SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

January 2020, Vol. 63 112202:1–112202:15
https://doi.org/10.1007/s11432-019-9890-5

Asymptotic properties of distributed social sampling algorithm

Qian LIU[1,2], Xingkang HE[3*] & Haitao FANG[1,2]

[1]*The Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;*
[2]*School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;*
[3]*ACCESS Linnaeus Centre, School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm SE-100 44, Sweden*

Received 18 January 2019/Revised 22 March 2019/Accepted 29 April 2019/Published online 23 December 2019

Abstract Social sampling is a novel randomized message passing protocol inspired by social communication for opinion formation in social networks. In a typical social sampling algorithm, each agent holds a sample from the empirical distribution of social opinions at initial time, and it collaborates with other agents in a distributed manner to estimate the initial empirical distribution by randomly sampling a message from current distribution estimate. In this paper, we focus on analyzing the theoretical properties of the distributed social sampling algorithm over random networks. First, we provide a framework based on stochastic approximation to study the asymptotic properties of the algorithm. Then, under mild conditions, we prove that the estimates of all agents converge to a common random distribution, which is composed of the initial empirical distribution and the accumulation of quantized error. Besides, by tuning algorithm parameters, we prove the strong consistency, namely, the distribution estimates of agents almost surely converge to the initial empirical distribution. Furthermore, the asymptotic normality of estimation error generated by distributed social sample algorithm is addressed. Finally, we provide a numerical simulation to validate the theoretical results of this paper.

Keywords social networks, opinion formation, social sampling, stochastic approximation, random networks, asymptotic normality

Citation Liu Q, He X K, Fang H T. Asymptotic properties of distributed social sampling algorithm. Sci China Inf Sci, 2020, 63(1): 112202, https://doi.org/10.1007/s11432-019-9890-5

1 Introduction

Social network analysis has attracted considerable interest in numerous different fields such as sociology, behaviorial science, economic, biology and cybernetics, because network perspective allows an effective leverage for modeling the complex dynamics induced by interpersonal interactions (see [1–4]). For a certain event or problem, individuals may hold different attitudes or behaviorial tendency initially, which then could be influenced by interaction or social influence. In social networks, the study of opinion formation is to model the fragmentation or merging of opinions among agents in a society. A large class of real world phenomena can be well interpreted with opinion dynamics, such as election forecasting [5], analysis of public opinions [6], and language evolution [7]. Over the past decades, with the development of network technology and increasing of social communication, more and more researchers have paid attention to the study of opinion formation in social networks as well as new approaches to distributed learning and estimation [8–15].

* Corresponding author (email: xingkang@kth.se)

The results in sociology on opinion formation are mainly based on empirical studies, which usually lack tools for quantitative analysis and numerical simulation of large-scale social groups. Thus, in recent years, the requirements for mathematical modeling and theoretical analysis are increasing, and network system appears naturally. Generally, in opinion dynamics, the communications between agents can be modeled by a network or a graph, whose edges represent channels through which one agent can share information with others. Naturally, the following issues are concerned: whether an equilibrium or consensus can be achieved via such social interactions? If the answer is yes, is this equilibrium influenced by network topology, initial opinions of agents or communication protocol? Otherwise, how can the opinion dynamics behave? There are quite a few studies considering the above problems. The early study [16] proposes a common paradigm called Friedkin and Johnson model, where agents are divided into stubborn agents and regular agents. In [17], the Friedkin and Johnson model is viewed as a randomized gossip algorithm inducing oscillations, which should be ergodic under some stable assumptions. Besides, a tractable communication model is developed in [18] to study the dynamics of belief formation and information aggregation. The authors propose asymptotic learning to describe that the fraction of agents can behave properly, and provide sufficient and necessary conditions to guarantee the asymptotic learning. A continuous-time opinion dynamic model with stochastic gossip process is proposed in [19] to investigate the generation of disagreement and fluctuation. It is shown that the society containing stubborn agents with different opinions keeps fluctuating in an ergodic manner. Ref. [20] deals with the situations that opinions are continuous variables while the opinions to be transmitted are discrete, which shows that in general the consensus cannot be controlled.

Most of existing results focus on modeling the dynamics of consensus, diversity or fluctuation in social networks, where opinions are represented as scaler variables. A nature extension is that how the agents can obtain a common knowledge of the global phenomenon, for example, the distribution of opinions across the whole society. In fact, probability distribution of opinions is a proper way to model the enormous opinions of agents on some certain subjects in a complex society, such as the election candidates they support or prefer. In what follows, we discuss the local reconstruction of the empirical distribution of initial opinions via social interactions. One should notice that the empirical distribution provides a sufficient statistic when the opinions of agents are independently sampled from an identical distribution. The aim of the agents is to estimate the discrete empirical distribution derived by the average of initial opinions through local interactions. Nevertheless, because of the limitation of communication cost, each agent cannot exchange their entire opinions completely. To deal with this problem, a communication scheme called social sampling is proposed [21]. The idea of this scheme is that each agent can only share a sample generated randomly by following its current distribution. The distributed social sampling algorithm given in [21] is a randomized approximation of consensus procedures, in which a group of agents aim to reach a common decision in a distributed way. Because the transferred message is quantized as an identical vector, the computation complexity is significantly reduced.

The applications of multi-agent and networked control to social network analysis will be a significant tendency. So far, there exist a wide range of results on networked consensus in terms of communication noise, time delay, network topology and so on [22–32]. In particular, consensus of network agents with noisy observation and fixed topologies has been discussed in [27], which establishes mean square consensus. It is worth noticing that with weaker conditions on network topology and measurement noise, strong consensus is achieved in [28] for the similar algorithm from [27]. Besides, consensus conditions for high-dimensional multi-agent systems with time-delay and additive or multiplicative measurement noise are explored in [30], while similar problem with time-varying topologies is investigated in [31], which cannot be applied to our problem directly because of the quantized noise and random network.

Actually the theoretical properties of distributed social sample algorithm in [21], such as the effects of network topology and the random sample process, have not been well investigated. Besides, one interesting theoretical problem is the asymptotic normality of estimation error. Ref. [33] considers such problem for certain cases of noisy communication in consensus schemes for scalars and finds a connection between network topology and covariance matrix of the error limitation distribution. However, such a result cannot be transferred to distributed social sampling algorithm because of the quantized error term induced by random sampling. We take advantage of the framework from [28] to ensure consensus addressed by the opinion formation problem. Compared with the analysis in [21], which describes the consensus conditions on the rearranged middle items, we extend the algorithm to random network and demonstrate how network topology and random sample protocol influence the convergence performance directly. In conclusion, we will complete the analysis and establish asymptotic normality for the social sampling scenario.

The main contributions of this paper are threefold:

(i) We provide a novel analysis framework based on stochastic approximation to study the asymptotic properties of the distributed social sampling algorithm over random networks. To ensure the convergence of the algorithm in an almost sure sense, we use the techniques of stochastic approximation [28], in which state space is decomposed into two parts: consensus part and vanishing part. Besides, some analysis methods provided in this paper can contribute to further related researches.

(ii) For consensus over random networks, the strong consensus is most desirable. Compared with [21], this is achieved in our work under the milder conditions on network structures and communication noise properties. We prove that the distribution estimates of agents reach consensus almost surely to the value related with true empirical distribution and the accumulation of quantized error. Besides, by tuning algorithm parameters, we prove the strong consistency, i.e., the distribution estimates of agents are almost surely convergent to the initial empirical distribution. On the other hand, unlike the fixed topology used in [21], the condition on network topology is relaxed to joint connectivity of mean digraphs for random networks.

(iii) We provide convergence rate of estimation of the social sampling protocol. Explicitly, we prove that the overall estimation errors of the algorithm are asymptotically normal with zero mean and known covariance matrix. The covariance matrix shows that how networks and quantized error influence the estimation performance of the distributed social sampling algorithm. Compared with [33], the conditions on communication noise in this work are more general.

The reminder of the paper is organized as follows. Section 2 provides some preliminary information about graph theory and the main problem we considered. In Section 3, we describe the social sampling protocol and the stochastic algorithm studied in this paper. Convergence analysis for the distributed algorithm is given in Section 4, while asymptotic properties are presented in Section 5. Section 6 shows a numerical simulation. Section 7 gives some concluding remarks.

Notations. Let $\boldsymbol{e}_i \in \mathbb{R}^M$ stand for the unit row vector whose i-th element equals to 1. $\mathbf{I}_{[\cdot]}$ denotes the indicator function and $\mathbf{1}$ stands for the proper dimensional column vector with elements all being 1. $\mathbf{I}_N$ denotes the N-dimension identity matrix. The superscript "T" represents the transpose. The abbreviation i.i.d. stands for independent identically distribution of random variables. $N(0, S)$ denotes the normal distribution with zero mean and covariance S. $\mathrm{E}[x]$ denotes the mathematical expectation of the stochastic variable x. Notation $\mathrm{diag}(\cdot)$ represents the diagonalization of scalar elements. $\mathbb{R}^n$ represents the n-dimension Euclidean space. Besides, we define $[N] = \{1, 2, \ldots, N\}$.

2 Preliminaries

2.1 Graph theory

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \boldsymbol{W})$ be a weighted digraph, where $\mathcal{V} = [N]$ is label set of N agents, $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the edge set. An ordered pair $(i, j) \in \mathcal{E}$ means that agent i can get information from agent j directly. The in-neighbor set is denoted by $\mathcal{N}_{\mathrm{in}}(i) = \{j \in \mathcal{V} |\, (i, j) \in \mathcal{E}\}$, while the out-neighbor set of agent i is denoted by $\mathcal{N}_{\mathrm{out}}(i) = \{j \in \mathcal{V} |\, (j, i) \in \mathcal{E}\}$. The graph is undirected if it is bidirectional; i.e., $(j, i) \in \mathcal{E}$ if and only if $(i, j) \in \mathcal{E}$. $\boldsymbol{W} = [w_{ij}] \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix of $\mathcal{G}$, where $w_{ij} > 0$ if $(i, j) \in \mathcal{E}$, and $w_{ij} = 0$ otherwise. The nonnegative matrix $\boldsymbol{W}$ is called row-wise stochastic if $\boldsymbol{W}\mathbf{1} = \mathbf{1}$, and is called column-wise stochastic if $\boldsymbol{W}^{\mathrm{T}}\mathbf{1} = \mathbf{1}$. We say $\boldsymbol{W}$ is doubly stochastic if it is both row-wise stochastic and column-wise stochastic.

