Jacobi-Style Iteration for Distributed Submodular Maximization

Bin Du[†], Kun Qian[‡], Christian Claudel[‡], and Dengfeng Sun[†]

***Abstract*—This paper presents a novel Jacobi-style iteration algorithm for solving the problem of distributed submodular maximization, in which each agent determines its own strategy from a finite set so that the global submodular objective function is jointly maximized. Building on the multi-linear extension of the global submodular function, we expect to achieve the solution from a probabilistic, rather than deterministic, perspective, and thus transfer the considered problem from a discrete domain into a continuous domain. Since it is observed that an unbiased estimation of the gradient of multi-linear extension function can be obtained by sampling the agents' local decisions, a projected stochastic gradient algorithm is proposed to solve the problem. Our algorithm enables the distributed updates among all individual agents and is proved to asymptotically converge to a desirable equilibrium solution. Such an equilibrium solution is guaranteed to achieve at least $1/2$-suboptimal bound, which is comparable to the state-of-art in the literature. Moreover, we further enhance the proposed algorithm by handling the scenario in which agents' communication delays are present. The enhanced algorithmic framework admits a more realistic distributed implementation of our approach. Finally, a movie recommendation task is conducted on a real-world movie rating data set, to validate the numerical performance of the proposed algorithms.**

I. Introduction

In this paper, we focus on the distributed maximization of submodular functions, involving a network of I agents, which aims at cooperatively solving the following problem,

$$\begin{aligned} \text{maximize} \quad & F(a_1, a_2, \cdots, a_I) \\ \text{subject to} \quad & a_i \in \mathcal{A}_i,\ i = 1, 2, \cdots, I. \end{aligned} \tag{P}$$

In problem (P), each agent $i \in \mathcal{I} := \{1, 2, \cdots, I\}$ within the network is expected to select a desirable local strategy a_i from the private finite set $\mathcal{A}_i$, such that the common objective function $F : \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_I \to \mathbb{R}$ is jointly maximized. Here, we assume that the objective function is submodular and additionally monotone; see definitions in Section II. In fact, such a (distributed) monotone submodular maximization problem has gained increasing attention in recent years, primarily due to the fact that it can be widely adopted in numerous applications, including resource allocation [1], [2], sensor placement [3], [4], data summarization [5], [6], information gathering [7], [8], to name a few.

[†]Bin Du is Ph.D. student, and Dengfeng Sun is Associate Professor, with the School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, {du185,dsun}@purdue.edu

[‡]Kun Qian is Ph.D. student, and Christian Claudel is Assistant Professor, with the Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, Austin, TX 78712, {kunqian,christian.claudel}@utexas.edu

Bin Du and Kun Qian contributed equally to this manuscript.

While the submodular maximization problem has successfully found so many applications, solving problem (P) is known to be NP-hard [9], even from a standard centralized perspective. Due to this fact, the approximation methods which are able to guarantee suboptimal solutions are broadly studied in the literature [10]–[14]. Among these various approximation methods, the greedy algorithm [12]–[14] attracts the most attentions by researchers. The key idea of this greedy algorithm is to determine the single best strategy at each time, by maximizing the marginal gain of the submodular function. It is shown in [13] that an $1/2$-suboptimal solution can be guaranteed by the greedy algorithm, i.e., the obtained solution A^g is ensured to have $F(A^g) \geq 1/2 \cdot F(A^\star)$ where $A^\star$ represents the global optimal solution. In particular, when the considered submodular maximization problem has the constraints of some specific forms [12], [14], the suboptimal bound can be further improved to $1 - 1/e$ with e being the natural constant. Nevertheless, it should be remarked that the greedy algorithm is inherently a sequential updating scheme, since the single best strategy needs to be determined one by one. On this account, the greedy algorithm may not be useful or even infeasible in many applications, especially when a large number of agents are involved in the problem.

In order to address such a sequential updating issue, recent papers [15]–[17] have developed distributed variants of the greedy algorithm in which the best single strategies can be determined simultaneously in a parallel architecture. However, as shown in [16], when it comes to a distributed setting (with limited information), the suboptimal bound needs to be degraded from $1/2$ to $1/(1+\beta)$ where β is a constant related to the multi-agent network topology. Although the best network topology is further designed in [17] to enhance performance of the distributed algorithm, the fact that $\beta \geq 1$ makes the obtained solution worse than $1/2$-suboptimal in nature.

It is worthy to note that some other distributed approaches have also been devised to solve the problem of maximizing submodular functions. The authors in [18] develop a new distributed method while considering the application of multi-agent task assignment. It is proved that the obtained solution is at least $1/2$-suboptimal. Moreover, the distributed submodular maximization problem is studied in both discrete and continuous settings [19], and the proposed algorithms are guaranteed to converge asymptotically to the $1 - 1/e$ suboptimal bound. A similar algorithm is also developed in [20] which further improves convergence performance. We remark that the problem considered in our paper is significantly different from the ones in [18]–[20], where a separable structure of the global objective function is specifically assumed. Precisely, in their

problem setups [18]–[20], each agent (or task) maintains a local utility function, and the global objective is to optimize the summation of local functions. This inherently makes the problem easier to solve. For instance, given that the global gradient is computed by summing all local gradients due to the specific structure of the function, the desired global information can be achieved by a standard consensus procedure as in [19], or more directly, the technique of gradient tracking as in [20]. However, in our problem, such an idea is not applicable due to the generality of the global objective function.

To sum up, while the $1/2$-suboptimal bound is known to be the best result that one can achieve in general, there is no existing distributed algorithm yet, which guarantees such a bound when solving problem (P). Motivated by this, it is exactly the purpose of this paper to devise such a distributed algorithm. Our contributions are summarized as follows. A novel Jacobi-style algorithm is proposed for solving the problem of distributed submodular maximization. Unlike existing works which are based on the greedy algorithm, we start from a probabilistic perspective, build on the multi-linear extension of the submodular function, and eventually transfer the considered problem from a discrete domain into a continuous domain. By leveraging the fact that an unbiased estimation of the gradient can be achieved by simply sampling the agents' local decisions, we develop a projected stochastic gradient algorithm. It is proved that our algorithm converges to an equilibrium solution which is guaranteed to be at least $1/2$-suboptimal. In addition, by handling the scenario where communication delays are present among agents, we further enhance the proposed algorithm to be implementable in a more realistic distributed architecture. The same convergence performance is proved for the enhanced distributed algorithm. Finally, the movie recommendation task is conducted on a real-world movie rating data set, and the simulation results validates the effectiveness of our algorithms.

The remainder of this paper is organized as follows. Sec. II formally defines the considered distributed submodular maximization problem. Sec. III develops the projected stochastic gradient algorithm, and Sec. IV further enhances the proposed algorithm by dealing with the communication delays. Numerical simulations are presented in Sec. V. Lastly, Sec. VI concludes this paper. For the reader's convenience, the proofs of propositions and theorems are provided in Appendix.

II. Problem Statement

Let us first formalize the considered distributed submodular maximization problem. For the sake of notational simplicity, we here assume that each agent's finite set $\mathcal{A}_i$ has the same size K, i.e., $|\mathcal{A}_i| = K, \forall i \in \mathcal{I}$. In addition, we stack all agents' local strategies as a vector $A = [a_1, a_2, \cdots, a_I]^\top \in \mathcal{A}$, where the entire searching space $\mathcal{A}$ is the Cartesian product of $\mathcal{A}_i$'s, i.e., $\mathcal{A} := \prod_{i\in\mathcal{I}} \mathcal{A}_i$ and $|\mathcal{A}| = K^I$. Based on the set of collected strategies, the objective function in problem (P) can be succinctly written as $F(A)$. Note that we allow each agent to choose the empty set as its strategy, i.e., $a_i = \emptyset$; and in particular, let $F(\emptyset) = 0$. With slight abuse of notations, we say that the set of strategies A' is contained in A, denoted as $A' \subseteq A$, if some component a'_i in A' has $a'_i = \emptyset$ and other components have $a'_j = a_j$, $\forall j \neq i$. In this case, we also denote $A = A' \cup \{a_i\}$. Furthermore, we restrict the objective function $F(A)$ to satisfy the following assumption.

Assumption 1: The function $F(A)$ is assumed to satisfy:

(A.1) (*Monotone*) If two sets A' and A have $A' \subseteq A$, then it implies $F(A') \leq F(A)$.

(A.2) (*Submodular*) If two sets A' and A have $A' \subseteq A$, then $F(A' \cup \{a\}) - F(A') \geq F(A \cup \{a\}) - F(A)$ for any a.

Clearly, the ultimate goal of problem (P) is to find the group of optimal strategies $A^\star = [a_1^\star, a_2^\star, \cdots, a_I^\star]^\top \in \mathcal{A}$ such that $F(A^\star)$ gives the maximal function value against all other possible groups of strategies. However, as mentioned before in Section I, achieving such a goal is NP-hard in general. The challenges mainly come from the following two aspects. First, the finite set $\mathcal{A}_i$ from which the agent chooses its strategy is inherently discrete. Thus, the well-developed techniques of continuous optimization cannot be adopted for solving the problem. Although the bright side of this fact is that one can apply some search-tree based approaches since the set $\mathcal{A}_i$ is anyway finite, the computational demand during such a searching procedure is often costly or even infeasible, especially when each individual searching space $\mathcal{A}_i$ is large-scale. This is also related to the second aspect of the challenges. Note that the objective function $F(A)$ can be evaluated only when all individual agents have decided their own strategies. That is to say, the function $F(A)$ mixes the decisions of each agent within the network. In this sense, the size of the entire searching space $\mathcal{A}$ also grows exponentially with respect to the number of agents. This undoubtedly prohibits the idea of using searching procedures to find the joint optimal strategies $A^\star$, when a large number of agents are involved in the problem.

In order to address the above two challenges, in this paper, our ideas are: 1) utilizing the multi-linear extension of the function $F(A)$; see Section III-A, and transferring the considered problem into a continuous domain, so that the techniques of continuous optimization can be exploited; and 2) developing the Jacobi-style iteration to decompose the mixing of individual agents' decisions. In particular, the algorithmic framework of our Jacobi-style iteration for each agent i can be abstracted as the following mapping $\mathcal{M}_i : \mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_I \to \mathcal{A}_i$ such that

$$a_i^+ = \mathcal{M}_i\Big(a_i^-, A_{-i}^r\Big). \tag{1}$$

Here, a_i^- is the i-th agent's previous decision of the desired strategy; $A_{-i}^r = [a_1^r, \cdots, a_{i-1}^r, a_{i+1}^r \cdots a_I^r]^\top \in \mathcal{A}_{-i}$ is the collection of the received decisions which have been made by the other agents $j \neq i$; and a_i^+ is the i-th agent's updated decision based on the previous a_i^- and the received information a_j^r's. It should be emphasized that, under such a framework (1), each agent only needs to take charge in its own decision, by receiving the information a_j^r from other agents. Thus, the computational complexity is expected to be primarily reduced, compared to the aforementioned searching procedure. In addition, another advantage of the framework is that individual agents can perform the update of decisions simultaneously, so that the overall processing time can be further saved. However,

it is also worthy to note that there are two potential issues regarding the framework. First, the mapping $\mathcal{M}_i$ suggests that an instantaneous all-to-all communication is required, i.e., each agent i needs to communicate with all other agents to receive the most updated information a_j^r's. Second, it is not clear that what kind of solution will be produced by the iteration (1). For the first concern, we here remark that the communication requirement will be eliminated later on; see Section IV, so that our algorithm is implementable in a distributed architecture with communication delays. For the second one, we are interested in finding an equilibrium solution A^e which is formally defined as below. Interestingly, it can be proved that such an equilibrium is guaranteed to be at least $1/2$-suboptimal which is comparable to the state-of-art in the literature; see Section I. Before defining the equilibrium A^e, let us first introduce another assumption related to the objective function $F(A)$.

Assumption 2 (Maximum Distinguishable): It is assumed that, once other agents $j \neq i$ have decided their strategies A_{-i}, the i-th agent's best strategy a_i, which gives the maximum function value $F(a_i; A_{-i})$, is unique.

Here, to concentrate on the effect of strategy a_i on the function value, $F(A)$ is expressed as the specific form $F(a_i; A_{-i})$. We note that the above Assumption 2 is easily satisfied in many applications, especially when some certain randomness is involved in the function values. In particular, we denote $\Delta^{\max}$ the maximum discrepancy of function values between two individual strategies a_i and a_i' when other strategies A_{-i} have been fixed, i.e.,

$$\Delta^{\max} = \max_{A_{-i} \in \mathcal{A}_{-i}} \Big\{ \max_{a_i, a_i' \in \mathcal{A}_i} \big\{ F(a_i; A_{-i}) - F(a_i'; A_{-i}) \big\} \Big\}. \quad (2)$$

It is trivial to see that $\Delta^{\max}$ has to be strictly greater than zero due to Assumption 2. Next, we formalize the definition of the equilibrium A^e.

