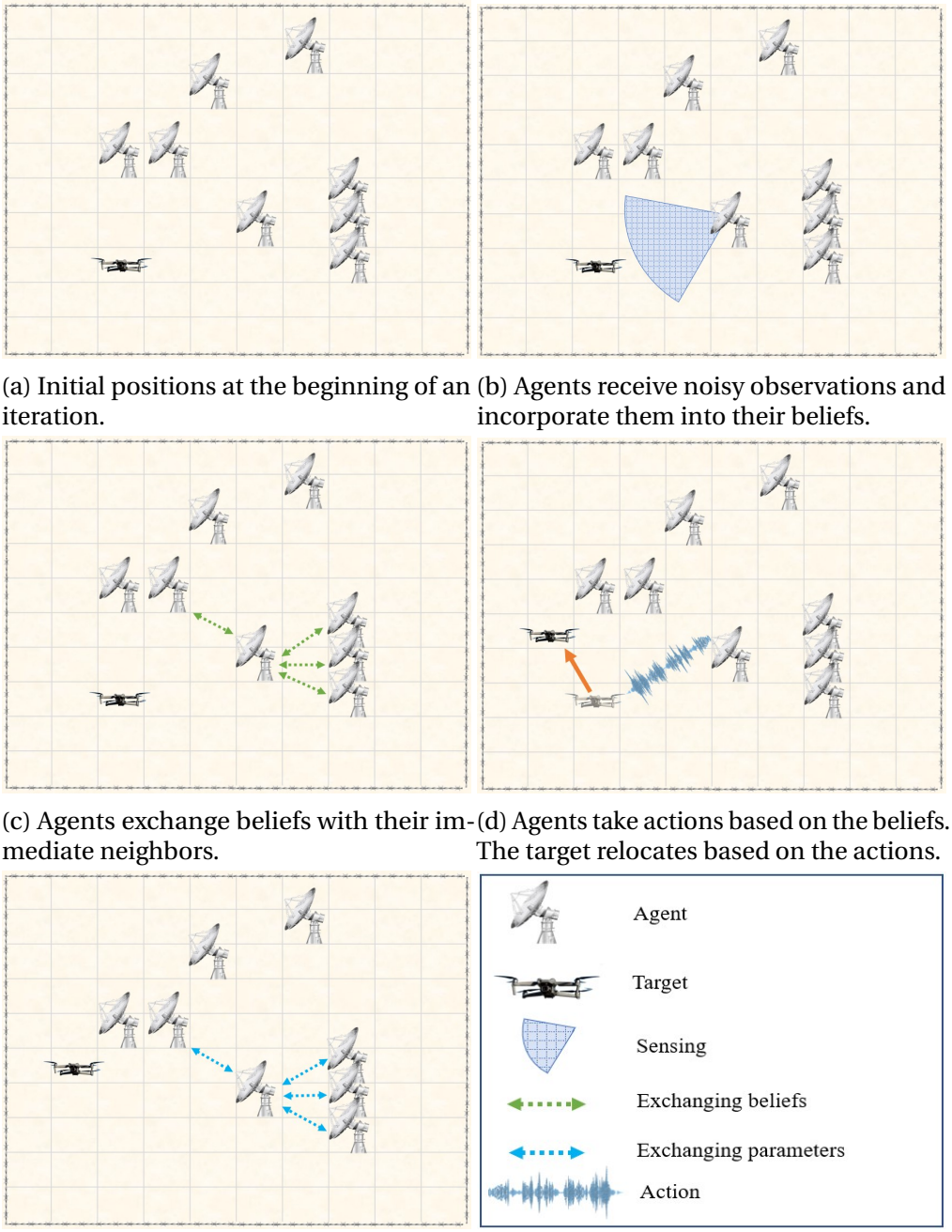
Theorem 6.3 implies that the disagreement between the network centroid, around which agents cluster, and the baseline strategy is on the order of B_{TV}. This means that if the local beliefs are similar to the centralized belief, agents get closer to the baseline parameter. In this regard, from the definition (6.50) for B_{TV}, it can be observed that B_{TV} gets smaller with increasing network connectivity (i.e., decreasing λ_2), as $\beta \to K$. In fact, it is equal to 0 for fully-connected networks with the choice of $\beta = K$ and $c_{\ell k} = 1/K$. Therefore, by changing β and $c_{\ell k}$, the fully decentralized strategy can match the value function estimates of a centralized training strategy that can gather all observations and actions in a fusion center. In the next section, by means of numerical simulations, we further compare the value function estimate accuracies of all Algorithms 6.1, 6.2 and 6.3 by using squared Bellman error (SBE).

6.5 Numerical Simulations

For numerical simulations, we consider a multi-agent target localization application. The implementation is available online[2]. We use a set of $K = 8$ agents and a moving target in a 10×10 two-dimensional grid world environment. The locations of the agents are fixed and their coordinates are randomly assigned at the beginning of the simulation. The goal of the agents is to cooperatively evaluate a given policy for hitting the target. Agents cannot observe the location (i.e., state) of the target accurately, but instead receive noisy observations based on how far they are from the real location of the target. The target is moving according to some pre-defined transition model that takes the actions (i.e., hits) of agents into account. Specifically, the target is trying to evade the hits of agents.

A possible scenario for this setting is a network of sensors and an intruder (e.g., a

[2] github.com/asl-epfl/DecPOMDP_Policy_Evaluation_w-Belief_Sharing



(a) Initial positions at the beginning of an iteration.

(b) Agents receive noisy observations and incorporate them into their beliefs.

(c) Agents exchange beliefs with their immediate neighbors.

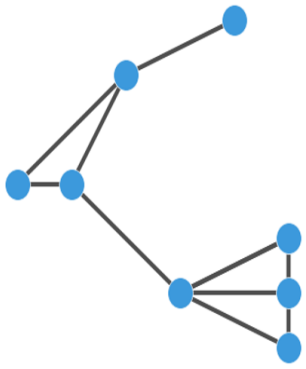
(d) Agents take actions based on the beliefs. The target relocates based on the actions.

(e) Agents update and exchange value function parameters.

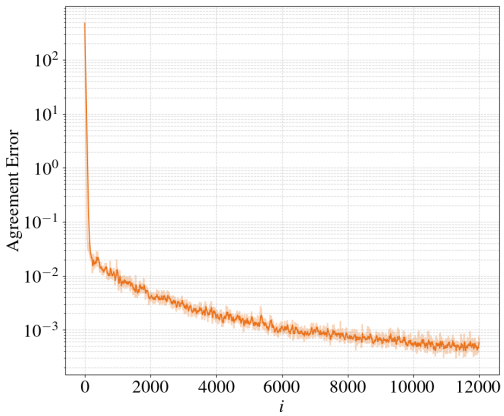
(f) Image credit for agent, target, and action: freepik.com

Figure 6.1: Experimental scenario. For visual purposes, the procedure is shown for only one agent. In fact, all agents execute the same procedure simultaneously.

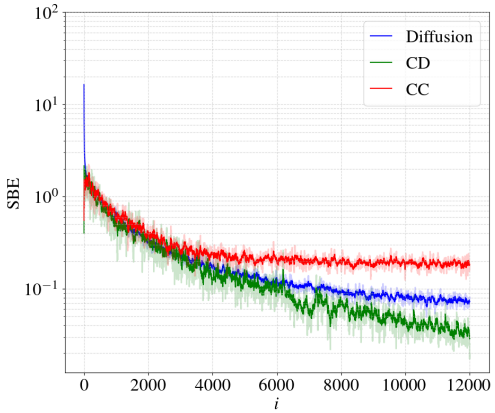
spy drone) — see Fig. 6.1. The sensors try to localize the intruder based on noisy measurements and belief exchanges. Moreover, in order to disrupt the communication between the intruder and its owner, each sensor sends a narrow sector jamming beam towards its target location estimate. However, the intruder is capable of detecting energy abnormalities and determines its next location by favoring distant locations from the jamming signals. We now describe the setting in more detail.



(a) Communication topology.



(b) Agreement error over time.



(c) SBE over time (running window of size 20) for CC (Alg. 6.1), Diffusion (Alg. 6.2), CD (Alg. 6.3).

Figure 6.2: Graph underlying the experimental scenario and the evolution of errors over time.

Combination matrix: The entries of the combination matrix are set such that they are inversely proportional to the ℓ_1-distance between the agents. That is to say, the further the agents are from each other, the smaller the value of the weight that is assigned to the edge connecting them. Weights smaller than some threshold are set to 0, which implies that agents that are too far from each other do not need to communicate. The

		initial position	
	ℓ_1-distance	≤ 4	> 4
average location	< 4	10	5
of agents' hits	≥ 4	100	50

Table 6.1: The table of scores used in the transition model. Each candidate state for next state (location) of the target gets a score based on the initial position of the target and the average action of agents.

resulting communication topology graph is illustrated in Fig. 6.2a.

Transition model: The target is moving between cells in a grid (i.e., states) randomly. The probability of a cell being the next location of the target depends on the current location of the target and the location of the agents' hits. Namely, each state in the grid is assigned a score based on its ℓ_1-distance to the current location of the target and to the average location of the agents' hits — see Table 6.1. For example, observe from Table 6.1 that the cells that are in the proximity of the target's current location and also far away from the agents' strikes are given the highest score. These scores are normalized to yield a probabilistic transition kernel.

Likelihood function: Agents cannot observe where the target is. They can only receive noisy observations. Each agent gets a more accurate observation of the target's position if the target is in close proximity to the agent. Otherwise, the larger the distance between the agent and the target, the higher the noise level. Depending on how close the target is to the agent, and in order to construct the likelihood function, we first assign scores to each cell in the grid that reflect how probable it is to find the target in that cell — see Table 6.2. Following that, the scores are normalized in order to yield a distribution function. For instance, if the target lies at an ℓ_1-distance that is less than 3 grid squares from the location of the agent, the actual position of the target gets a likelihood score of 400, cells within an ℓ_1 distance of 2 grid squares from the agent get a likelihood score of 200, and cells within an ℓ_1 distance of 4 grid squares from the agents get a likelihood score of 30.

Reward function: The reward function in the environment is such that an agent receives a reward of 1 if the agent is able to hit the position of the target. The agent also receives a reward of 0.2 if the ℓ_1-distance between the predicted location and the

		real location of target					
	ℓ_1-distance	$=0$	<3	<5	<7	<9	≥ 9
location	<3	400	200	30	1	1	1
of	≤ 6	200	180	100	1	1	1
agent	>6	25	25	25	25	4	1

Table 6.2: The table of scores used in the likelihood function model. Each state, when observed, gets a score that determines the likelihood of the presence of the target within the state, based on the position of the target and the average action of agents.

actual location of the target is less than 3 grid units. Otherwise, it gets 0 reward. Agents do not know the reward model, and use the instantaneous rewards instead.

Policy: We fix the policy that the agents evaluate as the maximum a-posteriori policy. Namely, agents detect (hit) a location if it corresponds to the maximum entry in their belief vector.

We use the belief vectors as the features directly, i.e., ϕ is an identity transformation. We set $\alpha = 0.1$, $\rho = 0.0001$, and $\beta = K = 8$, and average over 3 different realizations for all cases. In Fig. 6.2b, the average mean-square distance to the network centroid, i.e.,

$$\text{Agreement error} \triangleq \frac{1}{K}\sum_{k=1}^{K}\mathbb{E}\|\boldsymbol{w}_{k,i} - \boldsymbol{w}_{c,i}\|^2, \tag{6.71}$$

is plotted over time for the fully decentralized strategy. Confirming Theorem 6.2, it can be seen that agreement error rapidly decreases and converges to a small value.

In Fig. 6.2c, we plot the evolution of the average squared Bellman error (SBE) in the log domain, where the SBE expression is given by:

$$\text{SBE} \triangleq \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{\delta}_{k,i}^2, \tag{6.72}$$

and similarly for the centralized cases. It measures the network average of instantaneous TD-errors. It can be seen that all approaches converge, and in particular, diffusion strategy (Alg. 6.2) yields a comparable performance with CD (Alg. 6.3). This observation is in line with Theorem 6.3, which states that the disagreement between the fully decentralized strategy and the baseline centralized training for decentralized execution strategy is bounded. Notice also that CC (Alg. 6.1) results in a higher SBE

compared to the diffusion and CD, despite being a fully centralized strategy. This is because, CC evaluates a different policy, namely, the centralized execution policy. Therefore, as argued in Sec. 6.4.3, the SBE of CC is not a suitable baseline for the diffusion strategy.

6.6 Concluding Remarks

In this chapter, we proposed a policy evaluation algorithm for Dec-POMDPs over networks. We carried out a rigorous analysis that established: (i) the beliefs formed with local information and interactions have a bounded disagreement with the global posterior distribution, (ii) agents' value function parameters cluster around the network centroid, and (iii) the decentralized training can match the performance of the centralized training with appropriate parameters and increasing network connectivity.

There are two limitations of the current work that can be addressed in future work. First, we assume that agents know the local likelihood and transition models accurately. One possible question is if agents have approximation errors for the models, how would these affect the analytical results. Second, an implication of Theorem 6.3 is that there is necessity for regularization ($\rho > 0$). We leave the question of whether one can get bounds that does not require this, possibly with more assumptions on the model, to future work.

6.A Proof of Theorem 6.1

We can rewrite the risk function as

$$\begin{aligned} J_{k,i} &= \mathbb{E}_{\mathcal{F}_i} D_{\mathrm{KL}}(\boldsymbol{\mu}_i || \boldsymbol{\mu}_{k,i}) \\ &= \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{s \in \mathcal{S}} \boldsymbol{\mu}_i(s) \log \frac{\boldsymbol{\mu}_i(s)}{\boldsymbol{\mu}_{k,i}(s)} \Big] \\ &\overset{(a)}{=} \mathbb{E}_{\mathcal{F}_i} \Big[\sum_{s \in \mathcal{S}} \mathbb{P}(\boldsymbol{s}_i = s | \boldsymbol{\mathcal{F}}_i) \log \frac{\boldsymbol{\mu}_i(s)}{\boldsymbol{\mu}_{k,i}(s)} \Big] \\ &\overset{(b)}{=} \mathbb{E}_{\mathcal{F}_i} \Big[\mathbb{E}_{s_i | \mathcal{F}_i} \Big(\log \frac{\boldsymbol{\mu}_i(\boldsymbol{s}_i)}{\boldsymbol{\mu}_{k,i}(\boldsymbol{s}_i)} \Big) \Big] \\ &= \mathbb{E}_{\mathcal{F}_i, s_i} \Big[\log \frac{\boldsymbol{\mu}_i(\boldsymbol{s}_i)}{\boldsymbol{\mu}_{k,i}(\boldsymbol{s}_i)} \Big], \end{aligned} \tag{6.73}$$

where (a) follows from definition (6.7), (b) follows from the definition of conditional expectation with respect to $\boldsymbol{s}_i$ given $\boldsymbol{\mathcal{F}}_i$. Merging the diffusion adaptation step (6.33)

and the combination step (6.34) together yields the following form

$$\boldsymbol{\mu}_{k,i}(s) \propto \prod_{\ell \in \mathcal{N}_k} (L_\ell(\boldsymbol{\xi}_{\ell,i}|s))^{\beta c_{\ell k}} (\boldsymbol{\eta}_{\ell,i}(s))^{c_{\ell k}}, \tag{6.74}$$

which, combined with the update equation (6.20) for the centralized solution, results in

$$\begin{aligned} \log \frac{\boldsymbol{\mu}_i(s)}{\boldsymbol{\mu}_{k,i}(s)} = & \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \bigg(\log \frac{L(\boldsymbol{\xi}_i|s)}{(L_\ell(\boldsymbol{\xi}_{\ell,i}|s))^\beta} + \log \frac{\boldsymbol{\eta}_i(s)}{\boldsymbol{\eta}_{\ell,i}(s)} \bigg) \\ & + \log \sum_{s' \in \mathcal{S}} \bigg(\prod_{\ell \in \mathcal{N}_k} (L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}} \prod_{\ell \in \mathcal{N}_k} (\boldsymbol{\eta}_{\ell,i}(s'))^{c_{\ell k}} \bigg) - \log \boldsymbol{m}_i(\boldsymbol{\xi}_i). \end{aligned} \tag{6.75}$$

Here, we introduced the marginal distribution of the new observation given the past observations and actions:

$$\begin{aligned} \boldsymbol{m}_i(\xi_i) \triangleq \mathbb{P}(\boldsymbol{\xi}_i = \xi_i | \boldsymbol{\mathcal{F}}_{i-1}^a) &= \sum_{s \in \mathcal{S}} \mathbb{P}(\boldsymbol{\xi}_i = \xi_i, \boldsymbol{s}_i = s | \boldsymbol{\mathcal{F}}_{i-1}^a) \\ &= \sum_{s \in \mathcal{S}} L(\xi_i|s) \mathbb{P}(\boldsymbol{s}_i = s | \boldsymbol{\mathcal{F}}_{i-1}^a) \\ &= \sum_{s \in \mathcal{S}} L(\xi_i|s) \boldsymbol{\eta}_i(s). \end{aligned} \tag{6.76}$$

Observe that the expectation of the log-likelihood ratio terms in (6.75) satisfies:

$$\begin{aligned} & \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\xi_i, s_i} \bigg[\log \frac{L(\boldsymbol{\xi}_i|\boldsymbol{s}_i)}{(L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{s}_i))^\beta} \bigg] \\ & \quad \overset{(a)}{=} \mathbb{E}_{\xi_i, s_i} \bigg[\sum_{\ell=1}^K \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{s}_i) \bigg] - \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\xi_{\ell,i}, s_i} \Big[\beta \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{s}_i) \Big] \\ & \quad = \mathbb{E}_{\xi_i, s_i} \bigg[\sum_{\ell=1}^K (1 - \beta c_{\ell k}) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{s}_i) \bigg] \end{aligned} \tag{6.77}$$

where in (a) we used the spatial independency of the observations. Likewise, the expectation of the time-adjusted terms in (6.75) can be written as:

$$\begin{aligned} \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\mathcal{F}_i, s_i} \bigg[\log \frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)} \bigg] &\overset{(a)}{=} \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\mathcal{F}_i, s_i} \bigg[\log \frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)} + \log \frac{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)} \bigg] \\ &= \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\mathcal{F}_{i-1}^a, s_i} \bigg[\mathbb{E}_{\xi_i | \mathcal{F}_{i-1}^a, s_i} \bigg(\log \frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)} + \log \frac{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)} \bigg) \bigg] \\ &\overset{(b)}{=} \sum_{\ell \in \mathcal{N}_k} c_{\ell k} \mathbb{E}_{\mathcal{F}_{i-1}^a, s_i} \bigg[\log \frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)} + \log \frac{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)} \bigg] \end{aligned} \tag{6.78}$$

where in (a) we defined the agent-specific distribution:

$$\widetilde{\boldsymbol{\eta}}_{\ell,i}(s) \triangleq \sum_{s'\in\mathcal{S}} \mathbb{T}(s|s',\boldsymbol{a}_{i-1})\boldsymbol{\mu}_{\ell,i-1}(s'), \tag{6.79}$$

and (b) follows from the fact that the arguments are deterministic given the current state and the history of actions and observations. The first term of (6.78) can be written as a KL-divergence because of the following:

$$\begin{aligned}
\sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1},\boldsymbol{s}_i}\left[\log\frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}\right] &= \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1}}\left[\mathbb{E}_{\boldsymbol{s}_i|\boldsymbol{\mathcal{F}}^a_{i-1}}\left(\log\frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}\right)\right] \\
&= \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1}}\left[\sum_{s\in\mathcal{S}}\mathbb{P}(\boldsymbol{s}_i=s|\boldsymbol{\mathcal{F}}^a_{i-1})\log\frac{\boldsymbol{\eta}_i(s)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(s)}\right] \\
&\overset{(6.9)}{=} \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1}}\left[\sum_{s\in\mathcal{S}}\boldsymbol{\eta}_i(s)\log\frac{\boldsymbol{\eta}_i(s)}{\widetilde{\boldsymbol{\eta}}_{\ell,i}(s)}\right] \\
&= \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1}}\left[D_{\mathrm{KL}}(\boldsymbol{\eta}_i||\widetilde{\boldsymbol{\eta}}_{\ell,i})\right].
\end{aligned} \tag{6.80}$$

This expected KL-divergence can be bounded by using the strong-data processing inequality:

$$\sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\boldsymbol{\mathcal{F}}^a_{i-1}}\left[D_{\mathrm{KL}}(\boldsymbol{\eta}_i||\widetilde{\boldsymbol{\eta}}_{\ell,i})\right] \le \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\kappa(\mathbb{T})\underbrace{\mathbb{E}_{\boldsymbol{\mathcal{F}}_{i-1}}\left[D_{\mathrm{KL}}(\boldsymbol{\mu}_{i-1}||\boldsymbol{\mu}_{\ell,i-1})\right]}_{J_{\ell,i-1}}. \tag{6.81}$$

The second term of (6.78) arises due to transition model disagreement with the centralized belief. To bound it, we first introduce the LogSumExp function f with vector arguments $\nu\in\mathbb{R}^S$:

$$f(\nu) \triangleq \log\sum_{s\in\mathcal{S}}\exp\{\nu(s)\}. \tag{6.82}$$

Its gradient is given by

$$\nabla_\nu f(\nu) \triangleq \mathrm{col}\left\{\frac{\partial f(\nu)}{\partial\nu(s)}\right\}_{s\in\mathcal{S}} = \mathrm{col}\left\{\frac{\exp\{\nu(s)\}}{\sum_{s'}\exp\{\nu(s')\}}\right\}_{s\in\mathcal{S}}. \tag{6.83}$$

Observe that if we introduce the vectors

$$\widetilde{\boldsymbol{\nu}}_{\ell,i} \triangleq \mathrm{col}\left\{\log\Big(\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})\boldsymbol{\mu}_{\ell,i-1}(s)\Big)\right\}_{s\in\mathcal{S}} \tag{6.84}$$

and

$$\boldsymbol{\nu}_{\ell,i} \triangleq \mathrm{col}\left\{\log\Big(\mathbb{T}^\pi_\ell(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})\boldsymbol{\mu}_{\ell,i-1}(s)\Big)\right\}_{s\in\mathcal{S}}, \tag{6.85}$$

then, we can rewrite the second expression of (6.78) as follows:

$$\sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left[\log\frac{\widetilde{\boldsymbol{\eta}}_{\ell,i}(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)}\right] = \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\Big[f(\widetilde{\boldsymbol{\nu}}_{\ell,i}) - f(\boldsymbol{\nu}_{\ell,i})\Big]. \tag{6.86}$$

Applying mean value theorem to this difference yields

$$\begin{aligned}
&\mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\Big[f(\widetilde{\boldsymbol{\nu}}_{\ell,i}) - f(\boldsymbol{\nu}_{\ell,i})\Big] \\
&= \mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\Big[(\nabla_\nu f(\overline{\boldsymbol{\nu}}_{\ell,i}))^{\mathsf{T}}\cdot(\widetilde{\boldsymbol{\nu}}_{\ell,i}-\boldsymbol{\nu}_{\ell,i})\Big] \\
&\overset{(6.83)}{=} \mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left[\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}_{s\in\mathcal{S}}^{\mathsf{T}}\cdot(\widetilde{\boldsymbol{\nu}}_{\ell,i}-\boldsymbol{\nu}_{\ell,i})\right] \\
&\overset{(6.84),(6.85)}{=} \mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left[\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}_{s\in\mathcal{S}}^{\mathsf{T}}\cdot\mathrm{col}\Big\{\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_\ell^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})}\Big\}_{s\in\mathcal{S}}\right]
\end{aligned} \tag{6.87}$$

for some $\overline{\boldsymbol{\nu}}_{\ell,i}$ between $\widetilde{\boldsymbol{\nu}}_{\ell,i}$ and $\boldsymbol{\nu}_{\ell,i}$. The term in (6.87) is bounded as follows:

$$\begin{aligned}
&\left|\mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left[\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}_{s\in\mathcal{S}}^{\mathsf{T}}\cdot\mathrm{col}\Big\{\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_\ell^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})}\Big\}_{s\in\mathcal{S}}\right]\right| \\
&\overset{(a)}{\leq} \mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left|\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}_{s\in\mathcal{S}}^{\mathsf{T}}\cdot\mathrm{col}\Big\{\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_\ell^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})}\Big\}_{s\in\mathcal{S}}\right| \\
&\overset{(b)}{\leq} \mathbb{E}_{\mathcal{F}_{i-1}^a,s_i}\left[\left\|\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}_{s\in\mathcal{S}}\right\|_1\cdot\left\|\mathrm{col}\Big\{\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_\ell^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})}\Big\}_{s\in\mathcal{S}}\right\|_\infty\right] \\
&\overset{(c)}{=} \mathbb{E}_{s_i,a_{i-1}}\left\|\mathrm{col}\Big\{\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_\ell^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_\ell,i-1})}\Big\}_{s\in\mathcal{S}}\right\|_\infty
\end{aligned} \tag{6.88}$$

where (a) follows from the Jensen's inequality, (b) follows from the Hölder's inequality, and (c) follows from the fact that

$$\left\|\mathrm{col}\Big\{\frac{\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s)\}}{\sum_{s'\in\mathcal{S}}\exp\{\overline{\boldsymbol{\nu}}_{\ell,i}(s')\}}\Big\}\right\|_1 = 1. \tag{6.89}$$

Furthermore, due to Assumption 6.1 and to the fact that the number of maximum hops outside $\mathcal{N}_k$ is $(K-|\mathcal{N}_k|)$, we have

$$\left|\log\frac{\mathbb{T}(\boldsymbol{s}_i|s,\boldsymbol{a}_{i-1})}{\mathbb{T}_k^\pi(\boldsymbol{s}_i|s,\boldsymbol{a}_{\mathcal{N}_k,i-1})}\right| \leq (K-|\mathcal{N}_k|)\,\tau \leq (K-d_{\min})\,\tau. \tag{6.90}$$

If we combine (6.81), (6.86), and (6.90), the expectation of the time-adjusted terms in (6.75) can be bounded as:

$$\sum_{\ell\in\mathcal{N}_k} c_{\ell k}\mathbb{E}_{\mathcal{F}_{i-1}^a,\boldsymbol{s}_i}\left[\log\frac{\boldsymbol{\eta}_i(\boldsymbol{s}_i)}{\boldsymbol{\eta}_{\ell,i}(\boldsymbol{s}_i)}\right] \le (K-d_{\min})\,\tau + \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\kappa(\mathbb{T})J_{\ell,i-1} \tag{6.91}$$

Next, we bound the expectation of the remaining normalization terms in (6.75), which follows similar steps to what was done in the previous chapter:

$$\begin{aligned}
&\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell\in\mathcal{N}_k}(L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}}\prod_{\ell\in\mathcal{N}_k}(\boldsymbol{\eta}_{\ell,i}(s'))^{c_{\ell k}}\Big)\right]-\mathbb{E}_{\mathcal{F}_i}\Big[\log\boldsymbol{m}_i(\boldsymbol{\xi}_i)\Big]\\
&\overset{(a)}{\le}\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell\in\mathcal{N}_k}(L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}}\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]-\mathbb{E}_{\mathcal{F}_i}\Big[\log\boldsymbol{m}_i(\boldsymbol{\xi}_i)\Big]\\
&=\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell\in\mathcal{N}_k}(L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}}\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]\\
&\quad-\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|s')\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]\\
&\quad+\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|s')\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]-\mathbb{E}_{\mathcal{F}_i}\Big[\log\boldsymbol{m}_i(\boldsymbol{\xi}_i)\Big]\\
&\overset{(b)}{\le}\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}(L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}}\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]\\
&\quad-\mathbb{E}_{\mathcal{F}_i}\left[\log\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|s')\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)\right]
\end{aligned}\tag{6.92}$$

where (a) follows from the arithmetic-geometric mean inequality, (b) follows from:

$$\begin{aligned}
-\mathbb{E}_{\mathcal{F}_i}\left[\log\frac{\boldsymbol{m}_i(\boldsymbol{\xi}_i)}{\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|s')\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big)}\right]&=-\mathbb{E}_{\mathcal{F}_{i-1}^a}\mathbb{E}_{\boldsymbol{\xi}_i|\mathcal{F}_{i-1}^a}\left[\log\frac{\boldsymbol{m}_i(\boldsymbol{\xi}_i)}{\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i)}\right]\\
&=-\mathbb{E}_{\mathcal{F}_{i-1}^a}D_{\mathrm{KL}}(\boldsymbol{m}_i(\boldsymbol{\xi}_i)||\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i))\\
&\le 0
\end{aligned}\tag{6.93}$$

where we used the definition:

$$\boldsymbol{m}_i^\dagger(\boldsymbol{\xi}_i)\triangleq\sum_{s'\in\mathcal{S}}\Big(\prod_{\ell=1}^{K}L_\ell(\boldsymbol{\xi}_{\ell,i}|s')\sum_{\ell\in\mathcal{N}_k}c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s')\Big),\tag{6.94}$$

which is a density (or mass function if observations are discrete) since

$$\begin{aligned}\int_{\xi_i} \boldsymbol{m}_i^{\dagger}(\xi_i)d\xi_i &= \int_{\xi_i} \sum_{s'\in\mathcal{S}} \Big(\prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|s') \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big) d\xi_i \\ &= \sum_{s'\in\mathcal{S}} \Big[\underbrace{\int_{\xi_i} \prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|s')d\xi_i}_{1} \sum_{\ell=1}^{K} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big] \\ &= \sum_{s'\in\mathcal{S}} \Big[\sum_{\ell=1}^{K} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big] \\ &= \sum_{\ell=1}^{K} c_{\ell k} \Big[\sum_{s'\in\mathcal{S}} \boldsymbol{\eta}_{\ell,i}(s') \Big] = 1. \end{aligned} \tag{6.95}$$

Notice that the expression in (6.92) can be rewritten as

$$\begin{aligned}&\mathbb{E}_{\mathcal{F}_i} \Bigg[\log \sum_{s'\in\mathcal{S}} \Big(\prod_{\ell=1}^{K} (L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}} \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big) \Bigg] \\ &\qquad - \mathbb{E}_{\mathcal{F}_i} \Bigg[\log \sum_{s'\in\mathcal{S}} \Big(\prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|s') \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big) \Bigg] \\ &= \mathbb{E}_{\mathcal{F}_i} \Big[f(\boldsymbol{\vartheta}_{k,i}) \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[f(\widetilde{\boldsymbol{\vartheta}}_{k,i}) \Big], \end{aligned} \tag{6.96}$$

if we use the LogSumExp function f from (6.82) and use the definitions:

$$\boldsymbol{\vartheta}_{k,i} \triangleq \mathrm{col}\Big\{ \log \Big(\prod_{\ell=1}^{K} (L_\ell(\boldsymbol{\xi}_{\ell,i}|s'))^{\beta c_{\ell k}} \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big) \Big\}_{s'\in\mathcal{S}} \tag{6.97}$$

and

$$\widetilde{\boldsymbol{\vartheta}}_{k,i} \triangleq \mathrm{col}\Big\{ \log \Big(\prod_{\ell=1}^{K} L_\ell(\boldsymbol{\xi}_{\ell,i}|s') \sum_{\ell\in\mathcal{N}_k} c_{\ell k}\boldsymbol{\eta}_{\ell,i}(s') \Big) \Big\}_{s'\in\mathcal{S}}. \tag{6.98}$$

Following the steps in (6.87) and (6.88), this difference can be bounded as:

$$\mathbb{E}_{\mathcal{F}_i} \Big[f(\boldsymbol{\vartheta}_{k,i}) \Big] - \mathbb{E}_{\mathcal{F}_i} \Big[f(\widetilde{\boldsymbol{\vartheta}}_{k,i}) \Big] \leq \mathbb{E}_{\xi_i} \Big\| \mathrm{col}\Big\{ \sum_{\ell=1}^{K} (\beta c_{\ell k} - 1) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|s') \Big\}_{s'\in\mathcal{S}} \Big\|_\infty. \tag{6.99}$$

Again, similar to the previous chapter, by assumptions on the graph topology and on

the likelihood functions (6.41), this expression can be further bounded as:

$$\left\| \mathrm{col}\Big\{ \sum_{\ell=1}^{K} (\beta c_{\ell k} - 1) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|s') \Big\}_{s' \in \mathcal{S}} \right\|_\infty \leq \sqrt{K}\beta\lambda B \tag{6.100}$$

Subsequently, if we insert the bounds (6.77), (6.91), and (5.105) to (6.75), we arrive at the bound on the risk function:

$$\begin{aligned} J_{k,i} &\leq \mathbb{E}_{\xi_i, s_i} \Big[\sum_{\ell=1}^{K} (1 - \beta c_{\ell k}) \log L_\ell(\boldsymbol{\xi}_{\ell,i}|\boldsymbol{s}_i) \Big] \\ &\quad + \kappa(\mathbb{T}) \sum_{\ell \in \mathcal{N}_k} c_{\ell k} J_{\ell,i-1} + \sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau \\ &\overset{(5.105)}{\leq} \kappa(\mathbb{T}) \sum_{\ell \in \mathcal{N}_k} c_{\ell k} J_{\ell,i-1} + 2\sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau. \end{aligned} \tag{6.101}$$

Expanding this recursion over time yields:

$$\begin{aligned} J_{k,i} &\leq (2\sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau) \sum_{j=0}^{i-1} (\kappa(\mathbb{T}))^j + (\kappa(\mathbb{T}))^i \sum_{\ell=1}^{K} [C^i]_{\ell k} J_{\ell,0} \\ &= \frac{1 - (\kappa(\mathbb{T}))^i}{1 - \kappa(\mathbb{T})} (2\sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau) + (\kappa(\mathbb{T}))^i \sum_{\ell=1}^{K} [C^i]_{\ell k} J_{\ell,0}, \end{aligned} \tag{6.102}$$

which implies that if $\kappa(\mathbb{T}) < 1$, the risk function is bounded as $i \to \infty$:

$$\limsup_{i \to \infty} J_{k,i} \leq \frac{2\sqrt{K}\beta\lambda B + (K - d_{\min})\,\tau}{1 - \kappa(\mathbb{T})}. \tag{6.103}$$

By (6.91), this also implies that

$$\begin{aligned} \limsup_{i \to \infty} \widetilde{J}_{k,i} &\leq (K - d_{\min})\tau + \kappa(\mathbb{T}) \limsup_{i \to \infty} J_{k,i} \\ &\leq \frac{(K - d_{\min})\tau}{1 - \kappa(\mathbb{T})} + \kappa(\mathbb{T}) \frac{2\sqrt{K}\beta\lambda B}{1 - \kappa(\mathbb{T})}. \end{aligned} \tag{6.104}$$

6.B Proof of Corollary 6.1

In view of the Bretagnolle-Huber inequality [126], it holds that

$$\sum_{s \in \mathcal{S}} \Big| \boldsymbol{\mu}_i(s) - \boldsymbol{\mu}_{k,i}(s) \Big| \leq 2\Big(1 - \exp\{-D_{\mathrm{KL}}(\boldsymbol{\mu}_i || \boldsymbol{\mu}_{k,i})\}\Big)^{\frac{1}{2}}. \tag{6.105}$$

If we take the expectation of both sides, we get

$$\begin{aligned}\mathbb{E}\Big[\sum_{s\in\mathcal{S}}\Big|\boldsymbol{\mu}_i(s)-\boldsymbol{\mu}_{k,i}(s)\Big|\Big] &\leq 2\mathbb{E}\Big(1-\exp\{-D_{\mathrm{KL}}(\boldsymbol{\mu}_i||\boldsymbol{\mu}_{k,i})\}\Big)^{\frac{1}{2}}\\ &\overset{(a)}{\leq} 2\Big(1-\mathbb{E}\exp\{-D_{\mathrm{KL}}(\boldsymbol{\mu}_i||\boldsymbol{\mu}_{k,i})\}\Big)^{\frac{1}{2}}\\ &\overset{(b)}{\leq} 2\Big(1-\exp\{-J_{k,i}\}\Big)^{\frac{1}{2}},\end{aligned} \tag{6.106}$$

where (a) and (b) follow from Jensen's inequality. Together with Theorem 6.1, this implies that

$$\mathbb{E}\Big\|\boldsymbol{\mu}_i-\boldsymbol{\mu}_{k,i}\Big\|_1 \leq B_{\mathrm{TV}}, \tag{6.107}$$

where we use the definition (6.50). Furthermore, on account of the fact that the ℓ_2-norm of a vector is upper bounded by the ℓ_1-norm in $\mathbb{R}^S$, it is also true that

$$\mathbb{E}\Big\|\boldsymbol{\mu}_i-\boldsymbol{\mu}_{k,i}\Big\| \leq B_{\mathrm{TV}}. \tag{6.108}$$

With similar arguments, it can be shown that

$$\mathbb{E}\Big\|\boldsymbol{\eta}_i-\boldsymbol{\eta}_{k,i}\Big\| \leq \widetilde{B}_{\mathrm{TV}}, \tag{6.109}$$

where we use the definition (6.51).