For any $i, j \in [N]$, the in-degree of agent i is defined as $\deg_{\text{in}}(i) = \sum_{j=1}^{N} w_{ij}$ and the out-degree of agent i is defined as $\deg_{\text{out}}(i) = \sum_{j=1}^{N} w_{ji}$. We say $\mathcal{G}$ is a balanced digraph, if $\deg_{\text{in}}(i) = \deg_{\text{out}}(i)$ for any $i \in [N]$. The digraph $\mathcal{G}$ is strongly connected if for any pair $i, j \in \mathcal{V}$, there exists a directed sequence of nodes $i_1, i_2, \ldots, i_p \in \mathcal{V}$, such that $(i, i_1) \in \mathcal{E}, (i_1, i_2) \in \mathcal{E}, \ldots, (i_p, j) \in \mathcal{E}$.

The network topology in this paper is allowed to be time-varying; thus the weighted communication network at time k is denoted by $\mathcal{G}_k = (\mathcal{V}, \mathcal{E}_k, \boldsymbol{W}_k)$. The graph sequence $\{\mathcal{G}_k\}$ is called jointly connected, if there exists an integer $T > 0$, such that $(\mathcal{V}, \bigcup_{s=0}^{\mathrm{T}} \mathcal{E}_s)$ is strongly connected.

Besides, we introduce a definition to characterize the asymptotic behavior of the agents.

Definition 1 (Strong consensus [27])**.** The estimates of agents (i.e., $Q_{i,k}$) are said to reach strong consensus if there exists a random variable $\boldsymbol{q}^*$ such that, with probability 1 and for all $i \in [N]$, $\lim_{k\to\infty} Q_{i,k} = \boldsymbol{q}^*$. We also say that the estimates converge almost surely (a.s.).

2.2 Distributed learning of distributions

Consider a network of N agents. The communication relationship among agents is described by a sequence of directed graphs $\{\mathcal{G}_k\}$, where time is discrete and indexed by $k = \{0, 1, 2, \ldots\}$. Assume that opinions of agents on a certain event take values in a finite set, which is denoted by a label set $\chi = [M]$. At initial time $k = 0$, every agent has a single discrete sample $X_i \in \chi = [M]$ which represents its support opinion. The problem we consider here is that agents in a network need to learn the empirical distribution $\mathit{\Pi} \triangleq \{\mathit{\Pi}(x), x \in [M]\}$, where

$$\mathit{\Pi}(x) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{[X_i = x]} \boldsymbol{e}_x, \quad \forall x \in [M] = \{1, 2, \ldots, M\}. \tag{1}$$

It can be considered as the histogram of the initial distribution of the agent opinions over the network.

For illustration, we consider a motivating example. We wonder whether people in social network could know the exact approval ratings of M candidates over an electoral district through local communication. Assume every people has an opinion or preference on a certain candidate, which can be described as the initial sample X_i. Then the empirical distribution $\mathit{\Pi}$ in (1) demonstrates the approval ratings of all candidates. Distributed social sampling algorithm provides us a referenced answer to this question. Suppose that people in social networks will communicate with friends or neighbors so that they can get enough information about the approval ratings. In the traditional message protocol [10,11,13], the agents exchange their entire opinion histogram each time, which means that people will discuss the evaluations of every single candidate. However, this is not consistent with normal communication pattern, especially under the situation that there exist a number of candidates. In fact, people may only be willing to exchange their opinions about some of the candidates. A sample generated from the current estimate is transferred in social sampling protocol, which means that people simplify the communication process and exchanges information about one randomly selected candidate. The consensus analysis of this paper shows that individual in social network could get enough information about approval ratings of all candidates through the social sample communication.

3 Problem setup

Algorithm formulation. In this paper, we aim to estimate this histogram through a randomized algorithm called social sampling [21]. The algorithm is based on the sample generated from the current estimate $Q_{i,k}$ of the true distribution $\mathit{\Pi}$. At time k, each agent holds an internal estimate $Q_{i,k}$ of $\mathit{\Pi}$ with $Q_{i,0} = \boldsymbol{e}_{X_i}$. We treat $Q_{i,k}$ as a probability distribution of the elementary vectors $\{\boldsymbol{e}_m : m \in [M]\}$, so that $Q_{i,k}$ should be probability vector on $\chi = [M]$. Agent i could generate its message social sample $Y_{i,k}$ from the internal estimate $Q_{i,k}$ directly, or from a function of $Q_{i,k}$ denoted by $P_{i,k}$. Then agent i sends $Y_{i,k}$ to its out-neighbors and receives the in-neighbor messages $\{Y_{j,k} : j \in \mathcal{N}_i\}$. At each iteration, agent i uses the samples from its neighbors and current estimate to obtain the updated estimate $Q_{i,k+1}$.

More specifically, we assume $Y_{i,k} \in \mathcal{Y} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_M\}$, which can be viewed as a label of the opinion state space. So the opinion takes values from a finite, discrete value space. The random message $Y_{i,k} \in \mathcal{Y}$ of agent i at time k is generated according to the distribution $P_{i,k} \in \mathbb{P}(\mathcal{Y})$, which is a function of the internal estimate $Q_{i,k}$. More precisely, $P_{i,k}$ is an M-dimension row probability vector where the m-th element $P_{i,k}^m = \mathbb{P}(Y_{i,k} = \boldsymbol{e}_m)$.

Remark 1. We can choose $P_{i,k}$ properly and make it be a correction term associated to the internal estimate $Q_{i,k}$. For example, we set $P_{i,k} = 0$ when $Q_{i,k} < \alpha$, where α is a presetting bound. Under some complicated situations, such as the opinion space being extremely large or the histogram being far from uniform, the initial distribution is heavily concentrated on a few elements but still contains many elements with relatively low popularity. This kind of censoring can avoid inefficient communication. Of course, we can also choose $P_{i,k} = Q_{i,k}$ in some simple situations.

First we formulate the distributed social sampling algorithm for empirical distribution $\mathit{\Pi}$ as Algorithm 1. By designing update procedure like this, we can add some reasonable assumptions on the coefficients to guarantee that the internal estimate $Q_{i,k+1}$ is a probability vector on the opinion state space $\chi = [M]$ at any time for every $i \in [N]$.

Algorithm 1 Distributed social sampling algorithm

Step 1. Initialization: At initial time $k = 0$, let initial distribution estimate of agent $i \in [N]$ with sample data X_i be

$$Q_i(0) = \boldsymbol{e}_{X_i}, \tag{2}$$

where $X_i \in [M]$.

Step 2. Social sample: For agent i at time k, generate the random message $Y_{i,k} \in \mathcal{Y} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_M\}$ according to the distribution $P_{i,k}$,

$$\mathbb{P}(Y_{i,k} = \boldsymbol{e}_m) = P_{i,k}^m, \quad m \in [M], \tag{3}$$

where $P_{i,k}$ is an M-dimensional row probability vector which itself could be a function of the internal estimate $Q_{i,k}$.

Step 3. Consensus protocol: For agent i at time $k+1$, update the internal estimate $Q_{i,k+1}$ as follows:

$$Q_{i,k+1} = \left(1 - \delta_k a_{ii}^k\right) Q_{i,k} - \delta_k b_{ii}^k Y_{i,k} + \sum_{j \in \mathcal{N}_i(k)} \delta_k w_{ij}^k Y_{j,k}, \tag{4}$$

where a_{ii}^k, b_{ii}^k are communication coefficients subject to $a_{ii}^k \geqslant 0$, $b_{ii}^k \geqslant 0$, and $\mathcal{N}_i(k)$ represents neighbors of agent i at time k. $\boldsymbol{W}_k = [w_{ij}^k] \in \mathbb{R}^{N \times N}$ is the weighted adjacency matrix of the network topology and δ_k is the step size.

This paper focuses on solving the following two problems: (i) analyzing the conditions ensuring the convergence of distributed social sampling Algorithm 1 over random networks in the almost sure sense, and (ii) deriving the asymptotic normality of Algorithm 1 and characterizing the effect of random sampling protocol and network topology on the limit covariance matrix.

4 Consensus and consistency

In this section, we provide an analysis framework based on stochastic approximation to study the convergence of Algorithm 1. For analysis convenience, we rewrite the Algorithm 1 in a compact form. Define the social samples at time k as an NM-dimension column vector $\boldsymbol{Y}_k \triangleq (Y_{1,k}^{\mathrm{T}}, Y_{2,k}^{\mathrm{T}}, \ldots, Y_{N,k}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{NM}$, which is generated from the sampling function $\boldsymbol{P}_k \triangleq (P_{1,k}^{\mathrm{T}}, P_{2,k}^{\mathrm{T}}, \ldots, P_{N,k}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{NM}$. Similarly, we set $\boldsymbol{Q}_k \triangleq (Q_{1,k}^{\mathrm{T}}, Q_{2,k}^{\mathrm{T}}, \ldots, Q_{N,k}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{NM}$.

Let $\boldsymbol{A}_k \triangleq \mathrm{diag}\left(a_{11}^k, \ldots, a_{NN}^k\right)$, $\boldsymbol{B}_k \triangleq \mathrm{diag}\left(b_{11}^k, \ldots, b_{NN}^k\right)$; then we can rewrite (4) in a compact form:

$$\boldsymbol{Q}_{k+1} = \boldsymbol{Q}_k + \delta_k\{((\boldsymbol{W}_k - \boldsymbol{B}_k - \boldsymbol{A}_k) \otimes \mathbf{I}_M)\boldsymbol{Q}_k + ((\boldsymbol{W}_k - \boldsymbol{B}_k) \otimes \mathbf{I}_M)(\boldsymbol{Y}_k - \boldsymbol{Q}_k)\}, \tag{5}$$

where "$\otimes$" is the Kronecker product.

