Causal Impact Analysis for Asynchronous Decision Making

Mert Kayaalp*, Yunus İnan*, Visa Koivunen†, Ali H. Sayed*

*EPFL, Switzerland, {mert.kayaalp, yunus.inan, ali.sayed}@epfl.ch
†Aalto University, Finland, visa.koivunen@aalto.fi

***Abstract*—We consider a collaborative decision-making framework where heterogeneous agents receive streaming and partially informative observations. We consider two asynchronous scenarios that differ based on the agents' participation patterns and the fusion center's policies. By using hypothetical interventions on individual agents to conduct credit assignment, we attribute causal impact scores to each agent for the joint decision. By further employing these scores in a guided theoretical analysis, we compare the fusion center's two policies by evaluating their vulnerability to adversarial attacks, robustness against moderate deviations, and fairness.**

I. Introduction

Autonomous systems are usually equipped with sensing and computing capabilities to enable prompt decisions based on streaming observations. One example is self-driving vehicles, which need to react to changes in road conditions in real-time. In situations where a multitude of agents observe a common state of nature or phenomenon, then cooperative decision-making becomes beneficial because the individual sensor observations may only relay partial information about the event of interest. The main benefit of cooperative decision-making or inference is the diversity of the information provided by distinct agents. Yet, this very strength has its own challenges. For instance, empowering an outlier agent can make the system vulnerable to adversarial attacks, while discarding outliers may violate fairness. Therefore, understanding to what extent an agent impacts the quality of the decision in a multi-agent system is an important question for many applications.

To that end, in this paper, we examine how varying factors such as participation/absence patterns, fusion center policies, and data distribution affect the impact that an agent has on the aggregate decision. First, we revisit in Sec. II-A the distributed decision making setting proposed in [1]. Then, in Sec. II-B, we extend the framework by incorporating asynchronous behavior into the agents. We consider two distinct scenarios that vary based on the participation patterns of the agents during information sharing. In order to understand the influence of agents under these scenarios, we treat influence as a causal quantity in Sec. II-C and use hypothetical interventions [2] on agents to derive closed form expressions for the causal impacts of individual agents on the overall decision of the agents.

A. Related Work

We adopt the framework of [1], which is a special case of the non-Bayesian social learning paradigm from [3]–[7]. Under social learning, the objective is to perform cooperative decision-making through localized data processing and exchanges among agents. In the context of cooperative decision-making, the use of hypothetical causal interventions to understand influence was considered before in [8], [9] to assess the influence of agents on each other over a network. We focus here on using these causal impacts in the presence of *asynchronicity* to explore how the fusion center policy and the pattern of agent participation can influence dynamics. This is a "bottom-up" perspective in the sense that we examine the impact of individual agents on the joint decision. This is as opposed to "top-down" approaches in previous works, which employ objective functions or algorithms to promote robustness and fairness while safeguarding against outliers [10]–[13].

Application (**Spontaneous collaboration**). *In many applications, agents start cooperating spontaneously. For instance, intelligent vehicles on the same road can collaborate to better understand the road conditions. In these ad-hoc scenarios, it is impractical to assume synchronicity. As such, determining an agent's impact under asynchronicity is crucial for taking robust and fair joint decisions.* ■

Notation. Random variables are written in boldface letters. We use the "proportional to" symbol $\propto$ whenever the LHS of an equation is a proper normalization of the RHS. The KL divergence between two probability distributions p and q is denoted by $D_{\text{KL}}(p||q)$. Following the notation in [9], we use $\sim$ to denote the counterparts of variables after an intervention.