6.C Proof of Lemma 6.1

Inserting the definitions (6.54) and (6.58), the expected difference can be expanded as

$$\begin{aligned}\mathbb{E}\|\boldsymbol{H}_{k,i}-\boldsymbol{H}_i^{\star}\| &= \mathbb{E}\Big\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}}-\gamma\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}}\\ &\qquad -\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\mu}_i)^{\mathsf{T}}+\gamma\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}\Big\|\\ &\leq \mathbb{E}\Big\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}}-\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\mu}_i)^{\mathsf{T}}\Big\|\\ &\qquad +\gamma\mathbb{E}\Big\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}}-\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}\Big\|,\end{aligned} \tag{6.110}$$

where the last step follows from the triangle inequality. The first term can be bounded as

$$\begin{aligned}&\Big\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}}-\phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\mu}_i)^{\mathsf{T}}\Big\|\\ &\quad\leq \Big\|\phi(\boldsymbol{\mu}_{k,i})(\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}}-\phi(\boldsymbol{\mu}_i)^{\mathsf{T}})\Big\|+\Big\|(\phi(\boldsymbol{\mu}_{k,i})-\phi(\boldsymbol{\mu}_i))\phi(\boldsymbol{\mu}_i)^{\mathsf{T}}\Big\|\\ &\quad\leq \Big\|\phi(\boldsymbol{\mu}_{k,i})\Big\|\Big\|\phi(\boldsymbol{\mu}_{k,i})-\phi(\boldsymbol{\mu}_i)\Big\|+\Big\|\phi(\boldsymbol{\mu}_{k,i})-\phi(\boldsymbol{\mu}_i)\Big\|\Big\|\phi(\boldsymbol{\mu}_i)\Big\|\end{aligned}$$

$$\overset{(a)}{\leq} B_\phi L_\phi \|\boldsymbol{\mu}_{k,i} - \boldsymbol{\mu}_i\| + B_\phi L_\phi \|\boldsymbol{\mu}_{k,i} - \boldsymbol{\mu}_i\|, \tag{6.111}$$

where (a) follows from Assumption 6.2. Taking expectations and using (6.48) and (6.111), it follows that

$$\mathbb{E}\left\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\mu}_i)^{\mathsf{T}}\right\| \leq 2B_\phi L_\phi B_{\mathrm{TV}}. \tag{6.112}$$

Similarly, the second term in (6.110) can be bounded as

$$\begin{aligned}
&\left\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} - \phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}\right\| \\
&\quad\leq \left\|\phi(\boldsymbol{\mu}_{k,i})(\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} - \phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}})\right\| + \left\|(\phi(\boldsymbol{\mu}_{k,i}) - \phi(\boldsymbol{\mu}_i))\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}\right\| \\
&\quad\leq \left\|\phi(\boldsymbol{\mu}_{k,i})\right\|\left\|\phi(\boldsymbol{\eta}_{k,i+1}) - \phi(\boldsymbol{\eta}_{i+1})\right\| + \left\|\phi(\boldsymbol{\mu}_{k,i}) - \phi(\boldsymbol{\mu}_i)\right\|\left\|\phi(\boldsymbol{\eta}_{i+1})\right\| \\
&\quad\overset{(a)}{\leq} B_\phi L_\phi \|\boldsymbol{\eta}_{k,i+1} - \boldsymbol{\eta}_{i+1}\| + B_\phi L_\phi \|\boldsymbol{\mu}_{k,i} - \boldsymbol{\mu}_i\|
\end{aligned} \tag{6.113}$$

where (a) follows from Assumption 6.2. Using (6.48) and (6.49) we get:

$$\mathbb{E}\left\|\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} - \phi(\boldsymbol{\mu}_i)\phi(\boldsymbol{\eta}_{i+1})^{\mathsf{T}}\right\| \leq B_\phi L_\phi (B_{\mathrm{TV}} + \widetilde{B}_{\mathrm{TV}}). \tag{6.114}$$

Combining (6.112) and (6.114) in addition to the fact that $\widetilde{B}_{\mathrm{TV}} \leq B_{\mathrm{TV}}$ (since $\kappa(\mathbb{T}) < 1$) yields:

$$\mathbb{E}\|\boldsymbol{H}_{k,i} - \boldsymbol{H}_i^\star\| \leq 2B_\phi L_\phi B_{\mathrm{TV}}(1+\gamma). \tag{6.115}$$

6.D Proof of Theorem 6.2

For compactness of notation, it is useful to introduce the following quantities, which collect variables from across all agents:

$$\boldsymbol{\mathcal{W}}_i \triangleq \mathrm{col}\,\{\boldsymbol{w}_{1,i}, \ldots, \boldsymbol{w}_{K,i}\} \tag{6.116}$$

$$\mathcal{C} \triangleq C \otimes I_M \tag{6.117}$$

$$\boldsymbol{\mathcal{H}}_i \triangleq \mathrm{diag}\,\{\boldsymbol{H}_{k,i}\}_{k=1}^K \tag{6.118}$$

$$\boldsymbol{\mathcal{H}}_i^\star \triangleq I_K \otimes \boldsymbol{H}_i^\star \tag{6.119}$$

$$\boldsymbol{d}_i \triangleq \mathrm{col}\,\{\boldsymbol{d}_{k,i}\}_{k=1}^K \tag{6.120}$$

Then, equations (6.36)–(6.38) can be written as:

$$\boldsymbol{\mathcal{W}}_{i+1} = \mathcal{C}^{\mathsf{T}}\Big((I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i)\,\boldsymbol{\mathcal{W}}_i + \alpha\boldsymbol{d}_i \Big). \tag{6.121}$$

We define the K-times extended centroid vector:

$$\boldsymbol{\mathcal{w}}_{c,i} \triangleq \mathbb{1}_K \otimes \boldsymbol{w}_{c,i} = \left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\boldsymbol{\mathcal{w}}_i\,. \tag{6.122}$$

If we decompose $\boldsymbol{\mathcal{H}}_i$ into its centralized component $\boldsymbol{\mathcal{H}}_i^\star$ and the respective disagreement matrix $\boldsymbol{\Delta}_i \triangleq \boldsymbol{\mathcal{H}}_i - \boldsymbol{\mathcal{H}}_i^\star$, we obtain:

$$\begin{aligned}
&\boldsymbol{\mathcal{w}}_{i+1} - \boldsymbol{\mathcal{w}}_{c,i+1} \\
&\quad= \left(\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\Big(\left(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i\right)\boldsymbol{\mathcal{w}}_i + \alpha\boldsymbol{\mathcal{d}}_i\Big) \\
&\quad= \left(\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\Big(\left(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star - \alpha\boldsymbol{\Delta}_i\right)\boldsymbol{\mathcal{w}}_i + \alpha\boldsymbol{\mathcal{d}}_i\Big) \\
&\quad= \left(\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\Big(\left(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\right)\left(\boldsymbol{\mathcal{w}}_i - \boldsymbol{\mathcal{w}}_{c,i}\right) - \alpha\boldsymbol{\Delta}_i\,\boldsymbol{\mathcal{w}}_i + \alpha\boldsymbol{\mathcal{d}}_i\Big),
\end{aligned} \tag{6.123}$$

where the last step follows from the fact that

$$\mathcal{C}^{\mathsf{T}}\Big(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\Big)\boldsymbol{\mathcal{w}}_{c,i} = \left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\Big(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\Big)\boldsymbol{\mathcal{w}}_{c,i}\,. \tag{6.124}$$

Taking the norms of both sides in (6.123) leads to

$$\begin{aligned}
&\Big\|\boldsymbol{\mathcal{w}}_{i+1} - \boldsymbol{\mathcal{w}}_{c,i+1}\Big\| \\
&\le \left\|\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right\| \times \Big\|\left(I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\right)\left(\boldsymbol{\mathcal{w}}_i - \boldsymbol{\mathcal{w}}_{c,i}\right) - \alpha\boldsymbol{\Delta}_i\,\boldsymbol{\mathcal{w}}_i + \alpha\boldsymbol{\mathcal{d}}_i\Big\| \\
&\le \left\|\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right\|\,\|I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\|\|\,\boldsymbol{\mathcal{w}}_i - \boldsymbol{\mathcal{w}}_{c,i}\,\| \\
&\qquad + \alpha\left\|\mathcal{C}^{\mathsf{T}} - \frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right\|\Big(\|\boldsymbol{\Delta}_i\|\|\,\boldsymbol{\mathcal{w}}_i\,\| + \|\boldsymbol{\mathcal{d}}_i\|\Big).
\end{aligned} \tag{6.125}$$

Since the combination matrix C is a primitive stochastic matrix, it follows from the Perron-Frobenius theorem [7, 11] that its maximum eigenvalue is 1, and all other eigenvalues are strictly smaller than 1 in absolute value. Moreover, C is assumed to be symmetric, therefore its eigenvalue decomposition has the following form:

$$\begin{aligned}
C &= U\Lambda U^\top \\
&= \begin{bmatrix} u_1 & u_2 & \cdots & u_K \end{bmatrix}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{bmatrix}\begin{bmatrix} u_1^\top \\ u_2^\top \\ \vdots \\ u_K^\top \end{bmatrix}
\end{aligned} \tag{6.126}$$

where U is a unitary matrix of eigenvectors $\{u_k\}$, and Λ is the diagonal matrix of eigenvalues. Additionally, the powers of C converge (because it is primitive) to the

scaled all-ones matrix (because it is doubly-stochastic):

$$\lim_{i\to\infty} C^i = \begin{bmatrix} u_1 & u_2 & \cdots & u_K \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} u_1^\top \\ u_2^\top \\ \vdots \\ u_K^\top \end{bmatrix}$$
$$= \frac{1}{K} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \frac{1}{K} \mathbb{1}_K \mathbb{1}_K^\top \tag{6.127}$$

Therefore, the difference of these matrices becomes:

$$C - \frac{1}{K} \mathbb{1}_K \mathbb{1}_K^\top = U \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_K \end{bmatrix} U^\top, \tag{6.128}$$

which implies:

$$\left\| C - \frac{1}{K} \mathbb{1}_K \mathbb{1}_K^\top \right\| = \lambda_2 \tag{6.129}$$

where λ_2 is the second largest modulus eigenvalue of C. Moreover, the Kronecker product with the identity matrix does not change the spectral norm, hence:

$$\left\| \mathcal{C}^\mathsf{T} - \frac{1}{K} \mathbb{1}_K \mathbb{1}_K^\mathsf{T} \otimes I \right\| = \lambda_2 < 1. \tag{6.130}$$

Moreover, we know from Lemma 6.1 that

$$\mathbb{E}\|\boldsymbol{\Delta}_i\| \leq 2B_\phi L_\phi B_{\mathrm{TV}}(1+\gamma). \tag{6.131}$$

Additionally, in Appendix 6.F, we establish (6.132)–(6.135) which hold for any realization (with probability 1). From (6.158), note that:

$$\left\| I(1-2\alpha\rho) - \alpha \boldsymbol{\mathcal{H}}_i^\star \right\| < 1 \tag{6.132}$$

whenever $\rho > \gamma L_\phi B_\phi / \sqrt{2}$. Specifically, if $\rho \geq 0.75\gamma L_\phi B_\phi$, then

$$\left\| I(1-2\alpha\rho) - \alpha \boldsymbol{\mathcal{H}}_i^\star \right\| \leq (1 - 0.08\alpha\gamma L_\phi B_\phi). \tag{6.133}$$

In addition, we show in Lemma 6.2 that

$$\|\boldsymbol{\mathcal{W}}_i\| \leq \frac{\sqrt{K}R_{\max}}{0.08\gamma L_\phi} \tag{6.134}$$

and in expression (6.159) that

$$\|\boldsymbol{d}_i\| \leq \sqrt{K}R_{\max}B_\phi. \tag{6.135}$$

Inserting these results into (6.125) yields the following norm recursion:

$$\mathbb{E}\Big\|\,\boldsymbol{\mathcal{W}}_{i+1}-\boldsymbol{\mathcal{W}}_{c,i+1}\,\Big\| \leq \lambda_2(1-0.08\alpha\gamma B_\phi L_\phi)\mathbb{E}\Big\|\,\boldsymbol{\mathcal{W}}_i-\boldsymbol{\mathcal{W}}_{c,i}\,\Big\| + \alpha\lambda_2\sqrt{K}\epsilon. \tag{6.136}$$

Let us define the constant $\widetilde{\lambda}_2 \triangleq \lambda_2(1-0.08\alpha\gamma B_\phi L_\phi)$. Iterating (6.136) over time, we arrive at

$$\begin{aligned}
\mathbb{E}\,\|\boldsymbol{\mathcal{W}}_{i+1}-\boldsymbol{\mathcal{W}}_{c,i+1}\| &\leq \widetilde{\lambda}_2^{i+1}\|\,\mathcal{W}_0-\mathcal{W}_{c,0}\,\| + \alpha\lambda_2\sqrt{K}\epsilon\sum_{j=1}^{i+1}\widetilde{\lambda}_2^{i+1-j} \\
&\leq \widetilde{\lambda}_2^{i+1}\|\,\mathcal{W}_0-\mathcal{W}_{c,0}\,\| + \alpha\lambda_2\sqrt{K}\epsilon\frac{1}{1-\widetilde{\lambda}_2} \\
&\overset{(a)}{\leq} \alpha\lambda_2\sqrt{K}\epsilon\frac{1}{1-\widetilde{\lambda}_2} + O(\alpha^2)
\end{aligned} \tag{6.137}$$

where (a) holds whenever:

$$\begin{aligned}
&\widetilde{\lambda}_2^{i}\|\,\mathcal{W}_0-\mathcal{W}_{c,0}\,\| \leq c\alpha^2 \\
&\iff i\log\widetilde{\lambda}_2 \leq 2\log\alpha + \log c - \log\|\,\mathcal{W}_0-\mathcal{W}_{c,0}\,\| \\
&\iff i \geq \frac{2\log\alpha}{\log\widetilde{\lambda}_2} + O(1) = O(\log\alpha) = o(1/\alpha),
\end{aligned} \tag{6.138}$$

where c is an arbitrary constant.

6.E Proof of Theorem 6.3

Similar to [127–130], we begin by rewriting the baseline strategy recursion (6.67)–(6.68) in the form:

$$\boldsymbol{w}_{i+1}^\star = ((1-2\rho\alpha)I - \alpha\boldsymbol{H}_i^\star)\,\boldsymbol{w}_i^\star + \alpha\boldsymbol{d}_i^\star, \tag{6.139}$$

where $\boldsymbol{H}_i^\star$ is defined in (6.58), and

$$\boldsymbol{d}_i^\star \triangleq \Big(\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}_{k,i}\Big)\phi(\boldsymbol{\mu}_i). \tag{6.140}$$

We introduce the K-times extended versions of the vectors:

$$\boldsymbol{\mathcal{D}}_i^\star \triangleq \mathbb{1}_K \otimes \boldsymbol{d}_i^\star, \quad \boldsymbol{\mathcal{W}}_i^\star = \mathbb{1}_K \otimes \boldsymbol{w}_i^\star. \tag{6.141}$$

Then, the baseline recursion (6.139) transforms into

$$\boldsymbol{\mathcal{W}}_{i+1}^\star = ((1-2\rho\alpha)I - \alpha\boldsymbol{\mathcal{H}}_i^\star)\,\boldsymbol{\mathcal{W}}_i^\star + \alpha\boldsymbol{\mathcal{D}}_i^\star. \tag{6.142}$$

It follows from the extended network centroid definition (6.122) and (6.142) that

$$\begin{aligned}\boldsymbol{\mathcal{W}}_{i+1}^\star - \boldsymbol{\mathcal{W}}_{c,i+1} &= (I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star)\,(\boldsymbol{\mathcal{W}}_i^\star - \boldsymbol{\mathcal{W}}_{c,i}) \\ &\quad - \alpha\left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\boldsymbol{\Delta}_i\,\boldsymbol{\mathcal{W}}_i + \alpha\left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)(\boldsymbol{\mathcal{D}}_i^\star - \boldsymbol{d}_i)\end{aligned} \tag{6.143}$$

where we used the facts that

$$\left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\boldsymbol{\mathcal{D}}_i^\star = \boldsymbol{\mathcal{D}}_i^\star, \tag{6.144}$$

and

$$\left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)\boldsymbol{\mathcal{H}}_i^\star\,\boldsymbol{\mathcal{W}}_i = \boldsymbol{\mathcal{H}}_i^\star\,\boldsymbol{\mathcal{W}}_{c,i}\,. \tag{6.145}$$

Next, if we define the following average agent disagreement relative to the baseline term

$$\widetilde{\boldsymbol{d}}_i \triangleq \frac{1}{K}\sum_{k=1}^{K}(\boldsymbol{d}_i^\star - \boldsymbol{d}_{k,i}), \tag{6.146}$$

it holds that

$$\widetilde{\boldsymbol{\mathcal{D}}}_i \triangleq \mathbb{1}_K \otimes \widetilde{\boldsymbol{d}}_i = \left(\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right)(\boldsymbol{\mathcal{D}}_i^\star - \boldsymbol{d}_i). \tag{6.147}$$

Subsequently, taking the norm of both sides in (6.143) and applying the triangle inequality, we get

$$\begin{aligned}&\left\|\,\boldsymbol{\mathcal{W}}_{i+1}^\star - \boldsymbol{\mathcal{W}}_{c,i+1}\,\right\| \\ &\quad\le \left\|I(1-2\alpha\rho) - \alpha\boldsymbol{\mathcal{H}}_i^\star\right\|\left\|\,\boldsymbol{\mathcal{W}}_i^\star - \boldsymbol{\mathcal{W}}_{c,i}\,\right\| + \alpha\left\|\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right\|\left\|\boldsymbol{\Delta}_i\right\|\left\|\,\boldsymbol{\mathcal{W}}_i\,\right\| + \alpha\left\|\widetilde{\boldsymbol{\mathcal{D}}}_i\right\|.\end{aligned} \tag{6.148}$$

First, observe that

$$\left\|\frac{1}{K}\mathbb{1}_K\mathbb{1}_K^{\mathsf{T}} \otimes I\right\| = 1. \tag{6.149}$$

Moreover, from Assumption 6.2 and Corollary 6.1, it holds that

$$\mathbb{E}\|\widetilde{\boldsymbol{d}}_i\| = \mathbb{E}\Big\|\frac{1}{K}\sum_{k=1}^{K}\boldsymbol{r}_{k,i}(\phi(\boldsymbol{\mu}_i)-\phi(\boldsymbol{\mu}_{k,i}))\Big\| \leq R_{\max}L_\phi B_{\mathrm{TV}}, \tag{6.150}$$

and accordingly,

$$\mathbb{E}\Big\|\widetilde{\boldsymbol{\mathcal{D}}}_i\Big\| \leq \sqrt{K}R_{\max}L_\phi B_{\mathrm{TV}}. \tag{6.151}$$

By using the same bounds (6.132)–(6.135) from Appendix 6.D for the other terms (which are established in Lemmas 6.1 and 6.2, and equations (6.158) and (6.159)), we arrive at the recursion:

$$\mathbb{E}\Big\|\,\boldsymbol{\mathcal{W}}^\star_{i+1}-\boldsymbol{\mathcal{W}}_{c,i+1}\,\Big\| \leq (1-0.08\alpha\gamma B_\phi L_\phi)\mathbb{E}\Big\|\,\boldsymbol{\mathcal{W}}^\star_{i}-\boldsymbol{\mathcal{W}}_{c,i}\,\Big\| + \alpha\sqrt{K}\epsilon^\star, \tag{6.152}$$

where

$$\epsilon^\star \triangleq R_{\max}B_{\mathrm{TV}}\Big(\frac{2B_\phi(1+\gamma)}{0.08\gamma}+L_\phi\Big). \tag{6.153}$$

Iterating over time, we get:

$$\begin{aligned}\mathbb{E}\Big\|\boldsymbol{\mathcal{W}}^\star_{i+1}-\boldsymbol{\mathcal{W}}_{c,i+1}\Big\| &\leq (1-0.08\alpha\gamma B_\phi L_\phi)^{i+1}\|\,\mathcal{W}^\star_0-\mathcal{W}_{c,0}\,\| + \alpha\sqrt{K}\epsilon^\star\sum_{j=1}^{i+1}(1-0.08\alpha\gamma B_\phi L_\phi)^{i+1-j}\\ &\leq (1-0.08\alpha\gamma B_\phi L_\phi)^{i+1}\|\,\mathcal{W}^\star_0-\mathcal{W}_{c,0}\,\| + \frac{\sqrt{K}\epsilon^\star}{0.08\gamma B_\phi L_\phi}\\ &\overset{(a)}{\leq} \frac{\sqrt{K}\epsilon^\star}{0.08\gamma B_\phi L_\phi}+o(1)\end{aligned} \tag{6.154}$$

where (a) holds whenever

$$\begin{aligned}&(1-0.08\alpha\gamma B_\phi L_\phi)^{i+1}\Big\|\,\mathcal{W}^\star_0-\mathcal{W}_{c,0}\,\Big\| = o(1)\\ &\quad\Longleftrightarrow i\log(1-0.08\alpha\gamma B_\phi L_\phi) = o(1)\\ &\quad\Longleftrightarrow i \geq \frac{o(1)}{\log(1-0.08\alpha\gamma B_\phi L_\phi)} \geq o\left(\frac{1}{\alpha\gamma B_\phi L_\phi}\right).\end{aligned} \tag{6.155}$$

6.F Auxiliary Results

In the following lemma, we prove that the value function parameters are bounded in norm.

Lemma 6.2 (Bounded parameters). *For each agent $k \in \mathcal{K}$, the iterate $\boldsymbol{w}_{k,i}$ is bounded in norm if $\rho > \gamma B_\phi L_\phi/\sqrt{2}$, with probability 1. In particular, if $\rho \geq 0.75\gamma B_\phi L_\phi$, then*

$$\|\boldsymbol{\mathcal{W}}_i\| \leq \frac{\sqrt{K}R_{\max}}{0.08\gamma L_\phi} \tag{6.156}$$

after $i \geq i_0 = o\left(1/(\alpha\gamma B_\phi L_\phi)\right)$ iterations.

Proof. Taking the norms of both sides of (6.121) yields:

$$\begin{aligned}
\Big\| \boldsymbol{\mathcal{W}}_{i+1} \Big\| &= \Big\| \mathcal{C}^{\mathsf{T}} \Big(\left((1-2\alpha\rho)I - \alpha\boldsymbol{\mathcal{H}}_i\right) \boldsymbol{\mathcal{W}}_i + \alpha \boldsymbol{\mathcal{D}}_i \Big) \Big\| \\
&\leq \Big\| \mathcal{C}^{\mathsf{T}} \Big\| \Big\| \left((1-2\alpha\rho)I - \alpha\boldsymbol{\mathcal{H}}_i\right) \boldsymbol{\mathcal{W}}_i + \alpha \boldsymbol{\mathcal{D}}_i \Big\| \\
&\overset{(a)}{\leq} \Big\| \left((1-2\alpha\rho)I - \alpha\boldsymbol{\mathcal{H}}_i\right) \boldsymbol{\mathcal{W}}_i + \alpha \boldsymbol{\mathcal{D}}_i \Big\| \\
&\leq \Big\| (1-2\alpha\rho)I - \alpha\boldsymbol{\mathcal{H}}_i \Big\| \Big\| \boldsymbol{\mathcal{W}}_i \Big\| + \alpha \Big\| \boldsymbol{\mathcal{D}}_i \Big\|
\end{aligned} \tag{6.157}$$

where (a) follows from the fact $\mathcal{C}$ is doubly stochastic and symmetric with spectral radius equal to 1. Note that

$$\begin{aligned}
&\Big\| (1-2\alpha\rho)I - \alpha\boldsymbol{\mathcal{H}}_{k,i} \Big\| \\
&= \Big\| (1-2\alpha\rho)I - \alpha\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} + \alpha\gamma\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} \Big\| \\
&= \Big\| (1-2\alpha\rho)I - \alpha(1-\gamma)\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \alpha\gamma\phi(\boldsymbol{\mu}_{k,i})\Big(\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}}\Big) \Big\| \\
&\leq \Big\| (1-2\alpha\rho)I - \alpha(1-\gamma)\phi(\boldsymbol{\mu}_{k,i})\phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} \Big\| + \alpha\gamma\|\phi(\boldsymbol{\mu}_{k,i})\| \Big\| \phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} \Big\| \\
&\overset{(a)}{\leq} (1-2\alpha\rho) + \alpha\gamma\|\phi(\boldsymbol{\mu}_{k,i})\| \Big\| \phi(\boldsymbol{\mu}_{k,i})^{\mathsf{T}} - \phi(\boldsymbol{\eta}_{k,i+1})^{\mathsf{T}} \Big\| \\
&\overset{(b)}{\leq} (1-2\alpha\rho) + \alpha\gamma B_\phi L_\phi \Big\| \boldsymbol{\mu}_{k,i} - \boldsymbol{\eta}_{k,i+1} \Big\| \\
&\overset{(c)}{\leq} (1-2\alpha\rho) + \alpha\gamma B_\phi L_\phi \sqrt{2}
\end{aligned} \tag{6.158}$$

where (a) follows from the equality of spectral norm and maximum eigenvalue for symmetric matrices, (b) follows from Assumption 6.2, and (c) follows from the fact that the mean-square distance cannot exceed 2 over the probability simplex. The upper bound in (6.158) is smaller than 1 whenever $\rho > \gamma B_\phi L_\phi/\sqrt{2}$. Moreover,

$$\Big\| \boldsymbol{d}_{k,i} \Big\| = \Big\| \boldsymbol{r}_{k,i}\phi(\boldsymbol{\mu}_{k,i}) \Big\| \leq R_{\max} B_\phi. \tag{6.159}$$

As a result, if $\rho \geq 0.75\gamma B_\phi L_\phi$, we get:

$$\begin{aligned}\left\| \boldsymbol{w}_{i+1} \right\| &\overset{(6.158)}{\leq} (1 - 0.08\alpha\gamma B_\phi L_\phi)\left\| \boldsymbol{w}_i \right\| + \alpha\left\| \boldsymbol{d}_i \right\| \\ &\overset{(6.159)}{\leq} (1 - 0.08\alpha\gamma B_\phi L_\phi)\left\| \boldsymbol{w}_i \right\| + \alpha\sqrt{K}R_{\max}B_\phi. \end{aligned} \tag{6.160}$$

Iterating this recursion starting from $i = 0$ results in

$$\begin{aligned}\left\| \boldsymbol{w}_{i+1} \right\| &\leq \alpha\sqrt{K}R_{\max}B_\phi \sum_{j=1}^{i+1}(1 - 0.08\alpha\gamma B_\phi L_\phi)^{i+1-j} + (1 - 0.08\alpha\gamma B_\phi L_\phi)^{i+1}\left\| w_0 \right\| \\ &\leq \frac{\sqrt{K}R_{\max}}{0.08\gamma L_\phi} + (1 - 0.08\alpha\gamma B_\phi L_\phi)^{i+1}\left\| \boldsymbol{w}_0 \right\| \\ &= \frac{\sqrt{K}R_{\max}}{0.08\gamma L_\phi} + o(1), \end{aligned} \tag{6.161}$$

where the last step holds whenever

$$\begin{aligned}&(1 - 0.08\alpha\gamma B_\phi L_\phi)^{i+1}\left\| w_0 \right\| = o(1) \\ &\quad\iff i\log(1 - 0.08\alpha\gamma B_\phi L_\phi) = o(1) \\ &\quad\iff i \geq \frac{o(1)}{\log(1 - 0.08\alpha\gamma B_\phi L_\phi)} \geq o\left(\frac{1}{\alpha\gamma B_\phi L_\phi}\right). \end{aligned} \tag{6.162}$$

■

Part III Causal Influence

7 Causality in Social Networks

7.1 Introduction[1]

In this chapter, we aim to understand the influential agents in networked interactions. This a crucial problem with wide-ranging applications, such as detecting propaganda-sharing accounts [132] or selecting individuals to advertise to [133]. However, over any social network, information usually spreads and mixes, leading to ripple effects that make the discovery of influence rather challenging. For example, the information leaving a source agent may be altered and combined with comments/data from other agents along the path until it reaches its destination.

Most prior works measure influence through some network topology-based properties such as the eigenvector centrality of an agent [134], or through some descriptive importance factor depending on the problem at hand [135–137]. In comparison, this work treats influence as a *causal* quantity and approaches it from the perspective of structural causal models [138, 139]. More specifically, influence will be defined as the change in behavior of the network when interventions occur at individual agents. This is a useful method to discard spurious and non-causal associations, unlike other methods based, for example, on the use of Granger causality [140]. Obviously, conducting interventional experiments may not be always feasible over real world social networks. However, with the help of appropriate representative models, one can rely on the use of raw observational data [141].

To that end, in this chapter, we utilize the social learning models introduced earlier to examine *causal effects* over social graphs. To the best of our knowledge, this appears to be the first study to do so by studying the diffusion of influence over space and time from a causal inference perspective.

[1]The material in this chapter is based on [131].

7.1.1 Contributions

- We propose a novel causal impact criterion for dynamic social networks in Sec. 7.3. It quantifies how much an agent m affects another agent k for each agent pair (m, k).

- We derive in Secs. 7.4.1 and 7.4.2 closed-form expressions for these impact factors for two social network models: geometric non-Bayesian social learning (Algorithm 2.1) and adaptive social learning (Algorithm 2.3). Beyond general causal influence expressions, we also analyze useful special cases to illustrate how the causal effects depend on the model parameters.

- Analyzing the influence for every pair (m, k) results in a network causality matrix, which offers various options to rank agents for their overall influence. In Sec. 7.5, we propose a particular algorithm, CausalRank, to rank agents based on their overall influence on the network, while accounting for the fact that more impactful agents should be weighted more. This algorithm leverages the eigenvector centrality of the bipartite causal relations matrix introduced in Sec. 7.3.

- Additionally, we introduce a graph causality learning (GCL) algorithm in Sec. 7.6, designed to estimate causal influences using observational data comprised of social interactions and the graph topology underlying the agents.

- In Sec. 7.7.1, we illustrate our results with synthetic data. In Sec. 7.7.2, we apply these findings to real social media data, demonstrating their practical usefulness.

7.2 Challenges for Estimating Social Influence

Influence estimation over social networks faces several challenges.

Confounding factors. Understanding influence requires disentangling correlation from causation. For example, over a social network, it is often observed that individuals who are connected tend to have similar (correlated) opinions. However, this does not necessarily imply a causal relationship and there can also exist confounding factors. For instance, individuals may obtain information from similar external sources (say, the same TV channels), or they may be connected to others who share similar preferences (a.k.a. homophily) — see left panel in Fig. 7.1. Similar issues arise in distributed decision-making systems, such as networks of wireless sensors or robots. Devices that communicate with each other are often in spatial proximity, leading to correlated observations. Therefore, accounting for these confounding factors is crucial for discovering *true causal* relationships.

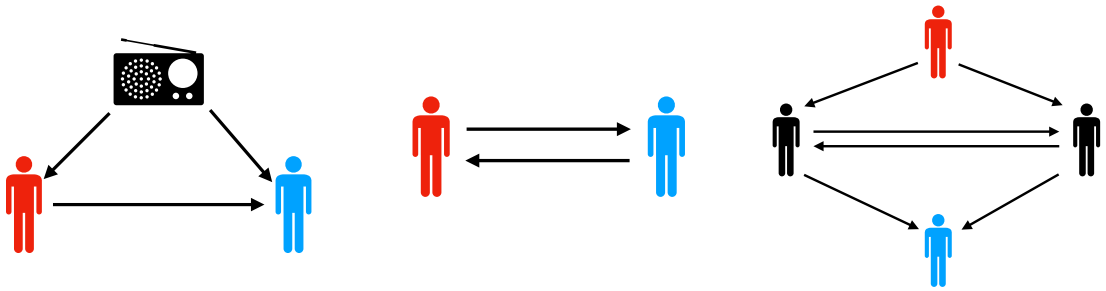


Figure 7.1: Some inherent challenges in estimating the causal influence of one agent on another, represented by the red and blue agents, respectively. *(Left):* There can be confounding factors that influence both parties and induce non-causal correlation. *(Middle):* Relationships within social networks are often bidirectional that are not instantaneous, but rather spread out over time. *(Right):* Information transmitted from the source is processed and potentially changed along the way.

Temporal dynamics. Social influence is not a one-time occurrence but rather a continuous process that unfolds over time. Therefore, we adopt a time-series approach to capture this dynamic nature. Unlike directed acyclic graph based models in the literature, we accommodate both cyclic networks and bidirectional links to capture feedback mechanisms — see middle panel in Fig. 7.1. Doing so is necessary in order to discover the propagation of influence over both space and time.

Mixing and diffusion of information. When examining the influence of an agent m on another agent k, it is essential to acknowledge that information leaving agent m can undergo alterations and become intertwined with the opinions of other agents in the network before reaching agent k — see right panel in Fig. 7.1. This phenomenon introduces additional complexity to the study of influence propagation over graphs.

In this work, the social network models we consider allow us to treat these challenges together in some detail. Specifically, the expressions we derive within a rigorous causal theoretical framework quantify how the instantaneous and direct social effects diffuse over time and space. In doing so, as we seek to understand the total and overall causal effects, we take into account potential hidden confounding factors.

7.3 Causal Effects in Social Learning

An intuitive and widely used assertion in defining causality is that manipulation of the causes should result in changes in the effect [142]. Based on this principle, interventions on a system, real or hypothetical, have been the primary tool for testing whether a variable causes another [139, Chapter 1]. In this work, in order to measure the causal influence strength between agents, we rely on the most basic intervention known as

atomic and *persistent* intervention [139, Chapter 3].

Specifically, in order to measure the social effect of an agent m on other agents, we analyze the belief evolution of these other agents if the belief of agent m is fixed to some particular constant belief vector, say, $\boldsymbol{\mu}_{m,i} := \mu_m$ for all time instants — see Fig. 7.2 for a visual depiction. In canonical causality notation, this intervention is denoted by $do(\boldsymbol{\mu}_{m,i} := \mu_m)$ [139].