Suppose that the σ-algebra $\mathcal{F}_k \triangleq \sigma\{Q_{i,0}, \boldsymbol{W}_t, \boldsymbol{B}_t, 1 \leqslant i \leqslant N, 0 \leqslant t \leqslant k\}$ is a filtration of the basic probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Hence $\boldsymbol{Q}_k$ is measurable with respect to $\mathcal{F}_k$. Given the update rule in (5), the consensus of the opinion dynamics is equivalent to the convergence of linear stochastic approximation algorithms. The linear matrix $(\boldsymbol{W}_k-\boldsymbol{B}_k-\boldsymbol{A}_k)\otimes\mathbf{I}_M$ represents the effect of network topology at time k, while $((\boldsymbol{W}_k-\boldsymbol{B}_k)\otimes\mathbf{I}_M)(\boldsymbol{Y}_k-\boldsymbol{Q}_k)$ is the error item caused by the quantized data.

As shown in (5), the opinion formation process can be considered as a linear regression case of stochastic approximation. Next, we will analyze the recursive iteration in (5) with stochastic approximation. To begin with, the following assumptions are given.

A1.

$$\delta_k \xrightarrow[k\to\infty]{} 0,\ \delta_k>0,\ \sum_{k=0}^{\infty}\delta_k=\infty,\ \sum_{k=0}^{\infty}\delta_k^2<\infty,\ \text{and}\ \frac{1}{\delta_{k+1}}-\frac{1}{\delta_k}\xrightarrow[k\to\infty]{}\delta\geqslant 0.$$

A2. (i) $\{\boldsymbol{W}_k\}_{k\geqslant 0}$ is an independent random sequence with expectation denoted by $\bar{\boldsymbol{W}}_k=\mathrm{E}[\boldsymbol{W}_k]$ and the adjacency matrix $\boldsymbol{W}_k=[w_{ij}^k]$ is double stochastic. Besides, there exists a uniform bound $\bar{w}_{ij}^k>\tau>0$, $\forall k>0$ for all nonzero $\bar{w}_{ij}^k\neq 0$.

(ii) There is an integer $T>0$, such that the mean graph $\bar{\mathcal{G}}_k=(\mathcal{V},\mathcal{E}_k,\bar{\boldsymbol{W}}_k)$ generated by $\{\bar{\boldsymbol{W}}_k\}$ is jointly connected in the fixed period $[k,k+T]$; i.e., there exits an integer $T>0$, such that the graph $(\mathcal{V},\bigcup_{s=0}^{\mathrm{T}}E\{\mathcal{E}_{k+s}\})$, $\forall k\geqslant 1$ is strongly connected.

In addition, we need extra assumptions on the mixed coefficients and the social sampling protocol.

A3.

$$\|\boldsymbol{P}_k-\boldsymbol{Q}_k\|^2\xrightarrow[k\to\infty]{}0,\ \text{a.s.}.$$

A4. The communication coefficients a_{ii}^k and b_{ii}^k are chosen properly such that $a_{ii}^k+b_{ii}^k=1$ for any $k\geqslant 0$, i.e., $\boldsymbol{A}_k+\boldsymbol{B}_k=\mathbf{I}_N$.

For convenience of analysis, we arrange (5) as follows:

$$\begin{aligned}\boldsymbol{Q}_{k+1}=&\boldsymbol{Q}_k+\delta_k\{((\boldsymbol{W}_k-\mathbf{I}_N)\otimes\mathbf{I}_M)\boldsymbol{Q}_k+((\boldsymbol{W}_k-\boldsymbol{B}_k)\otimes\mathbf{I}_M)(\boldsymbol{Y}_k-\boldsymbol{Q}_k)\}\\=&\boldsymbol{Q}_k+\delta_k\{\left(\left(\bar{\boldsymbol{W}}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\boldsymbol{Q}_k+\left(\left(\bar{\boldsymbol{W}}_k-\boldsymbol{B}_k\right)\otimes\mathbf{I}_M\right)(\boldsymbol{P}_k-\boldsymbol{Q}_k)\\&+\left(\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}_k\right)\otimes\mathbf{I}_M\right)\boldsymbol{P}_k+((\boldsymbol{W}_k-\boldsymbol{B}_k)\otimes\mathbf{I}_M)(\boldsymbol{Y}_k-\boldsymbol{P}_k)\}.\end{aligned}\tag{6}$$

Define

$$\begin{cases}\bar{\boldsymbol{H}}_k\triangleq\left(\bar{\boldsymbol{W}}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M,\\ \boldsymbol{C}_k\triangleq\left(\left(\bar{\boldsymbol{W}}_k-\boldsymbol{B}_k\right)\otimes\mathbf{I}_M\right)(\boldsymbol{P}_k-\boldsymbol{Q}_k),\\ \boldsymbol{M}_k\triangleq((\boldsymbol{W}_k-\boldsymbol{B}_k)\otimes\mathbf{I}_M)(\boldsymbol{Y}_k-\boldsymbol{P}_k)+\left(\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}_k\right)\otimes\mathbf{I}_M\right)\boldsymbol{P}_k;\end{cases}\tag{7}$$

then we have

$$\boldsymbol{Q}_{k+1}=\boldsymbol{Q}_k+\delta_k\left(\bar{\boldsymbol{H}}_k\boldsymbol{Q}_k+\boldsymbol{C}_k+\boldsymbol{M}_k\right).\tag{8}$$

Remark 2. Condition A1 can be automatically satisfied if $\delta_k=\frac{a}{k^\delta}$ with $a>0$, $\delta\in(\frac{1}{2},1]$. In fact, we can pick $\boldsymbol{P}_k=\boldsymbol{Q}_k$; i.e., we generate social sample from internal estimate directly without censoring, which means $\boldsymbol{C}_k=0$. The double stochastic assumption on $\bar{\boldsymbol{W}}_k$ means that the mean graph should be balanced. The lower bound τ in A2 for the nonzero elements is used to guarantee the stability of linear matrix sequence $\{\boldsymbol{H}_k\}$, which is easily satisfied in the case that the network is switched over a finite number of network topologies.

Before presenting the consensus results for the iteration in (8), we provide the following lemma.

Lemma 1 ([28])**.** Let $\{H_t\}$ be $n\times n$-matrices with $\sup_t\|H_t\|<\infty$. Assume that there is an $n\times n$-matrix $U>0$ and an integer $K>0$ such that for $\forall\ t\geqslant 0$,

$$UH_t+H_t^{\mathrm{T}}U\leqslant 0\ \text{and}\ \sum_{s=t}^{t+K}\left(UH_s^{\mathrm{T}}+H_sU\right)\leqslant-\beta I_n,\quad\beta>0.$$

If step-size $\{\delta_k\}$ satisfies A1 and ω_t can be expressed as $\omega_t=\mu_t+\nu_t$ where

$$\sum_{t=1}^{\infty}\delta_t\mu_{t+1}<\infty\ \text{ and }\ \nu_t\xrightarrow[t\to\infty]{}0,\ \text{a.s.}.$$

Then, for an arbitrary initial value x_0, the sequence $\{x_t\}$ generated by $x_{t+1} = x_t + \delta_t (H_t x_t + \omega_t)$ converges to zero almost surely.

Note that, in expression (8), we have separated the quantized error into two parts: $\boldsymbol{C}_k$ is a censoring item associated to the difference between the internal matrix $\boldsymbol{Q}_k$ and the sampling matrix $\boldsymbol{P}_k$, and $\boldsymbol{M}_k$ is a martingale difference sequence, which will be demonstrated in the following lemma.

Lemma 2. $(\boldsymbol{M}_k, \mathcal{F}_k)$ is a martingale difference sequence under A2.

Proof. See the proof in Appendix A.

As claimed above, for the opinion dynamic consensus we only need to show the convergence of $\boldsymbol{Q}_k$ given by (8) almost surely. This is given by the following theorem. It is shown that all estimates of agents will achieve consensus to a common estimate based on empirical distribution.

Theorem 1 (Consensus)**.** Let $\{\boldsymbol{Q}_k\}$ be generated by the recursive iteration in (8). Under the conditions A1, A2, A3 and A4, we have

$$\lim_{k\to\infty} \boldsymbol{Q}_k = \boldsymbol{Q}^*, \quad \text{a.s.}, \tag{9}$$

where $\boldsymbol{Q}^* = \mathbf{1} \otimes q^*$. Explicitly, $\lim_{k\to\infty} Q_{i,k} = q^*$, a.s., where

$$q^* \triangleq \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}} \otimes \mathbf{I}_M\right)\boldsymbol{Q}_0 + \frac{1}{N}\sum_{k=0}^{\infty}\delta_k \left(\mathbf{1}^{\mathrm{T}}(\boldsymbol{W}_k - \boldsymbol{B}_k)\otimes \mathbf{I}_M\right)(\boldsymbol{Y}_k - \boldsymbol{Q}_k), \quad \forall i \in [N].$$

Furthermore, if $\|\boldsymbol{P}_k - \boldsymbol{Q}_k\| = O(\delta_k)$, then $q^* < \infty$, a.s..

Proof. The mixed-product property of Kronecker product, $(A \otimes B)(C \otimes D) = AC \otimes BD$, will be frequently used in the following.

