Low Message-Cost Gossip for Welfare-Optimal Decentralized Decision Making

Extended Abstract

Vedic Sharma
Dept. of Computer Science, University of Toronto
Toronto, Canada
vedic.sharma@mail.utoronto.ca

Peter Marbach
Dept. of Computer Science, University of Toronto
Toronto, Canada
marbach@cs.toronto.edu

ABSTRACT

We study how agents in a multi-agent system can reach welfare-optimal consensus under strict communication limits. We propose a gossip-based protocol that allows agents with heterogeneous preferences to identify the welfare-optimal option while exchanging estimates for only $O(1)$ options, independent of the total number of options K. The algorithm converges almost surely on stochastically connected interaction graphs and has provably bounded communication complexity on stochastically fully connected graphs. In complex domains such as swarm coordination and decentralized marketplaces, the optimal collective solution often substantially outperforms most alternatives, producing a highly skewed social welfare distribution. In such heavy-tailed settings, communication complexity can even *decrease* as the decision space grows, showing that scalable, welfare-optimal coordination can emerge in large decentralized systems with minimal communication.

KEYWORDS

Multi-Agent Systems, Welfare-Optimal Consensus, Gossip Protocols, Low-Message Cost Gossip, Communication Complexity

ACM Reference Format:
Vedic Sharma and Peter Marbach. 2026. Low Message-Cost Gossip for Welfare-Optimal Decentralized Decision Making: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/JLER3835

1 INTRODUCTION

Decision making in multi-agent systems (MAS) is fundamentally a problem of collective choice under limited information exchange. As the number of agents and decision alternatives grows, achieving outcomes that are socially efficient and informationally feasible becomes increasingly difficult. Each agent holds private preferences or utilities, yet bandwidth and locality constraints prevent centralized aggregation. This tension—between social welfare optimality and communication efficiency—lies at the heart of collective intelligence, and is central to domains such as distributed resource allocation [1], swarm robotics [7], and autonomous negotiation among connected vehicles [4]. Designing mechanisms that approximate social welfare under strict message limits is thus a key challenge at the intersection of decentralized coordination, social choice, and multi-agent learning.

Classical approaches fall short in achieving socially optimal collective decisions under communication constraints. Algorithms such as *distributed averaging* [15] and *gossip protocols* [2] enable scalable information aggregation, but they require agents to exchange full utility or state vectors, causing message complexity to grow with the number of alternatives. Conversely, lightweight schemes such as *binary consensus* and *voter models* [13] preserve constant message size but are limited to majority or threshold decisions, and thus cannot represent heterogeneous preferences or welfare trade-offs.

In related learning-based settings, such as *policy consensus* [19] and *decentralized bandits* [11], agents face the same fundamental trade-off: rich decision spaces demand high communication, while communication-efficient methods collapse to low-expressivity choices. None of these approaches provide a mechanism that simultaneously achieves welfare-optimal aggregation and constant message complexity, which is a central goal for scalable collective intelligence.

2 PROBLEM DEFINITION

In this paper, we study the problem of welfare-optimal consensus in decentralized multi-agent systems: a setting in which agents must collectively identify the alternative that maximizes aggregate welfare. This framework unifies ideas from distributed consensus and welfare economics, asking how much social efficiency can be achieved when information transmission is fundamentally constrained.

Formally, let $\mathcal{V} = \{1, \ldots, N\}$ denote the set of agents and $\mathcal{Z} = \{1, \ldots, K\}$ the set of available options. Each agent $i \in \mathcal{V}$ has a utility vector $u_i \in \mathbb{R}^K$, where $u_i(z)$ represents the utility that agent i obtains from option $z \in \mathcal{Z}$. The social welfare of option z is then

$$S(z) = \sum_{i \in \mathcal{V}} u_i(z).$$

We assume, without loss of generality, that the options are ordered so that $S(1) \geq S(2) \geq \cdots \geq S(K)$, and that there is a unique optimal option, i.e. $S(1) > S(2)$.

The goal is for agents to identify the welfare-maximizing option $z^* = 1$ in a fully decentralized manner:

$$z^* = \arg\max_{z \in \mathcal{Z}} S(z). \tag{1}$$

We assume that agents have knowledge about their own utility $u_i(z)$ for each option $z \in \mathcal{Z}$, but they do not know the utilities of

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/JLER3835

other agents. As a result, agents need to communicate with each other in order to identify the welfare-optimal option z^*.

The goal of this paper is to show that gossip protocols can enable welfare-optimal decision making with low communication complexity, particularly when the social welfare distribution is highly skewed. This work provides a pathway toward scalable welfare-optimal decision making in large-scale, communication-limited multi-agent systems. It also opens avenues for future research, both theoretical (e.g., characterizing communication complexity for general stochastic graphs) and applied (e.g., adapting and fine-tuning the algorithm for practical applications).