Definition 1: A solution $A^e = [a_1^e, a_2^e, \cdots, a_I^e]^\top \in \mathcal{A}$ is said to be the equilibrium to problem (P), if and only if it satisfies the following condition:

$$a_i^e = \arg\max_{a_i \in \mathcal{A}_i} \; F(a_1^e, \cdots, a_{i-1}^e, a_i, a_{i+1}^e, \cdots, a_I^e), \; i \in \mathcal{I}. \quad (3)$$

Remark that the uniqueness of a_i^e in (3) is guaranteed by the maximum distinguishable assumption of the objective function $F(A)$. Moreover, by the definition of the equilibrium solution, it is clear that A^e is not unique for problem (P) and $A^\star$ is just a specific equilibrium which has the maximal function value. Next, we show, by the following proposition, that any equilibrium A^e satisfying (3) must be at least $1/2$-suboptimal to our problem.

Proposition 1: Suppose that the function $F(A)$ satisfies the conditions in Assumption 1. Let $A^\star$ be an optimal solution to problem (P) and A^e be an equilibrium solution following Definition 1, then it holds that

$$F(A^e) \geq \frac{1}{2} \cdot F(A^\star). \quad (4)$$

Proof: See Appendix A. ∎

III. Stochastic Gradient Based Method

To solve for the equilibrium solution A^e, we develop a stochastic gradient based solution method in this section. As an important building block of our method, the multi-linear extension of the function $F(A)$ is first introduced.

A. Multi-Linear Extension

Let us recall that $A = [a_1, a_2, \cdots, a_I]^\top$ is the set of I local strategies in which each a_i is chosen from the finite set $\mathcal{A}_i$. Now, instead of expecting the individual agents to seek the desired deterministic strategies from $\mathcal{A}_i$'s, we assign each agent a discrete probability distribution $\mathbf{p}_i = [p_i(a_i)]_{a_i \in \mathcal{A}_i} \in [0,1]^K$, where $p_i(a_i)$ represents the probability of choosing a_i as the i-th agent's strategy. On this account, we define the multi-linear extension of $F(A)$ as the function $f(P) : [0,1]^{I \cdot K} \to \mathbb{R}$,

$$f(P) = \sum_{a_1 \in \mathcal{A}_1} p_1(a_1) \sum_{a_2 \in \mathcal{A}_2} p_2(a_2) \cdots \sum_{a_I \in \mathcal{A}_I} p_I(a_I) \cdot F(A), \quad (5)$$

where the argument P is a compact vector which stacks all $\mathbf{p}_i$'s, i.e., $P = [\mathbf{p}_1^\top, \mathbf{p}_2^\top, \cdots, \mathbf{p}_I^\top]^\top \in [0,1]^{I \cdot K}$. We shall emphasize that a core property of the multi-linear extension (5) is its natural connection to the original function $F(A)$. That is,

$$f(P) = \mathbb{E}_{\tilde{\mathbf{a}}_i \sim \mathbf{p}_i, i \in \mathcal{I}} \big[F(\tilde{\mathbf{A}}) \big], \quad (6)$$

where the expectation is taken from $\tilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2, \cdots, \tilde{\mathbf{a}}_I]^\top$ and each $\tilde{\mathbf{a}}_i$ is an independent random variable following the discrete distribution $\mathbf{p}_i$. Moreover, consider that each $\mathbf{p}_i$ has to be subject to $p_i(a_i) \geq 0$ and $\mathbf{1}^\top \mathbf{p}_i = 1$, let us express those constraints as the following probability simplex $\mathcal{S}$, i.e.,

$$\mathbf{p}_i \in \mathcal{S} := \Big\{ \mathbf{p} \,|\, \mathbf{1}^\top \mathbf{p} = 1, \, \mathbf{p} \in [0,1]^K \Big\}. \quad (7)$$

In particular, we say $\mathbf{p}_i$ is a vertex of the simplex $\mathcal{S}$ if it has $\|\mathbf{p}_i\|_\infty = 1$. i.e., there is exactly one component $p_i(a_i)$ which equals one and all others are zeros.

With the help of this multi-linear extension function $f(P)$, our goal now becomes to seek the desired probability distribution $\mathbf{p}_i$'s such that $f(P)$ is optimized. Therefore, the submodular maximization problem can be equivalently written as,

$$\begin{aligned} \text{maximize} \quad & f(P) = \mathbb{E}_{\tilde{\mathbf{a}}_i \sim \mathbf{p}_i, i \in \mathcal{I}} \big[F(\tilde{\mathbf{A}}) \big] \\ \text{subject to} \quad & \mathbf{p}_i \in \mathcal{S}, \; i \in \mathcal{I}. \end{aligned} \quad (8)$$

Recall that, in the original problem (P), we are interested in seeking the equilibrium solution A^e which is defined by Definition 1. Now, following the same path, we introduce a similar equilibrium solution P^e to problem (8), based on the defined A^e.

Definition 2: A solution $P^e = [\mathbf{p}_1^{e\top}, \mathbf{p}_2^{e\top}, \cdots, \mathbf{p}_I^{e\top}]^\top$ where each $\mathbf{p}_i^e$ is a vertex of the probability simplex, i.e., there exists a_i^e such that $p_i^e(a_i^e) = 1$ and $p_i^e(a_i) = 0$ for $\forall a_i \neq a_i^e$, is said to be an equilibrium, if and only if $A^e = [a_1^e, a_2^e, \cdots, a_I^e]^\top$ is an equilibrium to problem (P).

In fact, combining the above Definition 1 and 2 together establishes the equivalence between problem (P) and (8). In other words, the desired equilibrium A^e to problem (P) can be easily resulted from the solution P^e by solving problem (8). Next, we develop the projected stochastic gradient algorithm to solve for the equilibrium solution P^e.

B. Projected Stochastic Gradient Algorithm

Before proceeding to the development of our algorithm, let us first investigate the gradient of function $\nabla f(P)$. Recall that the function $f(P)$ is a multi-linear extension of $F(A)$ and can be expressed as (5), thus the gradient of $f(P)$ with respect to each single component $p_i(a_i)$ can be represented as

$$\nabla_{p_i(a_i)} f(P) = \sum_{a_1 \in \mathcal{A}_1} p_1(a_1) \cdots \sum_{a_{i-1} \in \mathcal{A}_{i-1}} p_{i-1}(a_{i-1}) \sum_{a_{i+1} \in \mathcal{A}_{i+1}} p_{i+1}(a_{i+1}) \cdots \sum_{a_I \in \mathcal{A}_I} p_I(a_I) \cdot F(a_i; A_{-i}). \tag{9}$$

Since $F(A)$ has its expectation interpretation as shown in (6), the gradient $\nabla_{p_i(a_i)} f(P)$ can be also expressed as the following expectation form,

$$\nabla_{p_i(a_i)} f(P) = \mathbb{E}_{\tilde{\mathbf{a}}_j \sim \mathbf{p}_j, j \neq i} \big[F(a_i; \tilde{\mathbf{A}}_{-i}) \big]. \tag{10}$$

We remark that, to evaluate the gradient $\nabla_{p_i(a_i)} f(P)$ for the i-th agent in the form of (9), it is required to sum all possibilities that are governed by the probability distribution $\mathbf{p}_j$'s for all other agents $j \neq i$. However, due to its expectation form (10), a key observation here is that an unbiased estimation of the gradient can be obtained by sampling the strategies a_j based on $\mathbf{p}_j$ for $\forall j \neq i$. In this sense, we call $\nabla_{p_i(a_i)} f(P)$ the full gradient which is computed by (9), and meanwhile denote the following $\nabla_{p_i(a_i)} \hat{f}(P)$ as the sampled stochastic gradient with the sample-size $M \in \mathbb{N}_+$,

$$\nabla_{p_i(a_i)} \hat{f}(P) = \frac{1}{M} \cdot \sum_{s=1}^{M} F(a_i; \hat{A}_{-i}^s), \tag{11}$$

where $\hat{A}_{-i}^s = [\hat{a}_1^s, \cdots, \hat{a}_{i-1}^s, \hat{a}_{i+1}^s, \cdots, \hat{a}_I^s]^\top$ and each $\hat{a}_j^s, j \neq i$ is the independent and identically distributed (*i.i.d.*) sampled strategy based on the probability distribution $\mathbf{p}_j$. Taking advantage of this stochastic gradient $\nabla_{p_i(a_i)} \hat{f}(P)$, our projected stochastic gradient algorithm performs the following iteration, with index $k \in \mathbb{N}_+$,

$$\mathbf{p}_i^{k+1} = \Pi_{\mathcal{S}} \big(\mathbf{p}_i^k + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}(P^k) \big), \ \forall i \in \mathcal{I}. \tag{12}$$

In (12), P^k is the collection of $\mathbf{p}_i^k$'s for all agents at the k-th iteration, i.e., $P^k = [\mathbf{p}_1^{k\top}, \mathbf{p}_2^{k\top}, \cdots, \mathbf{p}_I^{k\top}]^\top$; $\nabla_{\mathbf{p}_i} \hat{f}(P^k)$ is a vector which stacks the stochastic gradients for all $a_i \in \mathcal{A}_i$, i.e., $\nabla_{\mathbf{p}_i} \hat{f}(P^k) = [\nabla_{p_i(a_i)} \hat{f}(P^k)]_{a_i \in \mathcal{A}_i}$; $\gamma > 0$ is a constant step-size; and the operator $\Pi_{\mathcal{S}}(\cdot) : \mathbb{R}^K \to \mathcal{S}$ defines the projection on the probability simplex $\mathcal{S}$, i.e.,

$$\Pi_{\mathcal{S}}(\mathbf{p}) := \arg\min_{\mathbf{x} \in \mathcal{S}} \ \|\mathbf{x} - \mathbf{p}\|^2. \tag{13}$$

Our ultimate goal here is to drive the sequence of P^k generated by the iteration (12) to the desired equilibrium solution P^e as defined in Definition 2. Before proceeding to the convergence analysis of our algorithm, let us first show, by the following proposition, an alternative way to characterize the equilibrium solution P^e.

Proposition 2: Under Assumption 2, the probability distribution P^e is an equilibrium solution to problem (8); see Definition 2, if and only if the following condition is satisfied,

$$\mathbb{E}\Big[\big\| \mathbf{p}_i^e - \Pi_{\mathcal{S}} \big(\mathbf{p}_i^e + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}(P^e) \big) \big\|^2 \Big] = 0, \ \forall i \in \mathcal{I}. \tag{14}$$

With the help of the above proposition, we are now in the position to analyze the convergence of our projected stochastic gradient algorithm.

Theorem 1: Suppose that Assumption 2 is satisfied, and let $\{P^k\}_{k \in \mathbb{N}_+}$ be the sequence generated by the iteration (12) with a small enough constant step-size γ and a large enough constant sample-size M. Then, it holds that

$$\lim_{k \to \infty} \mathbb{E}\Big[\big\| \mathbf{p}_i^k - \Pi_{\mathcal{S}} \big(\mathbf{p}_i^k + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}(P^k) \big) \big\|^2 \Big] = 0, \ \forall i \in \mathcal{I}, \tag{15}$$

and furthermore, the running average converges at the rate of $\mathcal{O}(1/T)$ where T is the number of iterations, i.e., there exists a constant $\eta_1 > 0$ such that

$$\frac{1}{T} \sum_{k=0}^{T} \mathbb{E}\Big[\big\| \mathbf{p}_i^k - \Pi_{\mathcal{S}} \big(\mathbf{p}_i^k + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}(P^k) \big) \big\|^2 \Big] \leq \frac{\eta_1}{T}. \tag{16}$$

Note that the proofs of both Proposition 2 and Theorem 1 are provided in Appendix B and C, respectively; in addition, the detailed conditions of the step-size γ and sample-size M are also specified in the proof. Now, combining the above Proposition 2 and Theorem 1 together, it has been shown that the projected stochastic gradient algorithm converges to the desired equilibrium P^e. To sum up, we outline our scheme as the following Algorithm 1 and provide a few remarks on it.