Firstly, we write $\boldsymbol{Q}_k$ as a sum of a vector in the consensus space and a disagreement vector by orthogonal decomposition. Let

$$T \triangleq \begin{bmatrix} T_1 \\ \frac{1}{\sqrt{N}}\mathbf{1}^{\mathrm{T}} \end{bmatrix}$$

be an orthogonal matrix; then we have $T_1\mathbf{1} = \mathbf{0}$ and $T_1 T_1^{\mathrm{T}} = \mathbf{I}_{N-1}$. Set $\varGamma \triangleq \mathbf{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}$; then $T\varGamma = \left[\begin{smallmatrix} T_1 \\ \mathbf{0}\end{smallmatrix}\right]$. Pre-multiplying (5) by $T\varGamma \otimes \mathbf{I}_M$ yields

$$\left(\begin{bmatrix} T_1 \\ \mathbf{0}\end{bmatrix} \otimes \mathbf{I}_M\right)\boldsymbol{Q}_{k+1} = \left(\begin{bmatrix} T_1 \\ \mathbf{0}\end{bmatrix} \otimes \mathbf{I}_M\right)\boldsymbol{Q}_k + \delta_k \Bigg\{ \left(\begin{bmatrix} T_1 \\ \mathbf{0}\end{bmatrix} \otimes \mathbf{I}_M\right)\bar{\boldsymbol{H}}_k\boldsymbol{Q}_k \tag{10}$$

$$+ \left[\begin{pmatrix} T_1 \\ \mathbf{0}\end{pmatrix} \otimes \mathbf{I}_M\right]\boldsymbol{C}_k + \left(\begin{bmatrix} T_1 \\ \mathbf{0}\end{bmatrix} \otimes \mathbf{I}_M\right)\boldsymbol{M}_k\Bigg\}. \tag{11}$$

Setting $\xi_k \triangleq (T_1 \otimes \mathbf{I}_M)\boldsymbol{Q}_k$, we obtain

$$(T\varGamma \otimes \mathbf{I}_M)\boldsymbol{Q}_k = \left[\xi_k^{\mathrm{T}}, 0\right]^{\mathrm{T}}, \tag{12}$$

and

$$\begin{aligned} \xi_{k+1} =& \xi_k + \delta_k \left\{(T_1 \otimes \mathbf{I}_M)\left((\bar{\boldsymbol{W}}_k - \mathbf{I}_N)\otimes \mathbf{I}_M\right)\left(T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right)\xi_k + (T_1 \otimes \mathbf{I}_M)(\boldsymbol{C}_k + \boldsymbol{M}_k)\right\} \\ =& \xi_k + \delta_k \left\{\left(\left(T_1\left(\bar{\boldsymbol{W}}_k - \mathbf{I}_N\right)T_1^{\mathrm{T}}\right)\otimes \mathbf{I}_M\right)\xi_k + (T_1 \otimes \mathbf{I}_M)\boldsymbol{C}_k + (T_1 \otimes \mathbf{I}_M)\boldsymbol{M}_k\right\}. \end{aligned} \tag{13}$$

Set $\boldsymbol{F}_k \triangleq T_1\left(\bar{\boldsymbol{W}}_k - \mathbf{I}_N\right)T_1^{\mathrm{T}} = T_1\bar{\boldsymbol{W}}_k T_1^{\mathrm{T}} - \mathbf{I}_{N-1}$. To use Lemma 1, we need to verify the stability of matrix sequence $\{\boldsymbol{F}_k \otimes \mathbf{I}_M\}$. Because the adjacency matrix $\boldsymbol{W}_k$ is doubly stochastic, i.e., $\mathbf{1}^{\mathrm{T}}\boldsymbol{W}_k = \mathbf{1}^{\mathrm{T}}$ and $\boldsymbol{W}_k\mathbf{1} = \mathbf{1}$, then $\bar{\boldsymbol{W}}_k$ has the single largest eigenvalue 1 by Perron's Theorem [34]. Hence, $\frac{(\bar{\boldsymbol{W}}_k + \bar{\boldsymbol{W}}_k^{\mathrm{T}})}{2}$ is a symmetric stochastic matrix which has the largest eigenvalue 1, and the eigenvector associated with 1 is $\mathbf{1} \in \mathbb{R}^N$. Now, for any nonzero column vector $z \in \mathbb{R}^N$,

$$z^{\mathrm{T}}\bar{\boldsymbol{W}}_k z = z^{\mathrm{T}}\frac{\left(\bar{\boldsymbol{W}}_k + \bar{\boldsymbol{W}}_k^{\mathrm{T}}\right)}{2}z \leqslant z^{\mathrm{T}}z.$$

Moreover, for any nonzero $u \in \mathbb{R}^{N-1}$,

$$u\left(T_1 \bar{\boldsymbol{W}}_k T_1^{\mathrm{T}} - \mathbf{I}_{N-1}\right) u^{\mathrm{T}} = (uT_1)\, \bar{\boldsymbol{W}}_k\, (uT_1)^{\mathrm{T}} - (uT_1)\,(uT_1)^{\mathrm{T}} \leqslant 0. \tag{14}$$

Similarly,

$$u\left(T_1 \bar{\boldsymbol{W}}_k^{\mathrm{T}} T_1^{\mathrm{T}} - \mathbf{I}_{N-1}\right) u^{\mathrm{T}} = (uT_1)\, \bar{\boldsymbol{W}}_k^{\mathrm{T}}\, (uT_1)^{\mathrm{T}} - (uT_1)\,(uT_1)^{\mathrm{T}} \leqslant 0. \tag{15}$$

By (14) and (15), it is easy to obtain that

$$\boldsymbol{F}_k + \boldsymbol{F}_k^{\mathrm{T}} \leqslant 0. \tag{16}$$

Via the jointly connectivity of the network defined in A2, $\frac{1}{T+1}\sum_{s=t}^{t+T} \bar{\boldsymbol{W}}_s$ and $\frac{1}{T+1}\sum_{s=t}^{t+T} \bar{\boldsymbol{W}}_s^{\mathrm{T}}$ are irreducible double stochastic matrices. Then, for any $z \in \mathbb{R}^N$, it can be obtained that

$$z\left(\frac{1}{2\,(T+1)}\sum_{s=t}^{t+T}\left(\bar{\boldsymbol{W}}_s + \bar{\boldsymbol{W}}_s^{\mathrm{T}}\right)\right) z^{\mathrm{T}} - zz^{\mathrm{T}} \leqslant 0,$$

where the equality holds if and only if $z = c\mathbf{1}$. As $u^{\mathrm{T}}T_1\mathbf{1} = 0$, $u^{\mathrm{T}}T_1$ cannot be expressed as $c\mathbf{1}$ for any constant c. Consequently, for any nonzero $u \in \mathbb{R}^{N-1}$, the following strict inequality must hold:

$$\left(u^{\mathrm{T}}T_1\right)\left(\frac{1}{2\,(T+1)}\sum_{s=t}^{t+T}\left(\bar{\boldsymbol{W}}_s + \bar{\boldsymbol{W}}_s^{\mathrm{T}}\right)\right)\left(u^{\mathrm{T}}T_1\right)^{\mathrm{T}} < \left(u^{\mathrm{T}}T_1\right)\left(u^{\mathrm{T}}T_1\right)^{\mathrm{T}}.$$

Note that $T_1T_1^{\mathrm{T}} = \mathbf{I}_{N-1}$, which implies that for any nonzero $u \in \mathbb{R}^{N-1}$,

$$\frac{1}{2\,(T+1)}\sum_{s=t}^{t+T} u^{\mathrm{T}}\left(T_1\left(\bar{\boldsymbol{W}}_s + \bar{\boldsymbol{W}}_s^{\mathrm{T}}\right)T_1^{\mathrm{T}} - 2I_{N-1}\right)u < 0.$$

As a result, $\sum_{s=t}^{t+T}\left(\boldsymbol{F}_s + \boldsymbol{F}_s^{\mathrm{T}}\right) < 0$. In addition, with the assumption on the uniform lower bound in A1, there is a constant $\beta > 0$ such that

$$\sum_{s=t}^{t+T}\left(\boldsymbol{F}_s + \boldsymbol{F}_s^{\mathrm{T}}\right) \leqslant -\beta\mathbf{I}_{N-1}. \tag{17}$$

By (16) and (17), we have verified the conditions on linear matrix sequence $\{H_t\}$ in Lemma 1.

Now, we analyze the noise terms $\boldsymbol{C}_k$ and $\boldsymbol{M}_k$ in the iteration in (8).

(a) According to Lemma 2, $\boldsymbol{M}_k$ is a martingale difference sequence. We obtain $\sum_{k=0}^{\infty}\delta_k\boldsymbol{M}_k < \infty$, a.s., via the martingale convergence theorem [35].

(b) By assumption A3, on the chosen scheme of correct function $\boldsymbol{P}_k$, we have that

$$\begin{aligned}
\|\boldsymbol{C}_k\|^2 =& \boldsymbol{C}_k^{\mathrm{T}}\boldsymbol{C}_k = (\boldsymbol{P}_k - \boldsymbol{Q}_k)^{\mathrm{T}}\left(\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)^{\mathrm{T}} \otimes \mathbf{I}_M\right)\left(\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right) \otimes \mathbf{I}_M\right)(\boldsymbol{P}_k - \boldsymbol{Q}_k) \\
=& (\boldsymbol{P}_k - \boldsymbol{Q}_k)^{\mathrm{T}}\left[\left(\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)^{\mathrm{T}}\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)\right) \otimes \mathbf{I}_M\right](\boldsymbol{P}_k - \boldsymbol{Q}_k) \\
\leqslant& \lambda_{\max}\left\{\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)^{\mathrm{T}}\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right) \otimes \mathbf{I}_M\right\}\|\boldsymbol{P}_k - \boldsymbol{Q}_k\|^2 \xrightarrow[k\to\infty]{} 0, \quad \text{a.s..}
\end{aligned} \tag{18}$$

Consequently, $\lim_{k\to\infty}\boldsymbol{C}_k = 0$, a.s..

