Controlled Sequential Estimation in Networked Data

Saurabh Sihag, Javad Heydari, and Ali Tajer

Abstract

This paper considers sequential estimation of a common random parameter in a network of interconnected agents constantly generating data over time. The objective is to design data-adaptive and sequential data-acquisition and decision-making processes that furnish all the agents with sufficiently reliable estimates of the shared parameter in the quickest fashion. A proper estimation cost function is adopted in order to signify the fidelity of the estimates to the ground truth, and to ensure consistency in the estimates of the interacting (neighboring) agents. By imposing practical constraints on the number of data points that the network affords to process at any time instant, the two intertwined data-acquisition and decision-making processes consist of (i) dynamically deciding about the stopping time of the process, at which point data-acquisition is terminated and the estimates are formed; (ii) dynamically selecting a subset of agents and collecting their measurements; and (iii) designing the optimal estimators. Different estimation rules are designed for different agents based on their local data and connectivity models. These estimators in conjunction with the proposed stopping time rule and a dynamic rule for selecting the agents over time are shown to admit weak asymptotic point-wise optimality.

1 Introduction

In this we paper we consider the problem of *sequentially* estimating a shared parameter in a large-scale and interconnected network of K agents. In this framework, the data is collected sequentially over time up to a stochastic stopping time, at which no further measurements are collected and the desired estimates are formed. Such a process aims to form reliable estimates with the minimal number of measurements (stopping time). In order to control the costs of collecting and processing data, at every given time the measurements from only a limited number of agents are processed. Due to varying degrees of connectivity of different agents as well as varying levels of interaction among them, the measurements from different agents are not expected to be equally informative about the shared parameter. As a result, selecting the sequence of agents whose data should be collected over time, which constitutes the control actions, are not necessarily statistically independent. Hence, the desired sequential estimation over time in such a network can be abstracted as a controlled sensing framework for an estimation objective, in which the control actions are possibly co-dependent.

The estimation objective of this framework consists of providing each agent with an estimate of the ground truth such that all the estimates exhibit a sound level of fidelity, and also the estimates obtained by interacting (neighboring) agents are more consistent compared to the entire range of estimates. To ensure both of these estimation objectives, we adopt an estimation cost function that assigns dedicated cost terms corresponding to each objective. Based on this cost function, we characterize estimators for all sensors as well as a decision rule for selecting and sampling agents over time, the combination of which in conjunction with a proposed stopping time is shown to achieve weak asymptotic point-wise optimality.

The existing literature on parameter estimation often focuses on settings in which the estimation strategy and data-acquisition processes are decoupled, and their focus is placed on forming the most reliable estimates based on a given set of measurements. In such strategies, the number of measurements to be collected, as well as their order should be specified. In contrast, the key element of the framework considered in this paper is the inclusion of dynamic agent selection, which can be modeled as co-dependent control actions. This is primarily motivated by controlling the costs associated with data acquisition and processing.

Forming an optimal estimate by using all the measurements collected in the network is computationally prohibitive, especially in a sequential setting. Moreover, under such a data-acquisition model, the problem simply reduces to existing ones analyzed in [1] and [2]. On the other extreme, forming local estimates at each agent solely based on their the local measurements available to the agent, while being computationally efficient, it is agnostic to the structure of the network, it is inherently suboptimal, and it incurs considerable estimation costs [3]. Therefore, this paper focuses on a setting that considers a balance between these two extremes. Specifically, in the framework considered, every agent constantly shares its estimate with the rest of the network. In this process, when a new measurement is collected by one agent, first that agent updates its estimate based on all the data it has received over time, as well as the most recent set of estimates it has received from other sensors.

When the sampling order is pre-specified, determining the optimal sampling strategy reduces to minimizing the number of measurements, which is well-investigated [2], [4], and [5]. However, incorporating dynamic decisions about the order of sampling, especially in networked data, is less-investigated. The sequential estimation of multiple unknown parameters by observing multiple sequences of independent random variables is studied in [6], where at each time instant, one of a finite number of independent actions (sequences) can be selected. In [6] each action depends only on one of the unknown parameters, while [7] generalizes the results to the setting in which actions depend on common unknown parameters. The problem considered in this paper is closely related to [7] where, at each time instant and based on the collected information up to that time, one of a finite number of possible actions is taken. However, the presence of multiple estimators, each one of which requiring to make reasonable estimate of the unknown parameter, makes the problem of this paper fundamentally different from [7]. Furthermore, the control actions in our framework are fundamentally co-dependent, while the framework in [7] considers independent control actions.