II. Problem Formulation

A. Synchronous Collaboration

We start by introducing the synchronous setting of [1]. Consider a setting where a group of K agents wish to discover the true state of nature θ° from a set of potential hypotheses $\Theta \triangleq \{\theta_1, \dots, \theta_H\}$, with the help of a fusion center. For instance, autonomous vehicles on the same road can be connected to a cloud with the objective of assessing the road conditions {crowded, accident, normal} — see Fig. 1 for a visual illustration. At each time instant i, each agent k receives an observation $\boldsymbol{\xi}_{k,i}$, which conveys partial information about θ°. Instead of directly transmitting the raw observations $\boldsymbol{\xi}_{k,i}$ to the central server, each agent k processes its data locally with its personalized likelihood model $L_k(\boldsymbol{\xi}_{k,i}|\theta)$ (e.g., a neural

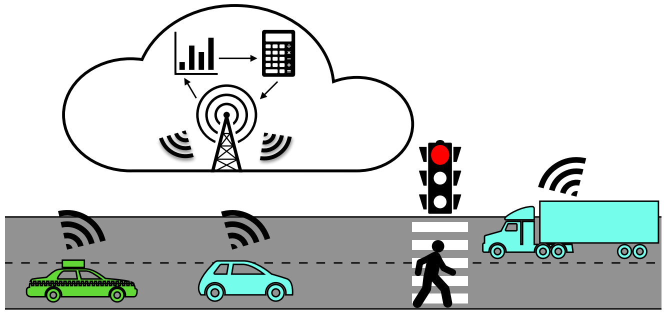


Fig. 1: Smart vehicles typically generate extensive raw sensory data. Exchanging soft-decisions instead of the raw data can be advantageous due to communication overhead.

network) in a Bayesian manner to obtain an intermediate belief (soft decision) about which hypothesis is the true one:

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

where $\boldsymbol{\mu}_{i-1}$ is the prior probability mass function (pmf). Subsequently, agent k forwards this intermediate pmf $\boldsymbol{\psi}_{k,i}$ to the fusion center (FC). The FC may lack knowledge about the system's joint likelihood, the observations at the agents, or the agents' likelihood models. Therefore, it employs a weighted geometric averaging of the received information in a non-Bayesian manner [3]–[7] for each $\theta \in \Theta$:

$$\boldsymbol{\mu}_i(\theta) \propto \prod_{k=1}^{K} (\boldsymbol{\psi}_{k,i}(\theta))^{\pi_k} \quad \text{(Combine).} \tag{2}$$

Here, $\pi \triangleq [\pi_1, \ldots, \pi_K]^\mathsf{T}$ denotes the vector of confidence weights $\pi_k \in (0,1)$ that the fusion center assigns to each agent k [14], [15], potentially formed from the previous interactions with the agents. They are assumed to be positive constants that sum up to 1. The server then sends the aggregated belief back to the agents. This procedure of updating and exchanging beliefs is executed repeatedly at every time instant.

B. Two Asynchronous Scenarios

Asynchronous behavior is common in many real-world applications of distributed systems. We consider two scenarios that are distinct based on the symmetry of communication between the agents and the fusion center. For both scenarios, we use the Bernoulli variable $\boldsymbol{q}_{k,i}$ to indicate if agent k is sharing its intermediate belief $\boldsymbol{\psi}_{k,i}$ with the server at time i, namely,

$$\boldsymbol{q}_{k,i} = \begin{cases} 1, & \text{with probability } p_k \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

We assume the process $\{\boldsymbol{q}_{k,i}\}$ is i.i.d. over time and also independent over space.

1) 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, non-informative soft decisions, or the lack of significant changes in intermediate statistics since the previous transmission. However, these agents can keep receiving the updates from the server. Another 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). In this case, the fusion center can fill the belief components of missing agents with its own prior while aggregating information. Therefore, the combination step (2) at the server side changes to

$$\boldsymbol{\mu}_i(\theta) \propto \prod_{k=1}^{K} \left(\boldsymbol{\psi}_{k,i}^{\boldsymbol{q}_{k,i}}(\theta)\boldsymbol{\mu}_{i-1}^{1-\boldsymbol{q}_{k,i}}(\theta)\right)^{\pi_k}. \tag{4}$$

Nevertheless, the adaptation step (1) at the agent side remains unchanged and agents continue to utilize the beliefs received from the server locally.