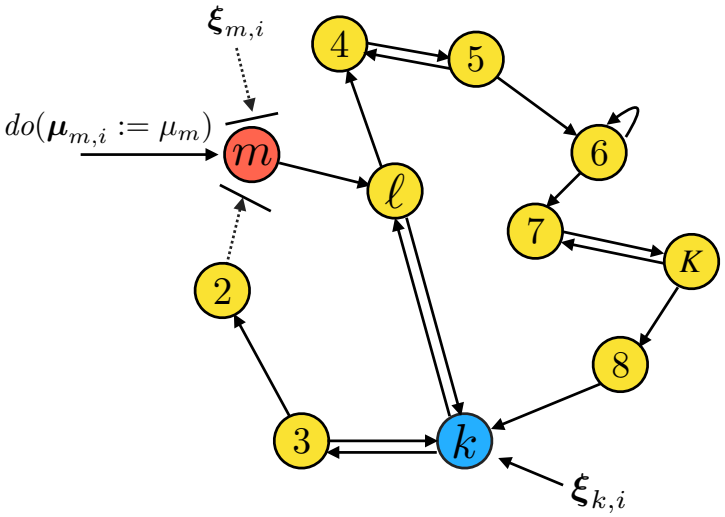


Figure 7.2: An illustration of an atomic and persistent intervention on agent $m = 1$.

Since we consider only this intervention in this work and there is no room for ambiguity, we will use the notation that the post-intervention counterparts of the variables in Chapter 2 are topped with the symbol '$\sim$'. For example, the log-belief ratio definition from (2.35) transforms into the following, under the intervention $do(\boldsymbol{\mu}_{m,i} := \mu_m)$:

$$\widetilde{\boldsymbol{\lambda}}_{k,i}(\theta) \triangleq \log \frac{\widetilde{\boldsymbol{\mu}}_{k,i}(\theta^\circ)}{\widetilde{\boldsymbol{\mu}}_{k,i}(\theta)}. \tag{7.1}$$

Causal influence strength. Intuitively, the amount of change in the effect following an intervention on the cause is expected to be related to the causal strength. Therefore, the difference between the post and pre-intervention distributions, or between appropriate functions of these distributions such as expectations, can be used to quantify the causal effect [139, Chapter 3]. In this work, we employ the following definition in order to measure the causal influence of agent m on agent k:

$$\boxed{C_{m\to k} \triangleq \mu_{k,\infty}(\theta^\circ) - \widetilde{\mu}_{k,\infty}(\theta^\circ)} \tag{7.2}$$

This formula measures the alteration of agent k as a consequence of an intervention on agent m. Specifically, it quantifies the magnitude of change of the expected asymptotic

belief of agent k on the true state θ°. In general, its value depends on the belief μ_m of the intervention $do(\boldsymbol{\mu}_{m,i} := \mu_m)$. Here, we use the following expression for the belief vector, which is explained in the sequel:

$$\mu_{k,\infty}(\theta^\circ) \triangleq \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\{-\lambda_{k,\infty}(\theta)\}}. \tag{7.3}$$

This expression is defined in terms of the expected asymptotic log-belief ratio:

$$\lambda_{k,\infty}(\theta) \triangleq \lim_{i\to\infty} \mathbb{E}[\boldsymbol{\lambda}_{k,i}(\theta)] \tag{7.4}$$

The variables topped with the symbol '$\sim$' for the intervention case are defined similarly (the existence of the limit for both NBSL and ASL under interventions will be discussed in the sequel). The transformation (7.3) is motivated by noting from (2.35) that

$$\exp\Big\{-\boldsymbol{\lambda}_{k,i}(\theta)\Big\} = \frac{\boldsymbol{\mu}_{k,i}(\theta)}{\boldsymbol{\mu}_{k,i}(\theta^\circ)}, \tag{7.5}$$

which implies

$$1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\Big\{-\boldsymbol{\lambda}_{k,i}(\theta)\Big\} = \frac{1}{\boldsymbol{\mu}_{k,i}(\theta^\circ)} \sum_{\theta\in\Theta} \boldsymbol{\mu}_{k,i}(\theta) = \frac{1}{\boldsymbol{\mu}_{k,i}(\theta^\circ)}, \tag{7.6}$$

which, in turn, yields

$$\boldsymbol{\mu}_{k,i}(\theta^\circ) = \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\{-\boldsymbol{\lambda}_{k,i}(\theta)\}}. \tag{7.7}$$

Here, if we replace log-belief ratio $\boldsymbol{\lambda}_{k,i}(\theta)$ with the expected asymptotic log-belief ratio $\lambda_{k,\infty}(\theta)$, we arrive at the definition (7.3) for $\mu_{k,\infty}(\theta^\circ)$. Defining $\mu_{k,\infty}(\theta^\circ)$ in terms of the expected log-belief ratios, as opposed to, say, expected beliefs (i.e., $\lim_{i\to\infty}\mathbb{E}[\boldsymbol{\mu}_{k,i}(\theta^\circ)]$), will enable us to obtain closed-form expressions for causality in terms of the informativeness of the agents, represented by the entries of $d(\theta)$, in Sec. 7.4. Next, we treat NBSL and ASL separately.

- **Non-Bayesian social learning**. In the idle case (i.e., no intervention) of NBSL, it is known from Theorem 2.1 that under global identifiability, $\boldsymbol{\mu}_{k,i}(\theta^\circ) \xrightarrow{\text{a.s.}} 1$ (i.e., $\boldsymbol{\lambda}_{k,i}(\theta) \xrightarrow{\text{a.s.}} +\infty$ for each $\theta \neq \theta^\circ$). Hence, for the NBSL case, the average (i.e., expected) causal influence (7.2) is given by

  $$C^{\text{NB}}_{m\to k} = 1 - \widetilde{\mu}_{k,\infty}(\theta^\circ). \tag{7.8}$$

  This immediately implies that $C^{\text{NB}}_{m\to k} \in [0,1]$, and it gets larger as the post-intervention belief diverges from the truth in expectation.

- **Adaptive social learning**. In a similar fashion, the causal effect strength for the

ASL case is given by the general expression (7.2)

$$C_{m\to k}^{\text{ASL}} = \mu_{k,\infty}(\theta^\circ) - \widetilde{\mu}_{k,\infty}(\theta^\circ), \tag{7.9}$$

where the pre-intervention asymptotic belief $\mu_{k,\infty}(\theta^\circ)$ can be found by inserting $\lambda_{k,\infty}(\theta)$ established in Corollary 2.1 in (7.3).

Remark 7.1 (Controllability). *The influence $C_{m\to k}$ of agent m on agent k can also be interpreted as the controllability or manipulability of agent k by agent m.*

7.4 Theoretical Results

In this section, we derive closed-form expressions for $\widetilde{\lambda}_{k,\infty} \triangleq [\widetilde{\lambda}_{k,\infty}(\theta_1), \ldots, \widetilde{\lambda}_{k,\infty}(\theta_H)]^{\mathsf{T}}$ in terms of the network topology and the informativeness of agents to obtain the causal strength measures $C_{m\to k}^{\text{NB}}$ and $C_{m\to k}^{\text{ASL}}$. For ease of notation and without loss of generality, we set $m = 1$. One can obtain $C_{m\to k}^{\text{NB}}$ by setting $\beta = 1$ and $\delta \to 0$ in $C_{m\to k}^{\text{ASL}}$ due to (2.7) and (2.16). Nevertheless, we first present the analysis for NBSL since it is easier to derive and provides useful insights for ASL. Subsequently, we provide the results for ASL with proofs deferred to the appendix.

7.4.1 Non-Bayesian Social Learning

The intervention $do(\boldsymbol{\mu}_{1,i} := \mu_1)$ ceases (or obstructs the use of) all incoming information at agent 1 from the neighbors $\mathcal{N}_1$ and the use of the streaming observations $\boldsymbol{\xi}_{1,i}$ from the environment. Consequently, we can model this effect by redefining the combination matrix and the LLR vector counterparts under the intervention:

$$\widetilde{A} \triangleq [\widetilde{a}_{\ell k}], \quad \widetilde{a}_{\ell k} \triangleq \begin{cases} 1, & \ell = k = 1 \\ 0, & \ell \neq k = 1 \\ a_{\ell k}, & \ell \neq 1, k \neq 1 \end{cases}, \quad \widetilde{\boldsymbol{x}}_i(\theta) \triangleq [0, \boldsymbol{x}_{2,i}(\theta), \ldots, \boldsymbol{x}_{K,i}(\theta)]^{\mathsf{T}}. \tag{7.10}$$

Observe that the effective combination matrix $\widetilde{A}$ can be obtained from A as follows:

$$A = \left[\begin{array}{c|c} a_{11} & r^{\mathsf{T}} \\ \hline a_{21} & \\ \vdots & R \\ a_{K1} & \end{array}\right] \Longrightarrow \widetilde{A} = \left[\begin{array}{c|c} 1 & r^{\mathsf{T}} \\ \hline 0 & R \\ & \end{array}\right], \tag{7.11}$$

for a $(K-1)\times 1$ dimensional vector r and and a $(K-1)\times(K-1)$ dimensional matrix R:

$$r \triangleq [a_{12} \quad a_{13} \quad \dots \quad a_{1K}]^{\mathsf{T}}, \qquad R \triangleq \begin{bmatrix} a_{22} & \dots & a_{2K} \\ \vdots & \ddots & \vdots \\ a_{K2} & \dots & a_{KK} \end{bmatrix}. \tag{7.12}$$

The matrix structure of $\widetilde{A}$ belongs to the class of *reducible* combination matrices [10, 143] that arise in the analysis of *weakly* connected networks [144–147]. As opposed to the strongly connected networks where information can flow thoroughly, in weakly connected networks information can flow only in one direction between certain parts of the network. In the current context, this corresponds to the fact that information continues to flow from agent m in the form of its belief vector fixed at μ_m, but no information is flowing into it in the sense that agent m ignores all signals arriving from its neighbors and does not use them to update its local belief. However, in contrast to these prior works that analyze opinion dynamics under weakly connected networks, we are interested in the effect of the intervention on the original *strongly* connected network. This alters the LLR $\boldsymbol{x}_i(\theta)$ as well — see (7.10). Similar to the original case in (2.36), we proceed by studying the log-belief ratio evolution that results from using $\widetilde{A}$:

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) = \widetilde{A}^{\mathsf{T}}(\widetilde{\boldsymbol{\lambda}}_{i-1}(\theta) + \widetilde{\boldsymbol{x}}_i(\theta)). \tag{7.13}$$

Recursive application of (7.13) across i iterations to the log-belief ratio $\widetilde{\boldsymbol{\lambda}}_i(\theta)$ yields

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) = \sum_{j=1}^{i} (\widetilde{A}^{i-j+1})^{\mathsf{T}} \widetilde{\boldsymbol{x}}_j(\theta) + (\widetilde{A}^{i})^{\mathsf{T}} \widetilde{\boldsymbol{\lambda}}_0(\theta). \tag{7.14}$$

To study (7.14), we need to evaluate the powers of the effective combination matrix:

$$\widetilde{A}^{i} = \left[\begin{array}{c|c} 1 & r_i'^{\mathsf{T}} \\ \hline 0 & R^i \end{array}\right], \qquad \widetilde{A}^{\infty} \overset{(a)}{=} \left[\begin{array}{c|ccc} 1 & 1 & \dots & 1 \\ \hline 0 & & 0 & \end{array}\right] \tag{7.15}$$

where (a) follows from the fact that the spectral radius of the matrix R is strictly smaller than 1, i.e., R is a stable matrix [146, Lemma 1]. For each time i and $0 < j \le i$, observe that

$$(\widetilde{A}^{i-j+1})^{\mathsf{T}} \widetilde{\boldsymbol{x}}_j(\theta) \overset{(7.10),(7.15)}{=} \left[\begin{array}{c|c} 1 & 0 \\ \hline r'_{i-j+1} & R^{i-j+1\mathsf{T}} \end{array}\right] \begin{bmatrix} 0 \\ \boldsymbol{x}_{2,j}(\theta) \\ \vdots \\ \boldsymbol{x}_{K,j}(\theta) \end{bmatrix} \tag{7.16a}$$

$$\implies [(\widetilde{A}^{i-j+1})^{\mathsf{T}} \widetilde{\boldsymbol{x}}_j(\theta)]_1 \;=\; 0 \tag{7.16b}$$

and

$$(\widetilde{A}^i)^{\mathsf{T}}\widetilde{\boldsymbol{\lambda}}_0(\theta) \overset{(7.10),(7.15)}{=} \left[\begin{array}{c|c} 1 & 0 \\ \hline r_i' & R^{i\mathsf{T}} \end{array}\right] \begin{bmatrix} \log\dfrac{\mu_1(\theta^\circ)}{\mu_1(\theta)} \\ \log\dfrac{\boldsymbol{\mu}_{2,0}(\theta^\circ)}{\boldsymbol{\mu}_{2,0}(\theta)} \\ \vdots \\ \log\dfrac{\boldsymbol{\mu}_{K,0}(\theta^\circ)}{\boldsymbol{\mu}_{K,0}(\theta)} \end{bmatrix} \tag{7.17a}$$

$$\Longrightarrow [(\widetilde{A}^i)^{\mathsf{T}}\widetilde{\boldsymbol{\lambda}}_0(\theta)]_1 \;=\; \log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)} \tag{7.17b}$$

Inserting these into (7.14) verifies that the intervention is fixed as intended for all time instants, since

$$[\widetilde{\boldsymbol{\lambda}}_i(\theta)]_1 \overset{(7.14)}{=} \sum_{j=1}^{i}[(\widetilde{A}^{i-j+1})^{\mathsf{T}}\widetilde{\boldsymbol{x}}_j(\theta)]_1 + [(\widetilde{A}^i)^{\mathsf{T}}\widetilde{\boldsymbol{\lambda}}_0(\theta)]_1. \tag{7.18a}$$

$$\widetilde{\boldsymbol{\lambda}}_{1,i}(\theta) \overset{(7.16b),(7.17b)}{=} \log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}, \quad \widetilde{\boldsymbol{\mu}}_{1,i}(\theta) = \mu_1(\theta). \tag{7.18b}$$

Moreover, if we take the expectation of both sides of (7.14), we get

$$\begin{aligned}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_i(\theta)] &= \sum_{j=1}^{i}(\widetilde{A}^{i-j+1})^{\mathsf{T}}\mathbb{E}[\widetilde{\boldsymbol{x}}_j(\theta)] + (\widetilde{A}^i)^{\mathsf{T}}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_0(\theta)] \\ &= \sum_{j=1}^{i}(\widetilde{A}^{i-j+1})^{\mathsf{T}}\widetilde{d}(\theta) + (\widetilde{A}^i)^{\mathsf{T}}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_0(\theta)].\end{aligned} \tag{7.19}$$

where we are using the definition $\widetilde{d}(\theta) \triangleq [0, d_2(\theta), \ldots, d_K(\theta)]^{\mathsf{T}}$. Hence, in the limit (the existence is guaranteed by the finiteness of LLRs and positive initial beliefs), it holds that

$$\begin{aligned}\lim_{i\to\infty}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_i(\theta)] &= \lim_{i\to\infty}\sum_{j=1}^{i}(\widetilde{A}^{i-j+1})^{\mathsf{T}}\widetilde{d}(\theta) + (\widetilde{A}^{\infty})^{\mathsf{T}}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_0(\theta)] \\ &= \sum_{j=1}^{\infty}(\widetilde{A}^{j})^{\mathsf{T}}\widetilde{d}(\theta) + (\widetilde{A}^{\infty})^{\mathsf{T}}\mathbb{E}[\widetilde{\boldsymbol{\lambda}}_0(\theta)].\end{aligned} \tag{7.20}$$

If we incorporate (7.15) into (7.20), this implies for the log-belief ratios of all agents except agent $m = 1$ that

$$\widetilde{\lambda}_{-m,\infty}(\theta) \triangleq [\widetilde{\lambda}_{2,\infty}(\theta), \ldots, \widetilde{\lambda}_{K,\infty}(\theta)]^{\mathsf{T}} = \sum_{j=1}^{\infty}(R^j)^{\mathsf{T}}d_{-m}(\theta) + \left(\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right)\mathbb{1}_{K-1} \tag{7.21}$$

where $d_{-m}(\theta)$ is the $(K-1)\times 1$ dimensional vector of local KL divergences of the remaining agents, i.e., $d_{-m}(\theta) \triangleq \text{col}\{d_\ell(\theta)\}_{\ell=2}^{K}$. Since R is a stable matrix, Eq. (7.21)

can alternatively be written as

$$\boxed{\widetilde{\lambda}_{-m,\infty}(\theta) = \Big((I - R^{\mathsf{T}})^{-1} - I\Big)d_{-m}(\theta) + \left(\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right)\mathbb{1}_{K-1}} \tag{7.22}$$

The causal influence of agent $m = 1$ on agent k can now be calculated by inserting $\widetilde{\lambda}_{k,\infty}(\theta)$ into (7.3) to find $\widetilde{\mu}_{k,\infty}(\theta)$, which is the input for (7.8) that yields $C^{\text{NB}}_{m\to k}$. More specifically, if we incorporate (7.22) into (7.3), we get

$$\widetilde{\mu}_{k,\infty}(\theta^\circ) = \frac{1}{1 + \sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}} \dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\left\{ - \left[\Big((I - R^{\mathsf{T}})^{-1} - I\Big)d_{-m}(\theta)\right]_k\right\}} \tag{7.23}$$

which, by (7.8), implies

$$C^{\text{NB}}_{m\to k} = 1 - \frac{1}{1 + \sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}} \dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\left\{ - \left[\Big((I - R^{\mathsf{T}})^{-1} - I\Big)d_{-m}(\theta)\right]_k\right\}}. \tag{7.24}$$

Equation (7.22) is a general result which shows that $C^{\text{NB}}_{m\to k}$ is a function of (i) the combination weights (via R), and (ii) the individual informativeness of each agent (via $d_{-m}(\theta)$).

Remark 7.2 (Generalization to sub-networks). *Note that expressions* (7.22)-(7.24) *can generalize to the influence of a sub-network with multiple agents rather than an individual agent* m*. In this case, upon intervening on all agents within the sub-network, the effective combination matrix becomes*

$$\widetilde{A} = \left[\begin{array}{c|c} I & r^{\mathsf{T}} \\ \hline 0 & R \end{array}\right] \tag{7.25}$$

where the first entry in (7.11) *is replaced by an identity matrix. Namely, if the sub-network under consideration has size* s*, the identity,* r *and* R *matrices would be of dimensions* $s\times s$*,* $(K-s)\times s$*, and* $(K-s)\times(K-s)$*, respectively. Similarly,* $d_{-m}(\theta)$ *can be replaced with local KL divergences of the agents that do not belong to the treated sub-network.*

Remark 7.3 (Finite-time spread of influence). *Expressions* (7.22)-(7.24) *reveal how the total overall effects depend on direct instantaneous effects in temporal networked interactions. For the spread of direct effects in a finite time instant* n*, we can modify* (7.21) *and* (7.22) *as*

$$\widetilde{\lambda}_{-m,n}(\theta) \triangleq [\widetilde{\lambda}_{2,n}(\theta),\ldots,\widetilde{\lambda}_{K,n}(\theta)]^{\mathsf{T}} = \sum_{j=1}^{n}(R^j)^{\mathsf{T}} d_{-m}(\theta) + \left(\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right)\mathbb{1}_{K-1} \tag{7.26}$$

due to (7.14)*, and consequently,*

$$\widetilde{\lambda}_{-m,n}(\theta) = \left((I-R^{\mathsf{T}})^{-1}(I-R^{n+1\,\mathsf{T}})-I\right) d_{-m}(\theta) + \left(\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right)\mathbb{1}_{K-1} \tag{7.27}$$

For ease of the presentation, we continue to describe the total causal effects of a single agent, even though our results can be extended to sub-network influence and finite time analysis in a straightforward manner as discussed in the above remarks.

In (7.24), $C^{\mathrm{NB}}_{m\to k}$ represents a dose-response curve, assuming different values for different intervention strengths (i.e., dose) μ_m. This is a typical situation in the context of continuous-valued interventions. In some applications, however, it proves beneficial to encapsulate the causal effect value with a single number. For this purpose, we may set μ_m to be uniform across all hypotheses, i.e., $\mu_m(\theta) = 1/H, \forall\theta\in\Theta$. This method of summarizing the causal effect is denoted as follows:

$$\overline{C}^{\mathrm{NB}}_{m\to k} \triangleq C^{\mathrm{NB}}_{m\to k}\Big|_{\mu_m(\theta)=1/H} \tag{7.28}$$

In Appendix 7.A, we show that this choice effectively parallels the process of determining the average causal *derivative* effect [148, Chapter 6], which is a method commonly used in the literature for summarizing the causal effect. Basically, it quantifies the extent of change in agent k in response to an infinitesimal variation in the intervention strength μ_m.

In the next section, we study two special network topologies that help illustrate the dependencies of the causal effect more explicitly.

Special cases

Fully-connected and federated architectures. In this example, we consider a fully-connected network (see Fig. 2.3) with a rank-one combination matrix and Perron

vector v, i.e.,

$$A = \begin{bmatrix} v_1 & v_1 & \cdots & v_1 \\ v_2 & v_2 & \cdots & v_2 \\ \vdots & \vdots & \ddots & \vdots \\ v_K & v_K & \cdots & v_K \end{bmatrix} = v\mathbb{1}_K^{\mathsf{T}}. \tag{7.29}$$

Recall that in terms of performance, this system is equivalent to a federated architecture in which (i) agents send their beliefs to a fusion center after local adaptation, (ii) the center averages the received beliefs in a weighted manner, and (iii) then broadcasts the combined belief to all agents — see Fig. 2.2. Under intervention $do(\boldsymbol{\mu}_{1,i} := \mu_1)$, we have

$$\widetilde{A} = \left[\begin{array}{c|ccc} 1 & v_1 & \cdots & v_1 \\ \hline 0 & v_2 & \cdots & v_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & v_K & \cdots & v_K \end{array}\right] \Longrightarrow R = \begin{bmatrix} v_2 & \cdots & v_2 \\ \vdots & \ddots & \vdots \\ v_K & \cdots & v_K \end{bmatrix} = v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}} \tag{7.30}$$

where $v_{-m} \triangleq \text{col}\{v_\ell\}_{\ell=2}^K$ is a $(K-1)\times 1$ dimensional vector consisting of all Perron entries except for agent $m=1$. Observe that

$$R^2 = v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}} v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}} \overset{(a)}{=} (1-v_1)v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}}, \tag{7.31}$$

where (a) follows from the fact that $\mathbb{1}_K^{\mathsf{T}} v = 1$ (Eq. (2.2)). Repeating the same arguments, it holds that

$$R^i = (1-v_1)^{i-1} v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}}. \tag{7.32}$$

Therefore,

$$(I-R^{\mathsf{T}})^{-1} - I = \sum_{i=1}^{\infty}(R^{\mathsf{T}})^i = \mathbb{1}_{K-1}v_{-m}^{\mathsf{T}} \sum_{i=1}^{\infty}(1-v_1)^{i-1} = \frac{1}{v_1}\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}}. \tag{7.33}$$

Inserting this into (7.22), we arrive at the following expression for each agent $k \neq m$:

$$\widetilde{\lambda}_{k,\infty}(\theta) = \frac{1}{v_m}\sum_{\ell=2}^{K} v_\ell d_\ell(\theta) + \log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{7.34}$$

Combining (7.3) and (7.8) with (7.34) yields the causal effect:

$$\begin{aligned} C_{m\to k}^{\text{NB}} &\overset{(7.3),(7.8)}{=} 1 - \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}}\exp\{-\widetilde{\lambda}_{k,\infty}(\theta)\}} \\ &\overset{(7.34)}{=} 1 - \frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}}\dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)}\exp\left\{-\dfrac{1}{v_m}\sum\limits_{\ell\neq m} v_\ell d_\ell(\theta)\right\}}. \end{aligned} \tag{7.35}$$

The effect of agent m on all other agents is the same, which is expected due to the symmetric nature of this special example. Furthermore, observe that the causal effect $C^{\text{NB}}_{m\to k}$ decreases with increasing $\widetilde{\lambda}_{k,\infty}(\theta)$. On that account, from (7.34) and (7.35) it can be seen that:

- Increasing the network centrality of agent $m=1$ (i.e., increasing v_m) decreases $\widetilde{\lambda}_{k,\infty}(\theta)$, and in turn increases the causal effect $C^{\text{NB}}_{m\to k}$. Therefore, an agent has more effect on other agents if it has a higher network centrality. In particular, if

$$v_m \to 0 \Longrightarrow \widetilde{\lambda}_{k,\infty}(\theta) \to +\infty \Longrightarrow C^{\text{NB}}_{m\to k} \to 0, \tag{7.36}$$

  which means that an agent with negligible network centrality has no causal effect on other agents.

- Increasing network centrality and informativeness of the other agents $\ell \neq m$ (i.e., increasing v_ℓ and $d_\ell(\theta)$) increases $\widetilde{\lambda}_{k,\infty}(\theta)$, and in turn decreases the causal effect $C^{\text{NB}}_{m\to k}$. In particular, if the most informative agents are equipped with the highest network centrality, then $\sum_{\ell=2}^{K} v_\ell d_\ell(\theta)$ is large and it is harder for agent $m=1$ to control other agents.

- If the fixed belief on the true hypothesis $\mu_m(\theta^\circ)$ decreases, then $\widetilde{\lambda}_{k,\infty}(\theta)$ decreases and the causal effect $C^{\text{NB}}_{m\to k}$ increases. This suggests that the further from the truth the information an agent supplies, the more effect that agent will have on other agents. In other words, agents supplying misinformation have more effect on the rest of the network. Specifically, observe that if the rest of the agents have a low informativeness average, i.e., if

$$\sum_{\ell=2}^{K} v_\ell d_\ell(\theta) \approx 0 \Longrightarrow C^{\text{NB}}_{m\to k} \approx 1-\mu_m(\theta^\circ). \tag{7.37}$$

  Therefore, the causal effect is proportional to the difference from the truth. It is maximized (i.e., $C^{\text{NB}}_{m\to k}=1$) when the fixed belief assigns 0 to the true hypothesis.

Ring architecture. In this example, we consider a unidirectional ring network where each agent has a self-confidence of α, and assigns a confidence of $1-\alpha$ to the preceding agent in the ring — see Fig. 7.3. Agents are indexed such that agent $k+1$ receives (or uses) information from agent k only. The combination matrix has the form:

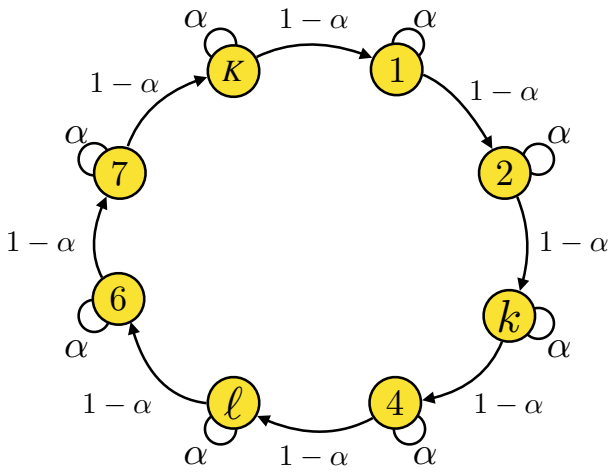


Figure 7.3: A unidirectional ring.

$$A = \begin{bmatrix} \alpha & 1-\alpha & 0 & \cdots & 0 \\ 0 & \alpha & 1-\alpha & \cdots & 0 \\ \vdots & 0 & \alpha & \cdots & \vdots \\ 0 & \vdots & \vdots & \ddots & 1-\alpha \\ 1-\alpha & 0 & 0 & \cdots & \alpha \end{bmatrix}. \tag{7.38}$$

Under intervention $do(\boldsymbol{\mu}_{1,i} := \mu_1)$, we have

$$\widetilde{A} = \left[\begin{array}{c|cccc} 1 & 1-\alpha & 0 & \cdots & 0 \\ \hline 0 & \alpha & 1-\alpha & \cdots & 0 \\ 0 & 0 & \alpha & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1-\alpha \\ 0 & 0 & 0 & \cdots & \alpha \end{array}\right] \Longrightarrow R = \begin{bmatrix} \alpha & 1-\alpha & \cdots & 0 \\ 0 & \alpha & \cdots & \vdots \\ \vdots & \vdots & \ddots & 1-\alpha \\ 0 & 0 & \cdots & \alpha \end{bmatrix}. \tag{7.39}$$

As a result,

$$(I - R^{\mathsf{T}}) = (1-\alpha)\begin{bmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -1 & 1 \end{bmatrix} \Longrightarrow (I - R^{\mathsf{T}})^{-1} = \frac{1}{1-\alpha}\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}, \tag{7.40}$$

which implies

$$((I - R^{\mathsf{T}})^{-1} - I)d_{-m}(\theta) = \frac{1}{1-\alpha}\Big[\alpha d_2(\theta), d_2(\theta) + \alpha d_3(\theta), \ldots, \sum_{\ell=2}^{K-1} d_\ell(\theta) + \alpha d_K(\theta)\Big]^{\mathsf{T}}. \tag{7.41}$$

Consequently, for agent k,

$$\widetilde{\lambda}_{k,\infty}(\theta) = \frac{1}{1-\alpha}\sum_{\ell=2}^{k-1} d_\ell(\theta) + \frac{\alpha}{1-\alpha} d_k(\theta) + \log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{7.42}$$

and, in addition, by definitions (7.3) and (7.8),

$$C^{\mathrm{NB}}_{m\to k} = 1 - \frac{1}{1 + \sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}} \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\Big\{ -\frac{1}{1-\alpha}\sum\limits_{\ell=2}^{k-1} d_\ell(\theta) - \frac{\alpha}{1-\alpha} d_k(\theta)\Big\}}. \tag{7.43}$$

As stated before, the causal effect $C^{\mathrm{NB}}_{m\to k}$ decreases with increasing $\widetilde{\lambda}_{k,\infty}(\theta)$. Therefore, the following remarks for (7.42) and (7.43) are in place:

- Since the KL divergence $d_\ell(\theta)$ is non-negative, $\widetilde{\lambda}_{k,\infty}(\theta)$ is monotonically increasing along the path $k = 2 \to \ldots \to k = K$. Therefore, the causal effect of agent $m = 1$ is monotonically decreasing along the same path: the closer agent m is to an agent, the higher its effect on that agent. This is intuitive because the effect that agent m has on agent $k+1$ is transferred via agent k in the ring structure. The difference between the causal effects of agent m on agents k and $k+1$ is proportional to the increase in $\widetilde{\lambda}_{k,\infty}(\theta)$, that is,

  $$\widetilde{\lambda}_{k+1,\infty}(\theta) - \widetilde{\lambda}_{k,\infty}(\theta) = d_k(\theta) + \frac{\alpha}{1-\alpha} d_{k+1}(\theta). \tag{7.44}$$

  This means that informative agents with high KL divergence on the path between agent m and agent k reduce the causal effect $C^{\mathrm{NB}}_{m\to k}$. In other words, the sphere of influence of an agent m is bigger if there are no other informative agents in the vicinity.

- For the immediate follower of agent $m = 1$, it follows that

  $$\widetilde{\lambda}_{2,\infty}(\theta) = \frac{\alpha}{1-\alpha} d_2(\theta) + \log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}. \tag{7.45}$$

  If agent 2 is not sufficiently informative itself, i.e., $d_2(\theta)$ is small, then $\widetilde{\lambda}_{2,\infty}(\theta)$ gets smaller and $C^{\mathrm{NB}}_{1\to 2}$ gets higher. In other words, an agent is more controllable if it is not knowledgeable.

- The limiting average $\widetilde{\lambda}_{k,\infty}(\theta)$ increases with increasing α. Therefore, if agents are

more self-confident, the causal strength is smaller and agents are less controllable by other agents.

7.4.2 Adaptive Social Learning

Similar to the modification in the NBSL case, the log-belief recursion (2.38) in ASL is modified as follows under intervention $do(\boldsymbol{\mu}_{1,i} := \mu_1)$:

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) = \widetilde{A}^{\mathsf{T}}((1-\delta)\widetilde{\boldsymbol{\lambda}}_{i-1}(\theta) + \beta\widetilde{\boldsymbol{x}}_i(\theta)). \tag{7.46}$$

The effective combination matrix $\widetilde{A}$ continues to be given by (7.10). However, the effective LLR is now given by

$$\widetilde{\boldsymbol{x}}_i(\theta) \triangleq \left[\frac{\delta}{\beta}\log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}, \boldsymbol{x}_{2,i}(\theta), \ldots, \boldsymbol{x}_{K,i}(\theta)\right]^{\mathsf{T}}, \tag{7.47}$$

where the first entry is different than the NBSL case. This is to compensate for the presence of the parameters δ and β. Observe from (7.46) that $\widetilde{A}$ from (7.10) and $\widetilde{\boldsymbol{x}}_i(\theta)$ from (7.47) verify $\widetilde{\boldsymbol{\lambda}}_{1,i}(\theta) = \log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}$ for all time instants, i.e., the intervention is fixed, by similar arguments to (7.18b). In Appendix 7.B, we derive the following expression for the limiting log-belief ratio expectations for the rest of the network:

$$\boxed{\widetilde{\lambda}_{-m,\infty}(\theta) = \left(\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right)(I-(1-\delta)R^{\mathsf{T}})^{-1}r + \frac{\beta}{1-\delta}\Big((I-(1-\delta)R^{\mathsf{T}})^{-1} - I\Big)d_{-m}(\theta)} \tag{7.48}$$

The causal effect $C^{\text{ASL}}_{m\to k}$ can be calculated by inserting post-intervention expression (7.48) and pre-intervention expression (2.40) into the definitions (7.3) and (7.9). Notice from (7.48) that similar to the NBSL case, the causal effect depends on the informativeness of agents, the network topology, and the strength of intervention via $d_{-m}(\theta), R$, and μ_m, respectively. In fact, if we set $\delta = 0$ and $\beta = 1$, Eq. (7.48) reduces to the NBSL expression (7.22) as expected. This is because the left-stochastic nature of A implies that

$$r + R^{\mathsf{T}}\mathbb{1}_{K-1} = \mathbb{1}_{K-1} \iff (I-R^{\mathsf{T}})^{-1}r = \mathbb{1}_{K-1}. \tag{7.49}$$

In addition, the causal effects in ASL are affected by the importance weighting parameters δ and β, as well as by the vector r that represents the confidence weights other agents assign to agent m. In (7.22), the entries of r implicitly influence the causal effect via R: the column-wise summation of the entries of r and R results in 1 at all columns due to the left-stochastic nature of A. In comparison, in ASL, both r and R impact $C^{\text{ASL}}_{m\to k}$ explicitly. If we take a closer look at the terms in (7.48), we can see that:

- The vector that scales the intervened log-belief ratio $\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}$ can be expanded

as

$$(I-(1-\delta)R^{\mathsf{T}})^{-1}r = r+(1-\delta)R^{\mathsf{T}}r+(1-\delta)^2R^{2\,\mathsf{T}}r+(1-\delta)^3R^{3\,\mathsf{T}}r+\ldots \quad (7.50)$$

On the RHS of this equation, the first r represents the scaling of the information transferred from agent m to the rest of the network *directly*. Namely, for an agent k, the scaling of the direct information is a_{mk} if m is an immediate neighbor ($m \in \mathcal{N}_k$); 0 if it is not ($m \notin \mathcal{N}_k$). On the other hand, the second term $(1-\delta)R^{\mathsf{T}}r$ describes the scaling of the information transferred from agent m to the rest of the network, which is then mixed with the other agents $\forall k \neq m$ (via R^{T}) and "forgotten" (i.e., lose its recency) for one time instant by a factor of $(1-\delta)$. The consecutive terms over time follow from the same scaling argument.