3 CONTRIBUTIONS

We propose a gossip-style algorithm that achieves welfare-optimal consensus with $O(1)$ message complexity, independent of the number of options K. The key idea is simple: rather than sharing information about all K options, each agent maintains and exchanges estimates only for a small subset of its *top-L* options—those currently believed to yield the highest welfare. By restricting communication to these L options ($L \ll K$), agents exchange a constant number of estimates per interaction, achieving $O(1)$ message complexity (in the number of estimates exchanged per interaction) while still ensuring convergence to the welfare-optimal option.

Intuitively, the process functions as a *focused gossip mechanism*. Each agent concentrates on its most promising alternatives, communicates them to others, and iteratively refines its beliefs as information diffuses through the network. This mirrors how humans naturally exchange information when identifying high-value choices among many—asking, for instance, "What is your top choice?" or "What are your top three?" Rather than exhaustively sharing every option, attention and communication focus on the most promising few, accelerating convergence while minimizing overhead. Our results formalizes this intuition.

A central analytical challenge is what we term *Option Neglect*: showing that, even if no agent initially ranks the welfare-optimal option z^* among its top-L set, the consensus protocol can still discover z^* and drive the population to select it. Our analysis formally shows that the proposed algorithm is indeed able to overcome option neglect.

Specifically, the main results of our analysis are as follows.

Novel Algorithm: We formulate welfare-optimal consensus as a decentralized decision-making problem, in which agents with heterogeneous preferences collaboratively identify the socially optimal option under strict communication limits, and provide a gossip-based protocol that enables agents to identify the welfare-optimal option by exchanging social welfare estimates of only $O(1)$ options per interaction.

Almost-Sure Convergence to Welfare Optimality: For any connected interaction graph and any $L \geq 1$, the protocol converges almost surely to consensus on the welfare-optimal option z^*, even if z^* is initially absent from all agents' top-L lists. This resolves the core information problem in decentralized welfare maximization: discovering the global optimum through purely local communication without access to others' utilities or global statistics, and it overcomes the option neglect problem by ensuring that z^* is eventually identified as the welfare-optimal option, even if z^* is initially absent from all agents' top-L lists.

Communication Complexity: For the minimal list size $L = 1$, the communication complexity of the proposed protocol scales as

$$O\Big(\frac{D(K) + EC(K)}{MS(K)}\Big),$$

where $MS(K) = \frac{S(z^*)-S(z^{(2)})}{N}$ is the per-agent welfare advantage of the optimal option z^* over the second-best $z^{(2)}$, $D(k)$ measures the dispersion of the private utilities $u_i(z^*)$ of the welfare-optimal option, and $EC(K)$ measures the initial "effective competition" from sub-optimal options.

Inverse Scaling Law for Heavy-Tailed Distributions: Specifically, when the distribution of the welfare values of the different options is heavy-tailed (e.g., follow a Fréchet distribution), the protocol exhibits an *inverse scaling law*: for $L = 1$, the communication complexity decreases as $O(K^{-2/\alpha})$ with the number of options K, where α is the tail exponent. This holds under the assumption that the variance of agents' private utilities for any fixed option remains constant as K grows. Intuitively, as K increases, the extremal separation between the welfare-optimal option z^* and other options grows faster than the local utility noise. The protocol exploits this structure and efficiently amplifies these extreme-value signals to identify z^* with decreasing communication cost.

Taken together, these results provide the first consensus protocol for MAS that combines optimality, decentralization, and constant message size.

4 RELATED WORK

This work builds on two main strands in distributed computing: averaging-based consensus and decentralized multi-option decision-making. Classical averaging algorithms, beginning with the DeGroot model [5], establish that agents connected through strongly connected networks converge through repeated averaging. Subsequent work extended these results to synchronous, asynchronous, and gossip-based communication settings [2, 14, 16]. While these methods provide robust consensus guarantees, they typically require exchanging full K-dimensional state vectors, resulting in $O(K)$ communication per interaction. Communication-efficient variants, including quantized consensus [3] and compressed distributed optimization [10], reduce message size but introduce approximation error and still scale linearly with K, limiting their ability to achieve exact welfare-optimal consensus.

Decentralized multi-agent bandit algorithms address distributed learning under heterogeneous rewards [11, 20], but primarily target regret minimization rather than welfare maximization or consensus formation. Best-arm identification methods [6, 8, 9, 12] often rely on centralized coordination or assume homogeneous rewards, while most decentralized bandit approaches require sharing statistics for all arms [18], incurring $O(K)$ communication. Federated approaches [17] reduce communication but depend on central aggregation. In contrast, our approach achieves welfare-optimal consensus with constant communication complexity independent of K while remaining fully decentralized.