It is worth noting the parallel between this scenario and the traditional distributed detection strategies [15]–[17]. Since the server knows $\boldsymbol{\mu}_{i-1}$, sharing $\boldsymbol{\psi}_{k,i}$ is essentially equivalent to sharing the observation likelihood $L_k(\boldsymbol{\xi}_{k,i}|\theta)$ due to (1). Similarly, agents (e.g., sensors) relay a sufficient statistics of their likelihoods to the fusion center in [15]–[17]. The difference is that in these works, the fusion center does not communicate any information back to the agents.

2) Symmetric communication: Another possibility is that an agent does not receive any update from the server if that agent does not transmit information to the central processor. In other words, the absence of communication is reciprocal. This particular communication pattern can be rationalized from an economic perspective. For instance, a server might strategically choose not to update agents that do not contribute information, hence incentivizing data sharing and promoting a give-and-take dynamics. In this scenario, the combination step at the server side is given by (4), whereas the adaptation step (1) at the agents becomes

$$\boldsymbol{\psi}_{k,i}(\theta) \propto \begin{cases} L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\mu}_{i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 1 \\ L_k(\boldsymbol{\xi}_{k,i}|\theta)\boldsymbol{\psi}_{k,i-1}(\theta), & \text{if } \boldsymbol{q}_{k,i-1} = 0 \end{cases}. \tag{5}$$

The rationale behind (5) is as follows. If agent k has shared information with the server (i.e., $\boldsymbol{q}_{k,i-1} = 1$), 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 next update.

C. Causal Impact Definition

We extend the causal effect definition from [9]. 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. 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. 2 for a visual representation. It can be shown 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 $i \to \infty$ [18]. When an intervention occurs at agent m, the steady-state belief vector will be denoted instead by $\widetilde{\mu}_\infty$ as opposed to μ_∞. As such, we can quantify the causal impact of agent m on the joint decision by using the difference:

$$C_m \triangleq 1 - \widetilde{\mu}_\infty(\theta^\circ). \tag{6}$$

Expression (6) 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 [9], we can express the average belief, $\widetilde{\mu}_\infty(\theta^\circ)$ in the form:

$$\widetilde{\mu}_\infty(\theta^\circ) \triangleq \frac{1}{1 + \sum_{\theta \neq \theta^\circ} \exp\{-\widetilde{\lambda}_\infty(\theta)\}}. \tag{7}$$

where

$$\widetilde{\lambda}_\infty(\theta) \triangleq \lim_{i\to\infty} \mathbb{E}[\widetilde{\boldsymbol{\lambda}}_i(\theta)] \tag{8}$$

represents the expected log-belief ratio under the intervention with the variables $\widetilde{\boldsymbol{\lambda}}_i(\theta)$ defined by

$$\widetilde{\boldsymbol{\lambda}}_i(\theta) \triangleq \log \frac{\widetilde{\boldsymbol{\mu}}_i(\theta^\circ)}{\widetilde{\boldsymbol{\mu}}_i(\theta)}. \tag{9}$$

Here, $\widetilde{\mu}_\infty(\theta^\circ)$ represents the average belief of the server at steady state under the intervention $do(\boldsymbol{\psi}_{m,i} := \mu_m)$.

III. Theoretical Results

We begin by reviewing the causal impact result from [9], which addresses synchronous communication. To that end, we first define the informativeness level of each agent k as

$$d_k(\theta) \triangleq D_{\text{KL}}\big(L_k(\cdot|\theta^\circ)||L_k(\cdot|\theta)\big) \tag{10}$$

which represents how informative agent k's observations are for distinguishing θ° from θ.