- In a similar manner, we can express the matrix that scales the vector of individual informativeness in the rest of the network $d_{-m}(\theta)$ as:

$$\frac{1}{1-\delta}\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\Big) = R^{\mathsf{T}}+(1-\delta)R^{2\,\mathsf{T}}+(1-\delta)^2R^{3\,\mathsf{T}}+\ldots \quad (7.51)$$

  Since there is no outgoing link from the rest of the network to agent $m=1$ under the intervention, the terms in this expression only depend on the combination matrix R. When new information arrives to the remaining agents, it is first mixed among these agents (corresponding to the first term R^{T} on RHS), and then in the next iteration, it is mixed again but also gets forgotten due to the time discount factor δ (corresponding to the second term $(1-\delta)R^{2\,\mathsf{T}}$ on RHS), and so on.

- Remember from (2.16) that $\beta > 0$ scales the likelihood of observations, reflecting the weight agents place on their own observations originating from out-of-network sources. As a result, notice that in (7.48), β scales the individual informativeness $d_\ell(\theta), \forall \ell \neq m$. In other words, it amplifies the effect of self observations on the state of nature.

Next, we analyze the special cases introduced in the NBSL case under ASL framework.

Special cases

Fully-connected and federated architectures. In Appendix 7.C, we prove that the additional δ and β parameters introduced for the ASL change the NBSL expression (7.34) to

$$\widetilde{\lambda}_{k,\infty}(\theta) = \frac{1}{1-(1-\delta)(1-v_m)}\left(\beta\sum_{\ell=2}^{K}v_\ell d_\ell(\theta)+v_m\log\frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right) \quad (7.52)$$

Notice that as $\delta \to 0$ and $\beta \to 1$, (7.52) recovers (7.34) as expected, and the following remarks from the NBSL case continue to hold here: (i) the influence of an agent $m = 1$ is identical for each agent $k \neq 1$ due to symmetry in the network topology, (ii) increasing the network centrality of agent m increases its causal influence, and (iii) increasing the network centrality and informativeness of the rest of the agents $\ell \neq 1$ decreases the causal effect of $m = 1$. Moreover, since causal effect $C^{\text{ASL}}_{m\to k}$ is a monotonic decreasing function of $\widetilde{\lambda}_{k,\infty}(\theta)$ by (7.9), Eq. (7.52) also implies the following conclusions:

- As stated after (7.48), β scales the informativeness of agents. Accordingly, (7.52) reveals that the causal effect is decreasing with increasing β. This is justifiable because as agents have greater reliance on their own observations about the state of nature, they are less influenced by other agents in the network. It is worth mentioning that in (7.52), the intervened log-belief ratio $\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}$ behaves as "pseudo-informativeness". It is scaled with the Perron entry v_m of agent $m = 1$, similar to how the rest of the agents' informativeness is scaled with their own Perron entries. The difference is that the other agents' informativeness are based on their log-likelihood ratios averaged with respect to the true distribution, whereas the intervened log-belief ratio can be arbitrary, possibly supplying misinformation.

- In the special case when the remaining agents have no informativeness (i.e., $d_\ell(\theta) = 0,\ \forall \ell \neq m$), the limiting mean log-belief ratio vector in (7.52) turns into

$$\widetilde{\lambda}_{k,\infty}(\theta) = \frac{1}{1 + \delta\left(\frac{1}{v_m} - 1\right)} \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}. \tag{7.53}$$

  This is in contast to the NBSL case, where, $\widetilde{\lambda}_{k,\infty}(\theta) = \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}$. In other words, in steady-state of NBSL, the average beliefs of all agents become equal to the intervened fixed belief μ_m, implying full controllability. In ASL, however, the controllability is reduced by a factor of $(1 + \delta(1/v_m - 1)) \geq 1$ as shown in (7.53). In particular, increasing the forgetting factor δ decreases controllability, especially when the network centrality of "controlling" agent $m = 1$ is small. However, if agent m is highly central, i.e., $v_m \to 1$, then the forgetting factor δ has negligible effect on controllability.

- Considering the general case (7.52), note that unlike β which only affects non-intervened observations, δ affects both intervened beliefs and non-intervened observations. Thus, to fully understand the impact of δ on the overall causal effect, we must consider the exact values of the relevant parameters.

Ring architecture. In Appendix 7.D, we prove that the additional δ and β parameters introduced for the ASL change the NBSL expression (7.42) to

$$\widetilde{\lambda}_{k,\infty}(\theta) = \left(\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right) \frac{(1-\delta)^{k-2}(1-\alpha)^{k-1}}{(1-(1-\delta)\alpha)^{k-1}} + \frac{\beta\alpha}{1-(1-\delta)\alpha} d_k(\theta) + \frac{\beta}{1-\delta} \sum_{\ell=2}^{k-1} \frac{(1-\delta)^{k-\ell}(1-\alpha)^{k-\ell}}{(1-(1-\delta)\alpha)^{k-\ell+1}} d_\ell(\theta) \tag{7.54}$$

from which we can make the following observations:

- As $\delta \to 0$ and $\beta \to 1$, the NBSL expression (7.42) for ring architectures is recovered. Similar to the earlier expressions, the causal effect decreases with increasing informativeness of agents along the path between $m = 1$ and k ($\ell = 2, \ldots, k-1$) (the last term on RHS), and also decreases with increasing informativeness of agent k. Informativeness is scaled by β as before.

- Recall that in NBSL, if the remaining agents have no informativeness, it holds that $\widetilde{\lambda}_{k,\infty}(\theta) = \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}$. In other words, agent $m = 1$ can fully control other agents' beliefs. Instead, in ASL, if $d_\ell(\theta) = 0,\ \forall \ell \neq 1$, it holds that

$$\begin{aligned} \widetilde{\lambda}_{k,\infty}(\theta) &= \frac{(1-\delta)^{k-2}(1-\alpha)^{k-1}}{(1-(1-\delta)\alpha)^{k-1}} \left(\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right) \\ &= \frac{1}{1-\delta} \left(1 - \frac{1}{\frac{1-\alpha}{\delta} + \alpha}\right)^{k-1} \left(\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}\right). \end{aligned} \tag{7.55}$$

  Observe that as the agent index k increases, the controllability decays at each hop by a factor of

$$\frac{\widetilde{\lambda}_{k+1,\infty}(\theta)}{\widetilde{\lambda}_{k,\infty}(\theta)} = 1 - \frac{1}{\frac{1-\alpha}{\delta} + \alpha} \quad \in [0,1]. \tag{7.56}$$

  The decrease is higher when δ is higher because the information from agent $m = 1$ gets "partially forgotten" at each hop as k (i.e., the distance to agent 1) increases. However, in general, the informativeness of agents along the path is not 0, and they have a shadowing effect on agent m's influence, as argued before. The forgetting factor δ decreases this shadowing effect as well, particularly for agents far from agent k.

7.5 Causal Ranking of Agents

In the previous sections, we examined the bipartite influence between agents, that is, how much an agent m affects another agent k in the network. By calculating this influence for any pair of agents (m,k), we can construct a $K \times K$ influence matrix C with entries $[C]_{mk} = C_{m\to k}$. One is often interested in the overall influence of agent m on the network rather than its effect on individual agents. To that end, in this section, we describe a procedure to use C for ranking and quantifying the agents' cumulative effect over the network.

Since C is constructed from intervened belief dependent entries $C_{m\to k}$, an ordering based on C would be valid for a particular intervention. For an intervention dose independent ranking of agents, one can consider the matrix $\overline{C}$, which is formed with dose independent causal effects $\overline{C}_{m\to k}$:

$$\overline{C}_{m\to k} \triangleq C_{m\to k}\Big|_{\mu_m(\theta)=1/H} \tag{7.57}$$

where we extend the definition (7.28) for the NBSL case to the general case. For simplicity of the presentation, in the sequel, we focus on the NBSL case, even though our arguments keep holding for a general $\overline{C}$ as well as an ordering based on intervention dose dependent matrix C, too. First, note that the causal effect for the NBSL case is given by

$$\begin{aligned} C^{\text{NB}}_{m\to k} &\overset{(7.3),(7.8)}{=} 1 - \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\{-\widetilde{\lambda}_{k,\infty}(\theta)\}} \\ &\overset{(7.22)}{=} 1 - \frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}} \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\Big\{ -\Big[\Big((I-R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta)\Big]_k\Big\}}. \end{aligned} \tag{7.58}$$

Here, setting $\mu_m(\theta) = 1/H$ for any $\theta \in \Theta$ based on (7.57) yields

$$\overline{C}^{\text{NB}}_{m\to k} = 1 - \frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\Big\{ -\Big[\Big((I-R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta)\Big]_k\Big\}}. \tag{7.59}$$

Since all KL divergences are assumed to be finite ($d_k(\theta) < \infty$) and the graph is strongly connected, this implies

$$\Big\|\Big((I-R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta)\Big\|_\infty < \infty. \tag{7.60}$$

Incorporating this into (7.59) implies that $\overline{C}^{\text{NB}}_{m\to k} > 0$, $\forall m \neq k$. Furthermore, regarding the diagonal elements of $\overline{C}$, it holds by definition that an intervention on agent m

implies

$$\widetilde{\lambda}_{m,\infty}(\theta) = \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{7.61}$$

As a result, if we set $\mu_m(\theta) = 1/H, \forall \theta \in \Theta$

$$\begin{aligned} \overline{C}^{\text{NB}}_{m\to m} &\overset{(7.3),(7.8)}{=} 1 - \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\{-\widetilde{\lambda}_{m,\infty}(\theta)\}} \\ &= 1 - \frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\{0\}} \\ &= 1 - \frac{1}{H} \\ &> 0 \end{aligned} \tag{7.62}$$

Consequently, all entries of $\overline{C}$ are positive, which implies that $\overline{C}$ is a primitive matrix [8]. Therefore, according to Perron's theorem [4, 8, 11], $\overline{C}$ has a unique, real and positive eigenvalue ρ that dominates all other eigenvalues in magnitude. Moreover, the eigenvector q corresponding to ρ is unique up to a scaling and all its entries are positive, i.e.,

$$\overline{C}q = \rho q, \quad q_k > 0, \quad \forall k = 1, \ldots, K. \tag{7.63}$$

The entry q_k is a measure of agent k's overall influence over the network. The agents can be ranked with respect to these entries. We name the resulting algoritm CausalRank which is summarized in Algorithm 7.1. Importantly, the vector q — which is the output of Alg. 7.1 — differs from the network centrality eigenvector v in general. While v is determined solely by the combination matrix A (see (2.2)), as shown in previous sections, causal influences and hence q depend on the informativeness of agents as well. More specifically, (7.63) computes a *causal* eigenvector centrality that attributes higher importance to exerting influence on agents who are themselves influential. A possible alternative approach (which we call *average influence ranking* (AIR)) can treat all agents with equal regard in the averaging process by assigning the following ranking score to each agent m:

$$\text{AIR}(m) = \frac{1}{K-1} \sum_{k\neq m} \overline{C}_{m\to k} \tag{7.64}$$

In contrast, rather than employing a simple averaging, CausalRank seeks the equilibrium vector by assigning significant weights to those agents that have a higher influence on other influential agents. This concept bears resemblance to other methodologies based on eigenvector centrality, such as the PageRank algorithm [149]. While ranking websites, PageRank gives preferential treatment to links from more central websites.

It is also worth mentioning that CausalRank is distinct from the causal ordering methods for directed acyclic graphical models [148] since we are dealing with cyclic graphs with bidirectional links due to our time-series setting. Furthermore, CausalRank is not only

Algorithm 7.1 CausalRank Algorithm

Input: a network of K agents with indices $\{1, 2, \ldots, K\}$, combination matrix A, set Θ of H hypotheses, informativeness vector $d(\theta)$
Initialize: $K \times K$-dimensional influence matrix $\overline{C}$
for each agent $m = 1, 2, \ldots, K$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    if $m = k$ **then**
      set $[\overline{C}]_{mm} := 1 - \dfrac{1}{H}$
    else
      compute the causal effect $\overline{C}_{m\to k}$ with (7.59)
      set $[\overline{C}]_{mk} := \overline{C}_{m\to k}$
    end if
  end for
end for
find the largest eigenvalue ρ of $\overline{C}$
Output: the eigenvector q satisfying $\overline{C}q = \rho q$

useful for ranking, but also provides information on the strength of agents' overall influence on others.

7.6 Causal Discovery from Observational Data

In Sec. 7.4, we derived the closed-form expressions (7.22) and (7.48) for the steady-state equilibrium of the network under interventions, which necessitate knowledge of the combination matrix A and the informativeness of agents $d(\theta)$. In practice, these parameters might not be readily available. The work [137] introduced the Graph Social Learning (GSL) algorithm, which can be used to recover A and $d(\theta)$ using a sequence of publicly shared intermediate beliefs $\{\boldsymbol{\psi}_{k,i}\}$ in the *observational* setting of the ASL algorithm. Using observational data only can be especially useful in social network contexts where conducting experiments is not feasible. Nonetheless, [137] acknowledge that the algorithm may not perform well in real-world scenarios, mainly due to the limitations of the social learning model in accurately describing the real world. However, in many applications, some information about the underlying combination matrix A may already be available. For instance, in Twitter, the publicly available adjacency matrix can provide information about which user follows which other users.

Taking these aspects into account, in this section, we propose an algorithm that utilizes the adjacency matrix and a temporal sequence of publicly shared intermediate beliefs $\{\boldsymbol{\psi}_{k,i}\}$ to estimate bipartite causal effects for both NBSL and ASL algorithms. Specifically, we leverage the graph of user connections to estimate combination matrix weights by using existing methods in the literature (e.g., averaging rule). Then, using the estimated combination matrix, we estimate the informativeness of agents by us-

ing belief update recursions. By inserting these to the closed-form expressions (7.22) and (7.48), we estimate the causal effects. To that end, observe that the intermediate log-belief ratios evolve based on a linear recursion due to (2.16):

$$\boldsymbol{\Lambda}_i = (1-\delta)A^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1} + \beta \boldsymbol{X}_i \tag{7.65}$$

where we are now defining the following $K \times H$ matrices over all agents $k \in \{1,\ldots,K\}$ and hypotheses $\theta_j \in \Theta$:

$$[\boldsymbol{\Lambda}_i]_{kj} \triangleq \log \frac{\boldsymbol{\psi}_{k,i}(\widehat{\boldsymbol{\theta}}^{\circ})}{\boldsymbol{\psi}_{k,i}(\theta_j)}, \quad [\boldsymbol{X}_i]_{kj} \triangleq \log \frac{L_k(\boldsymbol{\xi}_{k,i}|\widehat{\boldsymbol{\theta}}^{\circ})}{L_k(\boldsymbol{\xi}_{k,i}|\theta_j)}. \tag{7.66}$$

Here, $\widehat{\boldsymbol{\theta}}^{\circ}$ is an estimate for the latent state of nature θ° computed as follows after some time M:

$$\widehat{\boldsymbol{\theta}}^{\circ} \triangleq \arg\max_{\theta\in\Theta} \sum_{k=1}^{K} \boldsymbol{\psi}_{k,M}(\theta). \tag{7.67}$$

The rationale behind (7.67) is that under proper assumptions we know from Theorems 2.1 and 2.3 that agents learn the true hypothesis with more confidence as M grows. Our goal is to infer the true combination matrix A and informativeness vector $d(\theta)$ for each hypothesis from a sequence of $M+1$ matrices $\{\boldsymbol{\Lambda}_M, \boldsymbol{\Lambda}_{M-1}, \ldots, \boldsymbol{\Lambda}_0\}$ and the adjacency matrix of the agents.

We can estimate A by using existing procedures in the literature for forming combination matrices from adjacency matrices, e.g., by using the averaging or relative degree rules [8]. For instance, the averaging rule assigns the same weight to all neighbors of an agent, i.e.,

$$[\,\widehat{A}\,]_{\ell k} = \begin{cases} \dfrac{1}{|\mathcal{N}_k|}, & \text{if there is a link from } \ell \text{ to } k \text{ (i.e., } \ell \in \mathcal{N}_k) \\ 0, & \text{otherwise} \end{cases} \tag{7.68}$$

After forming the combination matrix estimate $\widehat{A}$, we can insert it into (7.65) and average over available M samples to estimate the average log-likelihood ratios that correspond to the informativeness of agents using

$$\widehat{\boldsymbol{D}} = \frac{1}{\beta M}\sum_{i=1}^{M}\Big(\boldsymbol{\Lambda}_i - (1-\delta)\widehat{A}^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1}\Big). \tag{7.69}$$

Then, one can replace A and $d_k(\theta_j)$ with $\widehat{A}$ and $[\widehat{\boldsymbol{D}}]_{kj}$ in Sec. 7.4 to obtain the causal effect estimate $\widehat{\boldsymbol{C}}_{m\to k}$. The complete procedure is summarized in Alg. 7.2. Essentially, it combines our causality results in Sec. 7.4 with a straightforward adjustment to the GSL algorithm from [137].

The graph causality learning (GCL) algorithm (Alg. 7.2) only requires a sequence of

Algorithm 7.2 Graph Causality Learning (GCL)

Input: a sequence of shared beliefs $\{\boldsymbol{\psi}_{k,i}\}$ for $M+1$ time instants, the graph topology of agents

Parameters: for NBSL $\delta = 0, \beta = 1$, for ASL $\delta \in (0,1), \beta > 0$

set true state of nature estimate: $\widehat{\boldsymbol{\theta}}^{\circ} = \arg\max_{\theta\in\Theta} \sum_{k=1}^{K} \boldsymbol{\psi}_{k,M}(\theta)$

form left-stochastic combination matrix estimate $\widehat{A}$ from input adjacency matrix, e.g., by (7.68)

for $i = 0, 1, \ldots, M$ **do**

  for each agent k and hypothesis θ_j, set the entry:

$$[\boldsymbol{\Lambda}_i]_{kj} = \log \frac{\boldsymbol{\psi}_{k,i}(\widehat{\boldsymbol{\theta}}^{\circ})}{\boldsymbol{\psi}_{k,i}(\theta_j)} \tag{7.70}$$

end for

estimate the informativeness:

$$\widehat{\boldsymbol{D}} = \frac{1}{\beta M} \sum_{i=1}^{M} \Big(\boldsymbol{\Lambda}_i - (1-\delta)\widehat{A}^{\mathsf{T}} \boldsymbol{\Lambda}_{i-1} \Big) \tag{7.71}$$

for any given agent pair (m,k), **compute approximate causal effect** $\widehat{\boldsymbol{C}}_{m\to k}$ by replacing A and $d_k(\theta_j)$ with $\widehat{A}$ and $[\widehat{\boldsymbol{D}}]_{kj}$ in the original expression (7.2) for $C_{m\to k}$ (which also requires using (7.22) for NBSL and (7.48) for ASL)

Output: $\widehat{\boldsymbol{C}}_{m\to k}$

shared intermediate beliefs (actions) and the knowledge of adjacency matrix. This enhances its practicality and makes it advantageous in terms of privacy for scenarios where only limited information is publicly accessible. For example, in a network of Twitter users, shared beliefs (opinions) in the form of tweets (posts) and the knowledge of who follows whom can usually be accessed by all users, while the external exposure to information (e.g., from mass media channels distinct from Twitter) may not be available. Therefore, the GCL algorithm can be useful for analyzing social media content while respecting privacy. In the next result, we provide a performance bound on the GCL algorithm.

Theorem 7.1 (**Causal influence estimation**). *For sufficiently small combination matrix estimation errors and δ values, the error in causal influence estimation decreases with increasing number of samples M in expectation, namely,*

$$\mathbb{E}\,\Big| C_{m\to k} - \widehat{\boldsymbol{C}}_{m\to k} \Big| = O(1/M) \tag{7.72}$$

for both NBSL and ASL under any intervention strength μ_m.

Proof. See Appendix 7.E. ■

We show the practical usefulness of Alg. 7.2 by means of a real-world application to Twitter data in Sec. 7.7.2. Furthermore, a detailed analysis of the time complexity of the algorithms discussed in this chapter is provided in Appendix 7.F.

7.7 Numerical Simulations

7.7.1 Synthetic Data

For our numerical simulations, we study a network of $K = 11$ agents, interconnected with the strongly connected graph topology in Fig. 7.4.

The agents observe data drawn from a Gaussian distribution and aim to distinguish the true state θ° from $H = 2$ possible hypotheses. Under the true state, each agent k observes data that follows a Gaussian distribution with zero mean and unit variance, expressed as:

$$L_k(\xi|\theta^\circ) = \frac{1}{\sqrt{2\pi}} \exp\Big\{ -\frac{1}{2}\xi^2 \Big\}. \tag{7.73}$$

Under the alternative hypothesis $\theta' \neq \theta^\circ$, we assume that the data still has unit variance for all agents, but the mean vector ν_k changes as shown in Table 7.1. Therefore, the informativeness of each agent, which is equal to the KL divergence between $L_k(\xi|\theta^\circ)$ and $L_k(\xi|\theta')$, is given in Table 7.1 and is calculated as follows:

$$d_k(\theta') = D_{\mathrm{KL}}\Big(L_k(\xi|\theta^\circ)||L_k(\xi|\theta')\Big) = \frac{1}{2}\nu_k^2. \tag{7.74}$$

Notably, agents 5 and 6 have no informativeness, that is, they are not able to learn the truth without cooperating with the other agents. Initially, we assume that the agents observe spatially independent data. In other words, the covariance matrix is an identity matrix.

We start with the NBSL case ($\delta = 0, \beta = 1$). The right panel in Fig. 7.5 shows the combination matrix that is derived from the averaging rule applied to the graph topology in Fig. 7.4. Notice that the averaging rule generates a matrix whose entries are constant column-wise. The left panel in Fig. 7.5 shows the matrix of bipartite causal effects where the entry in m-th row k-th column represents $\overline{C}^{\mathrm{NB}}_{m\to k}$ (see (7.59) for the explicit formula).

Upon comparing the two heat maps in Fig. 7.5, it becomes apparent that the combination matrix entries do not reveal the causal relationships directly. For example, despite the absence of a direct connection in the combination matrix (as indicated by 0 entries), agent 11 exerts significant influence on agents 2 and 8. This phenomenon

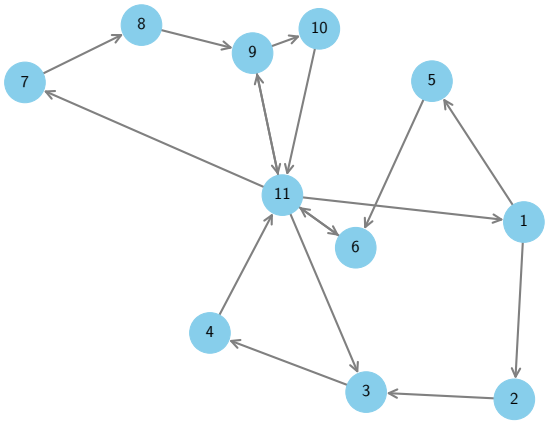


Figure 7.4: The strongly connected network architecture used in Sec. 7.7.1. Each agent also has a self-loop, which is omitted for visual simplicity.

Agent	ν_k	$d_k(\theta')$
1	0.8	0.32
2	0.6	0.18
3	0.2	0.02
4	0.6	0.18
5	0	0
6	0	0
7	0.4	0.08
8	0.4	0.08
9	0.2	0.02
10	0.6	0.18
11	0.8	0.32

Table 7.1: Mean ν_k of the observations under alternative hypothesis and the corresponding informativeness levels $d_k(\theta')$ for each agent k.

highlights the importance of taking the ripple effects over a network into account. Furthermore, the influence of agent 1 on agent 5 is notably high. Given the zero informativeness of agent 5, this finding aligns with our expectations, as low-informativeness agents are easier to control (remember the discussion in Sec. 7.4.1). Agent 5 being a low-informativeness agent also facilitates the propagation of influence from agent 1 to agent 6 via agent 5. Intriguingly, despite the absence of a direct connection between agents 1 and 6, this indirect influence is more substantial than the influence of agent 5 on agent 6. This shows that mixing of information over a network necessitates an understanding of causal influence beyond local interactions.

Next, in Fig. 7.6, by using the matrices in Fig. 7.5, we compare the overall influences of agents using three methods: CausalRank, AIR, and network eigenvector centrality. Notably, the CausalRank and AIR metrics yield similar results as they both use the bipartite causal relations matrix for causal ranking. For instance, agents 2 and 5 possess relatively low rankings in both of these metrics. The network eigenvector centrality, on the other hand, only relies on the combination matrix, and often deviates from these two metrics. Specifically, it assigns relatively higher scores to agents 2 and 5, and a comparatively lower score to agent 11. Moreover, an interesting distinction between AIR and CausalRank becomes apparent when considering the case of agent 9. We can see from the causal influence matrix in Fig. 7.5 that agent 9 has a substantial impact on agent 11 — the most influential agent (see Fig. 7.6). Consequently, agent 9's CausalRank score surpasses its AIR score. This can be attributed to CausalRank's consideration of the significance of influencing agent 11. Unlike AIR, which assigns uniform weights, CausalRank assigns a higher weight to influences on more influential agents.

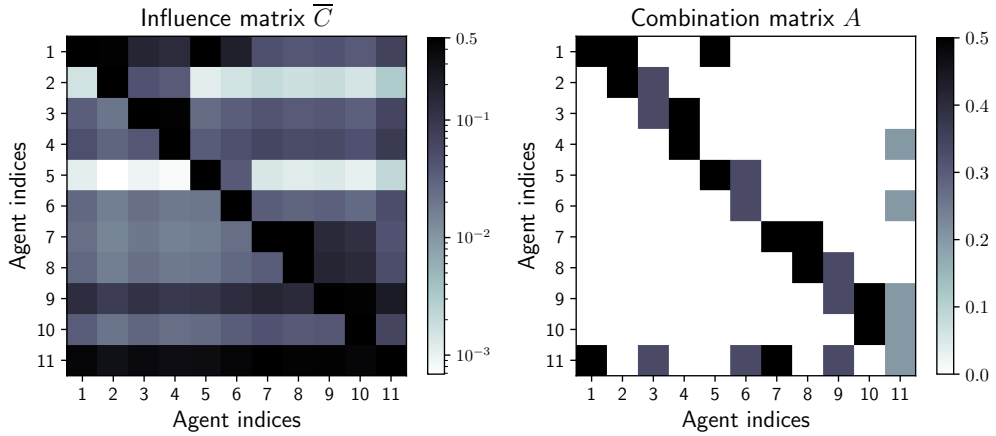


Figure 7.5: (*Left*): Bipartite causal influence matrix. (*Right*): Combination matrix corresponding to the network topology in Fig. 7.4 formed with averaging rule.

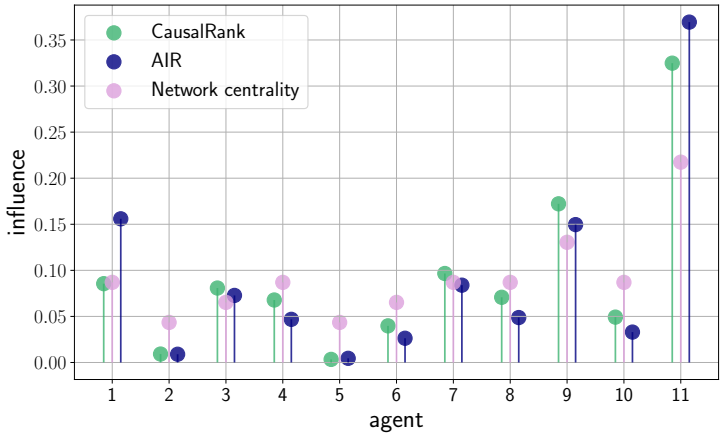


Figure 7.6: Ranking of agents based on their causal influence (for both CausalRank and AIR) and their network-topology based eigenvector centrality. All ranking scores are normalized to sum up to one.

To gain insights into the influence of the forgetting factor δ in the ASL case, we focus our attention on agent $k = 4$. In Fig. 7.7, we present the average influence exerted by agent 4's neighbors that are 1, 2, and 3 hops away. It is clear from Fig. 7.7 that the influence of distant agents diminishes with increasing δ. This is because increasing δ increases the significance of recent observations, and since information from distant agents loses its recency by the time it arrives at agent 4, this implies assigning less importance to information from those distant agents.

In the simulations conducted so far, we have considered dose-independent causal effects $\overline{C}_{m\to k}$. Figure 7.8 demonstrates the dependency of $C_{m\to k}$ on the intervened belief $\mu_m(\theta)$ for $\theta \neq \theta^\circ$, as described in equation (7.24). We use $m = 11$ and $k = 6$ for this plot, which illustrates the theoretical result presented in equation (7.24).

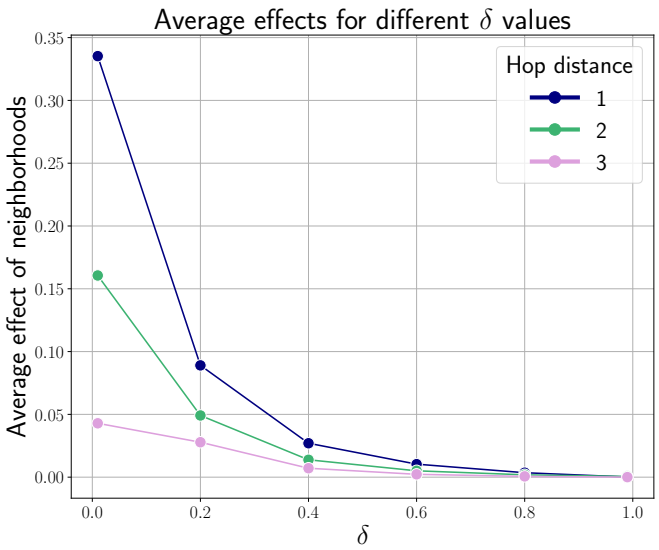


Figure 7.7: Average influence on agent $k = 4$ from its neighborhood of 1,2 and 3 hop distances with respect to changing values of forgetting factor δ.

Then, we fix $\delta = 0.1$, and use the GCL algorithm (Alg. 7.2) in order to estimate the causal effects using observational data (shared beliefs) as described in Sec. 7.6. The norm disagreement of the causal influence matrix formed with estimates and the true causal influences, averaged over 10 Monte Carlo simulations, is given in Fig. 7.9. Observe that the error is decreasing as the number of samples M increases, which supports Theorem 7.1.

Finally, we illustrate the distinction between causality and correlation by again considering agents $m = 11$ and $k = 6$. The joint distribution of their data is changed by introducing varying levels of correlation to the observations that these agents are receiving. Fig. 7.10 shows that as the correlation in data increases, the correlation of the asymptotic beliefs of these agents also changes. However, the causal effects (both the effect of agent 6 on 11 and that of agent 11 on 6) remain constant. This shows that external observations can act as a correlation inducing confounding factor. Yet, our method maintains consistent results, which shows its robustness against non-causal factors.

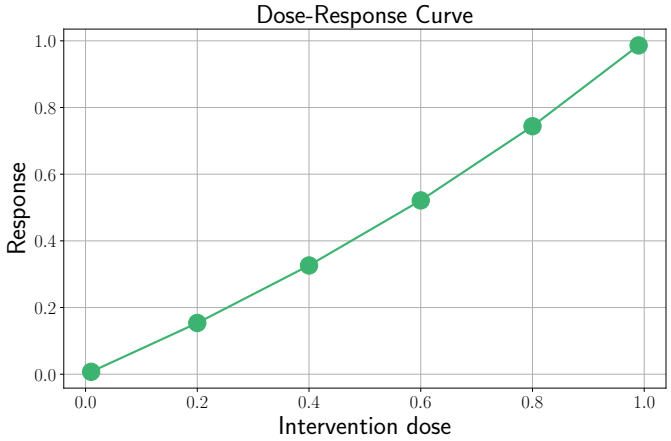


Figure 7.8: $C^{\mathrm{NB}}_{m\to k}$ value from (7.24) for $m = 11$ and $k = 6$ with respect to increasing intervention strength $\mu_m(\theta)$ for $\theta \neq \theta^\circ$.

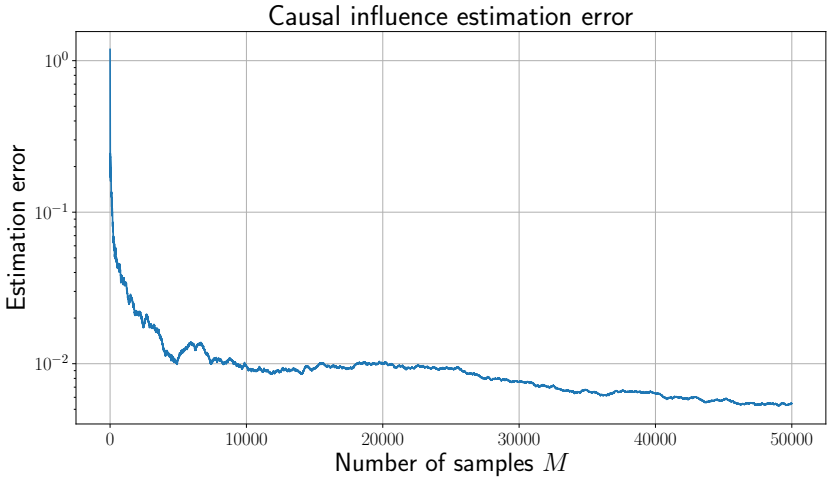


Figure 7.9: Causal influence estimation error with respect to increasing number of time samples M.

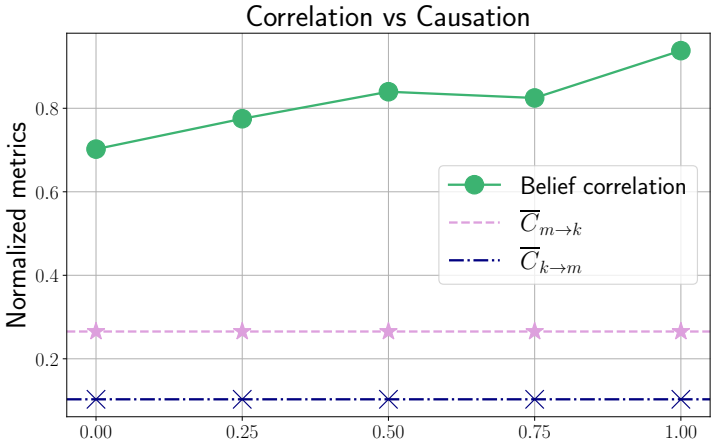


Figure 7.10: Correlation and causation between agents 6 and 11 with respect to varying dependence between the streaming observations they receive.

7.7.2 Application to Social Media Data

In this section, we use the GCL algorithm (Alg. 7.2) on real-world data to assess the influence of Twitter users. Our approach distinguishes itself from prior works [132, 137, 150, 151], which typically rely on some descriptive statistics to measure influence in Twitter. More specifically, in our approach,

- All input requirements are publicly available, i.e., publicly shared posts (tweets) by users and the information of who follows whom. This offers a significant advantage in terms of privacy, as we do not require any private feature about users.
- Going beyond providing a mere ranking of influential users, we also quantify the bipartite causal relations.
- We leverage natural language processing tools to extract meaningful information from the content of users' posts to form belief inputs, rather than relying on traditional simpler metrics such as the posting frequency. In contrast to the binary treatment approach, adopting the continuous treatment metric in (7.2) for our causality definition enables us to allow for the use of continuous variables such as opinions and sentiment scores.