Algorithm 1: Projected Stochastic Gradient Algorithm

Initialization: Each agent i initializes its own probability distribution $\mathbf{p}_i^0$ (not vertex), samples a set of M strategies $[\hat{a}_i^s]_{1 \leq s \leq M}$ based on $\mathbf{p}_i^0$, and sends it to all other agents. Set the maximum iteration K, and initialize index $k = 0$.

while $0 \leq k \leq K$ is satisfied **do**

Each sensor i **simultaneously** does

(S.1) Receive the sampled strategies $\hat{a}_j^s$ from all other agents j, and evaluate the stochastic gradient as (11);

(S.2) Update the distribution $\mathbf{p}_i^{k+1}$ as (12);

(S.3) Sample the M strategies $\hat{a}_i^s$'s based on the updated $\mathbf{p}_i^{k+1}$, and send it to other agents;

(S.4) Let $k \leftarrow k + 1$, and continue.

end

Remark 1: It is worth noting that the multi-linear extension function $f(P)$ is neither convex nor concave, thus the considered problem (8) belongs to the category of nonconvex optimization. Besides, our problem is also inherently nonsmooth, since the probability simplex constraints are present. In order to achieve the exact convergence for solving such nonsmooth nonconvex optimization (normally to the stationary point), typical stochastic gradient based approaches follow two paths: 1) increasing the sample-size with iteration numbers [21]; and 2) applying the techniques of variance reduction [22], [23]. One major novelty of our algorithm herein is the provable exact convergence with a constant sample-size. This primarily benefits from the fact that the stochastic gradient is not only bounded but also has finite possibilities in our problem.

Remark 2: We also remark that our algorithm is closely related to the well-known EXP3 algorithm [24]–[26] for solving the multi-armed bandit problems. In fact, the iteration (12) of our algorithm can be equivalently rewritten as

$$\mathbf{p}_i^{k+1} = \arg\min_{\mathbf{x}\in\mathcal{S}} \big\{ -\langle \nabla_{\mathbf{p}_i} \hat{f}(P^k),\, \mathbf{x}\rangle + \frac{1}{2\gamma}\|\mathbf{x}-\mathbf{p}_i^k\|^2 \big\}. \quad (17)$$

Once substituting the proximal regularization in (17) by the Kullback–Leibler divergence [27] regularization, it is straightforward to verify that our algorithm is equivalent to the EXP3 algorithm with full information feedback. To understand this connection, we can view each agent as a player choosing the desired arm from a bandit while the obtained reward is dynamically affected by all other players. In this sense, our algorithm drives all players to an equilibrium in which nobody can obtain more rewards by unilaterally changing its strategy. Other related works can be found in [28]–[31], which focus on the analysis of regret bound in the context of online learning.

IV. Distributed Algorithm with Communication Delays

As mentioned before, one major concern of the proposed Algorithm 1 is that individual agents are required to communicate with all others instantaneously, in order to received the sampled strategies $\hat{a}_i^s$'s based on the most updated distributions $\mathbf{p}_i^k$'s. This undoubtedly brings restrictions on the algorithm implementation. In this section, we relax such a requirement and further enhance the proposed algorithm by considering the scenario in which the communication delays are present.

Suppose that each individual agent can only receive others' strategies sampled from the time-delayed distributions. Concretely, let us assume, at the k-th iteration, each agent i receives the sampled strategy $\hat{a}_j^s$ from the agent j which is based on the distribution $\mathbf{p}_j^{k-\tau_{ij}}$. Note that τ_{ij} here represents the length of time-delays when the agent i receives the information from agent j. In addition, to ensure the informational flow between any pair of agents, we restrict, in the following assumption, that the time-delay τ_{ij} is bounded for any $i, j \in \mathcal{I}$.

Assumption 3: It is assumed that there exists a constant $D > 0$ such that $\tau_{ij} \leq D$ for $\forall i, j \in \mathcal{I}$.

We remark that the above Assumption 3 is quite standard in the study of algorithms with delayed communications. It inherently ensures that each agent receives others' information at least once within the time-window $k \leq t \leq k + D - 1$.

Since the agents' strategies are sampled from the delayed distributions, it is natural to see that the stochastic gradients are also subject to the time-delays. Let us denote the delayed stochastic gradient with respect to $p_i(a_i)$ as

$$\nabla_{p_i(a_i)} \hat{f}_\delta(P_i^{k-}) = \frac{1}{M} \cdot \sum_{s=1}^{M} F(a_i; \hat{A}_{-i}^s), \quad (18)$$

in which we use P_i^{k-} to represent the delayed distributions $\mathbf{p}_j^{k-\tau_{ij}}$'s associated with the agent i, and $\hat{A}_{-i}^s$ is the set of sampled strategies based on the delayed P_i^{k-}. As a result, the iteration of our projected stochastic gradient algorithm with communication delays becomes

$$\mathbf{p}_i^{k+1} = \Pi_{\mathcal{S}}\big(\mathbf{p}_i^k + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}_\delta(P_i^{k-}), \quad (19)$$

where $\nabla_{\mathbf{p}_i} \hat{f}_\delta(P_i^{k-})$ is the vector that stacks $\nabla_{p_i(a_i)} \hat{f}_\delta(P_i^{k-})$'s for all $a_i \in \mathcal{A}_i$.

Herein, let us refer to the scheme (19) as our Algorithm 2. As similar to Algorithm 1, in order to establish the convergence of Algorithm 2, we first characterize the condition of equilibrium solutions when communication delays are present.

Proposition 3: Suppose that Assumptions 2 and 3 hold, and let $\mathbf{P}_D^k = [P^k, P^{k+1}, \cdots, P^{k+D-1}]$ be a collection of probability distributions which are generated by Algorithm 2 within the time-window $k \leq t \leq k + D - 1$. Then, each P^t within the time-window is an equilibrium solution to problem (8), if the following condition is satisfied,

$$\sum_{t=k}^{k+D-1} \mathbb{E}\Big[\big\|\mathbf{p}_i^t - \Pi_{\mathcal{S}}\big(\mathbf{p}_i^t + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}_\delta(P_i^{t-})\big)\big\|^2\Big] = 0,\ \forall i \in \mathcal{I}. \quad (20)$$

Now, we establish the convergence of Algorithm 2 by the following theorem.

Theorem 2: Suppose that Assumptions 2 and 3 hold, and let $\{P^k\}_{k\in\mathbb{N}_+}$ be the sequence generated by Algorithm 2 with a small enough constant step-size γ and a large enough sample-size M. Then, it holds that, for $\forall i \in \mathcal{I}$,

$$\lim_{k\to\infty} \sum_{t=k}^{k+D-1} \mathbb{E}\Big[\big\|\mathbf{p}_i^t - \Pi_{\mathcal{S}}\big(\mathbf{p}_i^t + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}_\delta(P_i^{t-})\big)\big\|^2\Big] = 0, \quad (21)$$

and furthermore, the running average converges at the rate of $\mathcal{O}(1/T)$ where T is the number of iterations, i.e., there exists a constant $\eta_2 > 0$ such that

$$\frac{1}{T}\sum_{k=0}^{T} \mathbb{E}\Big[\big\|\mathbf{p}_i^t - \Pi_{\mathcal{S}}\big(\mathbf{p}_i^t + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}_\delta(P_i^{t-})\big)\big\|^2\Big] \leq \frac{\eta_2}{T}. \quad (22)$$

As earlier, we present the theoretical proofs of the above Proposition 3 and Theorem 2 in Appendix D and E, respectively; the detailed conditions of the step-size γ and sample-size M are also provided in the proof. In the end of this section, we make a few remarks on the implementation of the enhanced Algorithm 2 with delayed communications.

Remark 3: For the first D iterations, individual agents might not be able to receive all others' information, due to the presence of communication delays. Under such circumstance, our algorithm allows the agent to arbitrarily initialize the received strategies, and the convergence of algorithm will not be affected. For instance, each agent can choose the empty as the corresponding strategy if no information is received. In this sense, we remark that the delayed probability distribution P_i^{k-} is actually well-defined for all $k \geq 0$.

Remark 4: It is also noteworthy that, since the enhanced Algorithm 2 is robust against the communication delays, one can implement it in a fully distributed architecture where agents only need to communicate with their neighbors. Suppose that the communication channels among agents are governed by a peer-to-peer network, denoted as a general graph G. Then, the time-delay τ_{ij} in Algorithm 2 corresponds to the distance δ_{ij} between node i and j, i.e., the minimum number of edges that connect those two nodes. Therefore, as long as the graph G is

connected, meaning that δ_{ij} is bounded, our algorithm is still effective. In this case, however, the price is that each agent needs to maintain a memory buffer to store the information for all others within the network.

V. Simulation

In this section, we evaluate the effectiveness of the proposed algorithms by considering a real-world movie recommendation application [32], [33]. Our numerical simulations are conducted based on the well-known MovieLens dataset [34], which contains over 25 million ratings (ranging from 1 to 5) applied to $62,423$ movies by $162,541$ different users. In particular, we denote $r_{i,j}$ the rating submitted by the user i to the movie j, and say that the movie j is liked by the user i if $r_{i,j} \geq \bar{r}$ where $\bar{r}$ is some certain pre-defined threshold. The objective herein is to identify the top I movies, in the sense that those movies are liked by the maximum number of users. It should be noted that the considered problem is not trivial, since we count each user only once for all the chosen I movies. For example, suppose that the user i likes the movies j and j' at the same time and both of them are chosen as the top movies, then the user i will be counted only once, rather than twice, when counting the number of users for the top movies.

To formalize the above movie recommendation problem, let us denote $\mathcal{S}$ the set of all movies and $\mathcal{U}(j_s) := \{i \,|\, r_{i,j_s} \geq \bar{r}\}$ the set of users who like the movie j_s. Then, the considered movie recommendation application can be formulated as the following maximization problem,

$$\max_{j_s \in \mathcal{S}} \quad F(j_1, j_2, \cdots, j_I) := |\cup_{s=1}^{I} \mathcal{U}(j_s)|. \tag{23}$$

It can be verified that the objective function F is both monotone and submodular; see definitions in Assumption 1. Thus, (23) is a well-defined monotone submodular maximization problem which can be solved by the proposed algorithms.

In our simulations, we specify the rating threshold as $\bar{r} = 3$ and aim to identify the top $I = 10$ movies. In addition, to reduce the size of the candidate movie set $\mathcal{S}$, we pre-process the dataset and only consider the movies which are liked by no less than 300 users. As a result, totally 1,160 movies are picked up to comprise the candidate set $\mathcal{S}$. In the following, we conduct two separate simulations which are corresponding to the two proposed algorithms. Each simulation is compliant with a network composed of $I = 10$ agents. Therefore, each individual agent only needs to take charge in the determination of one (out of 1,160) movie, so that the entire network cooperatively finds the top 10 movies. In the first simulation, a fully-connected network is assumed, i.e., the all-to-all communications are available for each single agent, so that the Algorithm 1 can be implemented without time-delays. Additionally, in order to take into account the delayed communications, we assume a general but connected undirected network in the second simulation. In this case, agents only communicate with their neighbors. As mentioned in Remark 4, the length of time-delays τ_{ij} is governed by the distance between the agent i and j presented in the network.

The following Fig. 1 first plots the evolution of individual agents' probabilities of choosing the final top 10 movies, in the case that the fully-connected network is assumed. Note that the step-size and sample-size are set as $\gamma = 0.0005$ and $M = 3$. As one can observe from Fig. 1, each agent decides its own choice after around 350 iterations, and it is confirmed that the collection of the top 10 movies is an equilibrium solution to problem (23). More specifically, it turns out that the chosen top 10 movies are liked by $10,700$ users, while totally $11,842$ users are involved in all the 1,160 candidate movies. Although it is unknown how many users are covered by the optimal collection of the ten movies, given that this quantity has to be no larger than $11,842$, thus, it is immediately confirmed that the $1/2$-suboptimal bound is achieved.

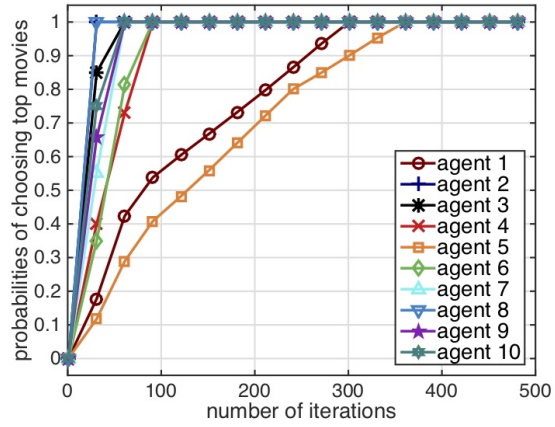


Fig. 1: Evolution of agents' probabilities (fully-connected graph).