Theorem 1 (**Synchronous collaboration [9]**). *Under synchronous collaboration described in Sec. II-A, the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_\infty(\theta) = \frac{1}{\pi_m} \sum_{k\neq m} \pi_k d_k(\theta) + \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{11}$$

Therefore, by (6)*, the causal impact of agent m on the joint decision is*

$$C_m = 1 - \frac{1}{1 + \sum_{\theta\neq\theta^\circ} \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\Big\{-\frac{1}{\pi_m}\sum_{k\neq m}\pi_k d_k(\theta)\Big\}} \tag{12}$$

■

Equations (11) and (12) imply that:

- An increase in the confidence π_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 (11) and (12) are dependent on the intervention strength μ_m. For an intervention dose-independent causal impact measure, setting μ_m to a uniform belief (i.e., setting the log-belief ratio $\log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)}$ to zero) is discussed in [9] along with its equivalence to causal derivative effect [19]. Next, we consider the causal impacts for the asynchronous scenarios we have introduced in Sec. II-B.

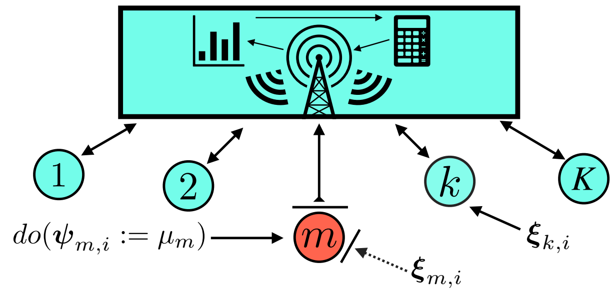


Fig. 2: Visual representation of a hypothetical intervention $do(\boldsymbol{\psi}_{m,i} := \mu_m)$. Agent m keeps sending information to the server with probability p_m, however, its belief is now fixed and is not dependent on any other variable.

Theorem 2 (**Asymmetric communication**). *Under the asymmetric communication protocol described in Sec. II-B, the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_\infty(\theta) = \frac{1}{\pi_m} \sum_{k\neq m} \pi_k p_k d_k(\theta) + p_m \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{13}$$

This implies by (6) *that the causal effect of agent m on the joint decision is given by*

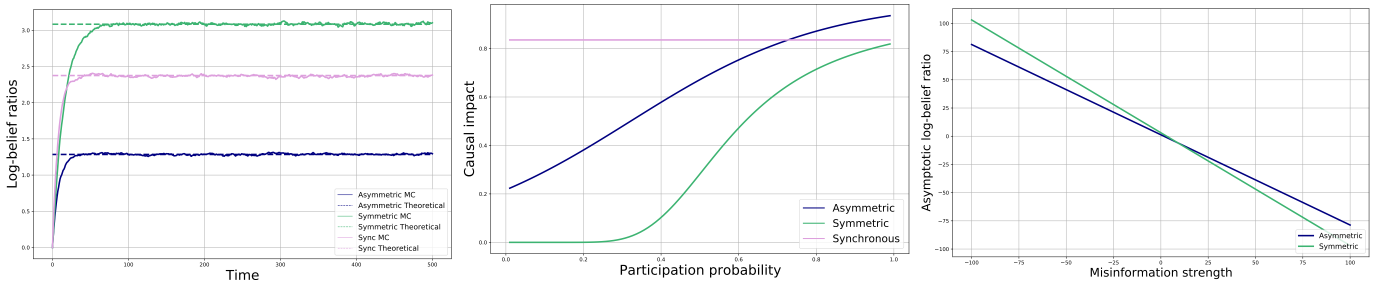
$$C_m = 1 - \frac{1}{1 + \sum_{\theta\neq\theta^\circ} \left(\frac{\mu_m(\theta)}{\mu_m(\theta^\circ)}\right)^{p_m} \exp\Big\{-\frac{1}{\pi_m}\sum_{k\neq m}\pi_k p_k d_k(\theta)\Big\}} \tag{14}$$

Proof. Omitted due to space limitations. Available in [18]. ■

Notice in Theorem 2 that as p_k approaches 1 for each agent k, i.e., when all agents participate synchronously at each iteration, we recover Theorem 1. Also notice that the essential difference from the synchronous scenario is the replacement of confidence weights π_k by $\pi_k p_k$. 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.