Note that we utilize Alg. 7.2 for the NBSL model ($\delta = 0, \beta = 1$) as-is, without employing additional techniques to enhance its accuracy for real-world modeling. Our intention is to demonstrate the practical usefulness of our algorithm rather than striving to develop the most advanced practical algorithm available.

Network structure. Performance evaluation of the influence estimation algorithms in real-world social networks is challenging due to the absence of ground truth regarding

influential users. There is also no ground truth reference for the confidence scores assigned by users to one another (i.e., combination weights) or for the information the users are obtaining from out-of-network resources (i.e., informativeness levels). Therefore, we utilize the framework from [137], namely, a sub-network consisting of $K = 20$ Twitter users, as illustrated in Fig. 7.11. Notably, this sub-network incorporates Elon Musk, a public figure with 140 million followers across Twitter, who is reportedly influential on cryptocurrency prices [152]. All users within the sub-network actively share posts related to cryptocurrencies and bitcoin-related topics. Furthermore, one user, who we will refer to as User 2, has 1,167 followers on Twitter and is notable for being followed by Elon Musk, as depicted in Fig. 7.11. Importantly, the sub-network exhibits a strong connectivity among its members.

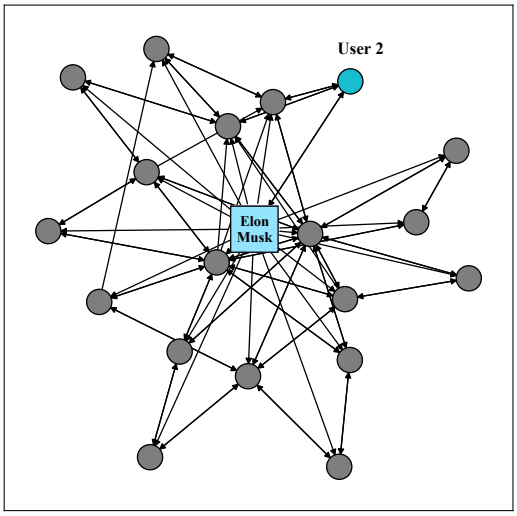


Figure 7.11: Sampled Twitter sub-network with $K = 20$ users. An arrow from a user k to ℓ means that user ℓ is following user k. The connectivity density of the graph is approximately 0.24.

Opinion processing. The Twitter API is leveraged in order to collect the posts (tweets) of users between 01.01.2017 and 01.05.2022 relevant to crypto-currency discussions, using query keywords such as "coin", "bitcoin", or "crypto-currency". To quantify the contextual information of these posts to form the input beliefs, sentiment analysis based on neural language models [153] is utilized. We refer to Fig. 7.12 for some illustrative examples. The sentiment scores obtained through natural language processing ranges from 0 to 1, signifying the degree of positive attitude towards Bitcoin. These scores correspond to the beliefs of the agents on the hypothesis of "Bitcoin is good/useful". We consider two hypotheses, i.e., $H = 2$, where the counter-hypothesis is "Bitcoin is bad/harmful".

We then integrated these beliefs obtained from users' tweets, along with the sub-network topology of who follows whom, into Alg. 7.2. Specifically, we employed NBSL modeling, i.e., $\delta = 0$ and $\beta = 1$. We binned the sequence of data to days, that is, each i in our algorithms corresponds to one day. Combination weights were estimated using an averaging rule on the sub-network topology.

Bipartite causality. Inserting the observational input into Alg. 7.2, the resulting average causal derivative effect matrix is shown in Fig. 7.13 in the form of a heat map. To facilitate comparison, we also include the adjacency matrix, which describes the connections between users. In these plots, the indices 1 and 2 correspond to specific

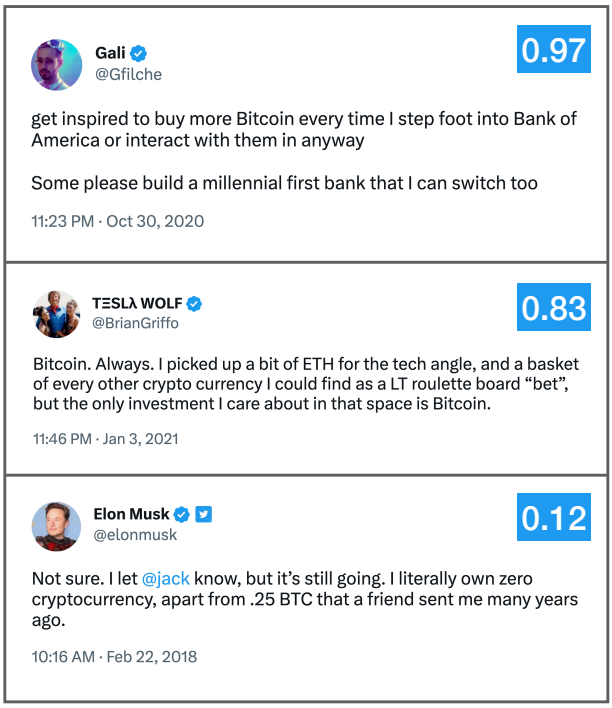


Figure 7.12: Sample tweets from the users in the sub-network. The number on the upper RHS quantifies the positive attitude towards Bitcoin.

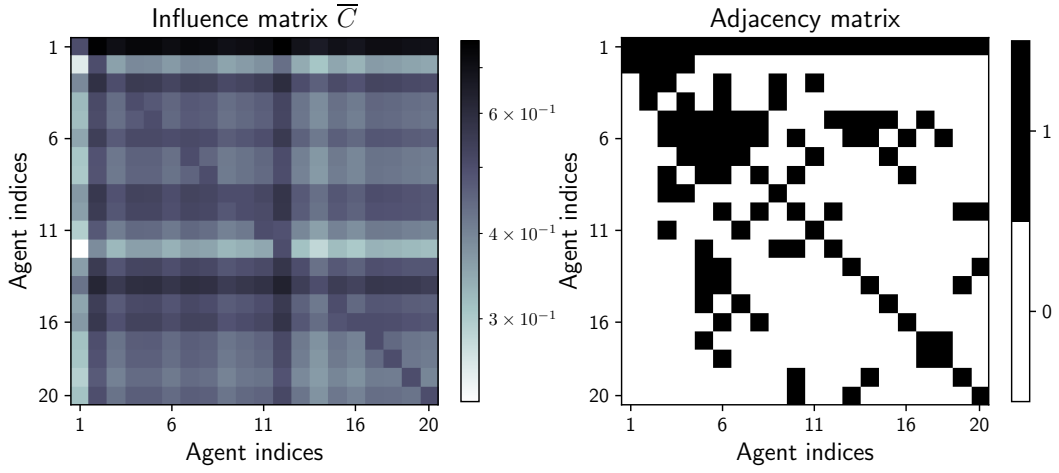


Figure 7.13: (*Left*): Bipartite causal influence matrix. (*Right*): Adjacency matrix corresponding to the sub-network topology in Fig. 7.11.

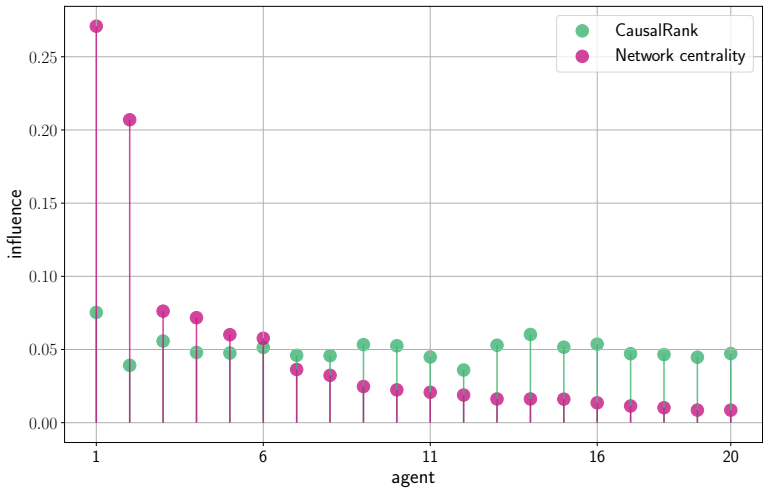


Figure 7.14: Ranking of agents based on their causal influence and their corresponding network eigenvector centrality. Both ranking scores are summing up to one.

users: Elon Musk (User 1) and User 2, respectively.

Upon observing the heat map, it is evident that Elon Musk holds significant influence over all other users, as indicated by the high values in the 1th row, which aligns with our expectations. However, notice that the adjacency matrix does not precisely mirror the causal relationships. For instance, User 2 is followed by Elon Musk, yet their influence on Elon Musk, as depicted in the heat map, is relatively low. On the other hand, User 14 exerts a substantial impact on User 2, despite not being directly followed by User 2. This fact may arise from the fact that User 14 holds one of the highest influences on Elon Musk (User 1) among all the users in this particular sub-network. These observations highlight the fact that the nature of influence dynamics within real-world social networks cannot simply be explained with direct follower relations.

Causal impact ranking. Once the influence matrix is determined, we apply the CausalRank algorithm (Alg. 7.1) to rank the agents based on their overall influence within the sub-network. The resulting plot is depicted in Fig. 7.14. Notably, Elon Musk emerges as the most influential agent, aligning with our initial expectations.

However, an intriguing observation can be made regarding User 2. Despite having a high eigenvector centrality, their causal impact score appears relatively small. This phenomenon arises because the causal effect is not solely determined by network centrality but also takes into account the informativeness of the agents. For instance, if a user primarily retweets (reposts) what their neighbors are tweeting, such users tend to possess low informativeness, decreasing their causal impact score. Thus, even though User 2 may have a high centrality within the considered sub-network primarily due to being followed by Elon Musk, their causal influence on their neighbors is low and does not propagate to other users, leading to a relatively small causal impact.

7.8 Concluding Remarks

In this chapter, we analyzed causal influences among agents that are connected over a network and whose interactions occur over time. Using social learning models, we derived expressions for the causal relationships between pairs of agents. These expressions offer key insights into the diffusion of influences across a social network. We also proposed the CausalRank algorithm for ranking the overall influence of agents, which allows discovering highly influential actors within a network. Furthermore, to enhance the practical usage of our results, we proposed the graph causality learning algorithm (GCL) that learns the necessary model parameters from raw observational data in order to estimate the causal effects. We demonstrated how GCL can be applied in practice through an application to real Twitter data.

The social learning models we considered in this work are useful for both modeling opinion formation over social networks as well as for designing distributed decision-making systems. Therefore, potential applications range from the analysis of human social networks, such as those on social media platforms, to cooperative decision-making processes of socially intelligent machines like networks of drones or sensors. In addition to these, our results can be useful for applications that involve time-series networked interactions, since they provide insights on the diffusion of influence across graphs.

7.A Connection of (7.28) to Average Causal Derivative Effect

In this appendix, we demonstrate how definition (7.28) for the causal effect summary $\overline{C}^{\mathrm{NB}}_{m\to k}$ can also be interpreted as the increment in causal effect $C^{\mathrm{NB}}_{m\to k}$ after an infinitesimal change in the intervention strength $\mu_m(\theta)$. In the literature, this is referred to as the average causal *derivative* effect, as it computes the derivative in causal effect with respect to the intervention strength [148, Chapter 6]. However, a simple derivative of $C^{\mathrm{NB}}_{m\to k}$ with respect to μ_m fails to produce a dose-independent summary, given that $C^{\mathrm{NB}}_{m\to k}$ is not a linear function of μ_m, which can be seen from (7.58).

We introduce the function

$$\begin{aligned} f(C^{\mathrm{NB}}_{m\to k}) &\triangleq \frac{C^{\mathrm{NB}}_{m\to k}}{1-C^{\mathrm{NB}}_{m\to k}} \\ &\overset{(7.58)}{=} \sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\Big\{-\Big[\Big((I-R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta)\Big]_k\Big\}, \end{aligned} \tag{7.75}$$

and notice that $f(C^{\mathrm{NB}}_{m\to k}) \in [0,\infty)$ is a linear function of the intervened belief ratio

vector s_m, defined as

$$s_m(\theta) \triangleq \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)}, \quad s_m \triangleq [s_m(\theta_1), \ldots, s_m(\theta_H)]^\mathsf{T}. \tag{7.76}$$

Here, the vector s_m quantifies the amount of relative misinformation the intervention μ_m produces. Therefore, the gradient

$$\nabla_{s_m} f(C^{\text{NB}}_{m\to k}) = \sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \frac{\partial f(C^{\text{NB}}_{m\to k})}{\partial s_m(\theta)} \overset{(7.75)}{=} \sum_{\theta\in\Theta\setminus\{\theta^\circ\}} \exp\Big\{ -\Big[\Big((I-R^\mathsf{T})^{-1}-I\Big)d_{-m}(\theta)\Big]_k\Big\} \tag{7.77}$$

is independent of the intervention dose μ_m, and satisfies

$$\nabla_{s_m} f(C^{\text{NB}}_{m\to k}) = f(\overline{C}^{\text{NB}}_{m\to k}). \tag{7.78}$$

In other words, to find the gradient of $f(C^{\text{NB}}_{m\to k})$ with respect to s_m, setting

$$\frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} = 1 \Longrightarrow \mu_m(\theta) = \frac{1}{H},\ \forall\theta\in\Theta \tag{7.79}$$

in (7.75) as it was done in (7.28) for finding $\overline{C}^{\text{NB}}_{m\to k}$, is sufficient. The reason we are interested in $\nabla_{s_m} f(C^{\text{NB}}_{m\to k})$ is the following. Notice from (7.75) that $f(C^{\text{NB}}_{m\to k})$ is a monotonic increasing function of $C^{\text{NB}}_{m\to k}$. Also, $s_m(\theta)$ is clearly a monotonic function of $\mu_m(\theta)$ — see (7.76). Therefore, $\nabla_{s_m} f(C^{\text{NB}}_{m\to k})$ can be considered as some proxy for the derivative of the causal effect $C^{\text{NB}}_{m\to k}$ with respect to $\mu_m(\theta)$. Consequently, we conclude that setting $\mu_m(\theta)$ as a uniform belief in $C^{\text{NB}}_{m\to k}$ effectively parallels finding the average causal derivative effect.

7.B Proof of (7.48)

Theorem 2.3, which establishes convergence in distribution for the pre-intervention ASL case does not require the strongly connected graph assumption. In fact, it holds as long as the observations that the agents receive are i.i.d. over time. This condition is satisfied under the post-intervention case as well. Hence, the log-beliefs under intervention converge in distribution, i.e.,

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) \xrightarrow{\text{dist.}} \beta\sum_{j=1}^{\infty}(1-\delta)^{j-1}(\widetilde{A}^\mathsf{T})^j\widetilde{\boldsymbol{x}}_j(\theta). \tag{7.80}$$

Then, following the same arguments from Corollary 2.1, the limiting expectation becomes

$$\widetilde{\lambda}_\infty(\theta) = \frac{\beta}{1-\delta}\Big((I-(1-\delta)\widetilde{A}^\mathsf{T})^{-1}-I\Big)\widetilde{d}(\theta) \tag{7.81}$$

where we define the vector of expected LLRs and the vector of limiting log-belief ratio expectations from across the network:

$$\widetilde{d}(\theta) \triangleq \left[\frac{\delta}{\beta}\log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}, d_2(\theta), \ldots, d_K(\theta)\right]^{\mathsf{T}}, \ \widetilde{\lambda}_\infty(\theta) \triangleq \left[\log\frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}, \lambda_{2,\infty}(\theta), \ldots, \lambda_{K,\infty}(\theta)\right]^{\mathsf{T}}. \tag{7.82}$$

Using the block matrix form of $\widetilde{A}$ from (7.11), we have

$$I-(1-\delta)\widetilde{A}^{\mathsf{T}} = \begin{bmatrix} \delta & 0 \\ -(1-\delta)r & I-(1-\delta)R^{\mathsf{T}} \end{bmatrix}, \tag{7.83}$$

which implies

$$(I-(1-\delta)\widetilde{A}^{\mathsf{T}})^{-1} = \begin{bmatrix} \dfrac{1}{\delta} & 0 \\ \dfrac{1-\delta}{\delta}(I-(1-\delta)R^{\mathsf{T}})^{-1}r & (I-(1-\delta)R^{\mathsf{T}})^{-1} \end{bmatrix}. \tag{7.84}$$

Inserting this into (7.81), we arrive at

$$\widetilde{\lambda}_\infty(\theta) = \frac{\beta}{1-\delta}\begin{bmatrix} \dfrac{1-\delta}{\delta} & 0 \\ \dfrac{1-\delta}{\delta}(I-(1-\delta)R^{\mathsf{T}})^{-1}r & (I-(1-\delta)R^{\mathsf{T}})^{-1}-I \end{bmatrix}\begin{bmatrix} \dfrac{\delta}{\beta}\log\dfrac{\mu_1(\theta^\circ)}{\mu_1(\theta)} \\ d_{-m}(\theta) \end{bmatrix} \tag{7.85}$$

which again verifies that $\widetilde{\lambda}_{1,\infty}(\theta) = \log\dfrac{\mu_1(\theta^\circ)}{\mu_1(\theta)}$ and proves relation (7.48).

7.C Proof of (7.52)

Recall from (7.29)–(7.30) that for this fully connected network:

$$A = v\mathbb{1}_K^{\mathsf{T}}, \quad R = v_{-m}\mathbb{1}_{K-1}^{\mathsf{T}}, \quad r = v_1\mathbb{1}_{K-1}. \tag{7.86}$$

As a result, it holds that

$$\begin{aligned}(I-(1-\delta)R^{\mathsf{T}})^{-1} &= \sum_{i=0}^{\infty}(1-\delta)^i(R^{\mathsf{T}})^i \\ &\overset{(7.32)}{=} I+(1-\delta)\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}}+(1-\delta)^2(1-v_1)\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}}+\ldots \\ &= I+(1-\delta)\Big(\sum_{i=0}^{\infty}(1-\delta)^i(1-v_1)^i\Big)\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}} \\ &= I+\frac{1-\delta}{1-(1-\delta)(1-v_1)}\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}}. \end{aligned} \tag{7.87}$$

This implies that

$$\begin{aligned}(I-(1-\delta)R^{\mathsf{T}})^{-1}r &= \Big(I+\frac{1-\delta}{1-(1-\delta)(1-v_1)}\mathbb{1}_{K-1}v_{-m}^{\mathsf{T}}\Big)v_1\mathbb{1}_{K-1} \\ &\overset{(a)}{=} v_1\Big(\mathbb{1}_{K-1}+\frac{1-\delta}{1-(1-\delta)(1-v_1)}(1-v_1)\mathbb{1}_{K-1}\Big) \\ &= \frac{v_1}{1-(1-\delta)(1-v_1)}\mathbb{1}_{K-1} \end{aligned} \tag{7.88}$$

where (a) follows from the fact that $\sum_{\ell=2}^{K}v_\ell = 1-v_1$. Eq. (7.87) also implies that

$$\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta) = \frac{1-\delta}{1-(1-\delta)(1-v_1)}\Big(\sum_{\ell=2}^{K}v_\ell d_\ell(\theta)\Big)\mathbb{1}_{K-1} \tag{7.89}$$

Incorporating (7.88) and (7.89) into (7.48) concludes the proof.

7.D Proof of (7.54)

Recall expressions (7.38) and (7.39) for A and R in the ring special case. Also observe that

$$r = \Big[1-\alpha, 0, \ldots, 0\Big]^{\mathsf{T}}. \tag{7.90}$$

Accordingly, it holds that

$$(I-(1-\delta)R^{\mathsf{T}}) = \begin{bmatrix} 1-(1-\delta)\alpha & 0 & \cdots & 0 & 0 \\ -(1-\delta)(1-\alpha) & 1-(1-\delta)\alpha & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1-(1-\delta)\alpha & 0 \\ 0 & 0 & \cdots & -(1-\delta)(1-\alpha) & 1-(1-\delta)\alpha \end{bmatrix} \tag{7.91}$$

The inverse of this Toeplitz matrix is given by the following lower diagonal matrix:

$$(I-(1-\delta)R^{\mathsf{T}})^{-1} = \begin{bmatrix} \frac{1}{1-(1-\delta)\alpha} & 0 & \cdots & 0 & 0 \\ \frac{(1-\delta)(1-\alpha)}{(1-(1-\delta)\alpha)^2} & \frac{1}{1-(1-\delta)\alpha} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{(1-\delta)^{K-3}(1-\alpha)^{K-3}}{(1-(1-\delta)\alpha)^{K-2}} & \frac{(1-\delta)^{K-4}(1-\alpha)^{K-4}}{(1-(1-\delta)\alpha)^{K-3}} & \cdots & \frac{1}{1-(1-\delta)\alpha} & 0 \\ \frac{(1-\delta)^{K-2}(1-\alpha)^{K-2}}{(1-(1-\delta)\alpha)^{K-1}} & \frac{(1-\delta)^{K-3}(1-\alpha)^{K-3}}{(1-(1-\delta)\alpha)^{K-2}} & \cdots & \frac{(1-\delta)(1-\alpha)}{(1-(1-\delta)\alpha)^2} & \frac{1}{1-(1-\delta)\alpha} \end{bmatrix}. \tag{7.92}$$

Then, the matrix-vector products in the general formula (7.48) become

$$(I-(1-\delta)R^{\mathsf{T}})^{-1}r = \left[\frac{1-\alpha}{1-(1-\delta)\alpha}, \frac{(1-\delta)(1-\alpha)^2}{(1-(1-\delta)\alpha)^2}, \ldots, \frac{(1-\delta)^{K-2}(1-\alpha)^{K-1}}{(1-(1-\delta)\alpha)^{K-1}}\right]^{\mathsf{T}} \tag{7.93}$$

and

$$\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta) = \begin{bmatrix} \frac{\alpha(1-\delta)}{1-(1-\delta)\alpha}d_2(\theta) \\ \frac{\alpha(1-\delta)}{1-(1-\delta)\alpha}d_3(\theta) + \frac{(1-\alpha)(1-\delta)}{(1-(1-\delta)\alpha)^2}d_2(\theta) \\ \vdots \\ \frac{\alpha(1-\delta)}{1-(1-\delta)\alpha}d_K(\theta) + \cdots + \frac{(1-\alpha)^{K-2}(1-\delta)^{K-2}}{(1-(1-\delta)\alpha)^{K-1}}d_2(\theta) \end{bmatrix} \tag{7.94}$$

Inserting these into the general expression (7.48) concludes the proof.

7.E Proof of Theorem 7.1

If we denote the error in estimating the combination matrix by $E \triangleq \widehat{A} - A$, then combining (7.65) and (7.69) yields:

$$\widehat{\boldsymbol{D}} = \frac{1}{\beta M}\sum_{i=1}^{M}\Big(\beta\mathbf{X}_i - (1-\delta)E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1}\Big). \tag{7.95}$$

Therefore, the error in informativeness can be decomposed as

$$\begin{aligned}\mathbb{E}\|\widehat{\boldsymbol{D}}-D\|_{\mathrm{F}}^2 &= \frac{1}{\beta^2M^2}\mathbb{E}\Big\|\sum_{i=1}^{M}\beta(\mathbf{X}_i-D)-(1-\delta)E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1}\Big\|_{\mathrm{F}}^2\\ &= \frac{1}{M^2}\mathbb{E}\Big\|\sum_{i=1}^{M}(\mathbf{X}_i-D)\Big\|_{\mathrm{F}}^2+\frac{(1-\delta)^2}{\beta^2M^2}\mathbb{E}\Big\|\sum_{i=1}^{M}E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1}\Big\|_{\mathrm{F}}^2\\ &\quad-\frac{2(1-\delta)}{\beta M^2}\mathbb{E}\left[\mathrm{Tr}\Big(\sum_{i=1}^{M}(\mathbf{X}_i-D)\Big)\Big(\sum_{i=1}^{M}(E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1})\Big)^{\mathsf{T}}\right] \qquad (7.96)\end{aligned}$$

where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm and D denotes the informativeness matrix with entries corresponding to $[D]_{kj}\triangleq d_k(\theta_j)$. Here, (i) because of the i.i.d. assumption on data over time, and (ii) under sufficiently small δ values that keep the probability of the event "$\widehat{\boldsymbol{\theta}}^{\circ}\neq\theta^{\circ}$" sufficiently small with increasing M (see Theorem 2.3 and also refer to [14]), the first term satisfies

$$\begin{aligned}\frac{1}{M^2}\mathbb{E}\Big\|\sum_{i=1}^{M}(\mathbf{X}_i-D)\Big\|_{\mathrm{F}}^2 &= \frac{1}{M}\mathbb{E}\big\|\mathbf{X}_i-D\big\|_{\mathrm{F}}^2+\frac{2}{M^2}\sum_{i=1}^{M}\sum_{j<i}\mathbb{E}\mathrm{Tr}\Big[\Big(\mathbf{X}_i-D\Big)\Big(\mathbf{X}_j-D\Big)^{\mathsf{T}}\Big]\\ &= \frac{1}{M}\mathrm{Tr}(\mathcal{R}), \qquad (7.97)\end{aligned}$$

where we also have defined the covariance matrix of the log-likelihood ratios

$$\mathcal{R}\triangleq\mathbb{E}\Big((\mathbf{X}_i-D)(\mathbf{X}_i-D)^{\mathsf{T}}\Big). \qquad (7.98)$$

Moreover, expanding recursion (7.65), it holds that

$$\sum_{i=1}^{M}\boldsymbol{\Lambda}_i=\sum_{i=1}^{M}(1-\delta)^i(A^i)^{\mathsf{T}}\boldsymbol{\Lambda}_0+\beta\sum_{i=1}^{M}\sum_{j=0}^{i-1}(1-\delta)^j(A^j)^{\mathsf{T}}\mathbf{X}_{i-j}. \qquad (7.99)$$

Furthermore, by the triangle inequality, it follows that

$$\begin{aligned}\Big\|\sum_{i=1}^{M}\boldsymbol{\Lambda}_i\Big\|_{\mathrm{F}} &\leq \Big\|\sum_{i=1}^{M}(1-\delta)^i(A^i)^{\mathsf{T}}\boldsymbol{\Lambda}_0\Big\|_{\mathrm{F}}+\beta\Big\|\sum_{i=1}^{M}\sum_{j=0}^{i-1}(1-\delta)^j(A^j)^{\mathsf{T}}\mathbf{X}_{i-j}\Big\|_{\mathrm{F}}\\ &= O\Big(\min\Big\{M^2,\frac{M}{\delta}\Big\}\Big). \qquad (7.100)\end{aligned}$$

Hence, the second term in (7.96) satisfies

$$\frac{(1-\delta)^2}{\beta^2M^2}\mathbb{E}\Big\|\sum_{i=1}^{M}E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1}\Big\|_{\mathrm{F}}^2=O\left(\|E\|^2\min\Big\{M^2,\frac{1}{\delta^2}\Big\}\right) \qquad (7.101)$$

and the third term in (7.96) satisfies

$$\frac{(1-\delta)}{\beta M^2}\mathbb{E}\left[\mathrm{Tr}\Big(\sum_{i=1}^{M}(\mathbf{X}_i - D)\Big)\Big(\sum_{i=1}^{M}(E^{\mathsf{T}}\boldsymbol{\Lambda}_{i-1})\Big)^{\mathsf{T}}\right] = O\left(\frac{\|E\|}{M}\min\left\{M, \frac{1}{\delta}\right\}\right) \tag{7.102}$$

Consequently, for sufficiently small combination matrix errors, namely,

$$E = o\Big(\max\left\{M^{-3/2}, \delta M^{-1/2}\right\}\Big), \tag{7.103}$$

it holds that

$$\mathbb{E}\|\widehat{\boldsymbol{D}} - D\|_{\mathrm{F}}^2 \leq \frac{1}{M}\mathrm{Tr}(\mathcal{R}) + o(1/M). \tag{7.104}$$

Next, recall (7.48) which establishes log-belief ratios under interventions for ASL. For the first term in (7.48), the matrix estimation error E implies errors for the matrix components R and r that are also proportional to E, which means

$$\begin{aligned}
&(I-(1-\delta)(R+O(E))^{\mathsf{T}})^{-1}(r+O(E)) \\
&\quad = (I-(1-\delta)R^{\mathsf{T}})^{-1}\Big(I-\Big(I-(1-\delta)R^{\mathsf{T}}\Big)^{-1}(1-\delta)O(E)\Big)^{-1}\Big(r+O(E)\Big) \\
&\quad \overset{(a)}{=} (I-(1-\delta)R^{\mathsf{T}})^{-1}\Big(I+(1-\delta)O(E)\Big)\Big(r+O(E)\Big) \\
&\quad = \Big((I-(1-\delta)R^{\mathsf{T}})^{-1}+(1-\delta)O(E)\Big)\Big(r+O(E)\Big) \\
&\quad = (I-(1-\delta)R^{\mathsf{T}})^{-1}r+O(E)
\end{aligned} \tag{7.105}$$

where (a) holds as long as

$$\begin{aligned}
E &= o\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}\Big) \\
&= o\Big(1/(1-(1-\delta)\rho(R))\Big) \\
&= o\Big(1/\Big(1-\rho(R)+\delta\rho(R)\Big)\Big) \\
&= o\Big(1/(1-\rho(R))\Big)
\end{aligned} \tag{7.106}$$

In (7.106) we assume that the graph topology is fixed (i.e., it is not changing with decreasing δ). By the same arguments, an error in the matrix expression of the second term of (7.48) implies

$$\begin{aligned}
\Big(\Big(I-(1-\delta)(R+O(E))^{\mathsf{T}}\Big)^{-1}-I\Big) &= \Big((I-(1-\delta)R^{\mathsf{T}})^{-1}+(1-\delta)O(E)-I\Big) \\
&= \Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\Big)+O(E),
\end{aligned} \tag{7.107}$$

and an error in the matrix expression of the pre-intervention case of (2.40) implies

$$\begin{aligned}\left((I-(1-\delta)\widehat{A}^{\mathsf{T}})^{-1}-I\right) &= \left((I-(1-\delta)(A+E)^{\mathsf{T}})^{-1}-I\right)\\ &= \left((I-(1-\delta)A^{\mathsf{T}})^{-1}-I\right)+O(E).\end{aligned} \tag{7.108}$$

Therefore, considering the fact that true causal effect can be written as

$$\begin{aligned}&C^{\text{ASL}}_{m\to k}\\ &=\mu_{k,\infty}(\theta^{\circ})-\widetilde{\mu}_{k,\infty}(\theta^{\circ})\\ &\overset{(7.3)}{=}\frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\exp\{-\lambda_{k,\infty}(\theta)\}}-\frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\exp\{-\widetilde{\lambda}_{k,\infty}(\theta)\}}\\ &\overset{(2.40),(7.48)}{=}\frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\exp\left\{-\left[\frac{\beta}{1-\delta}\Big((I-(1-\delta)A^{\mathsf{T}})^{-1}-I\Big)d(\theta)\right]_k\right\}}\\ &\quad-\frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\left(\frac{\mu_m(\theta)}{\mu_m(\theta^{\circ})}\right)^{[(I-(1-\delta)R^{\mathsf{T}})^{-1}r]_k}\exp\left\{-\left[\frac{\beta}{1-\delta}\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\Big)d_{-m}(\theta)\right]_k\right\}},\end{aligned} \tag{7.109}$$

by (7.105), (7.107), and (7.108), the causal effect estimate satisfies

$$\begin{aligned}&\widehat{C}^{\text{ASL}}_{m\to k}\\ &=\frac{1}{1+\sum_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\exp\left\{-\left[\frac{\beta}{1-\delta}\Big((I-(1-\delta)A^{\mathsf{T}})^{-1}-I+O(E)\Big)\widehat{\boldsymbol{d}}(\theta)\right]_k\right\}}\\ &\quad-\frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\left(\frac{\mu_m(\theta)}{\mu_m(\theta^{\circ})}\right)^{[(I-(1-\delta)R^{\mathsf{T}})^{-1}r]_k+O(E)}\exp\left\{-\left[\frac{\beta}{1-\delta}\Big((I-(1-\delta)R^{\mathsf{T}})^{-1}-I+O(E)\Big)\widehat{\boldsymbol{d}}_{-m}(\theta)\right]_k\right\}}\end{aligned} \tag{7.110}$$

Here, $\widehat{\boldsymbol{d}}(\theta)$ and $\widehat{\boldsymbol{d}}_{-m}(\theta)$ are the estimates for $d(\theta)$ and $d_{-m}(\theta)$ respectively, which can be formed from informativeness estimation matrix $\widehat{\boldsymbol{D}}$. In addition, the scalar function g with vector argument λ (over $\Theta\setminus\{\theta^{\circ}\}$) defined as

$$g(\lambda)\triangleq\frac{1}{1+\sum\limits_{\theta\in\Theta\setminus\{\theta^{\circ}\}}\exp\{-\lambda(\theta)\}} \tag{7.111}$$

has the partial derivatives

$$\frac{\partial g(\lambda)}{\partial \lambda(\theta)} = \frac{\exp\{-\lambda(\theta)\}}{\Big(1 + \sum\limits_{\theta \in \Theta \setminus \{\theta^\circ\}} \exp\{-\lambda(\theta)\}\Big)^2}, \tag{7.112}$$

which implies that

$$\begin{aligned} \|\nabla g(\lambda)\|_1 &= \sum_{\theta \in \Theta \setminus \{\theta^\circ\}} \left| \frac{\partial g(\lambda)}{\partial \lambda(\theta)} \right| \\ &= \sum_{\theta \in \Theta \setminus \{\theta^\circ\}} \frac{\exp\{-\lambda(\theta)\}}{\Big(1 + \sum\limits_{\theta \in \Theta \setminus \{\theta^\circ\}} \exp\{-\lambda(\theta)\}\Big)^2} \ \leq 1 \end{aligned} \tag{7.113}$$

By the mean-value theorem, this further implies that for any two vectors λ_1 and λ_2, there exists a vector λ_3 such that

$$g(\lambda_1) - g(\lambda_2) = \nabla g(\lambda_3)^\mathsf{T} (\lambda_1 - \lambda_2). \tag{7.114}$$

Then, using Hölder's inequality on (7.114) yields that

$$\Big| g(\lambda_1) - g(\lambda_2) \Big| \leq \Big\| \nabla g(\lambda_3) \Big\|_1 \Big\| \lambda_1 - \lambda_2 \Big\|_\infty \overset{(7.113)}{\leq} \Big\| \lambda_1 - \lambda_2 \Big\|_\infty. \tag{7.115}$$

Note that the norm $\|\cdot\|_\infty$ chooses the absolute maximum over all arguments induced from different hypotheses $\theta \in \Theta \setminus \{\theta^\circ\}$. If we apply (7.115) to the differences of first and second terms in (7.109) and (7.110), we get

$$\begin{aligned} &\Big| C^{\text{ASL}}_{m \to k} - \widehat{C}^{\text{ASL}}_{m \to k} \Big| \\ &\leq \frac{\beta}{1-\delta} \left\| \left[\Big((I - (1-\delta) A^\mathsf{T})^{-1} - I \Big) d(\theta) - \Big((I - (1-\delta) A^\mathsf{T})^{-1} - I + O(E) \Big) \widehat{\boldsymbol{d}}(\theta) \right]_k \right\|_\infty \\ &+ \left\| O(E) \log \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} + \frac{\beta}{1-\delta} \Big[\Big((I - (1-\delta) R^\mathsf{T})^{-1} - I \Big) d_{-m}(\theta) \right. \\ &\qquad \left. - \Big((I - (1-\delta) R^\mathsf{T})^{-1} - I + O(E) \Big) \widehat{\boldsymbol{d}}_{-m}(\theta) \Big]_k \right\|_\infty \\ &\leq \frac{\beta}{1-\delta} \left\| \left[\Big((I - (1-\delta) A^\mathsf{T})^{-1} - I \Big) \Big(d(\theta) - \widehat{\boldsymbol{d}}(\theta) \Big) + O(E) \widehat{\boldsymbol{d}}(\theta) \right]_k \right\|_\infty \\ &\quad + \left\| O(E) + \frac{\beta}{1-\delta} \Big[\Big((I - (1-\delta) R^\mathsf{T})^{-1} - I \Big) \Big(d_{-m}(\theta) - \widehat{\boldsymbol{d}}_{-m}(\theta) \Big) - O(E) \widehat{\boldsymbol{d}}_{-m}(\theta) \Big]_k \right\|_\infty \end{aligned}$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \frac{\beta}{1-\delta}\left\|\left[\left((I-(1-\delta)A^{\mathsf{T}})^{-1}-I\right)\left(d(\theta)-\widehat{\boldsymbol{d}}(\theta)\right)\right]_k\right\|_\infty \\
&\quad+\frac{\beta}{1-\delta}\left\|\left((I-(1-\delta)R^{\mathsf{T}})^{-1}-I\right)\left(d_{-m}(\theta)-\widehat{\boldsymbol{d}}_{-m}(\theta)\right)\right\|_\infty+O(E) \\
&\overset{(b)}{\leq} \frac{\beta}{1-\delta}\left\|(I-(1-\delta)A^{\mathsf{T}})^{-1}-I\right\|_\infty\left\|d(\theta)-\widehat{\boldsymbol{d}}(\theta)\right\|_\infty \\
&\quad+\frac{\beta}{1-\delta}\left\|(I-(1-\delta)R^{\mathsf{T}})^{-1}-I\right\|_\infty\left\|d_{-m}(\theta)-\widehat{\boldsymbol{d}}_{-m}(\theta)\right\|_\infty+O(E) \qquad (7.116)
\end{aligned}
$$

where (a) follows from the triangle inequality, and (b) follows from the sub-multiplicity of the matrix norms. Finally, since (7.104) implies

$$\mathbb{E}\left\|d(\theta)-\widehat{\boldsymbol{d}}(\theta)\right\|_\infty \leq \frac{1}{M}\mathrm{Tr}(\mathcal{R})+o(1/M) \qquad (7.117)$$

and

$$\mathbb{E}\left\|d_{-m}(\theta)-\widehat{\boldsymbol{d}}_{-m}(\theta)\right\|_\infty \leq \frac{1}{M}\mathrm{Tr}(\mathcal{R})+o(1/M), \qquad (7.118)$$

taking the expectation of both sides in (7.116) and incorporating (7.117) and (7.118) concludes that

$$\mathbb{E}\left|C^{\mathrm{ASL}}_{m\to k}-\widehat{\boldsymbol{C}}^{\mathrm{ASL}}_{m\to k}\right| = O(1/M). \qquad (7.119)$$

In the case of NBSL, the only source of disagreement between $C^{\mathrm{NB}}_{m\to k}$ and $\widehat{\boldsymbol{C}}^{\mathrm{NB}}_{m\to k}$ is at post-intervention, since pre-intervention steady state beliefs are fixed and identical (see (7.8)), thus their difference equals 0. Also, the proof for the post-intervention difference proceeds along the same lines as the ASL case above.