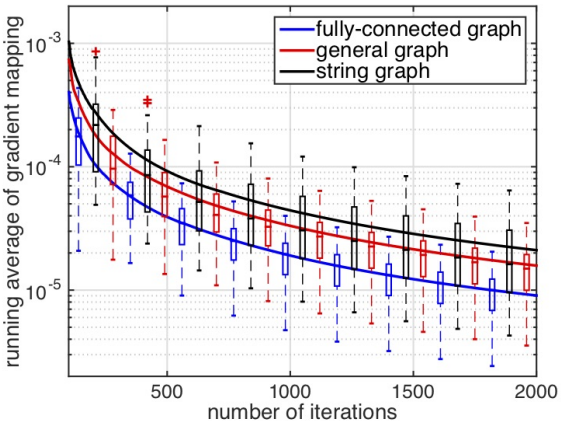
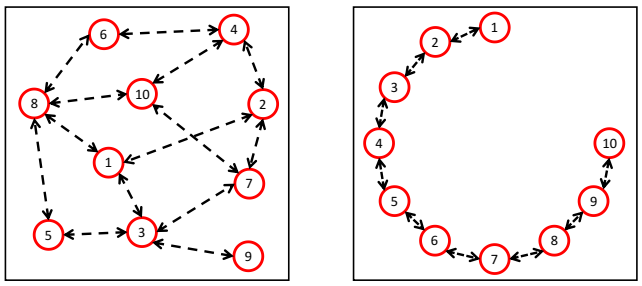


Fig. 2: Averaged gradient mapping with different graphs.

Furthermore, in order to evaluate the statistical performance of our stochastic gradient based algorithm, we carry out the Monte-Carlo simulation for 20 times. The simulation setting is the same as before. Fig. 2 plots the running average of the generated gradient mappings (blue curve), averaged by the 20 independent trials. Note that the running average of gradient mappings J^k at each iteration k is computed as

$$J^k := \frac{1}{k} \cdot \sum_{t=1}^{k} \sum_{i=1}^{I} \|\mathbf{p}_i^t - \mathbf{p}_i^{t-1}\|^2. \tag{24}$$

According to Proposition 2, it is implied that the algorithm converges to the desired equilibrium solution when $J^k \to 0$. Therefore, Fig. 2 validates the convergence of the proposed



(a) The general connected graph (b) The string graph

Fig. 3: Multi-agent network topologies.

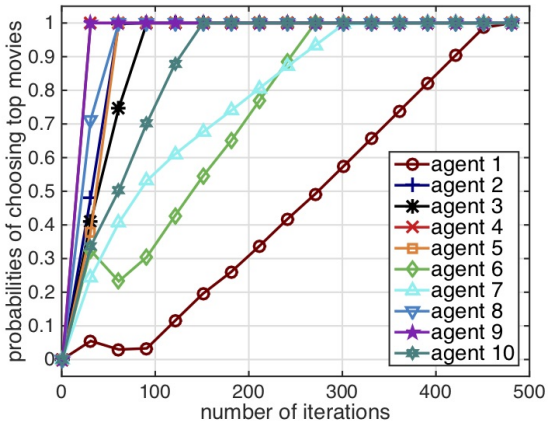


Fig. 4: Evolution of agents' probabilities (general graph).

Algorithm 1 and one can also observe a sublinear convergence rate as stated in Theorem 1.

In the second simulation, we run the proposed Algorithm 2 with delayed communications, under a general connected network as shown in Fig. 3(a). It can be seen that from the topology that the maximum distance between two nodes is four, and thus the communication delays are bounded by the constant $D = 3$. Fig. 4 shows the the evolution of agents' probabilities of choosing the top 10 movies, in the case that $\gamma = 0.0005$ and $M = 3$. Based on this figure, we conclude that Algorithm 2 also converges with the presence of delayed communications. However, compared to the case without time-delays as shown in Fig. 1, more iterations are needed to arrive at the final decisions of the top 10 movies.

Moreover, as similar to the first simulation, we conduct the Monte-Carlo simulation with 20 independent trails as well. Besides the general connected network as shown in Fig. 3(a), in this simulation, we additionally consider a specific string graph as shown in Fig. 3(b). It should be noted that, under such a string graph, the maximum distance between two nodes is nine and thus the communication delays are bounded by $D = 8$. The averaged J^k's in these two cases are also demonstrated in Fig. 2; see the red curve for the general connected graph and the black curve for the string graph. Based on Proposition 3, it is confirmed that Algorithm 2 converges to the equilibrium solution and a sublinear convergence rate is shown as expected. In addition, we also observe from the figure that a larger number of iterations are needed to obtain the solution when more communication delays are present in the network.

VI. Conclusion

In this paper, we developed a projected stochastic gradient algorithm for solving the distributed submodular maximization problem. Unlike the commonly-studied greedy algorithm, our approach enables the simultaneous updates among all individual agents. It is proved that the algorithm converges to an equilibrium solution, which is guaranteed to be at least $1/2$-suboptimal. Furthermore, we enhanced the proposed algorithm by handling the scenario in which agents' communication delays are present. The similar convergence result is proved for the enhanced distributed algorithm. Finally, a real-world movie recommendation application is considered to demonstrate the effectiveness of our algorithms. It should be also remarked that, compared to the brute-force searching scheme which has exponential complexity in terms of the number of agents I, the complexity of our distributed algorithms is primarily reduced. We leave the rigorous complexity analysis of the algorithms as our future work.

Appendix

A. Proof of Proposition 1

According to the definition of the function $F(A)$, let us first define its marginal function as

$$\delta(a \,|\, A) := F(\{a\} \cup A) - F(A), \tag{25}$$

It can be shown that the submodularity of function $F(A)$ implies that if $A' \subseteq A$, then

$$\delta(a \,|\, A') \geq \delta(a \,|\, A). \tag{26}$$

Recall that we use A_{-i} to denote the set of elements a_j's in A where $j \neq i$. In addition, we use $A_{<i}$ (or $A_{\leq i}$) to represent the set of elements a_j where $j < i$ (or $j \leq i$). Then, by the definition of the equilibrium A^e; see Definition 1, we can have

$$\delta(a_i^e \,|\, A_{-i}^e) \geq \delta(a_i \,|\, A_{-i}^e),\ \forall a_i \in \mathcal{A}_i, i \in \mathcal{I}. \tag{27}$$

Now, based on Assumption 1 of the function $F(A)$ (monotonicity and submodularity), it holds that

$$\begin{aligned}
F(A^\star) &\overset{(1a)}{\leq} F(A^\star \cup A^e) \\
&= F(A^e) + \sum_{i=1}^{I} \Big(F(A^\star_{\leq i} \cup A^e) - F(A^\star_{<i} \cup A^e) \Big) \\
&\overset{(1b)}{=} F(A^e) + \sum_{i=1}^{I} \delta(a_i^\star \,|\, A^\star_{<i} \cup A^e) \\
&\overset{(1c)}{\leq} F(A^e) + \sum_{i=1}^{I} \delta(a_i^\star \,|\, A^e_{-i}) \\
&\overset{(1d)}{\leq} F(A^e) + \sum_{i=1}^{I} \delta(a_i^e \,|\, A^e_{-i}) \\
&\overset{(1e)}{\leq} F(A^e) + \sum_{i=1}^{I} \delta(a_i^e \,|\, A^e_{<i}) \\
&\overset{(1f)}{=} F(A^e) + \sum_{i=1}^{I} \Big(F(A^e_{\leq i}) - F(A^e_{<i}) \Big) \\
&\overset{(1g)}{=} 2F(A^e).
\end{aligned} \tag{28}$$

Note that $(1a)$ comes from the monotonicity of $F(A)$; $(1b)$ and $(1f)$ is due to the definition of marginal function; $(1c)$ and $(1e)$ comes from the submodularity of $F(A)$; $(1d)$ is based on the inequality (27); and $(1g)$ is due to the fact that $F(\emptyset) = 0$. Therefore, the proof is completed.

B. Proof of Proposition 2

We start the proof by investigating properties of the projection on the probability simplex $\mathcal{S}$. According to the definition of the projection $\Pi_{\mathcal{S}}(\cdot)$, as shown in (13), it can be verified in [35], [36] that the projection is computed as

$$\Pi_{\mathcal{S}}(\mathbf{p}) = [\mathbf{p} - \lambda \mathbf{1}]^+, \tag{29}$$

where λ is the solution of the equation $\mathbf{1}^\top [\mathbf{p} - \lambda \mathbf{1}]^+ = 1$. Subsequently, we show the following two lemmas regarding the projection $\Pi_{\mathcal{S}}(\mathbf{p})$.

Lemma 1: Suppose that $\mathbf{p} \in [0,1]^K$ is a vertex of the simplex, i.e., $\|\mathbf{p}\|_\infty = 1$, and let us denote its non-zero component as $\mathbf{p}(n)$, i.e., $\mathbf{p}(n) = 1$ and $\mathbf{p}(k) = 0$ for $\forall k \neq n$. Then, one can have $\mathbf{p} = \Pi_{\mathcal{S}}(\mathbf{p} + \Delta_{\mathbf{p}})$ if and only if $\Delta_{\mathbf{p}}(n)$ is the maximum component of $\Delta_{\mathbf{p}}$, i.e., $\Delta_{\mathbf{p}}(n) \geq \Delta_{\mathbf{p}}(k)$ for $\forall k = 1, 2, \cdots, K$.

Proof: Let us recall that $\Pi_{\mathcal{S}}(\mathbf{p}+\Delta_{\mathbf{p}}) = [\mathbf{p}+\Delta_{\mathbf{p}} - \lambda \mathbf{1}]^+$. Suppose that $\mathbf{p} = \Pi_{\mathcal{S}}(\mathbf{p} + \Delta_{\mathbf{p}})$, since $\mathbf{p}$ is a vertex of the simplex and $\mathbf{p}(n) = 1$, then the n-th component of vector $\mathbf{p} + \Delta_{\mathbf{p}} - \lambda \mathbf{1}$ must be one, and all its other components must be non-positive. It means that $\mathbf{p}(n) + \Delta_{\mathbf{p}}(n) - \lambda = 1$ and $\mathbf{p}(k) + \Delta_{\mathbf{p}}(k) - \lambda \leq 0$ for $\forall k \neq n$. The former equality tells that $\Delta_{\mathbf{p}}(n) = \lambda$ and the latter yields $\Delta_{\mathbf{p}}(k) \leq \lambda$. Combining those two proves the first half of the statement.

Conversely, assume $\Delta_{\mathbf{p}}(n) \geq \Delta_{\mathbf{p}}(k)$ for $\forall k = 1, 2, \cdots, K$, then it holds that $\big(\mathbf{p}(n) + \Delta_{\mathbf{p}}(n)\big) - \big(\mathbf{p}(k) + \Delta_{\mathbf{p}}(k)\big) \geq 1$. Thus, to satisfy the equation $\mathbf{1}^\top [\mathbf{p}+\Delta_{\mathbf{p}} - \lambda \mathbf{1}]^+ = 1$, we must have $\lambda = \Delta_{\mathbf{p}}(n)$. On this account, we can further have that $\mathbf{p}(n)+\Delta_{\mathbf{p}}(n)-\lambda = 1$ and $\mathbf{p}(k)+\Delta_{\mathbf{p}}(k)-\lambda \leq 0$ for $\forall k \neq n$, and thus $[\mathbf{p} - \Delta_{\mathbf{p}} - \lambda \mathbf{1}]^+ = \mathbf{p}$. Therefore, the second part of the statement is proved. ■

Lemma 2: Suppose that $\mathbf{p} \in [0,1]^K$ is not a vertex of the simplex, and without loss of generality, let us assume its first n components ($2 \leq n \leq K$) to be non-zeros. Then, one can have $\mathbf{p} = \Pi_{\mathcal{S}}(\mathbf{p} + \Delta_{\mathbf{p}})$ if and only if there exists δ such that $\Delta_{\mathbf{p}}(k) = \delta$ for $k \leq n$ and $\Delta_{\mathbf{p}}(k) \leq \delta$ for $k > n$.

Proof: Recall again that $\Pi_{\mathcal{S}}(\mathbf{p}+\Delta_{\mathbf{p}}) = [\mathbf{p}+\Delta_{\mathbf{p}} - \lambda \mathbf{1}]^+$. Let us assume $\mathbf{p} = \Pi_{\mathcal{S}}(\mathbf{p}+\Delta_{\mathbf{p}})$, since the first n components of $\mathbf{p}$ are non-zeros, then we have $\mathbf{p}(k) = \mathbf{p}(k) + \Delta_{\mathbf{p}}(k) - \lambda$ for $k \leq n$ and $\Delta_{\mathbf{p}}(k) - \lambda \leq 0$ for $k > n$. Thus, simply taking $\delta = \lambda$ proves the first half of the statement.