Controlled inference has also been studied in [**?**, **?**, **?**, **?**] in domains other than sequential estimation. The studies in [**?**, **?**, **?**] focus on sequential hypothesis testing, with controlled actions. In [**?**], a controlled sensing based approach is adopted for graph classification in terms of its connectivity based on the data sampled at its nodes. In another direction, the problem of sequential joint detection and estimation without any controlled action is studied in [8] and [9]. The major distinction of this paper from [8] and [9] is that, besides having a detection action upon stopping, in these studies the observation process is fixed and the only dynamic of the sampling process is the stopping time. Furthermore, there exist two sequences to observe and both of them are being observed at each time, while in this paper multiple sequences are available and only one of them is observed at each time instance.

2 Controlled Sensing Model

2.1 Data Model

Consider a large-scale network of K interconnected agents forming an undirected graph $\mathcal{G}(\mathcal{U},\mathcal{E})$, where $\mathcal{U} \triangleq \{1,\dots,K\}$ and $\mathcal{E} \subseteq \mathcal{U}\times\mathcal{U}$ are the set of nodes and edges of the graph, respectively. Each agent is equipped with a agent that can constantly collect data over time with the ultimate objective of forming a reliable estimate for the common unknown parameter $X \in \mathbb{R}$. Due to the scale and the associated cost of sensing, at each time instant only ℓ number of agents can make a measurement[1]. Hence, the measurements are made sequentially over time. Define $u(n) \in \mathcal{U}$ as the index of agent that makes the measurement at time $n \in \mathbb{N}$. Hence, the measurement of agent $i \in \mathcal{U}$ at time $n \in \mathbb{N}$ is

$$Y_n^i = \begin{cases} X + N_n^i & \text{if } u(n) = i \\ \emptyset & \text{if } u(n) \neq i \end{cases} , \tag{1}$$

where the convention $\emptyset$ denotes lack of a measurement, and N_i^j for any $i \in \mathcal{U}$ accounts for independent and identically distributed noise. Accordingly, we define $\mathbf{Y}_n^i \triangleq [Y_1^i,\dots,Y_n^i]$ as the vector of measurements collected by agent $i \in \mathcal{U}$ upto time n. The probability distribution function (pdf) of the scalar parameter X is denoted by π, and when $Y_n^i \neq \emptyset$, we denote the pdf of Y_n^i by f_i.

2.2 Sensing Model

With the ultimate of objective of forming a reliable estimate for X, the sensing mechanism sequentially collects measurements. The information generated sequentially up to time t generates the filtration $\{\mathcal{F}_n : n \in \mathbb{N}\}$ where

$$\mathcal{F}_n \triangleq \sigma\big(Y_1^{u(1)}, Y_2^{u(2)}, \dots, Y_n^{u(n)}\big) . \tag{2}$$

Based on the data acquisition process at each time n, and based on filtration $\mathcal{F}_n$ consists in dynamically making three intertwined decisions. Specifically, it (i) decides to whether continue taking further measurements or stop, (ii) when the decision is to continue, specify the agent that should collect the next measurement, and (iii) when the decision is to stop, form an estimate for X.

To formalize this, we define $N \in \mathbb{N}$ as the stochastic stopping time of the process. Dynamic selection of agents at time n can be abstracted by a control policy, denoted by μ, that

[1]Without loss of generality and for the convenience in notations in this paper we assume $\ell = 1$

leverages all the past observations, denoted by $\mathcal{Y}^{n-1} \triangleq [\boldsymbol{Y}^1_{n-1}, \dots, \boldsymbol{Y}^K_{n-1}]$, and all the past control actions, denoted by $\mathcal{U}^{n-1} \triangleq \{u(1), \dots u(n-1)\}$ to determine $u(n)$, i.e.

$$\mu : \mathcal{U}^{n-1} \times \mathcal{Y}^{n-1} \to \mathcal{U}. \tag{3}$$

2.3 Estimation Model

Since all agents are estimating a shared parameter, at every time instant n by collecting a fresh measurement $Y_n^{u(n)}$, all agents can potentially update their estimates. We denote the estimate of X formed by agent i at time n by X_n^i. Accordingly, we define the set $\mathcal{X}_n \triangleq \{X_n^1, \dots, X_n^K\}$. We assume that every time the estimates are updated, each agent shares its estimate with the rest of the network. When at time n agent i is observed, i.e., $u(n) = i$, this agent leverages they data it has received over time up to time n, i.e., $\boldsymbol{Y}_n^i$, as well as the most recent set of estimates it has received from other agents, i.e., $\mathcal{X}_{n-1}$. Hence, the estimator at agent i can be abstracted by $g_i : \mathbb{R}^{K+|\boldsymbol{Y}_n^i|} \to \mathbb{R}$, i.e.,

$$X_n^i = g_i(\mathcal{X}_{n-1}, \boldsymbol{Y}_n^i) \ , \quad \forall i \in \mathcal{U} \ . \tag{4}$$

Once this estimate is formed, all other agents $j \neq i$, which have not received fresh data at time n, leverage X_n^i to update their estimates by using the estimator $h_j : \mathbb{R}^{K+1} \to \mathbb{R}$, i.e.,

$$X_n^j = h_j(\mathcal{X}_{n-1}, X_n^i) \ , \quad \forall j \in \mathcal{U} \backslash \{i\} \ . \tag{5}$$

Forming the optimal estimates over time can be abstracted by optimally designing the set of estimators $\{g_i : i \in \mathcal{U}\}$ and $\{h_i : i \in \mathcal{U}\}$.