7.F Discussion of Computational Complexity

First, assuming that the matrix A and the informativeness vector $d(\theta)$ for each hypothesis are known, the computational tasks for calculating the derived causal effect expressions (7.22) and (7.48) involve the following:

- Calculating $(I-(1-\delta)R^{\mathsf{T}})^{-1}$ for one agent can be achieved in $O(K^3)$ time with naive matrix operations, where K is the size of the network. Therefore, performing this operation for all agents in the network results in $O(K^4)$ worst case complexity.
- For the matrix-vector multiplications, the complexity for each agent is $O(K^2)$, which leads to a complexity of $O(K^3)$ when we consider all K agents.
- The computation to find the eigenvectors of the $K\times K$ causal influence matrix also involves a computation time of $O(K^3)$.

Therefore, the total computational complexity with known A and $d(\theta)$ is $O(K^4)$. Note that if one uses computationally efficient algorithms for matrix multiplication, such as the Strassen's algorithm [154], the complexity can potentially be reduced down to $O(K^w)$, where $3 < w < 4$.

If, however, A and $d(\theta)$ need to be estimated first as in Alg. 7.2, we need to perform the following additional operations:

- Estimating A, based on the averaging rule, is $O(K)$ or $O(K^2)$ depending on the utilized graph data structure.
- Estimating $d(\theta)$ involves learning from the average over M observations. Here, each computation involves a matrix multiplication between A and Λ (i.e., the observational data) which has $O(K^2)$ complexity.

Consequently, considering causal discovery from observational data phase, the total computational complexity for causal ranking of the network is $O(K^4 + M \cdot K^2)$.

8 Causality under Asynchronicity

8.1 Introduction[1]

In this chapter, we extend the causal influence concept from the previous chapter. Namely, we build upon the collaborative decision-making framework of [59], which involves heterogeneous agents exchanging beliefs (or soft-decisions) through a fusion center (FC) based on streaming observations. Furthermore, to better capture the real-world conditions, we will incorporate two asynchronicity scenarios to this framework; the scenarios differ in agent participation patterns and in FC policies.

8.1.1 Contributions

In a manner similar to the discussion in the previous chapter, by applying hypothetical interventions [139, 148] to our model, we implement a method to calculate causal impact scores for each agent's contribution to the joint decision. We also provide a theoretical analysis of participation patterns, FC policies, and data distribution on the decision-making process. We illustrate our theoretical findings with numerical simulations and also apply our methods to real-world data from a multi-camera crowd-size estimation application [157].

8.2 Problem Setting

Following the notation from the previous chapter, we use $\sim$ to denote the counterparts of variables after an intervention (e.g., $\widetilde{\lambda}$ represents the variable λ after an intervention).

[1] The material in this chapter is based on [155, 156].

8.2.1 Two Asynchronous Scenarios

Asynchronous behavior is common in many real-world distributed systems, and is particularly relevant for ad-hoc networks where time scheduling beforehand may not be plausible. To that end, we consider two scenarios that are distinct based on the symmetry of communication between the agents and the FC. For both scenarios, we use the Bernoulli variable $\boldsymbol{q}_{k,i}$ (with parameter p_k) to indicate if agent k is sharing its intermediate belief $\boldsymbol{\psi}_{k,i}$ with the server at time i, namely,

$$\boldsymbol{q}_{k,i} = \begin{cases} 1, & \text{with probability } p_k \\ 0, & \text{otherwise} \end{cases} . \tag{8.1}$$

We assume the process $\{\boldsymbol{q}_{k,i}\}$ is independent and identically distributed (i.i.d.) over time and independent over space.

Asymmetric communication

There can be instances when agents, despite being active, do not transmit information to the FC and remain idle in terms of data sharing. This non-engagement can be due to various factors, such as the need to conserve energy, particularly important for battery-operated agents where excessive transmission can lead to rapid battery depletion. Other reasons might include non-informative soft decisions, or the lack of significant changes in intermediate statistics since the previous transmission. However, these agents can keep receiving updates from the server. A possible reason for this disparity is that the uplink cost (from agent to server) is typically higher than the downlink cost (from server to agent). For instance, the downlink could be broadcast, i.e., the same message is transmitted to all agents without the need to exchange information separately with each individual agent.

In this case, the FC can fill the belief components of missing agents with its own prior while aggregating information. Therefore, the combination step at the server side changes to

$$\boldsymbol{\mu}_i(\theta) \propto \prod_{k=1}^{K} \left(\boldsymbol{\psi}_{k,i}^{\boldsymbol{q}_{k,i}}(\theta) \boldsymbol{\mu}_{i-1}^{1-\boldsymbol{q}_{k,i}}(\theta) \right)^{v_k} . \tag{8.2}$$

Nevertheless, agents continue to utilize the beliefs received from the server locally as in (2.7). The procedure under asymmetric asynchronicity is summarized in Algorithm 8.1.

It is worth noting the parallel between this scenario and the traditional distributed detection strategies [12, 158, 159]. Since the server knows the previous combined belief $\boldsymbol{\mu}_{i-1}$, the action of sharing intermediate beliefs $\{\boldsymbol{\psi}_{k,i}\}$ by agents is essentially equivalent to them sharing the observation likelihoods $\{L_k(\boldsymbol{\xi}_{k,i}|\theta)\}$. Similarly, nodes

(e.g., sensors) relay a sufficient statistics such as their local likelihoods or likelihood ratios to the FC in [12, 158, 159]. The difference is that in those works, the FC does not communicate any information back to the nodes.

Algorithm 8.1 Asymmetric communication

set initial prior to $\boldsymbol{\mu}_0(\theta) > 0$, $\forall \theta \in \Theta$ and $\forall k$
while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{i-1}(\theta) \tag{8.3}$$

    send the intermediate belief to FC *if* $\boldsymbol{q}_{k,i} = 1$ (with probability p_k)
  end for
  FC **combines** the received beliefs and its own prior:

$$\boldsymbol{\mu}_i(\theta) \propto \prod_{k=1}^{K} \left(\boldsymbol{\psi}_{k,i}^{\boldsymbol{q}_{k,i}}(\theta) \boldsymbol{\mu}_{i-1}^{1-\boldsymbol{q}_{k,i}}(\theta) \right)^{v_k} \tag{8.4}$$

  FC **broadcasts** combined belief $\boldsymbol{\mu}_i$ to *all* agents
  $i \leftarrow i + 1$
end while

Symmetric communication

Another possibility is that an agent does not receive any update from the server if that agent has not transmitted information to the central processor at that time instant. In other words, the absence of communication is reciprocal. This situation could arise in cases where the quality of the connection is not adequate for reliable communication. Alternatively, for various reasons, a server might strategically choose not to update agents that do not contribute information. By doing so, it can incentivize data sharing and promote a give-and-take dynamics. In this scenario, the combination step at the server side is given by (8.2), whereas the adaptation step at the agents becomes

$$\boldsymbol{\psi}_{k,i}(\theta) \propto \begin{cases} L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 1 \\ L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\psi}_{k,i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 0 \end{cases}. \tag{8.5}$$

The rationale behind (8.5) is as follows. If agent k has shared information with the server (i.e., $\boldsymbol{q}_{k,i-1} = 1$) at time $i-1$, the server returns the combined belief $\boldsymbol{\mu}_{i-1}$ to that agent. On the other hand, if the agent has not participated in the information exchange (i.e., $\boldsymbol{q}_{k,i-1} = 0$), then the server does not provide the updated belief and the agent resorts to its own belief $\boldsymbol{\psi}_{k,i-1}$ as a prior for the update at the next time instant i.

The procedure under symmetric asynchronicity is summarized in Algorithm 8.2.

Algorithm 8.2 Symmetric communication

set initial prior to $\boldsymbol{\mu}_0(\theta) > 0$, $\forall \theta \in \Theta$ and $\forall k$
while $i \geq 1$ **do**
  for each agent $k = 1, 2, \ldots, K$ **do**
    receive private observation $\boldsymbol{\xi}_{k,i}$
    adapt to obtain intermediate belief:

$$\boldsymbol{\psi}_{k,i}(\theta) \propto \begin{cases} L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 1 \\ L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\psi}_{k,i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 0 \end{cases} \tag{8.6}$$

    send the intermediate belief to FC *if* $\boldsymbol{q}_{k,i} = 1$ (with probability p_k)
  end for
  FC **combines** the received beliefs and its own prior:

$$\boldsymbol{\mu}_i(\theta) \propto \prod_{k=1}^{K} \left(\boldsymbol{\psi}_{k,i}^{\boldsymbol{q}_{k,i}}(\theta) \boldsymbol{\mu}_{i-1}^{1-\boldsymbol{q}_{k,i}}(\theta) \right)^{v_k} \tag{8.7}$$

  FC **sends** $\boldsymbol{\mu}_i$ *only* to agents that have participated in the cooperation in the current round (i.e., $\boldsymbol{q}_{k,i} = 1$)
  $i \leftarrow i + 1$
end while

8.3 Causal Influence

We extend the causal effect definition from the previous chapter. The main motivation for the definition is that the influence of an agent m on the collective decision should be proportional to the "amount" by which the outcome changes when this agent is intervened upon. In other words, the alteration of the outcome under a manipulation on the agent quantifies the causal influence. To this end, when an intervention occurs on agent m, we decouple its belief $\boldsymbol{\psi}_{m,i}$ from other beliefs and observations and *fix* it at some constant pmf, say, $\boldsymbol{\psi}_{m,i} = \mu_m$ — see Fig. 8.1 for a representation of an intervention on Fig. 2.2.

As an illustration, recall our recurring example on cooperative vehicular networks. Consider a scenario in which these vehicles navigate a dry road while receiving noisy data from their sensors. To measure the causal effect of an individual vehicle on the collective decision, we can ask the following question: How would the group's decision on road conditions change if a single vehicle, despite the actual conditions and data from other vehicles, consistently reported that the road is icy? A significant difference in the collective decision would imply an influential role for that vehicle. Conversely, a minimal alteration suggests a negligible causal effect. Furthermore, this effect is a *causal* effect since the hypothetical intervention we consider directly targets the

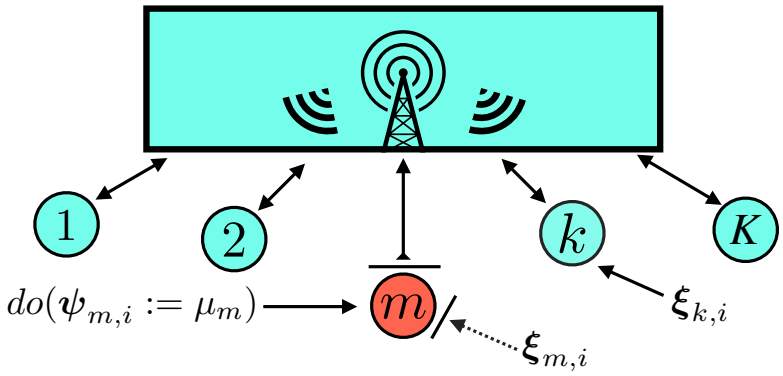


Figure 8.1: Visual representation of a hypothetical intervention $do(\boldsymbol{\psi}_{m,i} := \mu_m)$ on the graphical model in Fig. 8.1. Agent m keeps sending information to the server with probability p_m, however, its belief is now fixed and is not dependent on any other variable.

agent. Namely, it is irrespective of other environmental factors and vehicles that might otherwise induce non-causal correlations.

We establish in Theorem 8.1 that in the absence of any intervention, the belief vector $\boldsymbol{\mu}_i$ converges to a steady-state value μ_∞ that places a probability value of 1 on the true hypothesis θ° as $i \to \infty$. When an intervention occurs at agent m, the steady-state belief vector will be denoted by $\widetilde{\mu}_\infty$. As such, we can quantify the causal impact of agent m on the joint decision by using the difference:

$$C_m \triangleq 1 - \widetilde{\mu}_\infty(\theta^\circ). \tag{8.8}$$

Expression (8.8) measures the expected shift in the steady-state belief on the true hypothesis θ° due to an intervention on agent m. Note that as in previous chapter, we keep expressing the belief $\widetilde{\mu}_\infty(\theta^\circ)$ in the form:

$$\widetilde{\mu}_\infty(\theta^\circ) \triangleq \frac{1}{1 + \sum\limits_{\theta \neq \theta^\circ} \exp\{-\widetilde{\lambda}_\infty(\theta)\}}. \tag{8.9}$$

The variable $\widetilde{\lambda}_\infty(\theta)$ is defined as follows. First, we introduce the notation

$$\lambda_\infty(\theta) \triangleq \lim_{i \to \infty} \mathbb{E}[\boldsymbol{\lambda}_i(\theta)] \tag{8.10}$$

to represent the expected log-belief ratio in the limit with the variables $\boldsymbol{\lambda}_i(\theta)$ defined by

$$\boldsymbol{\lambda}_i(\theta) \triangleq \log \frac{\boldsymbol{\mu}_i(\theta^\circ)}{\boldsymbol{\mu}_i(\theta)}. \tag{8.11}$$

Then, we recall that we use $\sim$ to denote interventional counterparts of these variables,

which means

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) \triangleq \log \frac{\widetilde{\boldsymbol{\mu}}_i(\theta^\circ)}{\widetilde{\boldsymbol{\mu}}_i(\theta)} \tag{8.12}$$

represents the log-belief ratio under an intervention. Therefore, the variable

$$\widetilde{\lambda}_\infty(\theta) \triangleq \lim_{i\to\infty} \mathbb{E}[\widetilde{\boldsymbol{\lambda}}_i(\theta)] \tag{8.13}$$

represents the expected asymptotic log-belief ratio under an intervention.

8.4 Theoretical Results

Theorem 8.1 (**Pre-intervention**). *For the synchronous as well as the symmetric and asymmetric asynchronous communication protocols discussed in Sec. 8.2, the belief vector $\boldsymbol{\mu}_i$ converges to a steady-state probability mass function that places a value of 1 on the true hypothesis θ° almost surely:*

$$\lim_{i\to\infty} \boldsymbol{\mu}_i(\theta^\circ) = 1 \quad \textit{with probability 1.} \tag{8.14}$$

Proof. See Appendix 8.A. ■

The causal influence of an agent m on the joint decision is characterized by the shift of the overall decisions between pre and post-interventions. Therefore, we proceed to examine the beliefs under an intervention on agent m. We first recall the causal impact result (7.34) from previous chapter for the sake of completeness.

Theorem 8.2 (**Synchronous collaboration**). *Under synchronous collaboration, the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_\infty(\theta) = \frac{1}{v_m} \sum_{k\neq m} v_k d_k(\theta) + \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{8.15}$$

Therefore, by (8.8)*, the causal impact of agent m on the joint decision is*

$$C_m = 1 - \frac{1}{1 + \sum\limits_{\theta\neq\theta^\circ} \dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\left\{-\dfrac{1}{v_m} \sum\limits_{k\neq m} v_k d_k(\theta)\right\}} \tag{8.16}$$

Equations (8.15) and (8.16) imply that

- An increase in the confidence v_m by the fusion center increases the causal impact of agent m.

- Increasing the informativeness and confidence weights of the other agents decreases the impact of agent m.

Also, observe that (8.15) and (8.16) are dependent on the intervention strength μ_m. Next, we consider the causal influences for the asynchronous scenarios we have introduced in Sec. 8.2.1.

Theorem 8.3 (**Asymmetric communication**). *Under the asymmetric communication protocol described in Sec. 8.2.1, the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_{\infty}(\theta) = \frac{1}{v_m} \sum_{k \neq m} v_k p_k d_k(\theta) + p_m \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{8.17}$$

This implies by (8.8) *that the causal effect of agent* m *on the joint decision is given by*

$$C_m = 1 - \frac{1}{1 + \sum\limits_{\theta \neq \theta^\circ} \left(\frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \right)^{p_m} \exp\left\{ -\frac{1}{v_m} \sum\limits_{k \neq m} v_k p_k d_k(\theta) \right\}} \tag{8.18}$$

Proof. See Appendix 8.B. ■

Notice in Theorem 8.3 that as p_k approaches 1 for each agent k, i.e., when all agents participate synchronously at each iteration, we recover Theorem 8.2. Also notice that the essential difference from the synchronous scenario is the replacement of confidence weights v_k by $v_k p_k$. This is intuitive since more participation by an agent is expected to increase its influence on the joint decision, as if it had a higher confidence from the server. Similarly, more participation by the other agents decreases the overall impact of an agent on the joint decision, as the "relative" participation of the agent is decreasing compared to the others.

Theorem 8.4 (**Symmetric communication**). *Under the symmetric communication protocol described in Sec. 8.2.1 the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_{\infty}(\theta) = \frac{1}{v_m p_m} \sum_{k \neq m} \frac{v_k d_k(\theta)}{1 - v_k(1 - p_k)} + \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{8.19}$$

This implies by (8.8) *that the causal effect of agent* m *on the joint decision is given by*

$$C_m = 1 - \frac{1}{1 + \sum\limits_{\theta \neq \theta^\circ} \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\left\{ \frac{-1}{v_m p_m} \sum\limits_{k \neq m} \frac{v_k d_k(\theta)}{1 - v_k(1 - p_k)} \right\}} \tag{8.20}$$

Proof. See Appendix 8.C. ■

Similar to the asymmetric communication scenario in Theorem 8.3, as $p_k \to 1$ for all agents, Theorem 8.4 recovers the synchronous collaboration result of Theorem 8.2. Furthermore, as $p_m \to 0$, notice that $\widetilde{\lambda}_\infty(\theta) \to \infty$, which in turn implies $C_m \to 0$. In other words, if an agent does not participate in the decision making, it does not have any impact on the decision.

Next, we compare the causal impacts of agents under both asymmetric and symmetric communication schemes, given the same participation $\{p_k\}_{k=1}^K$, confidence weight $\{v_k\}_{k=1}^K$, and informativeness parameters $\{d_k(\theta)\}_{k=1}^K$.

Corollary 8.1 (**Comparison of asynchronous scenarios**). *Agent m exerts a stronger causal impact on the joint decision in the symmetric scenario compared to the asymmetric scenario if the misinformation strength (i.e., intervened belief) satisfies*

$$\log \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \geq \sum_{k \neq m} \frac{v_k d_k(\theta)}{v_m(1-p_m)} \left(\frac{1}{p_m(1-v_k(1-p_k))} - p_k \right) \tag{8.21}$$

Proof. Notice from (8.17) and (8.19) if agent m meets condition (8.21), then the $\widetilde{\lambda}_\infty(\theta)$ term in (8.17) exceeds that in (8.19). The result then follows by definition (8.9), since $\widetilde{\lambda}_\infty(\theta)$ is inversely proportional to the causal impact C_m. ■

Corollary 8.1 holds significant relevance for practical applications. For our problem setting, we can define misinformation as the ratio of the belief on a wrong hypothesis to the true hypothesis, i.e., $\dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)}$. Commonly, if misinformation is originating from *malfunctioning* agents, this ratio will have a moderate value. In contrast, *malicious* agents often supply adversarial misinformation that can be extreme. This suggests that the symmetric communication scenario is more vulnerable to highly outlying information potentially caused by adversarial agents, while asymmetric communication is more sensitive to moderate level misinformation that typically emerges from malfunctioning agents without harmful intentions. Furthermore, for a fair decision-making process that aims to account for all agents while remaining resilient against adversarial threats, asymmetric communication appears to be better in comparison to the symmetric case. This is because it allocates greater causal weight to moderate deviations from the nominal belief state while also reducing the influence of extreme misinformation, providing a safeguard against adversarial attacks.

8.5 Numerical Simulations

8.5.1 Synthetic Data

To illustrate our theoretical results, we first consider a binary hypothesis testing problem with $K = 12$ agents connected to a FC, each receiving observations that follow a Gaussian distribution. In other words, two possible hypotheses underlie streaming data with same variance Gaussian distributions but different means. Under the null hypothesis, the mean for all agents is assumed to be 0, while under the alternative hypothesis, it is 0.5 for odd-indexed agents and 1 for even-indexed agents. The probability of participation p_k, which is defined in (8.1) is set to 0.8 for each agent k with indices $1-3$, to 0.6 for agent indices $4-6$, 0.4 for agent indices $7-9$, and 0.2 for agent indices $10-12$. Furthermore, the confidence weight v_k assigned by the server to each agent k is 0.125 for agent indices $1-4$, 0.075 for agent indices $5-8$, and 0.05 for agent indices $9-12$, ensuring that the sum of all weights across the $K = 12$ agents equals 1.

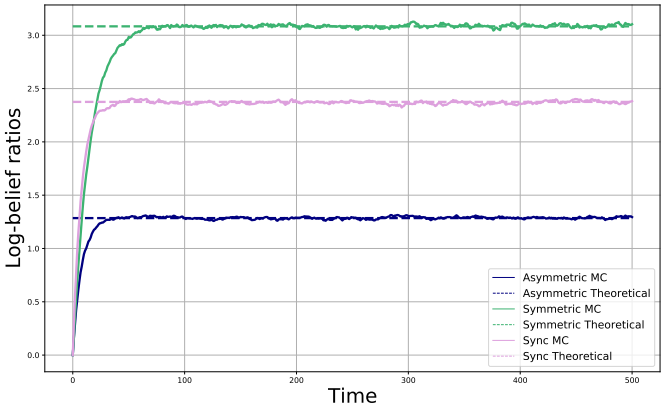


Figure 8.2: Simulated log-belief ratios averaged over 1000 Monte Carlo (MC) simulations and theoretical expressions over time overlap with each other, verifying the derivations in Theorems 8.2–8.4.

In the first experiment, we average 1000 simulations for three settings: the canonical synchronous setting, and the asymmetric and symmetric settings from Sec. 8.2.1. This is performed under an intervention on agent $m = 1$ with uniform beliefs ($\mu_m(\theta) = 0.5$). We plot the evolution of log-belief ratios over 500 time instants in Fig. 8.2, as well as the derived expressions for these values from Theorems 8.2, 8.3, and 8.4. Notice that the simulated log-belief ratios verify the derived analytical results since they closely align with the theoretical expressions.

In Fig. 8.3, we illustrate the causal impact of agent $m = 1$ on the joint decision with respect to changing participation probability p_m. We also include the synchronous setting where all agents participate with probability $\{p_k\}_{k=1}^K = 1$ as a reference. It is evident from this figure that increasing the frequency of information transmission by an agent increases its impact on the collaborative decision.

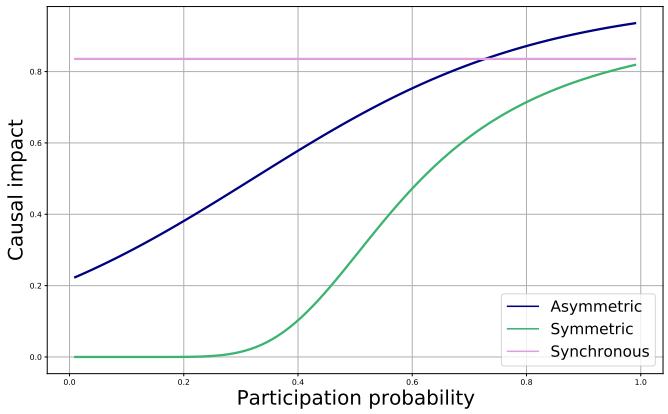


Figure 8.3: Causal impact of agent $m = 1$ on the joint decision with changing participation probability p_m. Note that p_m is constant for the synchronous case, and the corresponding constant line is also provided in the plot for comparison purposes.

Next, in Fig. 8.4, we plot the asymptotic log-belief ratios in relation to varying intervention strengths $\log \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)}$ on agent $m = 1$. Supporting our theoretical result in (8.21), the log-belief ratio in the asymmetric setting surpasses the one in the symmetric setting when the misinformation strength exceeds a certain threshold. As discussed before, this means that under conditions of high misinformation supply, the asymmetric communication framework assigns a relatively smaller causal impact compared to the symmetric communication framework.

Finally, in Fig. 8.5 we plot the causal impact of each agent on the joint decision which are normalized such that the sum of agents' impacts under each strategy equals to 1. This plot reveals that the asymmetric communication protocol results in a more uniform distribution of impacts, whereas the symmetric communication approach leads to a few agents having significant influence on the joint decision. This supports our discussions, suggesting that asymmetric communication fosters a fairer decision-making process that assigns a more uniform impact over participating agents under moderate deviations.

8.5.2 Application: Multi-Camera Crowd Counting

Next, we apply our results to a multi-view crowd-size estimation application using the WILDTRACK dataset from [157]. This dataset consists of synchronized video frames captured by seven static cameras (functioning as agents in our model, $K = 7$) with overlapping fields of view — see Fig. 8.6 for sample images.

The primary goal of the agents is to cooperatively track the dynamic size of the crowd in a specific overlapping region observed by all cameras. For this particular application, the aforementioned variables in the paper correspond to the following:

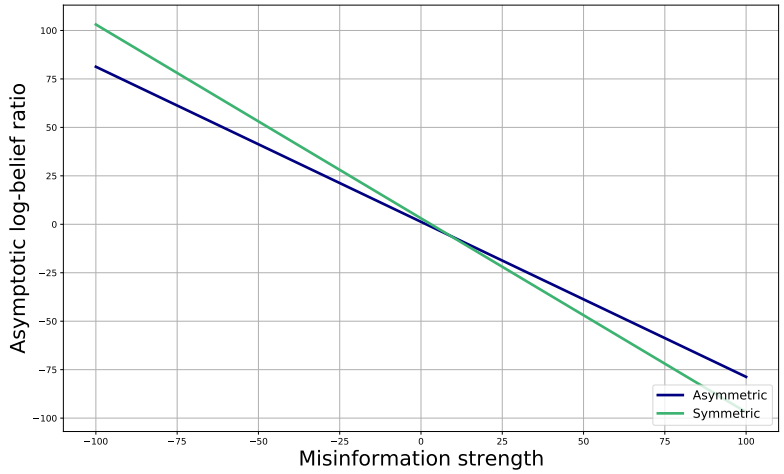


Figure 8.4: Asymptotic log-belief ratio with respect to misinformation strength the $\log \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)}$, verifying the derived threshold in Corollary 8.1.

- For each agent k (a camera), observation $\boldsymbol{\xi}_{k,i}$ corresponds to that agent's own recorded image frame at time instant i. Note that the cameras record the environment with 60 frames per second, that is, 60 time instants correspond to one second in total.

- A hypothesis $\theta \in \Theta$ is a possible integer for the crowd size, and $\Theta = \{0, 1, \ldots, 50\}$ is the set of all possible hypotheses. For the current application at hand, it is known that the number of people in the region of interest will not surpass 50.

- To apply our algorithms in this chapter, we equip the agents with the pre-trained crowd counting neural network from [160]. We then calibrate the likelihood functions of the agents by using the neural network estimates as well some dataset specific samples in order to obtain $L_k(\boldsymbol{\xi}_{k,i}|\theta)$ for each $\theta \in \Theta$ and for all agents.

- We set the weights FC assigns to the $K = 7$ agents uniform, i.e., $v_k = \frac{1}{7}$ for each agent k. Moreover, for both asynchronous scenarios, the participation probabilities p_k are set at 0.5.

Under these parameters, Fig. 8.7 illustrates the FC's crowd count estimates under all three scenarios, along with the actual number of people present (ground truth). Here, the estimates of FC represent the hypothesis θ that maximizes the belief $\boldsymbol{\mu}_i(\theta)$ at each time instant i. Furthermore, in Fig. 8.8, the normalized causal impact scores of each camera on the joint decision are presented for all three scenarios using uniform beliefs intervention. Notably, the score distribution in the symmetric case exhibits the highest level of skewness, which mirrors the conclusion with the synthetic data in Fig. 8.5.

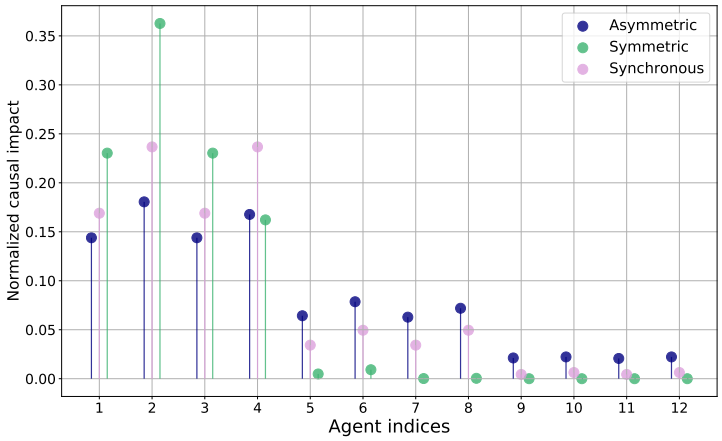


Figure 8.5: Causal impacts of each agent over three scenarios. The scores are normalized such that for each scenario, agents' scores sum up to one. It is clear that the distribution of the scores for the symmetric case has a higher skewness, as suggested by the theoretical results.

(a) (b) (c)

Figure 8.6: Images are part of the WILDTRACK dataset [157], which is acquired in front of the main building of ETH Zürich, Switzerland. In total, the dataset contains 7 simultaneous image sequences (with a rate of 60 frames per second), where each image has a resolution of 1920×1080 pixels. The sample figures (a), (b), and (c) from the dataset capture the same area simultaneously from different perspectives by agents 1, 3, and 7, respectively.

8.6 Concluding Remarks

In this chapter, we examined a collaborative prediction framework for identifying and quantifying causal impact of an agent where agents exchange their local inferences about a common target variable with a fusion center. We incorporated two asynchronicity scenarios that differ in terms of whether the fusion center updates the agents that do not provide information. Utilizing a causal theoretical framework, we derived expressions that describe how each agent's impact on the collective decision varies based on factors such as the distribution of data (via KL divergences representing the informativeness of data) received by the agents and their participation frequencies.

The results reveal that an agent has a stronger impact on the joint decision in the symmetric (reciprocal) communication protocol compared to the asymmetric communication protocol if the misinformation strength surpasses some threshold. This implies that asymmetric communication protocols are more robust in the face of

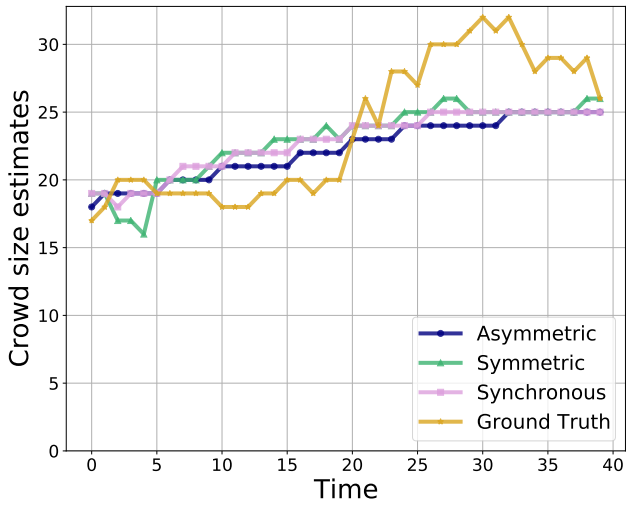


Figure 8.7: Crowd count estimates of FC under all three scenarios, along with the true number of people (ground truth). The estimates of FC correspond to the hypothesis that maximizes the belief at each time instant.

adversarial attacks. Nevertheless, symmetric communication offers greater resilience to moderate deviations from the usual, such as in the case of malfunctioning agents without harmful intentions.

Future directions include extending the causal impact analysis on this federated framework to decentralized peer-to-peer networks, and also examining different decision aggregation strategies at the server side such as median-based robust fusion [161].

8.A Proof of Theorem 8.1

In this section, we prove that under both scenarios considered for asynchronous behavior, and without any intervention, the expected beliefs at the agents place the value 1 on the true hypothesis as $i \to \infty$. Note that the proof for the synchronous communication case is a special case of Theorem 2.1. It can also be recovered from our novel derivations here by setting $p_k \to 1$ for each agent k.