Conversely, suppose that $\Delta_{\mathbf{p}}$ has $\Delta_{\mathbf{p}}(k) = \delta$ for $k \leq n$ and $\Delta_{\mathbf{p}}(k) \leq \delta$ for $k > n$. Since it is known that $\mathbf{1}^\top \mathbf{p} = 1$, in order to ensure the equation $\mathbf{1}^\top [\mathbf{p}+\Delta_{\mathbf{p}} - \lambda \mathbf{1}]^+ = 1$, we must have $\delta = \lambda$. Thus, it holds that $\Pi_{\mathcal{S}}(\mathbf{p} + \Delta_{\mathbf{p}}) = [\mathbf{p}]^+ = \mathbf{p}$. ■

With the help of the above two lemmas, we are now ready to prove the proposition in both directions separately.

Definition 2 ⇒ **Equation** (14):

Let us assume that $P^e = [\mathbf{p}_1^{e\top}, \mathbf{p}_2^{e\top}, \cdots, \mathbf{p}_I^{e\top}]^\top$ is an equilibrium following Definition 2. According to the definition, we know that each $\mathbf{p}_i^e$ must be a vertex of the simplex $\mathcal{S}$, i.e., there exists a_i^e such that $p_i^e(a_i^e) = 1$ and $p_i^e(a_i) = 0$ for $\forall a_i \neq a_i^e$; and in addition, the collection of a_i^e's, i.e., $A^e = [a_1^e, a_2^e, \cdots, a_I^e]^\top$, is the equilibrium following Definition 1. Provided that P^e is the collection of simplex vertices and the stochastic gradient $\nabla_{\mathbf{p}_i} \hat{f}(P^e)$ in (14) is sampled based on the probability distributions $\mathbf{p}_i^e$'s, thus it can be shown that the stochastic gradient has the following fixed form

$$\nabla_{\mathbf{p}_i} \hat{f}(P^e) = [F(a_i; A_{-i}^e)]_{a_i \in \mathcal{A}_i}, \tag{30}$$

and we can simply get rid of the expectation in (14). Moreover, by the definition of equilibrium A^e (see equation (3)), it holds that, for $\forall i \in \mathcal{I}$,

$$F(a_i^e; A_{-i}^e) \geq F(a_i; A_{-i}^e),\ \forall a_i \in \mathcal{A}_i. \tag{31}$$

On this account, we know that each gradient $\nabla_{\mathbf{p}_i} \hat{f}(P^e)$ has $F(a_i^e; A_{-i}^e)$ as its maximum component. Therefore, based on Lemma 1, it is proved that

$$\mathbf{p}_i^e = \Pi_{\mathcal{S}}\big(\mathbf{p}_i^e + \gamma \cdot \nabla_{\mathbf{p}_i} \hat{f}(P^e)\big), \tag{32}$$

and thus the proof of the first half is completed.

Equation (14) ⇒ **Definition** 2:

Suppose that the point P^e has already satisfied the condition (14). Next, we first show that each $\mathbf{p}_i^e$ in P^e has to be the vertex of the simplex $\mathcal{S}$. In fact, suppose that $\mathbf{p}_i^e$ is not a vertex and let $p_i^e(a_i^1)$ and $p_i^e(a_i^2)$ be the two non-zero components. Then, based on Lemma 2 and in order to ensure the condition (14), we must have that for $\forall a_i \in \mathcal{A}_i$,

$$\sum_{s=1}^{M} F(a_i^1; \hat{A}_{-i}^s) = \sum_{s=1}^{M} F(a_i^2; \hat{A}_{-i}^s) \geq \sum_{s=1}^{M} F(a_i; \hat{A}_{-i}^s), \tag{33}$$

where $\hat{A}_{-i}^s$ represents any possible sample of strategies. This clearly contradicts the maximum distinguishable assumption of the function $F(A)$ (see Assumption 2). As a result, we have proved that P^e must be the collection of simplex vertices, and the expectation in (14) can be removed. Next, let us assume that each $\mathbf{p}_i^e$ has $p_i^e(a_i^e) = 1$ and $p_i^e(a_i) = 0$ for $\forall a_i \neq a_i^e$, by the fact that it is simply a vertex. Applying Lemma 1 once again, it can be shown that the gradient $\nabla_{\mathbf{p}_i} \hat{f}(P^e)$ has its maximum component at $F(a_i^e; A_{-i}^e)$, i.e.,

$$a_i^e = \arg\max_{a_i \in \mathcal{A}_i}\ F(a_i; A_{-i}^e). \tag{34}$$

Thus, the second half of the proposition is proved.

C. Proof of Theorem 1

We begin the proof by recalling that the stochastic gradient $\nabla_{\mathbf{p}_i} \hat{f}(P^k)$ computed as (11) is an unbiased estimation of the full gradient $\nabla_{\mathbf{p}_i} f(P^k)$. In fact, this statement has been verified in the derivation of our projected stochastic gradient algorithm; see Section III-B. Thus, to facilitate the subsequent proof, we here extract the statement as the following lemma and omit the detailed proof.

Lemma 3: Suppose that the stochastic gradient $\nabla_{\mathbf{p}_i} \hat{f}(P)$ is defined as (11), then it holds that

$$\mathbb{E}_{\hat{a}_j \sim \mathbf{p}_j, j \neq i}\big[\nabla_{\mathbf{p}_i} \hat{f}(P)\big] = \nabla_{\mathbf{p}_i} f(P). \tag{35}$$

Next, let us introduce an additional notion, namely *gradient mapping*, which is defined as below,

$$\mathcal{G}_\gamma(\mathbf{g},\mathbf{p}) = \frac{1}{\gamma}\cdot\big(\mathbf{p}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})\big). \tag{36}$$

Here, $\mathbf{g}\in\mathbb{R}^K$ a general gradient, $\mathbf{p}\in\mathcal{S}$ is a general point from the probability simplex, and γ is a constant which represents the step-size. Recall that our projected stochastic gradient algorithm performs the iteration (12), it can be equivalently rewritten into the following gradient mapping form,

$$\mathbf{p}_i^{k+1} = \mathbf{p}_i^k - \gamma\cdot\mathcal{G}_\gamma(\nabla_{\mathbf{p}_i}\hat{f}(P^k),\mathbf{p}_i^k). \tag{37}$$

The iteration (37) can be interpreted as a standard line search algorithm with the constant step-size γ, while the searching direction is the gradient mapping $\mathcal{G}_\gamma(\nabla_{\mathbf{p}_i}\hat{f}(P^k),\mathbf{p}_i^k)$.

Associated with the gradient mapping, we next show the following two lemmas, which will play key roles in the proof of the theorem.

Lemma 4: Given the gradient mapping $\mathcal{G}_\gamma(\mathbf{g},\mathbf{p})$, for any $\mathbf{g}\in\mathbb{R}^K$, $\mathbf{p}\in\mathcal{S}$ and $\gamma\in\mathbb{R}_+$, it holds that

$$-\langle\mathbf{g},\ \mathcal{G}_\gamma(\mathbf{g},\mathbf{p})\rangle \geq \|\mathcal{G}_\gamma(\mathbf{g},\mathbf{p})\|^2. \tag{38}$$

Proof: Recall that the projection on the probability simplex $\Pi_{\mathcal{S}}(\cdot)$ is defined as (13). Thus, within the gradient mapping, the term $\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})$ can be computed as

$$\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}) = \mathop{\arg\min}_{\mathbf{x}\in\mathcal{S}}\ \{-\langle\mathbf{g},\ \mathbf{x}\rangle + \frac{1}{2\gamma}\|\mathbf{x}-\mathbf{p}\|^2\}. \tag{39}$$

By noticing that the optimization problem in (39) is convex, the optimality condition of solution $\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})$ ensures that, for $\forall\mathbf{x}\in\mathcal{S}$,

$$\Big\langle -\mathbf{g}+\frac{1}{\gamma}\big(\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})-\mathbf{p}\big),\ \mathbf{x}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})\Big\rangle \geq 0. \tag{40}$$

Now, let $\mathbf{x}=\mathbf{p}$, it can be shown that

$$-\big\langle\mathbf{g},\ \mathbf{p}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})\big\rangle \geq \frac{1}{\gamma}\|\mathbf{p}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g})\|^2. \tag{41}$$

Provided that the gradient mapping $\mathcal{G}_\gamma(\mathbf{g},\mathbf{p})$ is define as (14), thus the proof is completed. ∎

Lemma 5: Given the gradient mapping $\mathcal{G}_\gamma(\mathbf{g},\mathbf{p})$, for any $\mathbf{g}_1,\mathbf{g}_2\in\mathbb{R}^K$, $\mathbf{p}\in\mathcal{S}$ and $\gamma\in\mathbb{R}_+$, it holds that

$$\|\mathcal{G}_\gamma(\mathbf{g}_1,\mathbf{p})-\mathcal{G}_\gamma(\mathbf{g}_2,\mathbf{p})\| \leq \|\mathbf{g}_1-\mathbf{g}_2\|. \tag{42}$$

Proof: Applying again the optimality condition (40) with the gradient $\mathbf{g}$ substituted by $\mathbf{g}_1$ and $\mathbf{g}_2$ respectively, it yields that,

$$\Big\langle -\mathbf{g}_1+\frac{1}{\gamma}\big(\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_1)-\mathbf{p}\big),\ \mathbf{x}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_1)\Big\rangle \geq 0; \tag{43a}$$

$$\Big\langle -\mathbf{g}_2+\frac{1}{\gamma}\big(\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_2)-\mathbf{p}\big),\ \mathbf{x}-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_2)\Big\rangle \geq 0. \tag{43b}$$

Now, let $\mathbf{x}=\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_2)$ in (43a) and $\mathbf{x}=\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_1)$ in (43b), summing both inequalities gives that

$$\begin{aligned}&\big\langle\mathbf{g}_2-\mathbf{g}_1,\ \Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_2)-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_1)\big\rangle\\ &\geq\frac{1}{\gamma}\|\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_1)-\Pi_{\mathcal{S}}(\mathbf{p}+\gamma\mathbf{g}_2)\|^2\end{aligned} \tag{44}$$

By the definition of gradient mapping $\mathcal{G}_\gamma(\mathbf{g},\mathbf{p})$ and Cauchy-Schwartz inequality, the proof is completed. ∎

It should be remarked that, while Lemma 3 verifies that the stochastic gradient $\nabla_{\mathbf{p}_i}\hat{f}(P^k)$ is an unbiased estimation, Lemma 4 and 5 both characterize the properties of the gradient mapping. Next, let us show another lemma which investigates the variance of the stochastic gradient. Before stating the lemma, some more notations and a supporting lemma are needed to be first introduced. Let us simply denote $\nabla f(P)$ (also $\nabla\hat{f}(P)$) the stacked full (stochastic) gradient for each agent $i\in\mathcal{I}$, i.e., $\nabla f(P)=[\nabla_{\mathbf{p}_i}f(P)]_{i\in\mathcal{I}}$. Similarly, we use $\tilde{\mathcal{G}}_\gamma\big(\nabla f(P),P\big)$ and $\tilde{\Pi}_{\mathcal{S}}\big(P+\gamma\cdot\nabla\hat{f}(P)\big)$ to denote the stacked gradient mapping and also the updated probability distributions respectively, i.e.,

$$\begin{aligned}P^{k+1} &= \tilde{\Pi}_{\mathcal{S}}\big(P^k+\gamma\cdot\nabla\hat{f}(P^k)\big)\\ &= P^k-\gamma\cdot\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k),P^k\big).\end{aligned} \tag{45}$$

Lemma 6: Suppose that the current iterate P^k is a collection of simplex vertices but not an equilibrium. Let the step-size satisfy $\gamma<2/\Delta^{\max}$, then the next iterate P^{k+1} generated by Algorithm 1 must not be the collection of simplex vertices.

Proof: Let us first recall that the iterate P^k is a collection of $\mathbf{p}_i^k$'s, i.e., $P^k=[\mathbf{p}_1^{k\top},\mathbf{p}_2^{k\top},\cdots,\mathbf{p}_I^{k\top}]^\top$. Since it is assumed that each $\mathbf{p}_i^k$ is the simplex vertex, then we let $\mathbf{p}_i^k=\mathbf{e}_{n_i}$ with $\mathbf{e}_{n_i}\in\mathbb{R}^K$ being the unit vector whose n_i-th component is one and others are zeros. In addition, according to the iteration of Algorithm 1 and the fact that the projection $\Pi_{\mathcal{S}}$ can be computed as (29), thus we know

$$\mathbf{p}_i^{k+1} = \big[\mathbf{e}_{n_i}+\gamma\cdot F_{-i}-\lambda\mathbf{1}\big]^+, \tag{46}$$

where λ is governed by the equation $\mathbf{1}^\top\mathbf{p}_i^{k+1}=1$. Note that here the sampled gradient is deterministic since P^k is a collection of vertices, thus we use $F_{-i}\in\mathbb{R}^K$ to represent the sampled gradient based on the probability distribution $\mathbf{p}_j^k$, $j\neq i$. Furthermore, we denote $F_{-i}(n)$ the n-th component of the vector F_{-i}.