Theorem 3 (**Symmetric communication**). *Under the symmetric communication protocol described in Sec. II-B the expected log-belief ratio under intervention is given by*

$$\widetilde{\lambda}_\infty(\theta) = \frac{1}{\pi_m p_m} \sum_{k\neq m} \frac{\pi_k d_k(\theta)}{1 - \pi_k(1-p_k)} + \log \frac{\mu_m(\theta^\circ)}{\mu_m(\theta)} \tag{15}$$



(a) (b) (c)

Fig. 3: (a) Simulated log-belief ratios (averaged over 1000 Monte Carlo (MC) simulations and theoretical expressions over time, (b) Causal impact of agent $m = 1$ on the joint decision with changing participation probability p_m, (c) Asymptotic log-belief ratio with respect to misinformation strength $\log \dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)}$.

This implies by (6) *that the causal effect of agent m on the joint decision is given by*

$$C_m = 1 - \frac{1}{1 + \sum_{\theta \neq \theta^\circ} \dfrac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \exp\left\{ \dfrac{-1}{\pi_m p_m} \sum_{k \neq m} \dfrac{\pi_k d_k(\theta)}{1 - \pi_k(1 - p_k)} \right\}} \tag{16}$$

Proof. Omitted due to space limitations. Available in [18]. ■

Similar to the asymmetric communication scenario in Theorem 2, as $p_k \to 1$ for all agents, Theorem 3 recovers the synchronous collaboration result Theorem 1. 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 asynchronicity parameters. Notice from (13) and (15) that when the misinformation strength (defined as the ratio of an incorrect hypothesis belief to the correct hypothesis belief) from agent m meets the condition

$$\log \frac{\mu_m(\theta)}{\mu_m(\theta^\circ)} \geq \sum_{k \neq m} \frac{\pi_k d_k(\theta)}{\pi_m(1 - p_m)} \left(\frac{1}{p_m(1 - \pi_k(1 - p_k))} - p_k \right), \tag{17}$$

then the $\widetilde{\lambda}_\infty(\theta)$ term in (13) exceeds that in (15). Since by definition (7), $\lambda_\infty(\theta)$ is inversely proportional to the causal impact C_m, it also implies that agent m exerts a stronger causal impact on the joint decision in the symmetric scenario than in the asymmetric one if the misinformation strength surpasses the threshold specified in (17).

This observation holds significant relevance for practical applications. Commonly, misinformation (i.e., deviations from the norm) originating from malfunctioning agents is moderate. In contrast, malicious or Byzantine agents often supply adversarial misinformation that can be extreme. The discussion above suggests that the symmetric communication scenario is more vulnerable to adversarial attacks, while asymmetric communication is more sensitive to moderate level misinformation that typically emerges from malfunctioning agents without harmful intentions. Furthermore, for a decision-making process that aims to be both fair and 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 norm while also reducing the influence of extreme misinformation, providing a safeguard against adversarial attacks.

IV. Numerical Results

To verify our theoretical results, we consider a binary hypothesis testing problem with $K = 12$ agents connected to a fusion center, each receiving observations that follow a Gaussian distribution. 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 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 π_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.

In the first experiment, we average 1000 simulations for three settings: the synchronous setting from Sec. II-A, and the asymmetric and symmetric settings from Sec. II-B. This is performed under an intervention on agent $m = 1$ with uniform beliefs. We plot the evolution of log-belief ratios over 500 time instants in Fig. 3a, as well as the theoretical expressions for these values from Theorems 1, 2, and 3. Notice that the simulated log-belief ratios closely align with the theoretical expressions.

In Fig. 3b, we illustrate the causal impacts 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 a probability of 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 joint decision.