REFERENCES

[1] Stéphane Airiau and Ulle Endriss. 2010. Multiagent Resource Allocation with Sharable Items: Simple Protocols and Nash Equilibria. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (AAMAS '10)*. Toronto, Canada, 491–498.

[2] Stephen Boyd, Arpita Ghosh, Balaji Prabhakar, and Devavrat Shah. 2006. Randomized gossip algorithms. *IEEE Transactions on Information Theory* 52, 6 (2006), 2508–2530.

[3] Ruggero Carli, Fabio Fagnani, and Sandro Zampieri. 2008. Quantized average consensus via dynamic coding/decoding schemes. In *International Journal of Robust and Nonlinear Control*, Vol. 20. Wiley, 156–175.

[4] Dave De Jonge and Dongmo Zhang. 2016. Using GDL to Represent Domain Knowledge for Automated Negotiations. In *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (AAMAS '16)*. Singapore, Singapore, 134–153.

[5] Morris H. DeGroot. 1974. Reaching a consensus. *J. Amer. Statist. Assoc.* 69, 345 (1974), 118–121.

[6] Eyal Even-Dar, Shie Mannor, and Yishay Mansour. 2006. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. In *Journal of Machine Learning Research*, Vol. 7. 1079–1105.

[7] M.A. Hsieh, V. Kumar, and L. Chaimowicz. 2008. Decentralized Motion Control and Planning for Multiple Mobile Robots in the Presence of Obstacles. *Autonomous Agents and Multi-Agent Systems (JAAMAS)* 16 (2008), 3–34. Issue 1.

[8] Andrei Karpov and Yevgeny Seldin. 2022. Best arm identification under partial monitoring with distributed players. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*.

[9] Emilie Kaufmann, Shivaram Kalyanakrishnan, and Rémi Munos. 2016. On the complexity of best-arm identification in multi-armed bandit models. In *Journal of Machine Learning Research*, Vol. 17. 1–42.

[10] Anastasia Koloskova, Nicolas Loizou, Pavel Krejći, Martin Jaggi, and Sebastian U Stich. 2020. Decentralized deep learning with arbitrary communication compression. In *International Conference on Learning Representations (ICLR)*.

[11] Peter Landgren, Vaibhav Srivastava, and Naomi Ehrich Leonard. 2016. Distributed cooperative decision-making in multiarmed bandits: Frequentist and Bayesian algorithms. *IEEE Trans. Automat. Control* 61, 6 (2016), 1578–1589.

[12] Arnab Mitra, Ashutosh Sinha, Ittai Bistritz, Yuval Peres, and John Tsitsiklis. 2021. Collaborative best arm identification with heterogeneous preferences. In *Advances in Neural Information Processing Systems (NeurIPS)*.

[13] Mauro Mobilia. 2003. Does a single zealot affect an infinite group of voters? *Physical Review Letters* 91, 2 (2003), 028701.

[14] Angelia Nedic and Alex Olshevsky. 2018. Network topology and communication-computation tradeoffs in decentralized optimization. *Proc. IEEE* 106, 5 (2018), 953–976.

[15] Reza Olfati-Saber, J. Alex Fax, and Richard M. Murray. 2007. Consensus and Cooperation in Networked Multi-Agent Systems. *Proc. IEEE* 95, 1 (2007), 215–233.

[16] Alex Olshevsky and John N Tsitsiklis. 2009. Convergence speed in distributed consensus and averaging. *SIAM Journal on Control and Optimization* 48, 1 (2009), 33–55.

[17] Abbas Rezazadeh, Huan Yu, Xiaodong Liang, and Jun Wu. 2022. Federated deep reinforcement learning for collaborative energy management with heterogeneous preferences. In *IEEE Transactions on Smart Grid*, Vol. 13. 109–121.

[18] Kaushik Sankararaman, Akshay Krishnamurthy, and Vasilis Syrgkanis. 2021. Decentralized exploration in multi-armed bandits. In *International Conference on Artificial Intelligence and Statistics (AISTATS)*.

[19] L. Zhang et al. 2018. Fully distributed policy consensus in multi-agent systems. *IEEE Trans. Automat. Control* 63, 11 (2018), 3812–3819.

[20] Ying Zhang and Longbo Wang. 2021. Decentralized multi-player bandits with heterogeneous rewards. In *International Conference on Machine Learning (ICML)*. 12506–12515.