Given that P^k is not an equilibrium, thus there exist indices $i\in\mathcal{I}$ and $n_i^{\max}\neq n_i$, such that $F_{-i}(n_i^{\max})>F_{-i}(n_i)$ and

$$F_{-i}(n_i^{\max})\geq F_{-i}(n),\ \forall n=1,2,\cdots,K. \tag{47}$$

In fact, if $F_{-i}(n_i)$'s are the maximum components for $\forall i\in\mathcal{I}$, then P^k must be the equilibrium by definition. Consequently, according to the equation (46), we know that $\mathbf{p}_i^{k+1}(n_i^{\max})>0$ must be true. On this basis, in order to prove the lemma, it will suffice to show that $\mathbf{p}_i^{k+1}(n_i^{\max})<1$ if $\gamma<2/\Delta^{\max}$. Next, we prove this statement by contradiction.

Suppose that $\mathbf{p}_i^{k+1}(n_i^{\max})<1$ is false, i.e., $\mathbf{p}_i^{k+1}(n_i^{\max})=1$. Provided that $n_i^{\max}\neq n_i$, thus we have $\gamma\cdot F_{-i}(n_i^{\max})-\lambda=1$ and $1+\gamma\cdot F_{-i}(n_i)-\lambda\leq 0$. Substitute the former equation to the latter one and get rid of λ, it yields,

$$2+\gamma\cdot\big(F_{-i}(n_i)-F_{-i}(n_i^{\max})\big)\leq 0. \tag{48}$$

Recall the definition of $\Delta^{\max}$; see equation (2), and the fact that $F_{-i}(n_i^{\max})>F_{-i}(n_i)$, we have

$$0<F_{-i}(n_i^{\max})-F_{-i}(n_i)\leq\Delta^{\max}. \tag{49}$$

Combining both (48) and (49) shows that $\gamma \geq 2/\Delta^{\max}$. Thus, the proof is completed.

■

Now, we are ready to show the following lemma which characterizes the variance of stochastic gradients.

Lemma 7: Suppose that the sequence $\{P^k\}_{k\in\mathbb{N}_+}$ is the set of iterates generated by Algorithm 1 and the initialization P^0 is not a collection of simplex vertices. Let the step-size γ satisfy the condition $\gamma < 2/\Delta^{\max}$, then there exist constants $B_0 > 0$ and $B_1 > 0$ such that the following holds,

$$\sum_{k=0}^{T}\mathbb{E}\Big[\|\nabla\hat{f}(P^k)-\nabla f(P^k)\|^2\Big] \leq B_0 + B_1/M\cdot\sum_{k=0}^{T}\mathbb{E}\Big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k),P^k\big)\|^2\Big], \tag{50}$$

where the expectation is taken with respect to the sampling of stochastic gradient $\nabla\hat{f}(P)$ for all iterations $0 \leq k \leq T$ and M is the sample-size.

Proof: Before starting the proof, we first note that it is only needed to consider the case when none of $P^k,\ 0 \leq k \leq T$ is the equilibrium. In fact, it can be immediately verified that the algorithm will stay at the equilibrium P^e forever once it reaches the point. In addition, due to the fact that

$$\mathbb{E}\big[\|\nabla\hat{f}(P^e)-\nabla f(P^e)\|^2\big] = \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^e),P^e\big)\|^2\big] = 0, \tag{51}$$

thus we only need to prove the case in which the algorithm has not reach the equilibrium.

Now, let us begin the proof by introducing an additional notion, namely the reachable set of the iterates P^k. We define the reachable set $\mathcal{S}^k$ at each iteration k as follows,

$$\mathcal{S}^k := \big\{P^k \,|\, P^t = \tilde{\Pi}_{\mathcal{S}}\big(P^{t-1}+\gamma\cdot\nabla\bar{f}(P^{t-1})\big), 1\leq t\leq k\big\}. \tag{52}$$

Note that here the gradient $\nabla\bar{f}(P^{t-1})$ is any possible realization of the stochastic gradient $\nabla\hat{f}(P^{t-1})$. Due to the fact that each $\nabla\bar{f}(P^{t-1})$ only has finite possibilities and P^0 is well initialized, thus we know each $\mathcal{S}^k$ is also a finite set, but its cardinality grows quickly as the index k increases. Subsequently, let us divide each of the reachable sets $\mathcal{S}^k$ into two subsets, i.e., $\mathcal{S}^k = \bar{\mathcal{S}}^k \cup \bar{\mathcal{S}}^k_c$ where $\bar{\mathcal{S}}^k$ only contains the iterates P^k's which are collections of the simplex vertices and $\bar{\mathcal{S}}^k_c$ is the complement set. On this account, we next prove the following statements: there exists a constant $\epsilon > 0$ such that

1) if it is known that $P^{k+1} \in \bar{\mathcal{S}}^{k+1}_c$, then

$$\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k),P^k\big)\|^2\big] \geq \epsilon; \tag{53}$$

2) if it is known that $P^k \in \bar{\mathcal{S}}^k$, then

$$\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k),P^k\big)\|^2\big] \geq \epsilon. \tag{54}$$

Proof of statement 1): Recall again that the iterate P^k is the collection of $\mathbf{p}_i^k$'s. Since it is known that $P^{k+1} \in \bar{\mathcal{S}}^{k+1}_c$, then let us assume, without loss of generality, that $\mathbf{p}_i^{k+1}$ has two non-zero components $\mathbf{p}_i^{k+1}(u)$ and $\mathbf{p}_i^{k+1}(v)$ such that $\mathbf{p}_i^{k+1}(u) \geq \mathbf{p}_i^{k+1}(v)$. Then, the expectation term in (53) has,

$$\begin{aligned}\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k),P^k\big)\|^2\big] &= 1/\gamma^2\cdot\mathbb{E}\big[\|P^k-P^{k+1}\|^2\big]\\ &= 1/\gamma^2\cdot\mathbb{E}\Big[\sum_{i=1}^{I}\|\mathbf{p}_i^k-\mathbf{p}_i^{k+1}\|^2\Big]\\ &\geq 1/\gamma^2\cdot\mathbb{E}\big[\|\mathbf{p}_i^k-\mathbf{p}_i^{k+1}\|^2\big].\end{aligned} \tag{55}$$

According to the iteration of Algorithm 1 and the computation (29) of the projection $\Pi_{\mathcal{S}}$, then we have

$$\begin{cases}\mathbf{p}_i^{k+1}(u) = \mathbf{p}_i^k(u)+\gamma\cdot F_{-i}(u)-\lambda;\\ \mathbf{p}_i^{k+1}(v) = \mathbf{p}_i^k(v)+\gamma\cdot F_{-i}(v)-\lambda.\end{cases} \tag{56}$$

Note that $F_{-i}(u)$ and $F_{-i}(v)$ are the u-th and v-th components of the sampled gradient $\nabla\hat{f}(P^k)$. Since we have assumed that $\mathbf{p}_i^{k+1}(u) \geq \mathbf{p}_i^{k+1}(v) > 0$, it can be verified that $F_{-i}(u)$ has to be the maximum one against all other $F_{-i}(n)$'s. Next, based on the maximum distinguishable assumption, we know that $\mathbb{E}[F_{-i}(u)]$ has to be strictly greater than $\mathbb{E}[F_{-i}(v)]$. Then, let $\Delta = \mathbb{E}[F_{-i}(u)] - \mathbb{E}[F_{-i}(v)] > 0$, it holds that

$$\begin{aligned}&\mathbb{E}\big[\|\mathbf{p}_i^k-\mathbf{p}_i^{k+1}\|^2\big]\\ &\overset{(2a)}{\geq} \|\mathbb{E}\big[\mathbf{p}_i^k-\mathbf{p}_i^{k+1}\big]\|^2\\ &\geq \big(\gamma\cdot\mathbb{E}[F_{-i}(u)]-\mathbb{E}[\lambda]\big)^2+\big(\gamma\cdot\mathbb{E}[F_{-i}(v)]-\mathbb{E}[\lambda]\big)^2\\ &= \frac{1}{2}\big(2\gamma\mathbb{E}[F_{-i}(v)]+\gamma\Delta-2\mathbb{E}[\lambda]\big)^2+\frac{1}{2}\gamma^2\Delta^2\\ &\geq \frac{1}{2}\gamma^2\Delta^2.\end{aligned} \tag{57}$$

Note that $(2a)$ follows from the Jensen's inequality. Thus, the proof of statement 1) is completed.

Proof of statement 2): According to the above Lemma 6, we know that P^{k+1} must be not the collection of simplex vertices, if the step-size γ is choose under the condition and P^k is the collection of vertices. In other words, $P^k \in \bar{\mathcal{S}}^k$ implies $P^{k+1} \in \bar{\mathcal{S}}^{k+1}_c$, and conversely, $P^{k+1} \in \bar{\mathcal{S}}^{k+1}$ implies $P^k \in \bar{\mathcal{S}}^k_c$. Therefore, the proof of statement 2) can be done by following exactly the same path of statement 1).

Now, recall that the stochastic gradient $\nabla\hat{f}(P^k)$ is *i.i.d.* sampled with the sample-size M. Let us denote $\nabla\hat{f}_s(P^k)$ as the gradient decided by one single sample $s = 1, 2, \cdots, M$, thus we know

$$\begin{aligned}&\mathbb{E}\big[\|\nabla\hat{f}(P^k)-\nabla f(P^k)\|^2\big]\\ &= \mathbb{E}\Big[\Big\|\frac{1}{M}\sum_{s=1}^{M}\big(\nabla\hat{f}_s(P^k)-\nabla f(P^k)\big)\Big\|^2\Big]\\ &= \frac{1}{M^2}\cdot\sum_{s=1}^{M}\mathbb{E}\Big[\|\nabla\hat{f}_s(P^k)-\nabla f(P^k)\|^2\Big].\end{aligned} \tag{58}$$

Furthermore, by the definition of the function $f(P)$, it can be immediately verified that its gradient $\nabla f(P^k)$ is always bounded for $\forall P^k$, so is the *i.i.d.* sampled stochastic gradient $\hat{f}_s(P^k)$. Based on this, we can have that the variance term $\mathbb{E}\big[\|\nabla\hat{f}_s(P^k)-\nabla f(P^k)\|^2\big]$ is bounded. Therefore, the above two statements can further imply the following two conditions: there exists a constant B_1 such that,

1) if it is known that $P^{k+1} \in \bar{\mathcal{S}}_c^{k+1}$, then

$$\begin{aligned}&\mathbb{E}\big[\|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big]\\ &\leq B_1/M \cdot \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big];\end{aligned} \tag{59}$$

2) if it is known that $P^k \in \bar{\mathcal{S}}^k$, then

$$\begin{aligned}&\mathbb{E}\big[\|\nabla \hat{f}(P^{k-1}) - \nabla f(P^{k-1})\|^2 + \|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big]\\ &\leq B_1/M \cdot \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big].\end{aligned} \tag{60}$$

With the help of the above two inequalities (59) and (60), we are now ready to prove the statement in the lemma. Let us first denote q_k the probability that P^k is a collection of simplex vertices, i.e.,

$$q_k := \Pr(P^k \in \bar{\mathcal{S}}^k). \tag{61}$$

For the notational convenience, we denote

$$\begin{cases}\mathbf{I}^k = \mathbb{E}\Big[\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big] \,\big|\, P^k \in \bar{\mathcal{S}}^k\Big];\\ \mathbf{II}^k = \mathbb{E}\Big[\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big] \,\big|\, P^k \in \bar{\mathcal{S}}_c^k, P^{k+1} \in \bar{\mathcal{S}}_c^{k+1}\Big];\\ \mathbf{III}^k = \mathbb{E}\Big[\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big] \,\big|\, P^{k+1} \in \bar{\mathcal{S}}^{k+1}\Big].\end{cases} \tag{62}$$