In summary, we have verified all conditions in Lemma 1. Then we can obtain

$$\lim_{k\to\infty}\left(T\varGamma \otimes \mathbf{I}_M\right)\boldsymbol{Q}_k = \lim_{k\to\infty}\left[\xi_k^{\mathrm{T}}, 0\right]^{\mathrm{T}} = 0, \quad \text{a.s..}$$

Note that T is an orthogonal matrix, so

$$\varGamma \otimes \mathbf{I}_M\boldsymbol{Q}_k \xrightarrow[k\to\infty]{} 0, \text{ a.s.,}$$

where

$$\begin{aligned}(\Gamma \otimes \mathbf{I}_M)\boldsymbol{Q}_k &= \left[\left(\mathbf{I}_N - \frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\right)\otimes \mathbf{I}_M\right]\boldsymbol{Q}_k = \boldsymbol{Q}_k - \left(\frac{1}{N}\mathbf{1}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_k \\ &= \boldsymbol{Q}_k - (\mathbf{1}\otimes \mathbf{I}_M)\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_k \\ &= \boldsymbol{Q}_k - (\mathbf{1}\otimes \mathbf{I}_M)\left(\sum_{j=1}^{N}\frac{1}{N}Q_{j,k}\right);\end{aligned}$$

i.e., $Q_{i,k} - \frac{1}{N}\sum_{j=1}^{N} Q_{j,k} \xrightarrow[k\to\infty]{} 0$, a.s. for every agent $i \in [N]$. Pre-multiplying update (8) by $\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M$ yields

$$\begin{aligned}\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_{k+1} &= \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_k + \delta_k\left\{\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)(\bar{\boldsymbol{W}}_k - \mathbf{I}_k)\otimes \mathbf{I}_M\boldsymbol{Q}_k \right.\\ &\quad \left. + \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)(\boldsymbol{C}_k + \boldsymbol{M}_k)\right\} \\ &= \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_k + \delta_k\left\{\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)(\boldsymbol{C}_k + \boldsymbol{M}_k)\right\}, \end{aligned} \tag{19}$$

because row sum of the matrix $\bar{\boldsymbol{W}}_k - \mathbf{I}_N$ is zero. Summing (19) from $k = 0$ to ∞ yields

$$\begin{aligned}\lim_{k\to\infty}\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_{k+1} &= \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_0 + \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\sum_{t=0}^{\infty}\delta_k(\boldsymbol{C}_k + \boldsymbol{M}_k) \\ &= \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_0 + \frac{1}{N}\sum_{t=0}^{\infty}\delta_k\left(\mathbf{1}^{\mathrm{T}}(\boldsymbol{W}_k - \boldsymbol{B}_k)\otimes \mathbf{I}_M\right)(\boldsymbol{Y}_k - \boldsymbol{Q}_k) \\ &\triangleq \boldsymbol{q}^*.\end{aligned}$$

Therefore, $Q_{i,k} \xrightarrow[k\to\infty]{} q^*$, a.s., $\forall i \in [N]$.

In the following, we will verify $\boldsymbol{q}^* < \infty$ if $\|\boldsymbol{P}_k - \boldsymbol{Q}_k\| = O(\delta_k)$. Let

$$L_1 \triangleq \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\sum_{k=1}^{\infty}\delta_k\boldsymbol{C}_k = \frac{1}{N}\sum_{k=1}^{\infty}\delta_k\left(\mathbf{1}^{\mathrm{T}}\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)\otimes \mathbf{I}_M\right)(\boldsymbol{P}_k - \boldsymbol{Q}_k), \tag{20}$$

$$L_2 \triangleq \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\sum_{k=1}^{\infty}\delta_k\boldsymbol{M}_k; \tag{21}$$

then, according to $\|\boldsymbol{P}_k - \boldsymbol{Q}_k\| = O(\delta_k)$ and under condition A1, we have

$$\begin{aligned}\|L_1\| &= \left\|\frac{1}{N}\sum_{k=1}^{\infty}\delta_k\left(\mathbf{1}^{\mathrm{T}}\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)\otimes \mathbf{I}_M\right)(\boldsymbol{P}_k - \boldsymbol{Q}_k)\right\| \\ &\leqslant \frac{1}{N}\sum_{k=1}^{\infty}\delta_k\left\|\left(\mathbf{1}^{\mathrm{T}}\left(\bar{\boldsymbol{W}}_k - \boldsymbol{B}_k\right)\otimes \mathbf{I}_M\right)\right\|\cdot\|(\boldsymbol{P}_k - \boldsymbol{Q}_k)\| \\ &\leqslant \frac{M}{N}\sum_{k=1}^{\infty}\delta_k^2 < \infty.\end{aligned} \tag{22}$$

In addition, we have known that $\{\boldsymbol{M}_k, \mathcal{F}_k\}$ is a martingale difference sequence [35] in Lemma 2. Thus $\sum_{k=0}^{\infty}\delta_k\boldsymbol{M}_k < \infty$; i.e., $L_2 < \infty$. In conclusion, $q^* < \infty$.

Corollary 1 (Strong consistency)**.** Let $\boldsymbol{B}_k \equiv \mathbf{I}_N$ and A1–A4 hold. Then $\{\boldsymbol{Q}_k\}$ generated by the iteration in (8) converges to the true distribution; i.e.,

$$\begin{aligned}&\lim_{k\to\infty}(Q_{i,k} - Q_{j,k}) = 0, \quad \forall i,j \in [N], \quad \text{a.s.}, \\ &\lim_{k\to\infty}Q_{i,k} = q^* \triangleq \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes \mathbf{I}_M\right)\boldsymbol{Q}_0, \quad \text{a.s.}.\end{aligned}$$

Remark 3. Compared with the undirected graph in [21], the joint connectivity of directed graph pointed at A2 in Theorem 1 is a weaker condition. Besides, we have derived strong consensus to a finite limit q^*, which is almost identical with the true distribution Π.

5 Asymptotic normality

In this section, we establish asymptotic normality for estimate error $\boldsymbol{Q}_k - \boldsymbol{Q}^*$ of the distributed social sampling algorithm. The main tool for asymptotic normality analysis is shown in the following lemma.

Lemma 3 (Theorem 3.3.1 in [36])**.** Let H_k and H be $l \times l$-matrices, $\{x_k\}$ be given by $x_{k+1} = x_k + \delta_k \left(H_k x_k + e_{k+1} + v_{k+1}\right)$ with an arbitrarily given initial value. Assume that the step-size δ_k satisfies A1 and the following conditions hold:

C1. $H_k \xrightarrow[k\to\infty]{} H$ and $H + \frac{\delta}{2}\mathbf{I}$ is stable with the constant δ given in A1;

C2. $v_k = o(\sqrt{\delta_k})$;

C3. $\{e_k, \mathcal{F}_k\}$ is a martingale difference sequence of l-dimension which satisfies

$$
\begin{aligned}
&\mathrm{E}\left(e_k|\mathcal{F}_{k-1}\right) = 0 \text{ and } \sup_k \mathrm{E}\left(\|e_k\|^2|\mathcal{F}_{k-1}\right) \leqslant \sigma \text{ with } \sigma \text{ being a constant,}\\
&\lim_{k\to\infty} \mathrm{E}\left(e_k e_k^{\mathrm{T}}|\mathcal{F}_{k-1}\right) = \lim_{k\to\infty} \mathrm{E}e_k e_k^{\mathrm{T}} \triangleq S_0, \quad \text{a.s.,}\\
&\lim_{N\to\infty} \sup_k \mathrm{E}\|e_k\|^2 \mathbf{I}_{[\|e_k\|>N]} = 0.
\end{aligned}
$$

Then

$$
\frac{x_k}{\sqrt{\delta_k}} \xrightarrow[k\to\infty]{d} N\left(0, S\right),
$$

where $S = \int_0^\infty \mathrm{e}^{\left(H+\frac{\delta}{2}\mathbf{I}\right)t} S_0 \mathrm{e}^{\left(H^{\mathrm{T}}+\frac{\delta}{2}\mathbf{I}\right)t}\mathrm{d}t$.

For the case that the root set of the observation function $f\left(x\right) = \bar{\boldsymbol{H}}_k\left(x\right)$ consists of a singleton zero, we consider $\{\xi_k\}$ which is defined in (13). It has been verified that $\lim_{k\to\infty}\xi_k = 0$, a.s.. Rewrite (13) as

$$
\xi_{k+1} = \xi_k + \delta_k \left\{\left(T_1 \otimes \mathbf{I}_M\right) \bar{\boldsymbol{H}}_k \left(T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right) \xi_k + \left(T_1 \otimes \mathbf{I}_M\right)\left(\boldsymbol{C}_k + \boldsymbol{M}_k\right)\right\}, \tag{23}
$$

where $\bar{\boldsymbol{H}}_k$, $\boldsymbol{C}_k$ and $\boldsymbol{M}_k$ are given by (7). It can be seen that ξ_k is updated by a linear stochastic approximation algorithm to approach the sought root zero. Furthermore, we can investigate the asymptotic properties of (23).

Before describing the convergent rate of iteration in (23), we require the following assumptions. We keep A1 unchanged, but strengthen A3 to A3′ and change A2 to A2′ as follows.

A2′. (i) $\{\boldsymbol{W}_k\}$ is an i.i.d. sequence.

(ii) The mean graph $\bar{\mathcal{G}}_k = (\mathcal{V}, \mathcal{E}_k, \bar{\boldsymbol{W}})$ generated by $\bar{\boldsymbol{W}} = \mathrm{E}\left[\boldsymbol{W}_k\right]$ is strongly connected and the adjacency matrix $\boldsymbol{W}_k$ is doubly stochastic.

A3′. $\|\boldsymbol{P}_k - \boldsymbol{Q}_k\| = o\left(\delta_k\right)$.

A4′. Choose $\boldsymbol{B}_k \equiv \boldsymbol{B}$, where $\boldsymbol{B}$ is a constant matrix.

A5. For sampling $\boldsymbol{Y}_k$ under distribution $\boldsymbol{P}_k$, we assume that $\varSigma$ is a constant matrix almost surely, where $\varSigma \triangleq \lim_{k\to\infty} \mathrm{E}[(\boldsymbol{Y}_k - \boldsymbol{P}_k)\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)^{\mathrm{T}}|\mathcal{F}_{k-1}]$.