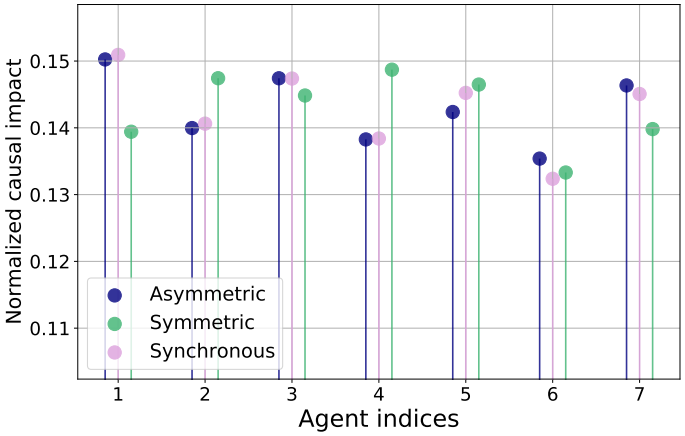


Figure 8.8: Normalized causal impact scores of each camera on the joint decision for all three scenarios. As in the Fig. 8.5 for synthetic data, distribution of the scores for the symmetric case has the highest skewness.

8.A.1 Asymmetric Communication

Recall that $\boldsymbol{q}_{k,i}$ is the Bernoulli random variable that is equal to 1 if agent k is connected to the FC at time i and is sending information, i.e.,

$$\boldsymbol{q}_{k,i} = \begin{cases} 1, & \text{if agent } k \text{ is sending information to FC at time } i \\ 0, & \text{else.} \end{cases} \tag{8.22}$$

Similar to the previous chapter, we define the scalar random variables

$$\boldsymbol{\lambda}_i(\theta) \triangleq \log \frac{\boldsymbol{\mu}_i(\theta^\circ)}{\boldsymbol{\mu}_i(\theta)}, \quad \boldsymbol{x}_{k,i}(\theta) \triangleq \log \frac{L_k(\boldsymbol{\xi}_{k,i}|\theta^\circ)}{L_k(\boldsymbol{\xi}_{k,i}|\theta)}. \tag{8.23}$$

In this case, to derive the recursion of log-belief ratios, observe that if an agent k:

- is not sending any information at time instant i (i.e., it is idle), then the FC uses its own log-belief ratio from the previous time instant $\boldsymbol{\lambda}_{i-1}(\theta)$ to fill the missing information of agent k during aggregation.
- Otherwise, if agent k is sending its intermediate belief to the server, then its contribution on the FC decision is $\boldsymbol{\lambda}_{i-1}(\theta) + \boldsymbol{x}_{k,i}(\theta)$.

In light of the observation above, the contribution of each agent k at time i is a function of $\boldsymbol{q}_{k,i}$ and can be written as

$$\boldsymbol{\lambda}_{i-1}(\theta), \quad \boldsymbol{q}_{k,i} = 0 \tag{8.24}$$

$$\boldsymbol{\lambda}_{i-1}(\theta) + \boldsymbol{x}_{k,i}, \quad \boldsymbol{q}_{k,i} = 1 \tag{8.25}$$

which is equivalent to

$$\boldsymbol{\lambda}_{i-1}(\theta)(1-\boldsymbol{q}_{k,i})+(\boldsymbol{\lambda}_{i-1}(\theta)+\boldsymbol{x}_{k,i})\boldsymbol{q}_{k,i}=\boldsymbol{\lambda}_{i-1}(\theta)+\boldsymbol{x}_{k,i}\boldsymbol{q}_{k,i}. \tag{8.26}$$

The fusion center update becomes

$$\boldsymbol{\lambda}_i(\theta)=\boldsymbol{\lambda}_{i-1}(\theta)+\mathbf{a}_i^{\mathsf{T}}\boldsymbol{x}_i(\theta), \tag{8.27}$$

where we are introducing the vectors

$$\mathbf{a}_i \triangleq [v_1\boldsymbol{q}_{1,i},\ldots,v_k\boldsymbol{q}_{K,i}]^{\mathsf{T}} \tag{8.28}$$

and

$$\mathbf{x}_i \triangleq [\boldsymbol{x}_{1,i},\ldots,\boldsymbol{x}_{K,i}]^{\mathsf{T}}. \tag{8.29}$$

Taking expectations over randomness of data yields

$$\begin{aligned}\lambda_i(\theta) &\triangleq \mathbb{E}[\boldsymbol{\lambda}_i(\theta)]\\ &\overset{(8.27)}{=} \mathbb{E}[\boldsymbol{\lambda}_{i-1}(\theta)+\mathbf{a}_i^{\mathsf{T}}\mathbf{x}_i(\theta)]\\ &= \lambda_{i-1}(\theta)+\mathrm{a}^{\mathsf{T}}\mathrm{d}(\theta),\end{aligned} \tag{8.30}$$

with the definitions:

$$\begin{aligned}\mathrm{a} &\triangleq [v_1p_1,\ldots,v_Kp_K]^{\mathsf{T}},\\ \mathrm{d}(\theta) &\triangleq \Big[d_1(\theta),\ldots,d_K(\theta)\Big]^{\mathsf{T}},\\ d_k(\theta) &\triangleq D_{\mathrm{KL}}\Big(L_k(\cdot|\theta^\circ)||L_k(\cdot|\theta)\Big).\end{aligned} \tag{8.31}$$

The global identifiability assumption ensures that for each $\theta\neq\theta^\circ$, there exists at least one agent $k^\star$ that satisfies $d_{k^\star}(\theta)>0$. Therefore, for each wrong hypothesis $\theta\neq\theta^\circ$, it holds that $\lambda_i(\theta)\to+\infty$ as long as $v_{k^\star}p_{k^\star}>0$, which in turn means $\mu_i(\theta^\circ)\to 1$ for the collective decision of agents as $i\to\infty$.

8.A.2 Symmetric Communication

In this scenario, agents do not get updated from the FC if they do not send information to the FC. Therefore, at each time instant, they can have different beliefs and priors. This situation necessitates the study of the evolution of a local belief (prior) $\boldsymbol{\mu}_{k,i}$ and the corresponding local log-belief ratio (LBR) $\boldsymbol{\lambda}_{k,i}$. If we define a vector consisting of all LBRs from all agents as

$$\boldsymbol{\Lambda}_i(\theta) \triangleq [\boldsymbol{\lambda}_{1,i}(\theta),\ldots,\boldsymbol{\lambda}_{K,i}(\theta)]^{\mathsf{T}}, \tag{8.32}$$

then, the LBR at the server side evolves according to the following dynamics:

$$\boldsymbol{\lambda}_i(\theta) = \mathbf{a}_i^{\mathsf{T}}(\boldsymbol{\Lambda}_{i-1}(\theta) + \mathbf{x}_i(\theta)) + (\bar{\mathbf{a}}_i^{\mathsf{T}}\mathbb{1}_K)\boldsymbol{\lambda}_{i-1}(\theta), \tag{8.33}$$

where we use the bar notation to denote additive complements of the random variables, i.e.,

$$\bar{\mathbf{a}}_i \triangleq v - \mathbf{a}_i = [v_1\bar{\boldsymbol{q}}_{1,i}, \ldots, v_K\bar{\boldsymbol{q}}_{K,i}]^{\mathsf{T}}, \;\; \bar{\boldsymbol{q}}_{k,i} \triangleq 1 - \boldsymbol{q}_{k,i} \tag{8.34}$$

and also use the definition

$$\mathbf{x}_i(\theta) \triangleq [\boldsymbol{x}_{1,i}(\theta), \boldsymbol{x}_{2,i}(\theta), \ldots, \boldsymbol{x}_{K,i}(\theta)]^{\mathsf{T}}. \tag{8.35}$$

Taking the expectation of both sides in (8.33) yields

$$\lambda_i(\theta) = \mathrm{a}^{\mathsf{T}}(\Lambda_{i-1}(\theta) + \mathrm{d}(\theta)) + (\bar{\mathrm{a}}^{\mathsf{T}}\mathbb{1}_K)\lambda_{i-1}(\theta). \tag{8.36}$$

On the other hand, the LBR evolution for an agent k depends on whether it provides information to the FC update in (8.33), and is given by

$$\boldsymbol{\lambda}_{k,i}(\theta) = \boldsymbol{\lambda}_i(\theta)\boldsymbol{q}_{k,i} + (\boldsymbol{\lambda}_{k,i-1}(\theta) + \boldsymbol{x}_{k,i}(\theta))\bar{\boldsymbol{q}}_{k,i}. \tag{8.37}$$

By taking expectations of both sides we arrive at

$$\lambda_{k,i}(\theta) = \mathbb{E}[\boldsymbol{\lambda}_i(\theta)\boldsymbol{q}_{k,i}] + (\lambda_{k,i-1}(\theta) + d_k(\theta))\bar{p}_k, \tag{8.38}$$

where $\bar{p}_k \triangleq 1 - p_k$. Note that in general,

$$\mathbb{E}[\boldsymbol{\lambda}_i(\theta)\boldsymbol{q}_{k,i}] \neq \mathbb{E}[\boldsymbol{\lambda}_i(\theta)]\mathbb{E}[\boldsymbol{q}_{k,i}] \tag{8.39}$$

due to the information sharing of server is conditioned on agents' information sharing, and that they are not independent. However, it holds that

$$\begin{aligned}\mathbb{E}[\boldsymbol{\lambda}_i(\theta)\boldsymbol{q}_{k,i}] &= \mathbb{E}\left[\boldsymbol{q}_{k,i}\left(\sum_{\ell=1}^{K} v_\ell(\boldsymbol{\lambda}_{\ell,i-1}(\theta) + \boldsymbol{x}_{\ell,i}(\theta))\boldsymbol{q}_{\ell,i} + \sum_{\ell=1}^{K} v_\ell\bar{\boldsymbol{q}}_{\ell,i}\boldsymbol{\lambda}_{i-1}(\theta)\right)\right] \\ &= p_k\left(v_k(\lambda_{k,i-1}(\theta) + d_k(\theta)) + \sum_{\ell\neq k} v_\ell(\lambda_{\ell,i-1}(\theta) + d_\ell(\theta))p_\ell + \sum_{\ell\neq k} v_\ell\bar{p}_\ell\lambda_{i-1}(\theta)\right)\end{aligned} \tag{8.40}$$

Next, we introduce the variables

$$\mathrm{s} \triangleq [s_1, \ldots, s_K]^{\mathsf{T}},\ s_k \triangleq p_k \sum_{\ell \neq k} v_\ell \bar{p}_\ell, \tag{8.41}$$

$$\mathrm{p} \triangleq [p_1, \ldots, p_K]^{\mathsf{T}},\ \bar{\mathrm{p}} \triangleq \mathbb{1}_K - \mathrm{p} \tag{8.42}$$

$$\sigma \triangleq \sum_{k=1}^{K} v_k \bar{p}_k = \bar{a}^{\mathsf{T}} \mathbb{1}_K \tag{8.43}$$

and the $K \times K$ diagonal matrices

$$\mathrm{A} \triangleq \mathrm{diag}(\mathrm{a}), \quad \mathrm{P} \triangleq \mathrm{diag}(\mathrm{p}), \quad \bar{\mathrm{P}} \triangleq \mathrm{diag}(\bar{\mathrm{p}}). \tag{8.44}$$

Accordingly, if we also define the $(K+1)$ dimensional extended LBR vector as

$$\bar{\Lambda}_i(\theta) \triangleq \begin{bmatrix} \Lambda_i(\theta) \\ \lambda_i(\theta) \end{bmatrix}, \tag{8.45}$$

then, by relations (8.36) and (8.38), we arrive at a linear recursion of the following form:

$$\bar{\Lambda}_i(\theta) = \mathrm{R}\bar{\Lambda}_{i-1}(\theta) + \mathrm{Ud}(\theta). \tag{8.46}$$

Here, we introduced the $(K+1) \times (K+1)$ dimensional matrix

$$\mathrm{R} \triangleq \left(\begin{array}{c|c} \bar{\mathrm{P}}\mathrm{A} + \bar{\mathrm{P}} + \mathrm{p}\mathrm{a}^{\mathsf{T}} & \mathrm{s} \\ \hline \mathrm{a}^{\mathsf{T}} & \sigma \end{array}\right) \tag{8.47}$$

and the $(K+1) \times K$ dimensional matrix

$$\mathrm{U} = \left(\begin{array}{c} \bar{\mathrm{P}}\mathrm{A} + \bar{\mathrm{P}} + \mathrm{p}\mathrm{a}^{\mathsf{T}} \\ \hline \mathrm{a}^{\mathsf{T}} \end{array}\right). \tag{8.48}$$

Since R is a stochastic matrix with nonzero entries, it is also a primitive matrix. Therefore, R has a Perron eigenvector v_{R} with all positive entries that corresponds to the largest magnitude eigenvalue. Hence, it holds that

$$\frac{1}{i}\bar{\Lambda}_i(\theta) \to v_{\mathrm{R}}\mathrm{Ud}. \tag{8.49}$$

Under global identifiability assumption that for at least one agent $d_k(\theta) > 0$, expected beliefs at the agents place the value 1 on the true hypothesis.

8.B Proof of Theorem 8.3

Under an intervention on agent m, the contribution of any agent $k \neq m$ at time i is the same as in the pre-intervention case and is given by (8.26). Furthermore, the

intervened agent m's one-step contribution is equal to the intervention strength if agent m is present, and equal to the FC's log-belief ratio (LBR) otherwise. In other words, agent m's contribution under an intervention becomes

$$\widetilde{\boldsymbol{\lambda}}_{i-1}(\theta)(1-\boldsymbol{q}_{m,i})+c\boldsymbol{q}_{m,i}. \tag{8.50}$$

In (8.50), we use $\sim$ on top of pre-intervention variables to denote the post-intervention counterparts of those variables and also note that we are defining

$$c \triangleq \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{8.51}$$

for brevity of notation. Aggregating the contributions (8.26) and (8.50) of each agent according to the geometric averaging rule yields the following update for the LBR of the fusion center:

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) = (1-v_m)\widetilde{\boldsymbol{\lambda}}_{i-1}(\theta) + \mathbf{a}_i^{\mathsf{T}}\widetilde{\mathbf{x}}_i(\theta), \tag{8.52}$$

where we introduced the log-belief ratio counterpart vector under intervention:

$$\widetilde{\mathbf{x}}_i(\theta) \triangleq [\boldsymbol{x}_{1,i}(\theta),\ldots,\boldsymbol{x}_{m-1,i}(\theta),c,\boldsymbol{x}_{m+1,i}(\theta),\ldots,\boldsymbol{x}_{K,i}(\theta)]^{\mathsf{T}} \tag{8.53}$$

where c is taken from (8.51). According to these definitions, for the expected LBR, it holds that

$$\begin{aligned}\widetilde{\lambda}_i(\theta) &\overset{(8.52)}{=} \mathbb{E}\left[(1-v_m)\widetilde{\boldsymbol{\lambda}}_{i-1}(\theta)+\mathbf{a}_i^{\mathsf{T}}\widetilde{\mathbf{x}}_i(\theta)\right]\\ &= (1-v_m)\widetilde{\lambda}_{i-1}(\theta)+\mathrm{a}^{\mathsf{T}}\mathbb{E}[\widetilde{\mathbf{x}}_i(\theta)]\\ &= (1-v_m)\widetilde{\lambda}_{i-1}(\theta)+\mathrm{a}^{\mathsf{T}}\widetilde{\mathrm{d}}(\theta)\end{aligned} \tag{8.54}$$

where the vector of KL divergences under an intervention on agent m is defined as

$$\widetilde{\mathrm{d}}(\theta) \triangleq [d_1(\theta),\ldots,d_{m-1}(\theta),c,d_{m+1}(\theta),\ldots,d_K(\theta)]^{\mathsf{T}}. \tag{8.55}$$

Consequently, in the limit, the LBR of the server under an intervention on agent m is given by

$$\lim_{i\to\infty}\widetilde{\lambda}_i(\theta) = \frac{1}{v_m}\mathrm{a}^{\mathsf{T}}\widetilde{\mathrm{d}}(\theta) = \frac{1}{v_m}\sum_{k\neq m} v_k p_k d_k(\theta) + p_m c. \tag{8.56}$$

8.C Proof of Theorem 8.4

For simplicity of notation and without loss of generality, we intervene on agent $m=1$. This means that whenever it is active ($\boldsymbol{q}_{1,i}=1$), the contribution in terms of the log-belief ratio will be equal to c defined in (8.51). Under such scheme, the linear recursion

from (8.46) transforms, under the intervention on $m = 1$, to

$$\widetilde{\Lambda}_i(\theta) = \widetilde{\mathrm{R}}\widetilde{\Lambda}_{i-1}(\theta) + \widetilde{\mathrm{U}}\widetilde{\mathrm{d}}(\theta) \tag{8.57}$$

where $\widetilde{\mathrm{R}}$ is the submatrix of R without the first column and row, $\widetilde{\mathrm{U}}$ is the submatrix of U without the first row, and $\widetilde{\mathrm{d}}(\theta) = [c, d_2(\theta), \ldots, d_k(\theta)]$. Note that the largest eigenvalue of $\widetilde{\mathrm{R}}$ is smaller than 1, hence we obtain

$$\begin{aligned}\widetilde{\Lambda}_\infty(\theta) &= (\mathrm{I} + \widetilde{\mathrm{R}} + \widetilde{\mathrm{R}}^2 \ldots)\widetilde{\mathrm{U}}\widetilde{\mathrm{d}}(\theta) \\ &= (\mathrm{I} - \widetilde{\mathrm{R}})^{-1}\widetilde{\mathrm{U}}\widetilde{\mathrm{d}}.\end{aligned} \tag{8.58}$$

We therefore need to invert $(\mathrm{I} - \widetilde{\mathrm{R}})$. We write $\mathrm{I} - \widetilde{\mathrm{R}}$ in block matrix form:

$$\begin{aligned}\mathrm{M} &\triangleq \mathrm{I} - \widetilde{\mathrm{R}} \\ &= \left(\begin{array}{c|c} \widetilde{\mathrm{P}} - (\mathrm{I} - \widetilde{\mathrm{P}})\widetilde{\mathrm{A}} - \widetilde{\mathrm{p}}\widetilde{\mathrm{a}}^\mathsf{T} & -\widetilde{\mathrm{s}} \\ \hline -\widetilde{\mathrm{a}}^\mathsf{T} & 1 - \sigma \end{array}\right) \\ &\triangleq \left(\begin{array}{c|c} \mathrm{M}_{11} & \mathrm{M}_{12} \\ \hline \mathrm{M}_{21} & \mathrm{M}_{22} \end{array}\right)\end{aligned} \tag{8.59}$$

where

$$\widetilde{\mathrm{p}} \triangleq [p_2, \ldots, p_K], \quad \widetilde{\mathrm{a}} \triangleq [v_2 p_2, \ldots, v_K p_K], \tag{8.60}$$

and $\mathrm{M}_{11}, \mathrm{M}_{12}, \mathrm{M}_{21}, \mathrm{M}_{22}$ are submatrices of dimensions $(K-1)\times(K-1), (K-1)\times 1, 1\times(K-1), 1\times 1$, respectively, and

$$\widetilde{\mathrm{P}} \triangleq \mathrm{diag}(\widetilde{\mathrm{p}}), \quad \widetilde{\mathrm{A}} \triangleq \mathrm{diag}(\widetilde{\mathrm{a}}). \tag{8.61}$$

Using the Schur complement of M [8, Chapter 1.4]

$$S \triangleq \mathrm{M}_{22} - \mathrm{M}_{21}(\mathrm{M}_{11})^{-1}\mathrm{M}_{12}, \tag{8.62}$$

which is scalar, we can write the last row of M^{-1} as

$$[-S^{-1}\mathrm{M}_{21}(\mathrm{M}_{11})^{-1} \quad | \quad S^{-1}] = S^{-1}[-\mathrm{M}_{21}(\mathrm{M}_{11})^{-1} \quad | \quad 1]. \tag{8.63}$$

Note that we are only interested in finding the last row of M^{-1} as only this row contributes to the FC's LBR in steady state, which is the last entry of $\widetilde{\Lambda}_\infty(\theta)$.

First, we find M_{11}^{-1}. Since M_{11} is the sum of a diagonal matrix and a rank-one matrix, we can calculate M_{11}^{-1} by the matrix inversion formula [8, Chapter 1.4]:

$$\mathrm{M}_{11}^{-1} = (\widetilde{\mathrm{P}} - (\mathrm{I} - \widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1} + \frac{(\widetilde{\mathrm{P}} - (\mathrm{I} - \widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}\widetilde{\mathrm{p}}\widetilde{\mathrm{a}}^\mathsf{T}(\widetilde{\mathrm{P}} - (\mathrm{I} - \widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}}{1 - \widetilde{\mathrm{a}}^\mathsf{T}(\widetilde{\mathrm{P}} - (\mathrm{I} - \widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}\widetilde{\mathrm{p}}} \tag{8.64}$$

Consequently,

$$-\mathrm{M}_{21}(\mathrm{M}_{11})^{-1} = \widetilde{\mathrm{a}}^{\mathsf{T}}(\widetilde{\mathrm{P}}-(\mathrm{I}-\widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1} + \frac{\widetilde{\mathrm{a}}^{\mathsf{T}}(\widetilde{\mathrm{P}}-(\mathrm{I}-\widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}\widetilde{\mathrm{p}}\widetilde{\mathrm{a}}^{\mathsf{T}}(\widetilde{\mathrm{P}}-(\mathrm{I}-\widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}}{1-\widetilde{\mathrm{a}}^{\mathsf{T}}(\widetilde{\mathrm{P}}-(\mathrm{I}-\widetilde{\mathrm{P}})\widetilde{\mathrm{A}})^{-1}\widetilde{\mathrm{p}}} \tag{8.65}$$

Observe that for $k>1$, the kth element of $-\mathrm{M}_{21}(\mathrm{M}_{11})^{-1}$ is given by

$$\frac{v_k}{1-\bar{p}_k v_k}\frac{1}{1-\sum\limits_{\ell\neq 1}\dfrac{v_\ell p_\ell}{1-\bar{p}_\ell v_\ell}} \tag{8.66}$$

If we define $s_k \triangleq p_k\sum_{\ell\neq k} v_\ell\bar{p}_\ell$, we get

$$\begin{aligned} S &= \mathrm{M}_{22}-\mathrm{M}_{21}(\mathrm{M}_{11})^{-1}\mathrm{M}_{12} \\ &= \sum_{k=1}^{K} v_k p_k - \sum_{k\neq 1}\frac{v_k s_k}{1-\bar{p}_k v_k}\frac{1}{1-\sum_{\ell\neq 1}\dfrac{v_\ell p_\ell}{1-\bar{p}_\ell v_\ell}} \end{aligned} \tag{8.67}$$

Now, we calculate $\widetilde{\mathrm{U}}\widetilde{\mathrm{d}}(\theta)$. Observe that

$$\widetilde{\mathrm{U}}\widetilde{\mathrm{d}}(\theta) = \left[u_2,\ldots,u_K \quad | \quad \Big(v_1p_1c+\sum_{k\neq 1}v_kp_kd_k(\theta)\Big)\right]^{\mathsf{T}} \tag{8.68}$$

where

$$u_k = p_k\Big(v_1p_1c+\sum_{\ell\neq 1}v_\ell p_\ell d_\ell(\theta)\Big)+\Big(\bar{p}_k v_k p_k+\bar{p}_k\Big)d_k(\theta). \tag{8.69}$$

As a result, it holds that

$$\begin{aligned} \widetilde{\lambda}_\infty(\theta) &= S^{-1}\Big[-\mathrm{M}_{21}(\mathrm{M}_{11})^{-1} \quad | \quad 1\Big]\Big(\widetilde{\mathrm{U}}\widetilde{\mathrm{d}}(\theta)\Big) \\ &= S^{-1}\Bigg(\sum_{k\neq 1}\frac{v_k d_k(\theta)(\bar{p}_k v_k p_k+\bar{p}_k)}{1-\bar{p}_k v_k}\frac{1}{1-\sum\limits_{\ell\neq 1}\dfrac{v_\ell p_\ell}{1-\bar{p}_\ell v_\ell}} \\ &\qquad + \sum_{k\neq 1}\frac{v_k p_k}{1-\bar{p}_k v_k}\frac{v_1p_1c+\sum\limits_{\ell\neq 1}v_\ell p_\ell d_\ell(\theta)}{1-\sum\limits_{\ell\neq 1}\dfrac{v_\ell p_\ell}{1-\bar{p}_\ell v_\ell}} + v_1p_1c+\sum_{\ell\neq 1}v_\ell p_\ell d_\ell(\theta)\Bigg) \\ &= \frac{\sum\limits_{k\neq 1}\dfrac{v_k d_k(\theta)(\bar{p}_k v_k p_k+\bar{p}_k)}{1-\bar{p}_k v_k}+v_1p_1c+\sum\limits_{\ell\neq 1}v_\ell p_\ell d_\ell(\theta)}{S\Bigg(1-\sum\limits_{\ell\neq 1}\dfrac{v_\ell p_\ell}{1-\bar{p}_\ell v_\ell}\Bigg)} \end{aligned}$$

$$= \frac{\displaystyle\sum_{k\neq 1} \frac{v_k d_k(\theta)(\bar{p}_k v_k p_k + \bar{p}_k)}{1-\bar{p}_k v_k} + v_1 p_1 c + \sum_{\ell \neq 1} v_\ell p_\ell d_\ell(\theta)}{\left(\displaystyle\sum_{k=1}^{K} v_k p_k\right)\left(1 - \displaystyle\sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k}\right) - \displaystyle\sum_{k\neq 1} \frac{v_k s_k}{1-\bar{p}_k v_k}}. \tag{8.70}$$

Observe that the term in the numerator is equivalent to

$$\sum_{k\neq 1} v_k d_k(\theta)\left(\frac{\bar{p}_k v_k p_k + \bar{p}_k}{1-\bar{p}_k v_k} + p_k\right) + v_1 p_1 c = \sum_{k\neq 1} \frac{v_k d_k(\theta)}{1-\bar{p}_k v_k} + v_1 p_1 c. \tag{8.71}$$

Furthermore, if we incorporate the following relation for s_k

$$s_k = p_k \sum_{\ell \neq k} v_\ell \bar{p}_\ell = p_k \left(\sum_{\ell=1}^{K} v_\ell \bar{p}_\ell - v_k \bar{p}_k\right) \tag{8.72}$$

to the term in the denominator of (8.70), we obtain

$$\begin{aligned}
&\left(\sum_{k=1}^{K} v_k p_k\right)\left(1 - \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k}\right) - \sum_{k\neq 1} \frac{v_k s_k}{1-\bar{p}_k v_k} \\
&= \left(\sum_{k=1}^{K} v_k p_k\right)\left(1 - \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k}\right) - \left(\sum_{k=1}^{K} v_k \bar{p}_k\right) \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k} + \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k} v_k \bar{p}_k \\
&= \left(1 - \sum_{k=1}^{K} v_k \bar{p}_k\right)\left(1 - \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k}\right) - \left(\sum_{k=1}^{K} v_k \bar{p}_k\right) \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k} + \sum_{k\neq 1} \frac{v_k p_k}{1-\bar{p}_k v_k} v_k \bar{p}_k \\
&= \sum_{k=1}^{K} v_k p_k - \sum_{k\neq 1} v_k p_k = v_1 p_1.
\end{aligned} \tag{8.73}$$

Thus, the expected LBR in steady state becomes

$$\widetilde{\lambda}_\infty(\theta) = \frac{1}{v_1 p_1} \sum_{k\neq 1} \frac{v_k d_k(\theta)}{1 - v_k(1-p_k)} + c = \frac{1}{v_1 p_1} \sum_{k\neq 1} \frac{v_k d_k(\theta)}{1 - v_k(1-p_k)} + \log \frac{\mu_1(\theta^\circ)}{\mu_1(\theta)}. \tag{8.74}$$

Since the choice of $m = 1$ was without loss of generality, replacing the subscripts 1 by m in (8.74), we arrive at the result.

9 Conclusion

In this thesis, we contributed to the field of collective decision-making processes within complex and uncertain environments by proposing variations of sequential Bayesian reasoning in the context of networked interactions. Here, we summarize the key findings and contributions across our investigations.

9.1 Summary of Results

In Chapter 3, we established a theoretical framework for comparing arithmetic and geometric averaging as methods for information fusion. Our results define quantitative bounds that link the effectiveness of these fusion rules to the network's connectivity and the diversity of information it contains. The findings offer practical guidelines for selecting fusion strategies that optimize learning rates in various network configurations.

In Chapter 4, we investigated the network evolution under randomness in communication patterns. We identified conditions under which networked agents can either successfully learn the truth or, alternatively, mislearn. Our results provide guidelines on how the distribution and frequency of communication can impact the networked inference processes.

In Chapter 5, we utilized HMMs to model the changes in the state of nature. We expanded existing algorithms of static belief formation to track dynamic states. We provided a comparison of theoretical performance between a centralized strategy, which is maximum-a-posteriori (MAP) optimal, and our proposed decentralized approach.

In Chapter 6, we explored the implications of our partially observable multi-agent framework to Markov decision processes. We proposed a policy evaluation algorithm for agents over a graph under stochastic observations. The contribution in this chapter is relevant to the practical deployment of multi-agent reinforcement learning when

agents operate without full knowledge of their environment.

In Chapter 7, trying to understand the causal relationships within social interactions over time, we developed a theoretical framework to capture the causal impact of agents. In addition to deriving expressions for causal impact between each agent pair, we also proposed algorithms for ranking agents based on their influence and for discovering causal relationships from observational data.

In Chapter 8, we extended our discussion on causality to consider the effect of asynchronous communication. Our findings revealed how different participation patterns and fusion center policies can change the strength of causal influences.

9.2 Future Directions

Finally, we discuss possible extensions to our work.

Beyond categorical variables. A direction for future research is to explore the applicability of our results to distributed estimation [68, 69, 162], which focuses on the inference of continuous variables. The algorithms we propose in this dissertation could be adapted for continuous state estimation with the caveat that the summations are replaced with possibly intractable integrals. In some scenarios, evaluation of the integrals could still be numerically feasible under conditions, such as likelihood functions following exponential family of distributions. Furthermore, in the context of our discussion in Chapter 3, information fusion for continuous variables generally relies on fusing point estimates. However, it is known that these can also be understood as fusions of probability density functions [35]. This leads us to question whether the results in Chapter 3 for finite-sized hypothesis sets could extend to the distributed parameter estimation problems.

Finite sample guarantees. The current dissertation primarily provided asymptotic theoretical guarantees. A possible extension would be to explore theoretical guarantees in finite sample or time regimes, as in [26, 163]. For example, by extending our results in Chapter 3 in a subsequent work [59], we established the asymptotically normal behaviors of arithmetic and geometric rule based federated inference algorithms. Building on this result, employing the Berry-Esseen theorem [60] and similar techniques could provide deeper insights into the limited sample behaviors of these algorithms.

Network interference. At the intersection of causal inference and networks, there exists a body of work focusing on network interference [164–166]. These studies examine causal inference when interventions on certain agents can affect others within the network, and aim to design randomized experiments that take this fact into account. However, these works typically assume that only immediate neighbors of an agent can impact an individual's response. Although our settings in Chapters 7 and 8 differ,

our analytical contributions may still prove useful for this area of research, since we discover how influence and interference propagate throughout the network in Chapter 7. For example, our findings allow us to quantify how an agent's influence diminishes with increasing distance from other agents in the graph.

Causal discovery in networked time-series. In Chapters 7 and 8, we have studied causal influence attributions in temporal interactions, with an emphasis on the effect of the underlying communication network. A promising future direction is to incorporate other causality methodologies in time-series analysis [167, 168] to our framework. This could involve exploring alternative metrics [169, 170] that capture higher-order effects beyond expected value differences for the causal impact definition. Additionally, we can combine our approach with recent advancements in causal discovery from time series data [171–175]. Moreover, going beyond identifying influence, we could aim to address counterfactual questions.

Heterogeneous environments. In this dissertation, we have primarily focused on scenarios where the network of agents shares a common true state of nature. A natural question that arises is what happens when agents receive observations from different hypotheses. There are recent studies in the literature that explore this problem from various angles. Some works consider agents to be cooperative regardless of the observations they receive and investigate how network topology – such as community structures – affects overall network behavior [176–178]. Other research adopts a game-theoretic approach and models the agents to be strategic and self-interested who aim to identify others following the same hypothesis [179, 180]. A direction for future work is to combine our models and analysis with these studies.

Information sharing variations. In Chapters 3 and 4, we have focused on specific methods for information fusion and sharing. There are other procedures for information exchange worth exploring. For instance, in Chapter 4, we addressed the issue of filling missing belief components using proportional assignments in (4.8). An alternative approach, as suggested by [181], is to directly incorporate the received information into the beliefs without normalization. Moreover, our model in Chapter 4 assumes that the trending hypothesis $\boldsymbol{\tau}_i$ is the same for all agents and is independent and identically distributed over time. Future research could examine scenarios where each agent k possibly shares a different hypothesis $\boldsymbol{\tau}_{k,i}$ or where the trending hypothesis $\boldsymbol{\tau}_i$ evolves according to a Markov chain. Additionally, we have considered the case that agents communicate with all of their neighbors at each time instant. In some real-world applications, individuals can exchange beliefs with only a subset of their contacts, as discussed in [182, 183]. Future work could involve integrating our models and analyses with these studies to more accurately reflect real-world information exchange over networks.

Combining data-driven and model-based designs. Throughout this dissertation,

we provided both theoretical insights and practical guidance for decision-making networks. In practice, the models we consider might not fully encapsulate real-world conditions. Therefore, a possible future direction could be to blend these models with data-driven approaches. This way, not only we can use our results on these models for explainability, but also can enhance the capacity of our models to mirror real-world conditions. By doing so, we can achieve a balance between interpretability and performance that results in a more trustworthy and transparent approach, as opposed to fully data-driven approaches with black-box models.

Robustness to model uncertainties. Another key issue is assessing the sensitivity of the algorithms and analyses in this dissertation to approximation errors in the models. In practice, model parameters may be estimated or subject to uncertainties. For example, in Chapter 5, investigating the impact of inaccuracies in the state transition model could reveal how deviations from the assumed dynamics affect the agents' ability to track the true state.

Causality for systems. In fields like signal processing, control, and communications, systems often have a modular structure that allows for a clear understanding of the data-generating process. This contrasts with fields such as healthcare, where the challenge lies in learning causal representations from data [184]. However, this inherent understanding of causal structures is often underutilized. By more explicitly incorporating causal reasoning into these fields, we can improve the design and analysis of general systems, as demonstrated in Chapters 7 and 8 to understand the behavior of networked systems.

Bibliography

[1] T. Bayes and R. Price, "An essay towards solving a problem in the doctrine of chances," *Philosophical Transactions of the Royal Society of London*, pp. 370–418, Dec. 1763.

[2] M. J. Wainwright and M. I. Jordan, "Graphical models, exponential families, and variational inference," *Foundations and Trends in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.

[3] D. Shah, "Gossip algorithms," *Foundations and Trends in Networking*, vol. 3, no. 1, pp. 1–125, 2009.

[4] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.

[5] K. P. Murphy, *Probabilistic Machine Learning: Advanced Topics*. MIT Press, 2023.

[6] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

[7] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Foundations and Trends in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, July 2014.