It should be remarked that, in (62), while the inner expectation is taken with respect to the stochastic gradient $\nabla \hat{f}(P^k)$, the outer expectation is taken with respect to the randomness of P^k and P^{k+1}. Consequently, it holds that, for $\forall k \geq 1$,

$$\begin{aligned}&\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big]\\ &= q_k \cdot \mathbf{I}^k + (1 - q_k - q_{k+1}) \cdot \mathbf{II}^k + q_{k+1} \cdot \mathbf{III}^k\\ &\overset{(3a)}{\geq} q_k M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^{k-1}) - \nabla f(P^{k-1})\|^2\big]\\ &\quad + (1 - q_{k+1}) M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big].\end{aligned} \tag{63}$$

Note that $(3a)$ is due to the inequalities (59), (60) and the fact that $\mathbf{III}^k \geq 0$. According to (63), we have

$$\begin{aligned}&\sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big]\\ &\overset{(4a)}{=} (1 - q_1) \cdot \mathbf{II}^0 + q_1 \cdot \mathbf{III}^0\\ &\quad + \sum_{k=1}^{T} \Big(q_k \cdot \mathbf{I}^k + (1 - q_k - q_{k+1}) \cdot \mathbf{II}^k + q_{k+1} \cdot \mathbf{III}^k\Big)\\ &\overset{(4b)}{\geq} (1 - q_1) M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^0) - \nabla f(P^0)\|^2\big]\\ &\quad + \sum_{k=1}^{T} q_k M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^{k-1}) - \nabla f(P^{k-1})\|^2\big]\\ &\quad + \sum_{k=1}^{T} (1 - q_{k+1}) M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big]\\ &= M/B_1 \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big]\\ &\quad - q_{T+1} M/B_1 \cdot \mathbb{E}\big[\|\nabla \hat{f}(P^T) - \nabla f(P^T)\|^2\big]\\ &\overset{(4c)}{\geq} M/B_1 \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\nabla \hat{f}(P^k) - \nabla f(P^k)\|^2\big] - MB_0/B_1.\end{aligned} \tag{64}$$

Note that $(4a)$ is due to the fact that the initialization P^0 is not the collection of vertices, i.e. $q_0 = 0$; $(4b)$ comes from the inequality (63); and $(4c)$ is based on the fact that the variance term $\mathbb{E}\big[\|\nabla \hat{f}(P^T) - \nabla f(P^T)\|^2\big]$ can be upper bounded by the constant B_0. Rearranging the inequality (64) and noticing the definition (45) of the stacked gradient mapping complete the proof of the lemma. ∎

After showing the above lemmas, we are now in the position to prove the theorem. Since the Hessian of the function $f(P)$ is always bounded, it can be immediately verified that the gradient of $f(P)$ is Lipschitz continuous, so is the gradient of $-f(P)$. Thus, there exists a constant $L > 0$ such that,

$$\begin{aligned}&- f(P^{k+1})\\ &\leq -f(P^k) + \big\langle -\nabla f(P^k),\ P^{k+1} - P^k\big\rangle + \frac{L}{2}\|P^{k+1} - P^k\|^2\\ &\overset{(5a)}{=} -f(P^k) + \gamma\Big\langle \nabla f(P^k) \pm \nabla \hat{f}(P^k),\ \tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\Big\rangle\\ &\quad + \frac{\gamma^2 L}{2}\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\\ &\overset{(5b)}{\leq} -f(P^k) + (\frac{\gamma^2 L}{2} - \gamma)\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\\ &\quad + \gamma\Big\langle \nabla f(P^k) - \nabla \hat{f}(P^k),\\ &\qquad \tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big) \pm \tilde{\mathcal{G}}_\gamma\big(\nabla f(P^k), P^k\big)\Big\rangle\\ &\overset{(5c)}{\leq} -f(P^k) + (\frac{\gamma^2 L}{2} - \gamma)\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\\ &\quad + \gamma\Big\langle \nabla f(P^k) - \nabla \hat{f}(P^k),\ \tilde{\mathcal{G}}_\gamma\big(\nabla f(P^k), P^k\big)\Big\rangle\\ &\quad + \gamma\|\nabla f(P^k) - \nabla \hat{f}(P^k)\|^2.\end{aligned} \tag{65}$$

Note that $(5a)$ is due to the definition of gradient mapping; $(5b)$ comes from Lemma 4; and $(5c)$ is due to the Cauchy-Schwartz inequality and Lemma 5. Now, let us take expectation on the inequality (65), with respect to the random sampling of stochastic gradient $\nabla \hat{f}(P^k)$ by given the probability distribution P^k. Since $\nabla \hat{f}(P^k)$ is the unbiased estimation of the full gradient $\nabla f(P^k)$ according to Lemma 3, it holds that

$$\begin{aligned}\mathbb{E}\big[f(P^{k+1})\big] - f(P^k) \geq &(\gamma - \frac{\gamma^2 L}{2}) \cdot \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big]\\ &- \gamma \cdot \mathbb{E}\big[\|\nabla f(P^k) - \nabla \hat{f}(P^k)\|^2\big].\end{aligned} \tag{66}$$

Consequently, summing up the above inequality (66) for all $0 \leq k \leq T$ and taking the expectation with respect to the random sampling for all iterations, we have

$$\begin{aligned}&\mathbb{E}\big[f(P^{T+1})\big] - f(P^0)\\ &\geq (\gamma - \frac{\gamma^2 L}{2}) \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big]\\ &\quad - \gamma \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\nabla f(P^k) - \nabla \hat{f}(P^k)\|^2\big]\\ &\overset{(6a)}{\geq} -\gamma B_0 + (\gamma - \frac{B_1 \gamma}{M} - \frac{\gamma^2 L}{2}) \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla \hat{f}(P^k), P^k\big)\|^2\big].\end{aligned} \tag{67}$$

Note that $(6a)$ follows from Lemma 7. Now, suppose that $P^\star$ is the optimal solution for solving problem (8), i.e.,

$\mathbb{E}\big[f(P^{T+1})\big] \leq f(P^\star)$, $\forall T \in \mathbb{R}_+$. Then, the above inequality (67) implies that, if the sample-size M and step-size γ are chosen satisfying $M > B_1$ and $\gamma - B_1\gamma/M - \gamma^2 L/2 > 0$, the non-negative sequence $\big\{\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k), P^k\big)\|^2\big]\big\}_{k\in\mathbb{N}_+}$ is summable, i.e.,

$$\sum_{k=0}^{\infty}\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k), P^k\big)\|^2\big] \leq \frac{f(P^\star) - f(P^0) + \gamma B_0}{\gamma - B_1\gamma/M - \gamma^2 L/2}. \tag{68}$$

Thus, $\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k), P^k\big)\|^2\big]$ converges to zero; and furthermore, its running average converges at the rate of $\mathcal{O}(1/T)$.

D. Proof of Proposition 3

Let us first remark that the condition (20) simply implies that the following equation holds for all $k \leq t \leq k + D - 1$,

$$\mathbb{E}\Big[\big\|\mathbf{p}_i^t - \Pi_{\mathcal{S}}\big(\mathbf{p}_i^t + \gamma\cdot\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{t-})\big)\big\|^2\Big] = 0,\ \forall i \in \mathcal{I}. \tag{69}$$

According to the iteration (19), it can be immediately verified that $P^{t+1} = P^t$ is true for all $k \leq t \leq k + D - 2$. Therefore, to prove the statement in Proposition 3, it will suffice to show that P^{k+D-1} is an equilibrium solution.

Since each P^t is identical within the entire time-window $k \leq t \leq k + D - 1$, then let $t = k + D - 1$, the equation (69) implies that, for all $i \in \mathcal{I}$,

$$\mathbb{E}\Big[\big\|\mathbf{p}_i^{k+D-1} - \Pi_{\mathcal{S}}\big(\mathbf{p}_i^{k+D-1} + \gamma\cdot\nabla_{\mathbf{p}_i}\hat{f}(P^{k+D-1})\big)\big\|^2\Big] = 0. \tag{70}$$

Note that, in (70), the stochastic gradient $\nabla_{\mathbf{p}_i}\hat{f}(P^{k+D-1})$ is evaluated without time-delays. As a result of the above Proposition 2, we know that P^{k+D-1} has to be an equilibrium. Therefore, the proof is completed.

E. Proof of Theorem 2

As similar to the previous proof, let us first mention that the iteration (19) of Algorithm 2 can be compactly expressed as,

$$P^{k+1} = P^k - \gamma\cdot\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big), \tag{71}$$

where $\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)$ is the stacked gradient mapping and $\mathbf{P}_D^{k-} = [P_i^{k-}]_{i\in\mathcal{I}}$ collects the delayed distributions P_i^{k-} for all $i \in \mathcal{I}$. Now, according to the Lipschitz continuous gradient of the function$-f(P)$, we invoke the descent lemma again and it holds that,

$$\begin{aligned}
&-f(P^{k+1})\\
&\leq -f(P^k) + \big\langle -\nabla f(P^k),\ P^{k+1} - P^k\big\rangle + \frac{L}{2}\|P^{k+1} - P^k\|^2\\
&= -f(P^k) + \gamma\Big\langle\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\quad + \gamma\Big\langle\nabla f(P^k) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\quad + \frac{\gamma^2 L}{2}\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\\
&\overset{(7a)}{\leq} -f(P^k) + (\frac{\gamma^2 L}{2} - \gamma)\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\\
&\quad + \gamma\Big\langle\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\quad + \gamma\Big\langle\nabla f_\delta(\mathbf{P}_D^{k-}) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle.
\end{aligned} \tag{72}$$

Note that the inequality $(7a)$ is due to Lemma 4; furthermore, we use $\nabla f_\delta(\mathbf{P}_D^{k-})$ to denote the full gradient which is based on the delayed probability distributions $\mathbf{P}_D^{k-}$. It should be emphasized that the sampled gradient $\nabla\hat{f}_\delta(\mathbf{P}_D^{k-})$ is an unbiased estimation of the full gradient $\nabla f_\delta(\mathbf{P}_D^{k-})$. Thus, according to the Cauchy-Schwartz inequality and Lemma 5, the above (72) can be continued as

$$\begin{aligned}
&-f(P^{k+1})\\
&\leq -f(P^k) + (\frac{\gamma^2 L}{2} - \gamma)\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\\
&\quad + \gamma\Big\langle\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\quad + \gamma\Big\langle\nabla f_\delta(\mathbf{P}_D^{k-}) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-}),\\
&\qquad\qquad \tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big) \pm \tilde{\mathcal{G}}_\gamma\big(\nabla f_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\leq -f(P^k) + (\frac{\gamma^2 L}{2} - \gamma)\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\\
&\quad + \gamma\big\|\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-})\big\|\big\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\big\|\\
&\quad + \gamma\Big\langle\nabla f_\delta(\mathbf{P}_D^{k-}) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-}),\ \tilde{\mathcal{G}}_\gamma\big(\nabla f_\delta(\mathbf{P}_D^{k-}), P^k\big)\Big\rangle\\
&\quad + \gamma\big\|\nabla f_\delta(\mathbf{P}_D^{k-}) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-})\big\|^2.
\end{aligned} \tag{73}$$

Now, taking the expectation on both sides and summing up the inequalities for all $0 \leq k \leq T$, it holds that

$$\begin{aligned}
&\mathbb{E}\big[f(P^{T+1})\big] - f(P^0)\\
&\geq (\gamma - \frac{\gamma^2 L}{2})\cdot\sum_{k=0}^{T}\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\big]\\
&\quad \underbrace{-\gamma\cdot\sum_{k=0}^{T}\mathbb{E}\big[\|\nabla f_\delta(\mathbf{P}_D^{k-}) - \nabla\hat{f}_\delta(\mathbf{P}_D^{k-})\|^2\big]}_{:=\mathcal{T}_1}\\
&\quad \underbrace{-\gamma\cdot\sum_{k=0}^{T}\mathbb{E}\big[\|\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-})\|\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|\big]}_{:=\mathcal{T}_2}.
\end{aligned} \tag{74}$$

As shown in the above inequality, let us denote the last two summation terms as $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively. Next, we prove the following two lemmas which upper bound $\mathcal{T}_1$ and $\mathcal{T}_2$ by the summation of gradient mappings.