Remark 4. We consider an example of A5 for illustration. Recall that the random message $\boldsymbol{Y}_k \triangleq (Y_{1,k}^{\mathrm{T}}, Y_{2,k}^{\mathrm{T}}, \ldots, Y_{N,k}^{\mathrm{T}})^{\mathrm{T}}$ is generated from the sampling function $\boldsymbol{P}_k \triangleq (P_{1,k}^{\mathrm{T}}, P_{2,k}^{\mathrm{T}}, \ldots, P_{N,k}^{\mathrm{T}})^{\mathrm{T}}$. Let $Y_{i,k} = \boldsymbol{e}_m$ with probability $P_{i,k}^m$, where $m \in [M]$. Then

$$
\varSigma_k \triangleq \mathrm{E}\left[\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)^{\mathrm{T}}|\mathcal{F}_{k-1}\right] = \begin{pmatrix} \varSigma_{1,k} & 0 & \cdots & 0 \\ 0 & \varSigma_{2,k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \varSigma_{N,k} \end{pmatrix}, \tag{24}
$$

where

$$\Sigma_{i,k} = \mathrm{E}\left[\left(Y_{i,k} - P_{i,k}\right)\left(Y_{i,k} - P_{i,k}\right)^{\mathrm{T}} | \mathcal{F}_{k-1}\right] = \begin{pmatrix} \left(1 - P_{i,k}^{1}\right) P_{i,k}^{1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \left(1 - P_{i,k}^{M}\right) P_{i,k}^{M} \end{pmatrix}.$$

From $\boldsymbol{P}_k - \boldsymbol{Q}_k \xrightarrow[k\to\infty]{} 0$ a.s., $\boldsymbol{Q}_k - \boldsymbol{Q}^* \xrightarrow[k\to\infty]{} 0$ a.s., we have

$$\Sigma_k \xrightarrow[k\to\infty]{} \Sigma = \left\{\left(\mathbf{I}_M - \mathrm{diag}\left(q^*\right)\right) \mathrm{diag}\left(q^*\right)\right\} \otimes \mathbf{I}_N.$$

The following lemma considers the martingale difference sequence part.

Lemma 4. Under A2′ and A4′, by choosing $\boldsymbol{B} \equiv \mathbf{I}_N$, we have that

$$\varepsilon_k \triangleq \left(T_1 \otimes \mathbf{I}_M\right) \boldsymbol{M}_k = \left(T_1\left(\boldsymbol{W}_k - \boldsymbol{B}\right) \otimes \mathbf{I}_M\right)\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right) + \left(T_1\left(\boldsymbol{W}_k - \bar{\boldsymbol{W}}\right) \otimes \mathbf{I}_M\right) \boldsymbol{P}_k \tag{25}$$

is a martingale difference sequence satisfying

$$\mathrm{E}\left(\varepsilon_k | \mathcal{F}_{k-1}\right) = 0, \quad \sup_k \mathrm{E}\left(\|\varepsilon_k\|^2 | \mathcal{F}_{k-1}\right) \leqslant \sigma \text{ with } \sigma \text{ being a constant,} \tag{26}$$

$$\lim_{N\to\infty} \sup_k \mathrm{E}\|\varepsilon_k\|^2 \mathbf{I}_{[\|\varepsilon_k\| > N]} = 0. \tag{27}$$

Proof. See the proof in Appendix B.

Now we can establish the asymptotic normality of the distributed social sampling algorithm (23).

Theorem 2 (Asymptotic normality)**.** Let A1, A2′, A3′, A4′ and A5 hold. Then $\xi_k = \left(T_1 \otimes \mathbf{I}_M\right) \boldsymbol{Q}_k$ is asymptotically normal; i.e., the distribution of $\frac{1}{\sqrt{\delta_k}}\xi_k$ converges to a normal distribution:

$$\frac{\xi_k}{\sqrt{\delta_k}} \xrightarrow[k\to\infty]{d} N\left(\mathbf{0}, \boldsymbol{S}\right), \tag{28}$$

where

$$\begin{aligned} &\boldsymbol{S} = \int_0^\infty \mathrm{e}^{\left(\bar{\boldsymbol{F}} + \frac{\delta}{2}\mathbf{I}_{(N-1)M}\right)t} \boldsymbol{S}_0 \mathrm{e}^{\left(\bar{\boldsymbol{F}} + \frac{\delta}{2}\mathbf{I}_{(N-1)M}\right)^{\mathrm{T}} t} \mathrm{d}t, \\ &\boldsymbol{S}_0 \triangleq \mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k - \mathbf{I}_N\right) \otimes \mathbf{I}_M\right)\right] \cdot \Sigma \cdot \mathrm{E}\left[\left(\boldsymbol{W}_k - \mathbf{I}_N\right)^{\mathrm{T}} T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right], \\ &\bar{\boldsymbol{F}} \triangleq \left(T_1\left(\bar{\boldsymbol{W}} - \mathbf{I}_N\right) T_1^{\mathrm{T}}\right) \otimes \mathbf{I}_M, \quad \Sigma \triangleq \lim_{k\to\infty} \mathrm{E}\left[\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)^{\mathrm{T}} | \mathcal{F}_{k-1}\right]. \end{aligned}$$

Proof. To use Lemma 3, we have to validate conditions C1 and C2.

First, we consider C1. Let $\bar{\boldsymbol{F}}_k \triangleq \left(T_1 \otimes \mathbf{I}_M\right) \bar{\boldsymbol{H}}_k \left(T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right)$, by A2′ on the weighted matrix $\boldsymbol{W}_k$, we have

$$\begin{aligned} \bar{\boldsymbol{F}} + \frac{\delta}{2}\mathbf{I}_{(N-1)M} &= \left(T_1 \otimes \mathbf{I}_M\right)\left(\left(\bar{\boldsymbol{W}} - \mathbf{I}_N\right) \otimes \mathbf{I}_M\right)\left(T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right) + \frac{\delta}{2}\mathbf{I}_{(N-1)M} \\ &= \left(T_1\left(\bar{\boldsymbol{W}} - \mathbf{I}_N\right) T_1^{\mathrm{T}}\right) \otimes \mathbf{I}_M + \frac{\delta}{2}\mathbf{I}_{(N-1)M}. \end{aligned}$$

Under assumption A2′, $\bar{\boldsymbol{W}} = \mathrm{E}\left[\boldsymbol{W}_k\right]$ is a stochastic matrix and its largest eigenvalue is 1 via Perron's Theorem [34]. Let the second largest eigenvalue of $\bar{\boldsymbol{W}}$ be λ_2. We can choose step-size δ_k properly such that $\lambda_2 < 1 - \frac{\delta}{2}$, where the linear matrix in (23) satisfies the stable assumption in C1.

Now we analyze C2 item by item. First $v_k \triangleq \left(T_1 \otimes \mathbf{I}_M\right) \boldsymbol{C}_k$, where $\boldsymbol{C}_k$ is defined in (7). We can obtain $v_k = o\left(\delta_k\right)$ according to A3′ and (18). According to (25), the martingale difference part of noise has

$$\begin{aligned} \varepsilon_k \varepsilon_k^{\mathrm{T}} = &\left(T_1\left(\boldsymbol{W}_k - \bar{\boldsymbol{W}}\right) \otimes \mathbf{I}_M\right) \boldsymbol{P}_k \boldsymbol{P}_k^{\mathrm{T}} \left(\left(\boldsymbol{W}_k - \bar{\boldsymbol{W}}\right)^{\mathrm{T}} T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right) \\ &+ \left(T_1\left(\boldsymbol{W}_k - \bar{\boldsymbol{W}}\right) \otimes \mathbf{I}_M\right) \boldsymbol{P}_k \left(\boldsymbol{Y}_k - \boldsymbol{P}_k\right)^{\mathrm{T}} \left(\left(\boldsymbol{W}_k - \mathbf{I}_N\right)^{\mathrm{T}} T_1^{\mathrm{T}} \otimes \mathbf{I}_M\right) \end{aligned}$$

$$
\begin{aligned}
&+\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\boldsymbol{P}_k^{\mathrm{T}}\left(\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right)\\
&+\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}\left(\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right)\\
\triangleq&S_1+S_2+S_3+S_4.
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\boldsymbol{P}_k=&\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{P}_k-\boldsymbol{Q}_k\right)+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Q}_k-\boldsymbol{Q}^*\right)\\
&+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\boldsymbol{Q}^*\\
=&\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{P}_k-\boldsymbol{Q}_k\right)+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Q}_k-\boldsymbol{Q}^*\right)\\
&+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\mathbf{1}\otimes\left(\frac{1}{N}\mathbf{1}^{\mathrm{T}}\otimes\mathbf{I}_M\boldsymbol{Q}_0\right)\right)\\
=&\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{P}_k-\boldsymbol{Q}_k\right)+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Q}_k-\boldsymbol{Q}^*\right)\\
&+\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\mathbf{1}\right)\otimes\left(\frac{1}{N}\mathbf{I}_M\left(\mathbf{1}^{\mathrm{T}}\otimes\mathbf{I}_M\right)\boldsymbol{Q}_0\right).
\end{aligned}
$$

Because $\boldsymbol{W}_k$ and $\bar{\boldsymbol{W}}$ all are double stochastic matrix and

$$
\|\boldsymbol{P}_k-\boldsymbol{Q}_k\|\xrightarrow[k\to\infty]{}0,\ \text{a.s.},\boldsymbol{Q}_k-\boldsymbol{Q}^*\xrightarrow[k\to\infty]{}0,\ \text{a.s.},
$$

we have

$$
\left(T_1\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes\mathbf{I}_M\right)\boldsymbol{P}_k\xrightarrow[k\to\infty]{}0,\quad\text{a.s..}\tag{29}
$$

Therefore, we can obtain $\lim_{k\to\infty}\mathrm{E}\left[S_1|\mathcal{F}_{k-1}\right]=0$, a.s., $\lim_{k\to\infty}\mathrm{E}\left[S_2|\mathcal{F}_{k-1}\right]=0$, a.s., and $\lim_{k\to\infty}\mathrm{E}\left[S_3|\mathcal{F}_{k-1}\right]=0$, a.s..