3 Problem Formulation

We aim to design a data acquisition policy μ, the set of estimators $\{g_i : i \in \mathcal{U}\}$, $\{h_i : i \in \mathcal{U}\}$, and the stopping time N. Designing the optimal strategy for achieving reliable estimates involves striking a balance between two opposing figures of merit pertinent to the quality of the estimates on the one hand, and the aggregate number of measurements made, on the other hand. The quality of the process is captured by the reliability of the estimates at individual agents. Also, we assume that each agent desires to have its estimate remains consistent with and in close proximity of those of its neighbors. To accommodate this, we propose to use the following cost function for quantifying the estimation quality of the network at any time n.

$$\begin{aligned} J(n, \mu) \triangleq & \sum_{i=1}^{K} \mathbb{E}[(X - X_n^i)^2 \mid \boldsymbol{Y}_n^i, \, \mathcal{U}^n] \\ & + \frac{1}{2} \sum_{i=1}^{K} \sum_{j \in B_i} ||X_n^i - X_n^j||^2 \ , \end{aligned} \tag{6}$$

where $B_i \triangleq \{j \,:\, (i,j) \in \mathcal{E}\}$ is the set of neighbors of agent i. Based on this estimation cost function, the term $\mathbb{E}[(X - X_n^i)^2 \mid \boldsymbol{Y}_n^i, \mathcal{U}^n]$ ensures that each agent forms an estimate of X with high fidelity, and $||X_n^i - X_n^j||^2$ guarantees that the estimates of neighboring agents stay consistent and close. By integrating both estimation and the sampling costs into a unified cost function, for the unified cost when the stopping time is $N = n$ we have

$$J_2(n, \mu, c) \triangleq J(n, \mu) + nc \;, \tag{7}$$

where $c > 0$ is cost per sample and controls the balance between the quality and the agility of the process. In this paper, we utilize the results on the asymptotic efficiency of Bayesian estimators and maximum likelihood estimators, as well as the asymptotic normality of their posterior distributions. For this purpose, define $I_i(X)$ as the Fisher information corresponding to the likelihood of measurements at agent i, i.e.,

$$I_i(X) \triangleq -\mathbb{E}\left[\frac{\partial^2 \log f_i(Y \mid X)}{\partial X^2}\right] \;. \tag{8}$$

In this paper, we use several results from [2] and [7] that require asymptotic consistency and asymptotic normality of the local maximum likelihood estimators to hold. Define $X^i_{n,\mathsf{mle}}$ as the maximum likelihood estimate corresponding to agent $i \in \{1, \dots K\}$ at time n, i.e.,

$$X^i_{n,\mathsf{mle}} \triangleq \arg\max_X f_i(\boldsymbol{Y}_n^i \mid X, \mathcal{U}^n) \;. \tag{9}$$

We assume that all the assumptions for the asymptotic normality and consistency of maximum likelihood estimators listed in [7] are satisfied. The notion of asymptotically point-wise optimal sequential procedures are explored for simpler settings in [2] and [6], which is further extended to weak asymptotically point-wise optimality in [7]. In this paper, the measurements collected in the network are conditionally independent and can have different distributions depending on the choice of agents determined by the control actions. The data model in this paper is similar to the one in [7], where the distribution of the conditionally independent measurements is dependent on the control actions. Therefore, we focus on the design of sequential decision rules that provide similar guarantees on the optimality as the ones developed in [7]. We first characterize the properties of the stopping time and control policy that conform to the notion of weak asymptotically point-wise optimality.

Definition 1. *The control policy μ and the stopping rule N are said to be weak asymptotically point-wise optimal if*

$$\lim_{c \to 0} \mathbb{P}\left(\frac{J_2(N, \mu, c)}{J_2(N', \mu', c)} \leq 1 + \epsilon\right) \to 1 \;, \forall \epsilon > 0 \;, \mu',\; N' \;. \tag{10}$$

In this paper, we use [7, Theorem 3.1] to establish the sufficient condition for weak asymptotically point-wise optimality of the stopping rule and control policy. Theorem 1 provides the sufficient conditions for weak asymptotically point-wise optimal decision rules.