[8] ——, *Inference and Learning from Data*. Cambridge University Press, 2022, 3 vols.

[9] ——, "Adaptive networks," *Proc. IEEE*, vol. 102, no. 4, pp. 460–497, 2014.

[10] S. I. Resnick, *Adventures in Stochastic Processes*. Birkhäuser, 2002.

[11] S. U. Pillai, T. Suel, and S. Cha, "The Perron-Frobenius theorem: some of its applications," *IEEE Signal Processing Magazine*, vol. 22, no. 2, pp. 62–75, 2005.

[12] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer, 2012.

[13] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. Cambridge University Press, 2011.

[14] V. Bordignon, V. Matta, and A. H. Sayed, "Adaptive social learning," *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 6053–6081, 2021.

[15] C. P. Chamley, *Rational Herds: Economic Models of Social Learning*. Cambridge University Press, 2003.

[16] P. M. Djurić and Y. Wang, "Distributed Bayesian learning in multiagent systems: Improving our understanding of its capabilities and limitations," *IEEE Signal Processing Magazine*, vol. 29, no. 2, pp. 65–76, 2012.

[17] C. Chamley, A. Scaglione, and L. Li, "Models for the diffusion of beliefs in social networks: An overview," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 16–29, 2013.

[18] V. Krishnamurthy and H. V. Poor, "Social learning and Bayesian games in multiagent signal processing: how do local and global decision makers interact?" *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 43–57, 2013.

[19] E. Mossel and O. Tamuz, "Opinion exchange dynamics," *Probability Surveys*, vol. 14, pp. 155–204, 2017.

[20] D. Acemoglu, M. A. Dahleh, I. Lobel, and A. Ozdaglar, "Bayesian learning in social networks," *The Review of Economic Studies*, vol. 78, no. 4, pp. 1201–1236, 2011.

[21] A. Jadbabaie, P. Molavi, A. Sandroni, and A. Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, vol. 76, no. 1, pp. 210–225, 2012.

[22] X. Zhao and A. H. Sayed, "Learning over social networks via diffusion adaptation," in *Proc. Asilomar Conference*, 2012, pp. 709–713.

[23] A. Nedić, A. Olshevsky, and C. A. Uribe, "Fast convergence rates for distributed non-Bayesian learning," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 5538–5553, 2017.

[24] A. Lalitha, T. Javidi, and A. D. Sarwate, "Social learning and distributed hypothesis testing," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6161–6179, 2018.

[25] R. Parasnis, M. Franceschetti, and B. Touri, "Non-Bayesian social learning on random digraphs with aperiodically varying network connectivity," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 3, pp. 1202–1214, 2022.

[26] P. Hu, V. Bordignon, M. Kayaalp, and A. H. Sayed, "Performance of social machine learning under limited data," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1–5.

[27] J. Hazla, A. Jadbabaie, E. Mossel, and M. A. Rahimian, "Bayesian decision making in groups is hard," *Operations Research*, vol. 69, no. 2, pp. 632–654, 2021.

[28] H. A. Simon, "Bounded rationality," in *Utility and Probability*, 1990, pp. 15–18.

[29] J. Conlisk, "Why bounded rationality?" *Journal of Economic Literature*, vol. 34, no. 2, pp. 669–700, 1996.

[30] K. Friston, "A theory of cortical responses," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 360, no. 1456, pp. 815–836, 2005.

[31] M. Oaksford and N. Chater, *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*. Oxford University Press, 2007.

[32] A. Zellner, "Optimal information processing and Bayes's theorem," *The American Statistician*, vol. 42, no. 4, pp. 278–280, 1988.

[33] C. Genest, "A characterization theorem for externally Bayesian groups," *The Annals of Statistics*, vol. 12, no. 3, pp. 1100–1105, 1984.

[34] M. Stone, "The opinion pool," *The Annals of Mathematical Statistics*, vol. 32, no. 4, pp. 1339–1342, 1961.

[35] G. Koliander, Y. El-Laham, P. M. Djurić, and F. Hlawatsch, "Fusion of probability density functions," *Proc. IEEE*, vol. 110, no. 4, pp. 404–453, 2022.

[36] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman filtering based on covariance intersection," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 891–902, 2012.

[37] T. Li, H. Fan, J. García, and J. M. Corchado, "Second-order statistics analysis and comparison between arithmetic and geometric average fusion: Application to multi-sensor target tracking," *Information Fusion*, vol. 51, pp. 233–243, 2019.

[38] M. H. DeGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, no. 345, pp. 118–121, 1974.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.

[40] D. Williams, *Probability with Martingales*. Cambridge University Press, 1991.

[41] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *The Annals of Mathematical Statistics*, vol. 23, no. 4, pp. 493–507, 1952.

[42] M. Kayaalp, Y. Inan, E. Telatar, and A. H. Sayed, "On the arithmetic and geometric fusion of beliefs for distributed inference," *IEEE Transactions on Automatic Control*, vol. 69, no. 4, pp. 2265–2280, 2024.

[43] T. Li, Y. Xin, Y. Song, E. Song, and H. Fan, "Some statistic and information-theoretic results on arithmetic average density fusion," *arXiv:2110.01440*, 2021.

[44] L. Gao, G. Battistelli, and L. Chisci, "Multiobject fusion with minimum information loss," *IEEE Signal Processing Letters*, vol. 27, pp. 201–205, 2020.

[45] K. C. Chang, C.-Y. Chong, and S. Mori, "Analytical and computational evaluation of scalable distributed fusion algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 46, no. 4, pp. 2022–2034, 2010.

[46] K. Dedecius and P. M. Djurić, "Bayesian approach to collaborative inference in networks of agents," in *Cooperative and Graph Signal Processing*, 2018, vol. 2, pp. 131–145.

[47] A. Jadbabaie, P. Molavi, and A. Tahbaz-Salehi, "Information heterogeneity and the speed of learning in social networks," *Columbia Business School Research Paper*, no. 13-28, 2013.

[48] V. Krishnamurthy, *Partially Observed Markov Decision Processes: From Filtering to Controlled Sensing*. Cambridge University Press, 2016.

[49] R. Bellman, "Limit theorems for non-commutative operations. I." *Duke Mathematical Journal*, vol. 21, no. 3, pp. 491 – 500, 1954.

[50] H. Furstenberg and H. Kesten, "Products of random matrices," *The Annals of Mathematical Statistics*, vol. 31, no. 2, pp. 457–469, 1960.

[51] J. F. C. Kingman, "Subadditive ergodic theory," *The Annals of Probability*, vol. 1, no. 6, pp. 883 – 899, 1973.

[52] M. Fekete, "Über die verteilung der wurzeln bei gewissen algebraischen gleichungen mit ganzzahligen koeffizienten," *Mathematische Zeitschrift*, vol. 17, no. 1, pp. 228–249, 1923.

[53] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.

[54] M. Hairer, "Convergence of Markov processes," *Lecture notes*, 2021. [Online]. Available: http://www.hairer.org/notes/Convergence.pdf

[55] Y. Polyanskiy and Y. Wu, "Strong data-processing inequalities for channels and Bayesian networks," in *Convexity and Concentration*. Springer, 2017, pp. 211–249.

[56] R. Ahlswede and P. Gacs, "Spreading of sets in product spaces and hypercontraction of the Markov operator," *The Annals of Probability*, vol. 4, no. 6, pp. 925 – 939, 1976.

[57] S.-Y. Tu and A. H. Sayed, "Diffusion strategies outperform consensus strategies for distributed estimation over adaptive networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 12, pp. 6217–6234, 2012.

[58] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, Jun 1953.

[59] M. Kayaalp, Y. İnan, V. Koivunen, E. Telatar, and A. H. Sayed, "On the fusion strategies for federated decision making," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, 2023, pp. 270–274.

[60] A. C. Berry, "The accuracy of the Gaussian approximation to the sum of independent variates," *Transactions of the American Mathematical Society*, vol. 49, no. 1, pp. 122–136, 1941.

[61] M. Kayaalp, V. Bordignon, and A. H. Sayed, "Social opinion formation and decision making under communication trends," *IEEE Transactions on Signal Processing*, vol. 72, pp. 506–520, 2024.

[62] V. Bordignon, V. Matta, and A. H. Sayed, "Partial information sharing over social learning networks," *IEEE Transactions on Information Theory*, vol. 69, no. 3, pp. 2033–2058, 2023.

[63] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and disagreement in social networks," *Mathematics of Operations Research*, vol. 38, no. 1, pp. 1–27, 2013.

[64] E. Yildiz, A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione, "Binary opinion dynamics with stubborn agents," *ACM Transactions on Economics and Computation*, vol. 1, no. 4, pp. 1–30, 2013.

[65] S. D. Lena, "Non-Bayesian social learning and the spread of misinformation in networks," Department of Economics, University of Venice Ca' Foscari, Working Papers 2019:09, 2019, https://ssrn.com/abstract=3355245.

[66] D. Vial and V. Subramanian, "Local non-Bayesian social learning with stubborn agents," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 3, pp. 1178–1188, 2022.

[67] C. A. Uribe, A. Olshevsky, and A. Nedić, "Nonasymptotic concentration rates in cooperative learning—part II: Inference on compact hypothesis sets," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 3, pp. 1141–1153, 2022.

[68] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3575–3605, 2012.

[69] K. Dedecius and P. M. Djurić, "Sequential estimation and diffusion of information over networks: A Bayesian approach with exponential family of distributions," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1795–1809, 2017.

[70] E. Cesàro, "Sur la convergence des séries," *Nouvelles Annales de Mathématiques: Journal des Candidats aux Écoles Polytechnique et Normale*, vol. 7, pp. 49–59, 1888.

[71] M. Kayaalp, V. Bordignon, S. Vlaski, and A. H. Sayed, "Hidden Markov modeling over graphs," in *Proc. IEEE Data Science and Learning Workshop (DSLW)*, Singapore, Singapore, 2022, pp. 1–6.

[72] M. Kayaalp, V. Bordignon, S. Vlaski, V. Matta, and A. H. Sayed, "Distributed Bayesian learning of dynamic states," *arXiv:2212.02565*, Dec. 2022.

[73] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260–269, 1967.

[74] T. Heskes, "Selecting weighting factors in logarithmic opinion pools," in *Advances in Neural Information Processing Systems*, vol. 10, Denver, CO, USA, 1997, pp. 1–7.

[75] S. Shahrampour, A. Rakhlin, and A. Jadbabaie, "Distributed detection: Finite-time analysis and impact of network topology," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3256–3268, 2016.

[76] R. L. Dobrushin, "Central limit theorem for nonstationary Markov chains. I," *Theory of Probability & Its Applications*, vol. 1, no. 1, pp. 65–80, 1956.

[77] P. Chigansky, R. Liptser, and R. Van Handel, "Intrinsic methods in filter stability," in *Handbook of Nonlinear Filtering*, D. Crisan and B. Rozovskii, Eds. Oxford University Press, 2009.

[78] L. Shue, B. Anderson, and S. Dey, "Exponential stability of filters and smoothers for hidden Markov models," *IEEE Transactions on Signal Processing*, vol. 46, no. 8, pp. 2180–2194, 1998.

[79] C. McDonald and S. Yüksel, "Exponential filter stability via Dobrushin's coefficient," *Electronic Communications in Probability*, vol. 25, pp. 1 – 13, 2020.

[80] L. Shue, S. Dey, B. Anderson, and F. De Bruyne, "On state-estimation of a two-state hidden Markov model with quantization," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 202–208, 2001.

[81] T. Tao, *An Introduction to Measure Theory*. American Mathematical Society, 2011, vol. 126.

[82] M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods*. John Wiley & Sons, 2009.

[83] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics*. Springer, 1998, vol. 97.

[84] H. B. Mann and A. Wald, "On stochastic limit and order relationships," *The Annals of Mathematical Statistics*, vol. 14, no. 3, pp. 217–226, 1943.

[85] R. Axelrod, "The dissemination of culture: A model with local convergence and global polarization," *The Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997.

[86] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis, "On Krause's multi-agent consensus model with state-dependent connectivity," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2586–2597, 2009.

[87] S. Bandyopadhyay and S. Chung, "Distributed Bayesian filtering using logarithmic opinion pool for dynamic sensor networks," *Automatica*, vol. 97, pp. 7–17, 2018.

[88] M. Kayaalp, F. Ghadieh, and A. H. Sayed, "Policy evaluation in decentralized POMDPs with belief sharing," *IEEE Open Journal of Control Systems*, vol. 2, pp. 125–145, 2023.

[89] L. Busoniu, R. Babuska, and B. De Schutter, "A comprehensive survey of multiagent reinforcement learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 2, pp. 156–172, 2008.

[90] K. Zhang, Z. Yang, and T. Başar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," *Handbook of Reinforcement Learning and Control*, pp. 321–384, 2021.

[91] F. A. Oliehoek and C. Amato, *A Concise Introduction to Decentralized POMDPs*. Springer, 2016.

[92] S. Omidshafiei, J. Pazis, C. Amato, J. P. How, and J. Vian, "Deep decentralized multi-task multi-agent reinforcement learning under partial observability," in *Proc. International Conference on Machine Learning*, 2017, pp. 2681–2690.

[93] J. K. Gupta, M. Egorov, and M. Kochenderfer, "Cooperative multi-agent control using deep reinforcement learning," in *Proc. AAMAS*, 2017, pp. 66–83.

[94] J. Foerster, G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson, "Counterfactual multi-agent policy gradients," in *Proc. AAAI Conference on Artificial Intelligence*, 2018, pp. 2974–2982.

[95] P. Moreno, J. Humplik, G. Papamakarios, B. A. Pires, L. Buesing, N. Heess, and T. Weber, "Neural belief states for partially observed domains," in *NeurIPS Workshop on Reinforcement Learning under Partial Observability*, 2018, pp. 1–5.

[96] K. Gregor, G. Papamakarios, F. Besse, L. Buesing, and T. Weber, "Temporal difference variational auto-encoder," in *Proc. International Conference on Learning Representations*, 2019, pp. 1–17.

[97] S. Kar, J. M. F. Moura, and H. V. Poor, "$\mathcal{QD}$-learning: A collaborative distributed strategy for multi-agent reinforcement learning through $\mathrm{consensus}+\mathrm{innovations}$," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, pp. 1848–1862, 2013.

[98] K. Zhang, Z. Yang, H. Liu, T. Zhang, and T. Basar, "Fully decentralized multi-agent reinforcement learning with networked agents," in *Proc. International Conference on Machine Learning*, 2018, pp. 5872–5881.

[99] L. Cassano, K. Yuan, and A. H. Sayed, "Multiagent fully decentralized value function learning with linear convergence rates," *IEEE Transactions on Automatic Control*, vol. 66, no. 4, pp. 1497–1512, 2021.

[100] S. V. Macua, I. Davies, A. Tukiainen, and E. M. De Cote, "Fully distributed actor-critic architecture for multitask deep reinforcement learning," *The Knowledge Engineering Review*, vol. 36, pp. 1–30, 2021.

[101] X. Sha, J. Zhang, K. You, K. Zhang, and T. Başar, "Fully asynchronous policy evaluation in distributed reinforcement learning over networks," *Automatica*, vol. 136, p. 110092, 2022.

[102] Y. Lin, G. Qu, L. Huang, and A. Wierman, "Multi-agent reinforcement learning in stochastic networked systems," in *Advances in Neural Information Processing Systems*, vol. 34, 2021, pp. 7825–7837.

[103] G. Wang, S. Lu, G. Giannakis, G. Tesauro, and J. Sun, "Decentralized TD tracking with linear function approximation and its finite-time analysis," in *Advances in Neural Information Processing Systems*, 2020, pp. 13 762–13 772.

[104] J. Sun, G. Wang, G. B. Giannakis, Q. Yang, and Z. Yang, "Finite-time analysis of decentralized temporal-difference learning with linear function approximation," in *Proc. International Conference on Artificial Intelligence and Statistics*, 2020, pp. 4485–4495.

[105] Q. Lin and Q. Ling, "Decentralized TD(0) with gradient tracking," *IEEE Signal Processing Letters*, vol. 28, pp. 723–727, 2021.

[106] A. Mahajan and M. Mannan, "Decentralized stochastic control," *Annals of Operations Research*, vol. 241, no. 1-2, pp. 109–126, 2016.

[107] A. A. Malikopoulos, "On team decision problems with nonclassical information structures," *IEEE Transactions on Automatic Control*, vol. 68, no. 7, pp. 3915–3930, 2023.

[108] S. Yuksel, "Stochastic nestedness and the belief sharing information pattern," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2773–2786, 2009.

[109] A. Nayyar, A. Mahajan, and D. Teneketzis, "Decentralized stochastic control with partial history sharing: A common information approach," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1644–1658, 2013.

[110] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 2018.

[111] R. S. Sutton, "Learning to predict by the methods of temporal differences," *Machine Learning*, vol. 3, no. 1, pp. 9–44, 1988.

[112] J. Z. Kolter and A. Y. Ng, "Regularization and feature selection in least-squares temporal difference learning," in *Proc. International Conference on Machine Learning*, 2009, pp. 521–528.

[113] M. W. Hoffman, A. Lazaric, M. Ghavamzadeh, and R. Munos, "Regularized least squares temporal difference learning with nested ℓ_2 and ℓ_1 penalization," in *Proc. European Workshop on Reinforcement Learning*. Springer, 2011, pp. 102–114.

[114] J. Farebrother, M. C. Machado, and M. Bowling, "Generalization and regularization in DQN," *arXiv:1810.00123*, 2018.

[115] K. Cobbe, O. Klimov, C. Hesse, T. Kim, and J. Schulman, "Quantifying generalization in reinforcement learning," in *Proc. International Conference on Machine Learning*, vol. 97, 09–15 Jun 2019, pp. 1282–1289.

[116] E. J. Sondik, "The optimal control of partially observable Markov processes over the infinite horizon: Discounted costs," *Operations Research*, vol. 26, no. 2, pp. 282–304, 1978.

[117] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. A. Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 4909–4926, 2022.

[118] E. Bıyık, D. P. Losey, M. Palan, N. C. Landolfi, G. Shevchuk, and D. Sadigh, "Learning reward functions from diverse sources of human feedback: Optimally integrating demonstrations and preferences," *The International Journal of Robotics Research*, vol. 41, no. 1, pp. 45–67, 2022.

[119] P. Moreno, E. Hughes, K. R. McKee, B. A. Pires, and T. Weber, "Neural recursive belief states in multi-agent reinforcement learning," *arXiv:2102.02274*, 2021.

[120] D. Muglich, L. M. Zintgraf, C. A. S. De Witt, S. Whiteson, and J. Foerster, "Generalized beliefs for cooperative AI," in *Proc. International Conference on Machine Learning*, vol. 162, Jul 2022, pp. 16 062–16 082.

[121] W. Mao, K. Zhang, E. Miehling, and T. Basar, "Information state embedding in partially observable cooperative multi-agent reinforcement learning," in *Proc. IEEE CDC*, 2020, pp. 6124–6131.

[122] D. Bertsekas and J. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 2015.

[123] M. R. Garey and D. S. Johnson, *Computers and Intractability*. San Francisco: Freeman, 1979, vol. 174.

[124] J. Hazla, A. Jadbabaie, E. Mossel, and M. A. Rahimian, "Bayesian decision making in groups is hard," *Operations Research*, vol. 69, no. 2, pp. 632–654, 2021.

[125] E. Jorge, M. Kågebäck, F. D. Johansson, and E. Gustavsson, "Learning to play guess who? and inventing a grounded language as a consequence," *arXiv:1611.03218*, 2016.

[126] J. Bretagnolle and C. Huber, "Estimation des densités : Risque minimax," in *Séminaire de Probabilités XII*, C. Dellacherie, P. A. Meyer, and M. Weil, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1978, pp. 342–363.

[127] S. Vlaski and A. H. Sayed, "Distributed learning in non-convex environments—part I: Agreement at a linear rate," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1242–1256, 2021.

[128] ——, "Distributed learning in non-convex environments— part II: Polynomial escape from saddle-points," *IEEE Transactions on Signal Processing*, vol. 69, pp. 1257–1270, 2021.

[129] M. Kayaalp, S. Vlaski, and A. H. Sayed, "Dif-MAML: Decentralized multi-agent meta-learning," *IEEE Open Journal of Signal Processing*, vol. 3, pp. 71–93, 2022.

[130] ——, "Distributed meta-learning with networked agents," in *Proc. European Signal Processing Conference (EUSIPCO)*, 2021, pp. 1361–1365.

[131] M. Kayaalp and A. H. Sayed, "Causal influences over social learning networks," *arXiv:2307.09575*, July 2023.

[132] S. T. Smith, E. K. Kao, E. D. Mackin, D. C. Shah, O. Simek, and D. B. Rubin, "Automatic detection of influential actors in disinformation networks," *Proc. National Academy of Sciences*, vol. 118, no. 4, 2021.

[133] P. Lagrée, O. Cappé, B. Cautis, and S. Maniu, "Algorithms for online influencer marketing," *ACM Transactions on Knowledge Discovery from Data*, vol. 13, no. 1, pp. 1–30, 2018.

[134] F. Dablander and M. Hinne, "Node centrality measures are a poor substitute for causal inference," *Scientific Reports*, vol. 9, no. 1, p. 6846, 2019.

[135] D. Kempe, J. Kleinberg, and É. Tardos, "Maximizing the spread of influence through a social network," in *Proc. ACM SIGKDD*, 2003, pp. 137–146.

[136] A. Banerjee, A. G. Chandrasekhar, E. Duflo, and M. O. Jackson, "The diffusion of microfinance," *Science*, vol. 341, no. 6144, p. 1236498, 2013.

[137] V. Shumovskaia, M. Kayaalp, M. Cemri, and A. H. Sayed, "Discovering influencers in opinion formation over social graphs," *IEEE Open Journal of Signal Processing*, vol. 4, pp. 188–207, 2023.

[138] S. Wright, "The method of path coefficients," *The Annals of Mathematical Statistics*, vol. 5, no. 3, pp. 161–215, 1934.

[139] J. Pearl, *Causality*. Cambridge University Press, 2009.

[140] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, pp. 424–438, 1969.

[141] J. Pearl and D. Mackenzie, *The Book of Why: The New Science of Cause and Effect*. Basic Books, 2018.

[142] J. Woodward, *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, 2004.

[143] S. T. Smith, E. K. Kao, K. D. Senne, G. Bernstein, and S. Philips, "Bayesian discovery of threat networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5324–5338, 2014.

[144] P. Molavi, A. Jadbabaie, K. R. Rad, and A. Tahbaz-Salehi, "Reaching consensus with increasing information," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 358–369, 2013.

[145] E. Mossel, A. Sly, and O. Tamuz, "Strategic learning and the topology of social networks," *Econometrica*, vol. 83, no. 5, pp. 1755–1794, 2015.

[146] B. Ying and A. H. Sayed, "Information exchange and learning dynamics over weakly connected adaptive networks," *IEEE Transactions on Information Theory*, vol. 62, no. 3, pp. 1396–1414, 2016.

[147] H. Salami, B. Ying, and A. H. Sayed, "Social learning over weakly connected graphs," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 2, pp. 222–238, 2017.

[148] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press, 2017.

[149] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Computer Networks and ISDN Systems*, vol. 30, no. 1, pp. 107–117, 1998.

[150] D. Quercia, J. Ellis, L. Capra, and J. Crowcroft, "In the mood for being influential on Twitter," in *Proc. IEEE International Conference on Social Computing*, Boston, MA, 2011, pp. 307–314.

[151] S. Jain and A. Sinha, "Identification of influential users on Twitter: A novel weighted correlated influence measure for Covid-19," *Chaos, Solitons & Fractals*, vol. 139, pp. 1–8, 2020.

[152] L. Ante, "How Elon Musk's Twitter activity moves cryptocurrency markets," *Technological Forecasting and Social Change*, vol. 186, pp. 1–14, 2023.

[153] D. Loureiro, F. Barbieri, L. Neves, L. E. Anke, and J. Camacho-Collados, "TimeLMs: Diachronic language models from Twitter," in *Proc. Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 2022, pp. 251–260.

[154] V. Strassen, "Gaussian elimination is not optimal," *Numerische Mathematik*, vol. 13, no. 4, pp. 354–356, 1969.

[155] M. Kayaalp, Y. Inan, V. Koivunen, and A. H. Sayed, "Causal impact analysis for asynchronous decision making," *Proc. IEEE International Symposium on Information Theory (ISIT)*, pp. 1641–1645, July 2024.

[156] ——, "Causal influence in federated edge inference," *arXiv:2405.01260*, May 2024.

[157] T. Chavdarova, P. Baqué, A. Maksai, S. Bouquet, C. Jose, L. Lettry, F. Fleuret, P. Fua, and L. V. Gool, "WILDTRACK: A multi-camera hd dataset for dense unscripted pedestrian detection," in *Proc. IEEE/CVF CVPR*, Salt Lake City, UT, June 2018, pp. 5030–5039.

[158] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals and Systems*, vol. 1, no. 2, pp. 167–182, 1988.

[159] Y. Inan, M. Kayaalp, A. H. Sayed, and E. Telatar, "A fundamental limit of distributed hypothesis testing under memoryless quantization," in *Proc. IEEE International Conference on Communications (ICC)*, 2022, pp. 4824–4829.

[160] W. Liu, M. Salzmann, and P. Fua, "Context-aware crowd counting," in *Proc. IEEE/CVF CVPR*, Long Beach, CA, June 2019.

[161] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*. Cambridge University Press, 2018.

[162] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[163] P. Hu, V. Bordignon, M. Kayaalp, and A. H. Sayed, "Non-asymptotic performance of social machine learning under limited data," *arXiv:2306.09397*, June 2023.

[164] P. Toulis and E. Kao, "Estimation of causal peer influence effects," in *Proc. ICML*, 2013, pp. 1489–1497.

[165] D. L. Sussman and E. M. Airoldi, "Elements of estimation theory for causal effects in the presence of network interference," *arXiv:1702.03578*, 2017.

[166] A. Agarwal, S. Cen, D. Shah, and C. L. Yu, "Network synthetic interventions: A framework for panel data with network interference," *arXiv:2210.11355*, 2022.

[167] M. Eichler and V. Didelez, "On Granger causality and the effect of interventions in time series," *Lifetime data analysis*, vol. 16, pp. 3–32, 2010.

[168] M. Eichler, "Causal inference in time series analysis," *Causality: Statistical Perspectives and Applications*, pp. 327–354, 2012.

[169] J. Etesami and N. Kiyavash, "Directed information graphs: A generalization of linear dynamical graphs," in *Proc. American Control Conference*, 2014, pp. 2563–2568.

[170] C. J. Quinn, N. Kiyavash, and T. P. Coleman, "Directed information graphs," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6887–6909, 2015.

[171] J. Etesami, N. Kiyavash, K. Zhang, and K. Singhal, "Learning network of multivariate Hawkes processes: a time series approach," in *Proc. Uncertainty in Artificial Intelligence*, 2016, p. 162–171.

[172] S. W. Mogensen, D. Malinsky, and N. R. Hansen, "Causal learning for partially observed stochastic dynamical systems." in *Proc. Uncertainty in Artificial Intelligence*, 2018, pp. 350–360.

[173] W. Trouleau, J. Etesami, M. Grossglauser, N. Kiyavash, and P. Thiran, "Learning Hawkes processes under synchronization noise," in *Proc. International Conference on Machine Learning*, vol. 97, 09–15 Jun 2019, pp. 6325–6334.

[174] M. Divernois, J. Etesami, D. Filipovic, and N. Kiyavash, "Analysis of large market data using neural networks: A causal approach," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 833–847, 2023.

[175] A. Santos, D. Rente, R. Seabra, and J. M. F. Moura, "Learning the causal structure of networked dynamical systems under latent nodes and structured noise," in *Proc. AAAI*, vol. 38, no. 13, 2024, pp. 14 866–14 874.

[176] V. Shumovskaia, M. Kayaalp, and A. H. Sayed, "Distributed decision-making for community structured networks," in *Proc. IEEE ICASSP*, April 2024, pp. 9316–9320.

[177] ——, "Detection of malicious agents in social learning," *arXiv:2403.12619*, March 2024.

[178] ——, "Social learning in community structured graphs," *IEEE Transactions on Signal Processing*, vol. 72, pp. 2812–2826, 2024.

[179] K. Ntemos, V. Bordignon, S. Vlaski, and A. H. Sayed, "Social learning with disparate hypotheses," in *Proc. EUSIPCO*, 2022, pp. 2171–2175.

[180] ——, "Self-aware social learning over graphs," *IEEE Transactions on Information Theory*, vol. 69, no. 8, pp. 5299–5317, 2023.

[181] M. Cirillo, V. Bordignon, V. Matta, and A. H. Sayed, "Memory-aware social learning under partial information sharing," *IEEE Transactions on Signal Processing*, vol. 71, pp. 2833–2848, 2023.

[182] Y. Inan, M. Kayaalp, E. Telatar, and A. H. Sayed, "Social learning under randomized collaborations," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2022, pp. 115–120.

[183] M. Cemri, V. Bordignon, M. Kayaalp, V. Shumovskaia, and A. H. Sayed, "Asynchronous social learning," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1–5.

[184] B. Schölkopf, F. Locatello, S. Bauer, N. R. Ke, N. Kalchbrenner, A. Goyal, and Y. Bengio, "Toward causal representation learning," *Proc. IEEE*, vol. 109, no. 5, pp. 612–634, 2021.

Curriculum Vitae

MERT KAYAALP

EDUCATION

Ph.D., École Polytechnique Fédérale de Lausanne (EPFL) 2019-Present
Computer and Communication Sciences Lausanne, Switzerland
Thesis: Sequential Reasoning with Socially Caused Beliefs Advisor: Prof. Ali H. Sayed

B.S., Koç University 2015-2019
Electrical and Electronics Engineering Istanbul, Turkey
GPA: 4.14 / 4.30, Semahat Arsel Scholar & Vehbi Koç Honor Awardee

Ankara High School of Science 2011-2015
Physics Olympiad Team, IPhO medalist Ankara, Turkey

HONORS AND AWARDS

Best Student Paper Award – IEEE Statistical Signal Processing Workshop Hanoi, Vietnam, Jul 2023
Bronze Medal - International Physics Olympiad (IPhO) Mumbai, India, Jul 2015
Silver Medal - Turkish National Physics Olympiad (TUBITAK) Dec 2014, Dec 2013
Turkish National University Entrance Examination (LYS) Jun 2015
1st out of 1.8 million participants (with IPhO points) Score: 586/560

TEACHING

Teaching assistant

Information, Computation and Communication (BS)	Fall 2021, 2023
Adaptation and Learning (MS)	Spring 2022
Advanced Machine Learning (MS)	Spring 2021
Large-Scale Data Science for Real-World Data (MS)	Spring 2020
Analysis I (BS)	Fall 2020

PUBLICATIONS

JOURNAL

1. V. Shumovskaia, M. Kayaalp, and A. H. Sayed, "Detection of malicious agents in social learning", *IEEE Signal Processing Letters*, vol. 31, pp. 1745-1749, July 2024, arXiv:2403.12619 [cs.SI]
2. V. Shumovskaia, M. Kayaalp, and A. H. Sayed, "Social learning in community structured graphs", *IEEE Transactions on Signal Processing*, vol. 72, pp. 2812-2826, May 2024, arXiv:2312.12186 [cs.SI]
3. M. Kayaalp, V. Bordignon, and A. H. Sayed, "Social opinion formation and decision making under communication trends", *IEEE Transactions on Signal Processing*, vol. 72, pp. 506-520, January 2024, arXiv:2203.02466 [eess.SP]
4. M. Kayaalp, Y. Inan, E. Telatar, and A. H. Sayed, "On the arithmetic and geometric fusion of beliefs for distributed inference", *IEEE Transactions on Automatic Control*, November 2023, arXiv:2204.13741 [eess.SP]
5. M. Kayaalp, F. Ghadieh, and A. H. Sayed, "Policy evaluation in decentralized POMDPs with belief sharing", *IEEE Open Journal of Control Systems*, vol. 2, pp. 125-145, June 2023
6. V. Shumovskaia, M. Kayaalp, M. Cemri, and A. H. Sayed, "Discovering influencers in opinion formation over social graphs", *IEEE Open Journal of Signal Processing*, vol. 4, pp. 188-207, March 2023
7. M. Kayaalp, S. Vlaski, and A. H. Sayed, "Dif-MAML: Decentralized multi-agent meta-learning", *IEEE Open Journal of Signal Processing*, vol. 3, pp. 71-93, January 2022

PRE-PRINTS

1. M. Kayaalp, Y. Inan, V. Koivunen, and A. H. Sayed, "Causal influence in federated edge inference", arXiv:2405.01260 [cs.LG]
2. M. Kayaalp and A. H. Sayed, "Causal influences over social learning networks", arXiv:2307.09575 [cs.SI]
3. P. Hu, V. Bordignon, M. Kayaalp, and A. H. Sayed, "Non-asymptotic performance of social machine learning under limited data", arXiv:2306.09397 [cs.LG]
4. M. Kayaalp, V. Bordignon, S. Vlaski, V. Matta, and A. H. Sayed, "Distributed Bayesian learning of dynamic states", arXiv:2212.02565 [eess.SP]

CONFERENCE

1. M. Kayaalp, Y. Inan, V. Koivunen, and A. H. Sayed, "Causal impact analysis for asynchronous decision making", *IEEE Int. Symp. Information Theory (ISIT)*, Athens, Greece, July 2024
2. V. Shumovskaia, M. Kayaalp, and A. H. Sayed, "Distributed decision-making for community structured networks", *IEEE ICASSP*, Seoul, Korea, April 2024
3. V. Bordignon, M. Kayaalp, V. Matta, and A. H. Sayed, "Social learning with non-Bayesian local updates", *EUSIPCO*, Helsinki, Finland, September 2023
4. M. Kayaalp, Y. Inan, V. Koivunen, E. Telatar, and A. H. Sayed, "On the fusion strategies for federated decision making", *IEEE Statistical Signal Processing Workshop (SSP)*, Hanoi, Vietnam, July 2023
5. V. Shumovskaia, M. Kayaalp, and A. H. Sayed, "Identifying opinion influencers over social networks", *IEEE ICASSP*, Rhodes Island, Greece, June 2023
6. P. Hu, V. Bordignon, M. Kayaalp, and A. H. Sayed, "Performance of social machine learning under limited data", *IEEE ICASSP*, Rhodes Island, Greece, June 2023
7. M. Cemri, V. Bordignon, M. Kayaalp, V. Shumovskaia, and A. H. Sayed, "Asynchronous social learning", *IEEE ICASSP*, Rhodes Island, Greece, June 2023
8. Y. Inan, M. Kayaalp, E. Telatar, and A. H. Sayed, "Social learning under randomized collaborations", *IEEE Int. Symp. Information Theory (ISIT)*, Helsinki, Finland, June 2022
9. M. Kayaalp, V. Bordignon, S. Vlaski, and A. H. Sayed, "Hidden Markov modeling over graphs", *IEEE Data Science and Learning Workshop (DSLW)*, Singapore, May 2022
10. Y. Inan, M. Kayaalp, A. H. Sayed, and E. Telatar, "A fundamental limit of distributed hypothesis testing under memoryless quantization", *IEEE ICC*, Seoul, South Korea, May 2022
11. M. Kayaalp, S. Vlaski, and A. H. Sayed, "Distributed meta-learning with networked agents", *EUSIPCO*, Dublin, Ireland, August 2021