On the other hand, because of A5, $\varSigma\triangleq\lim_{k\to\infty}\mathrm{E}[(\boldsymbol{Y}_k-\boldsymbol{P}_k)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}|\mathcal{F}_{k-1}]$ is a constant matrix almost surely. Besides, $\boldsymbol{W}_k$ is independent with $\mathcal{F}_{k-1}$ in A2′; then

$$
\begin{aligned}
&\lim_{k\to\infty}\mathrm{E}\left[S_4|\mathcal{F}_{k-1}\right]\\
&=\lim_{k\to\infty}\mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}\left(\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right)|\mathcal{F}_{k-1}\right]\\
&=\lim_{k\to\infty}\mathrm{E}\left[\mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}\left(\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right)|\mathcal{F}_{k-1},\boldsymbol{W}_k\right]|\mathcal{F}_{k-1}\right]\\
&=\lim_{k\to\infty}\mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\mathrm{E}\left[\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}|\mathcal{F}_{k-1},\boldsymbol{W}_k\right]\left(\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right)|\mathcal{F}_{k-1}\right]\\
&=\lim_{k\to\infty}\mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)|\mathcal{F}_{k-1}\right]\mathrm{E}\left[\mathrm{E}\left[\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)^{\mathrm{T}}|\mathcal{F}_{k-1},\boldsymbol{W}_k\right]|\mathcal{F}_{k-1}\right]\\
&\quad\times\mathrm{E}\left[\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M|\mathcal{F}_{k-1}\right]\\
&=\mathrm{E}\left[\left(T_1\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes\mathbf{I}_M\right)\right]\cdot\varSigma\cdot\mathrm{E}\left[\left(\boldsymbol{W}_k-\mathbf{I}_N\right)^{\mathrm{T}}T_1^{\mathrm{T}}\otimes\mathbf{I}_M\right]\triangleq\boldsymbol{S}_0.
\end{aligned}
$$

According to Lemma 4, we have verified C2.

In summary, we have verified all conditions in Lemma 3; thus the conclusion of this theorem holds.

Corollary 2. Suppose A1, A2′, A3′, A4′ and A5 hold; then

$$
\frac{\boldsymbol{Q}_k-\boldsymbol{Q}^*}{\sqrt{\delta_k}}\xrightarrow[k\to\infty]{d}N\left(\mathbf{0},\widetilde{\boldsymbol{S}}\right),
$$

where $\widetilde{\boldsymbol{S}}\triangleq\left(T\varGamma\otimes\mathbf{I}_M\right)^{\mathrm{T}}\left(\begin{smallmatrix}\boldsymbol{S}&0\\0&0\end{smallmatrix}\right)\left(T\varGamma\otimes\mathbf{I}_M\right)$.

Remark 5. Theorem 2 and Corollary 2 establish that the error between estimates generated by recursive iteration in (5) and true empirical distribution is asymptotically normal, and the asymptotic covariance is characterized by network topology and quantized protocol. Our analysis results are more detailed and profound than that in [21], which only gives bounds on the expected squared error.

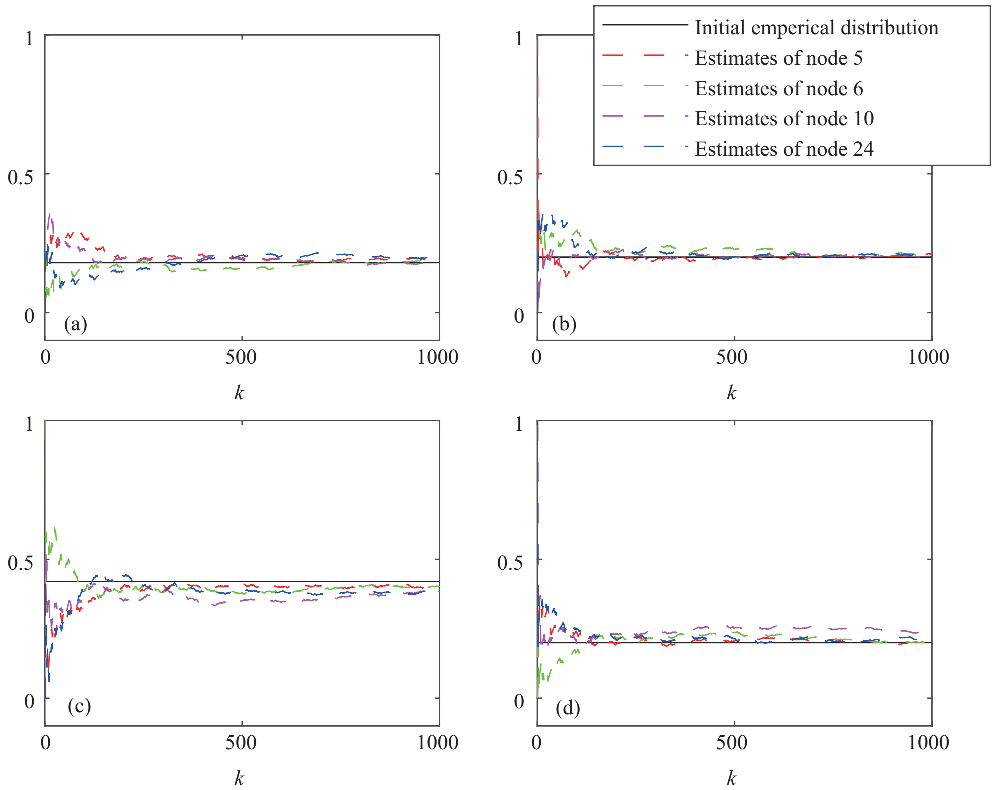

Figure 1 (Color online) Trace of estimate of $Q_{i,k}$ for $i \in [N]$ over every single opinion state $m \in [M]$ with $M = 4$. (a) $m = 1$, the estimate of the first element of true empirical distribution q^*; (b) $m = 2$, the estimate of the second element of q^*; (c) $m = 3$, the estimate of the third element of q^*; (d) $m = 4$, the estimate of the last element of q^*.

6 Numerical simulation

In this section, we provide a numerical simulation for the distribution social sampling algorithm considered in Algorithm 1. Let $N = 50$ with the underlying graph being fully connected. Each agent holds an initial opinion $Q_{i,0} = \boldsymbol{e}_{X_i}$, $\forall i \in [N]$, which is drawn i.i.d. from [0.2 0.3 0.4 0.1]. It means that agents' opinions are divided into $M = 4$ kinds, the probability of $\boldsymbol{e}_{X_i} = \boldsymbol{e}_1$ equals to 0.2, the probability of $\boldsymbol{e}_{X_i} = \boldsymbol{e}_2$ equals to 0.3, the probability of $\boldsymbol{e}_{X_i} = \boldsymbol{e}_3$ equals to 0.4, and the probability of $\boldsymbol{e}_{X_i} = \boldsymbol{e}_4$ equals to 0.1. The aim of distributed social sampling algorithm is to estimate the empirical distribution Π in (1). At each time k, agent i generates its random message $Y_{i,k} \in \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4\}$ based on the internal estimate $Q_{i,k}$ directly; i.e., we choose $P_{i,k} = Q_{i,k}$ and do not make corrections. Setting the mixed coefficients $a_{ii}^k = 0$, $b_{ii}^k = 1$, step size $\delta_k = \frac{1}{k^{0.75}}$, we update the internal estimate sequence $\{Q_{i,k}\}$ according to Algorithm 1.

The trace of estimate sequence of empirical distribution Π for some selected agents is shown in Figure 1, where each subgraph presents a state in $[M]$. As shown in Theorem 1, the estimated sequence generated by social sampling procedure converges to the true empirical distribution $q^* \triangleq \left(\frac{1}{N}\mathbf{1}^{\mathrm{T}} \otimes \mathbf{I}_M\right)\boldsymbol{Q}_0$. We have calculated Algorithm 1 for 1000 times independently. The histograms for each component of $(Q_{i,k} - q^*)/\sqrt{\delta_k}$ at $k = 500$ are shown in Figure 2. It is shown that the data fit the normal distribution well.

7 Conclusion

In this paper, convergence of distributed social sampling algorithm toward a common distribution has been established over random networks based on stochastic approximation. We have proved that the distribution estimates derived by agents' local interaction reached consensus almost sure to a value, which is related with the true empirical distribution and accumulation of quantized error. Furthermore, the error between estimates and true empirical distribution has been shown to be asymptotically normal

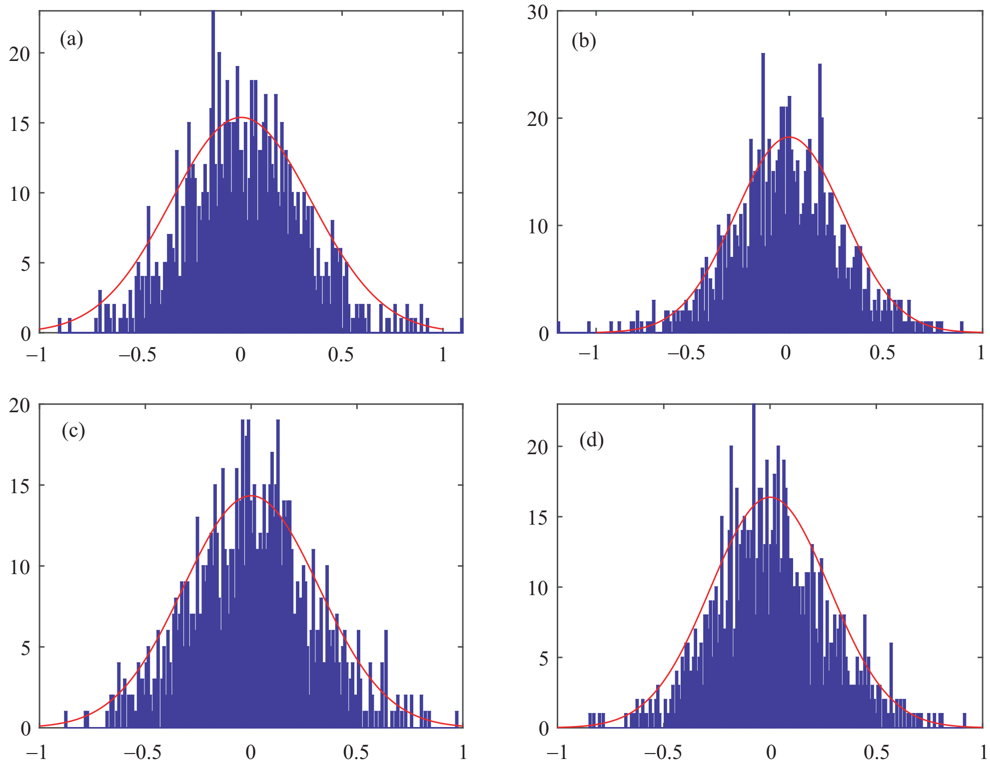

Figure 2 (Color online) Histogram and limit distribution for $\left(Q_{i,k}-q^*\right)/\sqrt{\delta_k}$ at $k=500$. (a) Node 5; (b) node 6; (c) node 10; (d) node 24.

with zero mean and known covariance, which is characterized by network topology and the social sampling protocol.