Lemma 8: Suppose that the sequence $\{P^k\}_{k\in\mathbb{N}_+}$ is the set of iterates generated by Algorithm 2 and the initialization P^0 is not a collection of simplex vertices. Let the step-size γ satisfy the condition $\gamma < 2/\Delta^{\max}$, then there exist constants $C_0 > 0$ and $C_1 > 0$ such that the following holds,

$$\mathcal{T}_1 \leq C_0 + C_1/M\cdot\sum_{k=0}^{T}\mathbb{E}\Big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\Big]. \tag{75}$$

Proof: This proof can be done by following exactly the similar path of Lemma 7, and thus we omit the details. ■

Lemma 9: Suppose that the conditions on Lemma 8 are satisfied, then there exist constants $C_2 > 0$ and $C_3 > 0$ such that the following holds,

$$\mathcal{T}_2 \leq \gamma C_2 + \gamma C_3\cdot\sum_{k=0}^{T}\mathbb{E}\Big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\Big]. \tag{76}$$

Proof: We first recall that $\nabla f_\delta(\mathbf{P}_D^{k-})$ represents the stacked full gradient with respect to the delayed probability distributions $\mathbf{P}_D^{k-}$. Precisely, let us denote

$$\nabla f_\delta(\mathbf{P}_D^{k-}) = [\nabla_{\mathbf{p}_i} f(P_i^{k-})]_{i\in\mathcal{I}}, \tag{77}$$

where P_i^{k-} captures all the delayed distributions associated with the i-th agent. On this account, we can have

$$\begin{aligned}
&\|\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-})\|^2 \\
&= \sum_{i=1}^{I} \|\nabla_{\mathbf{p}_i} f(P^k) - \nabla_{\mathbf{p}_i} f(P_i^{k-})\|^2 \\
&\overset{(8a)}{\leq} \sum_{i=1}^{I} L_i^2 \|P^k - P_i^{k-}\|^2 \\
&= \sum_{i=1}^{I}\sum_{j\neq i}^{I} L_i^2 \|\mathbf{p}_j^k - \mathbf{p}_j^{k-\tau_{ij}}\|^2 \\
&\overset{(8b)}{\leq} \sum_{i=1}^{I}\sum_{j\neq i}^{I} L_i^2 \sum_{t=0}^{\tau_{ij}-1} \|\mathbf{p}_j^{k-t} - \mathbf{p}_j^{k-t-1}\|^2,
\end{aligned} \tag{78}$$

where $(8a)$ is due to the fact that each gradient $\nabla_{\mathbf{p}_i} f(P)$ is L_i-Lipschitz continuous and $(8b)$ comes from the triangle inequality. In addition, according to Lemma 4 and Cauchy-Schwartz inequality, it holds that

$$\begin{aligned}
&\|\mathcal{G}_\gamma\big(\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-}), \mathbf{p}_i^k\big)\|^2 \\
&\leq -\Big\langle \nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-}),\ \mathcal{G}_\gamma\big(\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-}), \mathbf{p}_i^k\big)\Big\rangle \\
&\leq \|\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-})\| \cdot \|\mathcal{G}_\gamma\big(\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-}), \mathbf{p}_i^k\big)\|,
\end{aligned} \tag{79}$$

and thus for $\forall k \in \mathbb{N}_+$,

$$\begin{aligned}
\|\mathbf{p}_i^{k+1} - \mathbf{p}_i^k\|^2 &= \gamma^2 \|\mathcal{G}_\gamma\big(\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-}), \mathbf{p}_i^k\big)\|^2 \\
&\leq \gamma^2 \|\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-})\|^2.
\end{aligned} \tag{80}$$

Note that the above inequality also shows that the gradient mapping is always bounded by the stochastic gradient. Next, based on the inequalities (78), (80) and the facts that the gradient $\nabla_{\mathbf{p}_i}\hat{f}_\delta(P_i^{k-})$ is bounded and $\tau_{ij} \leq D,\ \forall i,j \in \mathcal{I}$, we know that there must exist a constant $\beta > 0$ such that

$$\|\nabla f(P^k) - \nabla f_\delta(\mathbf{P}_D^{k-})\| \leq \gamma\beta. \tag{81}$$

Consequently, the summation term $\mathcal{T}_2$ can be bounded by

$$\mathcal{T}_2 \leq \gamma\beta \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|\big]. \tag{82}$$

Now, the rest of the proof follows the similar path of the proof in Lemma 7 (or Lemma 8). It can be shown that there exist two constants $\rho_0 > 0$ and $\rho_1 > 0$ such that

$$\begin{aligned}
&\sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\big] \\
&\geq \frac{1}{\rho_1} \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|\big] - \frac{\rho_0}{\rho_1}.
\end{aligned} \tag{83}$$

Combining the inequalities (82) and (83), we can have

$$\mathcal{T}_2 \leq \gamma\beta\rho_0 + \gamma\beta\rho_1 \cdot \sum_{k=0}^{T} \mathbb{E}\Big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\Big]. \tag{84}$$

Therefore, let $C_2 = \beta\rho_0$ and $C_3 = \beta\rho_1$ respectively, the proof is completed. ■

Next, we prove the statement in Theorem 2. Taking into account Lemma 8 and Lemma 9 together, the inequality (74) can be continued as

$$\begin{aligned}
&\mathbb{E}\big[f(P^{T+1})\big] - f(P^0) \\
&\geq (\gamma - \frac{\gamma C_1}{M} - \gamma^2 C_3 - \frac{\gamma^2 L}{2}) \cdot \sum_{k=0}^{T} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\big] \\
&\quad - \gamma C_0 - \gamma^2 C_2.
\end{aligned} \tag{85}$$

As a result, it holds that,

$$\begin{aligned}
&\sum_{k=0}^{\infty} \mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}_\delta(\mathbf{P}_D^{k-}), P^k\big)\|^2\big] \\
&\leq \frac{f(P^\star) - f(P^0) + \gamma C_0 + \gamma^2 C_2}{\gamma - \gamma C_1/M - \gamma^2 C_3 - \gamma^2 L/2}.
\end{aligned} \tag{86}$$

Therefore, if the sample-size M and step-size γ are chosen satisfying $M > C_1$ and $\gamma - \gamma C_1/M - \gamma^2 C_3 - \gamma^2 L/2 > 0$, then we can have that $\mathbb{E}\big[\|\tilde{\mathcal{G}}_\gamma\big(\nabla\hat{f}(P^k), P^k\big)\|^2\big]$ converges to zero; and furthermore, its running average converges at the rate of $\mathcal{O}(1/T)$.

REFERENCES

[1] Jason R Marden. The role of information in distributed resource allocation. *IEEE Transactions on Control of Network Systems*, 4(3):654–664, 2016.

[2] Matthew Streeter and Daniel Golovin. An online algorithm for maximizing submodular functions. In *Advances in Neural Information Processing Systems*, pages 1577–1584, 2009.

[3] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes: theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9:235–284, 2008.

[4] Syed Talha Jawaid and Stephen L Smith. Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems. *Automatica*, 61:282–288, 2015.

[5] Baharan Mirzasoleiman, Amin Karbasi, Rik Sarkar, and Andreas Krause. Distributed submodular maximization: identifying representative elements in massive data. In *Advances in Neural Information Processing Systems*, pages 2049–2057, 2013.

[6] Ashwinkumar Badanidiyuru, Baharan Mirzasoleiman, Amin Karbasi, and Andreas Krause. Streaming submodular maximization: massive data summarization on the fly. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 671–680, 2014.

[7] Nikolay Atanasov, Jerome Le Ny, Kostas Daniilidis, and George J Pappas. Decentralized active information acquisition: Theory and application to multi-robot slam. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4775–4782. IEEE, 2015.

[8] Brent Schlotfeldt, Dinesh Thakur, Nikolay Atanasov, Vijay Kumar, and George J Pappas. Anytime planning for decentralized multirobot active information gathering. *IEEE Robotics and Automation Letters*, 3(2):1025–1032, 2018.

[9] László Lovász. Submodular functions and convexity. In *Mathematical Programming the State of the Art*, pages 235–257. Springer, 1983.

[10] Maxim Sviridenko. A note on maximizing a submodular set function subject to a knapsack constraint. *Operations Research Letters*, 32(1):41–43, 2004.

[11] Niv Buchbinder, Moran Feldman, Joseph Seffi, and Roy Schwartz. A tight linear time (1/2)-approximation for unconstrained submodular maximization. *SIAM Journal on Computing*, 44(5):1384–1402, 2015.

[12] George L Nemhauser, Laurence A Wolsey, and Marshall L Fisher. An analysis of approximations for maximizing submodular set functions—i. *Mathematical programming*, 14(1):265–294, 1978.

[13] Marshall L Fisher, George L Nemhauser, and Laurence A Wolsey. An analysis of approximations for maximizing submodular set functions—ii. In *Polyhedral Combinatorics*, pages 73–87. Springer, 1978.

[14] Gruia Calinescu, Chandra Chekuri, Martin Pal, and Jan Vondrák. Maximizing a monotone submodular function subject to a matroid constraint. *SIAM Journal on Computing*, 40(6):1740–1766, 2011.

[15] Bahman Gharesifard and Stephen L Smith. Distributed submodular maximization with limited information. *IEEE Transactions on Control of Network Systems*, 5(4):1635–1645, 2017.

[16] David Grimsman, Mohd Shabbir Ali, Joao P Hespanha, and Jason R Marden. The impact of information in greedy submodular maximization. *IEEE Transactions on Control of Network Systems*, 6(4):1334–1343, 2019.

[17] Haoyuan Sun, David Grimsman, and Jason R Marden. Distributed submodular maximization with parallel execution. *arXiv preprint arXiv:2003.04364*, 2020.

[18] Guannan Qu, Dave Brown, and Na Li. Distributed greedy algorithm for multi-agent task assignment problem with submodular utility functions. *Automatica*, 105:206–215, 2019.

[19] Aryan Mokhtari, Hamed Hassani, and Amin Karbasi. Decentralized submodular maximization: Bridging discrete and continuous settings. In *International Conference on Machine Learning*, pages 3616–3625, 2018.

[20] Jiahao Xie, Chao Zhang, Zebang Shen, Chao Mi, and Hui Qian. Decentralized gradient tracking for continuous DR-submodular maximization. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 2897–2906, 2019.

[21] Saeed Ghadimi, Guanghui Lan, and Hongchao Zhang. Mini-batch stochastic approximation methods for nonconvex stochastic composite optimization. *Mathematical Programming*, 155(1-2):267–305, 2016.

[22] Sashank J Reddi, Suvrit Sra, Barnabas Poczos, and Alexander J Smola. Proximal stochastic methods for nonsmooth nonconvex finite-sum optimization. In *Advances in Neural Information Processing Systems*, pages 1145–1153, 2016.

[23] Zhize Li and Jian Li. A simple proximal stochastic gradient method for nonsmooth nonconvex optimization. In *Advances in Neural Information Processing Systems*, pages 5564–5574, 2018.

[24] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 322–331. IEEE, 1995.

[25] Gilles Stoltz. *Incomplete Information and Internal Regret in Prediction of Individual Sequences*. PhD thesis, 2005.

[26] Sébastien Bubeck, Nicolo Cesa-Bianchi, and Sham M Kakade. Towards minimax policies for online linear optimization with bandit feedback. In *Conference on Learning Theory*, pages 41–1, 2012.

[27] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[28] Tim Roughgarden. Intrinsic robustness of the price of anarchy. In *Proceedings of the Forty-First Annual ACM symposium on Theory of Computing*, pages 513–522, 2009.

[29] Vasilis Syrgkanis, Alekh Agarwal, Haipeng Luo, and Robert E Schapire. Fast convergence of regularized learning in games. In *Advances in Neural Information Processing Systems*, pages 2989–2997, 2015.

[30] Thodoris Lykouris, Vasilis Syrgkanis, and Éva Tardos. Learning and efficiency in games with dynamic population. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 120–129. SIAM, 2016.

[31] Dylan J Foster, Zhiyuan Li, Thodoris Lykouris, Karthik Sridharan, and Eva Tardos. Learning in games: Robustness of fast convergence. In *Advances in Neural Information Processing Systems*, pages 4734–4742, 2016.

[32] Jianxin Ma, Chang Zhou, Peng Cui, Hongxia Yang, and Wenwu Zhu. Learning disentangled representations for recommendation. In *Advances in Neural Information Processing Systems*, pages 5711–5722, 2019.

[33] Yongfeng Zhang, Xu Chen, et al. Explainable recommendation: A survey and new perspectives. *Foundations and Trends® in Information Retrieval*, 14(1):1–101, 2020.

[34] F Maxwell Harper and Joseph A Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems (TIIS)*, 5(4):1–19, 2015.

[35] John Duchi, Shai Shalev-Shwartz, Yoram Singer, and Tushar Chandra. Efficient projections onto the $\mathcal{L}_1$-ball for learning in high dimensions. In *Proceedings of the 25th International Conference on Machine learning*, pages 272–279, 2008.

[36] Weiran Wang and Miguel A Carreira-Perpinán. Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application. *arXiv preprint arXiv:1309.1541*, 2013.