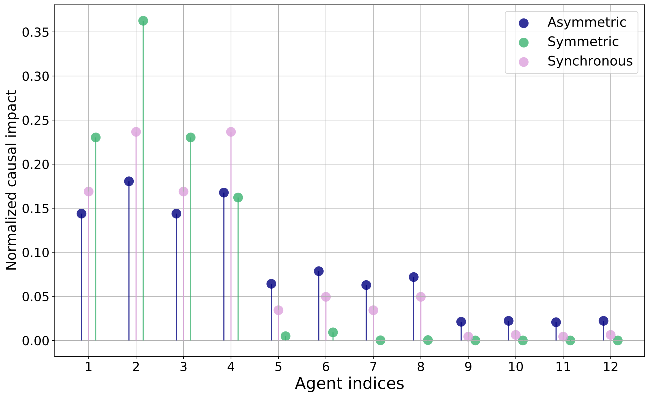


Fig. 4: Normalized causal impacts of each agent over three frameworks.

Next, in Fig. 3c, we plot the asymptotic log-belief ratios in relation to varying intervention strengths on agent $m = 1$. Supporting our theoretical finding in (17), 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 impacts compared to the symmetric communication framework.

Finally, in Fig. 4 we present 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 earlier theoretical findings, suggesting that asymmetric communication fosters a fairer decision-making process.

V. Concluding Remarks

In this paper, we examined two distinct asynchronous decision-making models, which differ in terms of whether the fusion center updates the agents that do not provide information. Utilizing a causal theoretical framework, we explained how each agent's impact on the collective decision varies based on factors such as the distribution of data received by the agents and their participation frequencies. These results revealed that symmetric (reciprocal) communication offers greater resilience to moderate deviations from the usual, whereas asymmetric communication protocols are more effective against adversarial attacks and better for fairness.

Future directions include extending this federated framework to decentralized peer-to-peer networks, and also examining different combination strategies at the server side such as median-based robust fusion [11].

References

[1] 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)*, Hanoi, Vietnam, July 2023, pp. 270–274. [Online]. Available: arXiv:2303.06109

[2] J. Pearl, *Causality*. Cambridge University Press, 2009.

[3] 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.

[4] X. Zhao and A. H. Sayed, "Learning over social networks via diffusion adaptation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, 2012, pp. 709–713.

[5] 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.

[6] 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.

[7] V. Bordignon, V. Matta, and A. H. Sayed, "Adaptive social learning," *IEEE Transactions on Information Theory*, vol. 67, no. 9, pp. 6053–6081, 2021.

[8] 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.

[9] M. Kayaalp and A. H. Sayed, "Causal influences over social learning networks," *arXiv:2307.09575*, July 2023.

[10] A. Vempaty, L. Tong, and P. K. Varshney, "Distributed inference with Byzantine data: State-of-the-art review on data falsification attacks," *IEEE Signal Processing Magazine*, vol. 30, no. 5, pp. 65–75, 2013.

[11] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*. Cambridge University Press, 2018.

[12] M. Zimmer, C. Glanois, U. Siddique, and P. Weng, "Learning fair policies in decentralized cooperative multi-agent reinforcement learning," in *Proc. ICML*, vol. 139, 18–24 Jul 2021, pp. 12 967–12 978.

[13] C. A. Schroth, S. Vlaski, and A. M. Zoubir, "Robust M-estimation based distributed expectation maximization algorithm with robust aggregation," in *Proc. IEEE ICASSP*, 2023, pp. 1–5.

[14] 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.

[15] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer, 2012.

[16] 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.

[17] 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.

[18] M. Kayaalp, Y. Inan, V. Koivunen, and A. H. Sayed, "Causal influence in federated edge inference," *arXiv:2405.01260*, May 2024.

[19] J. Peters, D. Janzing, and B. Schölkopf, *Elements of Causal Inference: Foundations and Learning Algorithms*. The MIT Press, 2017.