In fact, the randomized sample procedure is fairly general to be used in other problems, such as distributed optimization over large data sets. As the messages are quantized as identical vectors, the computation complexity is significantly reduced. In the future work, we will dig deeper about this random message passing protocol.

Acknowledgements This work was supported by National Key Research and Development Program of China (Grant No. 2016YFB0901900) and National Natural Science Foundation of China (Grant No. 61573345).

References

1 Anderson B D O, Ye M. Recent advances in the modelling and analysis of opinion dynamics on influence networks. Int J Autom Comput, 2019, 16: 129–149
2 Flache A, Mäs M, Feliciani T, et al. Models of social influence: towards the next frontiers. J Artif Soc Social Simulat, 2017, 20: 2
3 Proskurnikov A V, Tempo R. A tutorial on modeling and analysis of dynamic social networks. Part I. Annu Rev Control, 2017, 43: 65–79
4 Proskurnikov A V, Tempo R. A tutorial on modeling and analysis of dynamic social networks. Part II. Annu Rev Control, 2018, 45: 166–190
5 Holley R A, Liggett T M. Ergodic theorems for weakly interacting infinite systems and the voter model. Ann Probab, 1975, 3: 643–663
6 Acemoglu D, Dahleh M A, Lobel I, et al. Bayesian learning in social networks. Rev Economic Studies, 2011, 78: 1201–1236
7 Narayanan H, Niyogi P. Language evolution, coalescent processes, and the consensus problem on a social network. J Math Psychol, 2014, 61: 19–24
8 Xiao Y P, Li X X, Liu Y N, et al. Correlations multiplexing for link prediction in multidimensional network spaces. Sci China Inf Sci, 2018, 61: 112103
9 Friedkin N E, Proskurnikov A V, Tempo R, et al. Network science on belief system dynamics under logic constraints. Science, 2016, 354: 321–326
10 Hegselmann R, Krause U. Opinion dynamics and bounded confidence models, analysis, and simulation. J Artif Soc Social Simulat, 2002, 5: 1–33

11 Zhang J, Hong Y. Opinion evolution analysis for short-range and long-range Deffuant-Weisbuch models. Physica A-Stat Mech Appl, 2013, 392: 5289–5297
12 Pineda M, Toral R, Hernández-García E. Noisy continuous-opinion dynamics. J Stat Mech, 2009, 2009: P08001
13 Boyd S, Ghosh A, Prabhakar B, et al. Randomized gossip algorithms. IEEE Trans Inform Theor, 2006, 52: 2508–2530
14 Lou Y C, Strub M, Li D, et al. Reference point formation in social networks, wealth growth, and inequality. SSRN J, 2017. doi: 10.2139/ssrn.3013124
15 Frasca P, Ishii H, Ravazzi C, et al. Distributed randomized algorithms for opinion formation, centrality computation and power systems estimation: a tutorial overview. Eur J Control, 2015, 24: 2–13
16 Friedkin N E, Johnsen E C. Social influence networks and opinion change. Adv Group Process, 1999, 16: 1–29
17 Ravazzi C, Frasca P, Tempo R, et al. Ergodic randomized algorithms and dynamics over networks. IEEE Trans Control Netw Syst, 2015, 2: 78–87
18 Acemoglu D, Bimpikis K, Ozdaglar A. Dynamics of information exchange in endogenous social networks. Theor Economics, 2014, 9: 41–97
19 Acemoğlu D, Como G, Fagnani F, et al. Opinion fluctuations and disagreement in social networks. Math Ope Res, 2013, 38: 1–27
20 Ceragioli F, Frasca P. Consensus and disagreement: the role of quantized behaviors in opinion dynamics. SIAM J Control Opt, 2018, 56: 1058–1080
21 Sarwate A D, Javidi T. Distributed learning of distributions via social sampling. IEEE Trans Automat Contr, 2015, 60: 34–45
22 Degroot M H. Reaching a consensus. J Am Stat Assoc, 1974, 69: 118–121
23 Borkar V, Varaiya P P. Asymptotic agreement in distributed estimation. IEEE Trans Automat Contr, 1982, 27: 650–655
24 Tsitsiklis J N, Athans M. Convergence and asymptotic agreement in distributed decision problems. IEEE Trans Automat Contr, 1984, 29: 42–50
25 Olfati-Saber R, Murray R M. Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans Automat Contr, 2004, 49: 1520–1533
26 Kar S, Moura J M F. Distributed consensus algorithms in sensor networks with imperfect communication: link failures and channel noise. IEEE Trans Signal Process, 2009, 57: 355–369
27 Huang M, Manton J H. Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior. SIAM J Control Opt, 2009, 48: 134–161
28 Fang H, Chen H F, Wen L. On control of strong consensus for networked agents with noisy observations. J Syst Sci Complex, 2012, 25: 1–12
29 Leblanc H J, Zhang H, Koutsoukos X, et al. Resilient asymptotic consensus in robust networks. IEEE J Sel Areas Commun, 2013, 31: 766–781
30 Zong X F, Li T, Zhang J F. Consensus conditions of continuous-time multi-agent systems with time-delays and measurement noises. Automatica, 2019, 99: 412–419
31 Zong X F, Li T, Zhang J F. Consensus conditions of continuous-time multi-agent systems with additive and multiplicative measurement noises. SIAM J Control Opt, 2018, 56: 19–52
32 Wang Y H, Lin P, Hong Y G. Distributed regression estimation with incomplete data in multi-agent networks. Sci China Inf Sci, 2018, 61: 092202
33 Rajagopal R, Wainwright M J. Network-based consensus averaging with general noisy channels. IEEE Trans Signal Process, 2011, 59: 373–385
34 Meyer C D. Matrix Analysis and Applied Linear Algebra. Philadelphia: SIAM, 2000
35 Durrett R. Probability Theory and Examples. Cambridge: Cameberidge Press, 2010
36 Chen H F. Stochastic Approximation and Its Applications. New York: Kluwer Academic Publishers, 2003

Appendix A Proof of Lemma 2

According to the protocol of social sampling, we have $\mathrm{E}\left[Y_{i,k}\right] = \sum_{m=1}^{M} \mathbb{P}\left(Y_{i,k} = \boldsymbol{e}_m\right) = \sum_{m=1}^{M} \boldsymbol{e}_m P_{i,k}^{m} = P_{i,k}$. Thus, $\mathrm{E}\left[\boldsymbol{Y}_k\right] = \boldsymbol{P}_k$. Taking conditional expectation given $\mathcal{F}_{k-1}$ over both sides of (7), we obtain

$$\begin{aligned}\mathrm{E}\left[\boldsymbol{M}_k|\mathcal{F}_{k-1}\right] =& \mathrm{E}\left[\left(\left(\boldsymbol{W}_k-\boldsymbol{B}_k\right)\otimes \mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)|\mathcal{F}_{k-1}\right] + \mathrm{E}\left[\left(\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}_k\right)\otimes \mathbf{I}_M\right)\boldsymbol{P}_k|\mathcal{F}_{k-1}\right]\\ =& \mathrm{E}\left[\left(\boldsymbol{W}_k-\boldsymbol{B}_k\right)\otimes \mathbf{I}_M\right]\mathrm{E}\left[\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)|\mathcal{F}_{k-1}\right] + \mathrm{E}\left[\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}_k\right)\otimes \mathbf{I}_M\right]\boldsymbol{P}_k = 0.\end{aligned}$$

Hence, $\{\boldsymbol{M}_k, \mathcal{F}_k\}$ is a martingale difference sequence.

Appendix B Proof of Lemma 4

As we have verified that $(\boldsymbol{M}_k, \mathcal{F}_k)$ is a martingale difference sequence and T_1 is a constant matrix, $(\varepsilon_k, \mathcal{F}_k)$ given by (25) is also a martingale difference sequence. At first, we demonstrate the boundedness of ε_k. The random message $Y_{i,k} \in \mathcal{Y} = \{\mathbf{0}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_M\}$ of agent i at time k is generated according to the M-dimension row probability vector $P_{i,k} \in \mathbb{P}(\mathcal{Y})$. Besides, $\boldsymbol{W}_k$ is a double stochastic matrix by A2′; then we obtain

$$\begin{aligned}\|\varepsilon_k\| &\leqslant \|\left(T_1\otimes \mathbf{I}_M\right)\left(\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes \mathbf{I}_M\right)\left(\boldsymbol{Y}_k-\boldsymbol{P}_k\right)\| + \|\left(T_1\otimes \mathbf{I}_M\right)\left(\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes \mathbf{I}_M\right)\boldsymbol{P}_k\|\\ &\leqslant \|T_1\otimes \mathbf{I}_M\|\cdot\|\left(\boldsymbol{W}_k-\mathbf{I}_N\right)\otimes \mathbf{I}_M\|\cdot\|\boldsymbol{Y}_k-\boldsymbol{P}_k\| + \|T_1\otimes \mathbf{I}_M\|\cdot\|\left(\boldsymbol{W}_k-\bar{\boldsymbol{W}}\right)\otimes \mathbf{I}_M\|\cdot\|\boldsymbol{P}_k\|\\ &\leqslant cN^2M^2.\end{aligned}$$

Hence, the noise sequence ε_k is bounded. We can derive (26) and (27) directly.