Theorem 1. *For a sampling policy μ^*, define the stopping rule as*

$$N^*(c) = \inf\{n : J(n, \mu^*) \leq (n+1)c\} \ . \tag{11}$$

The sampling policy μ^* and stopping rule N^* are concluded to be weak asymptotically pointwise optimal as defined in Definition 1, if the following conditions are satisfied:

$$\mathbb{P}(n \inf_{\mu} J(n, \mu) > 0) = 1 \ , \tag{12}$$

and

$$\lim_{n\to\infty} \mathbb{P}(n \inf_{\mu} J(n, \mu) \geq Q(X_0) - \epsilon) \xrightarrow{\mathsf{P}(X_0)} 1 \ , \quad \forall \epsilon > 0 \ , \tag{13}$$

and the sampling policy μ^* satisfies

$$nJ(n, \mu^*) \xrightarrow{\mathsf{P}(X_0)} Q(X_0) \ , \tag{14}$$

where $Q(X_0)$ is a positive random variable and the convergence is under $\mathsf{P}(X_0)$, which is the probability distribution of measurements when the true value of X is X_0.

Proof. Proof follows directly from the proof of [7, Theorem 3.1]. □

4 Sequential Decision Rules

In this section, we first provide the design of the estimators $\mathcal{X}_n$. Under this estimation strategy, we also provide the control policy μ^* and establish its weak asymptotically pointwise optimality.

4.1 Optimal Estimators

We provide the design of the optimal estimators under the assumption that any agent cannot evaluate the joint distribution of the observations from other agents and its own observations[2]. In this paper, we assume that every agent shares its estimates with all other agents, which implies that every agent has access to $\mathcal{X}_{n-1}$ when the sample at time n from the agent determined by the control action $u(n)$. Given $u(n) = i$, the estimate X_n^i is first updated by using the function g_i. The optimal structure of g_i is characterized by a linear function of the estimates $\mathcal{X}_{n-1}$ aggregated with the local posterior mean of the observations at agent i. For every agent $j \in \mathcal{U}\backslash\{i\}$, the estimate X_n^j is updated by the function h_j, where the optimal h_j is characterized by a linear combination of the estimates $\mathcal{X}_{n-1}$ and the updated estimate X_n^i. Theorem 2 provides the optimal design of the functions g_i and h_j.

[2]Optimal estimates for any reasonable estimation cost function without this assumption will be a function of the joint likelihood of observations available at all agents, leading to the same drawbacks as mentioned before.

Theorem 2. *The optimal design of the estimators* $g_i(\mathcal{X}_{n-1}, \mathbf{Y}_n^i), i = u(n)$ *and* $h_j(\mathcal{X}_{n-1}, X_n^i), \forall j \in \mathcal{U}\backslash\{i\}$ *that minimize the cost* $J(n,\mu)$ *for* $i = u(n)$ *is given by*

$$X_n^i = \sum_{\substack{m=1\\m\neq i}}^{K} a_i^m((|B_m|+1)X_{n-1}^m - \sum_{l\in B_m} X_{n-1}^l) \tag{15}$$

$$+ a_i^i\mathbb{E}[X \mid \mathbf{Y}_n^i, \mathcal{U}^n] , \tag{16}$$

and for all $j \in \mathcal{U}\backslash\{i\}$ *is given by*

$$X_n^j = \sum_{\substack{l=1\\l\neq i}}^{K} \left(a_j^l - \frac{a_j^i}{a_i^i}a_i^l\right)((|B_l|+1)X_{n-1}^l - \sum_{m\in B_l} X_{n-1}^m) + \frac{a_j^i}{a_i^i}X_n^i , \tag{17}$$

where a_i^j *is the* (i,j)*-th term of the matrix* A^{-1}*, where* A *is a* $K\times K$ *matrix, such that,*

$$\begin{cases} A_{kk} = (|B_k|+1), & \forall k \in \mathcal{U} \\ A_{kl} = -1 & \textit{iff } l \in B_k, \quad \forall k,l \in \mathcal{U} \\ A_{kl} = 0, & \textit{otherwise} \end{cases} , \tag{18}$$

where A_{kl} *is the* (k,l)*-th element of* A.

Proof. See Appendix A. □

4.2 Stopping Rule and Control Policy

In this paper, we adopt the stopping rule used in Theorem 1, in which we compare the estimation cost $J(n,\mu)$ with the total sampling cost $(n+1)c$ and decide to stop taking new measurements according to the rule given by

$$N^*(c) = \inf\{n : J(n,\mu^*) \le (n+1)c\} . \tag{19}$$

Next, we establish that the condition (13) in Theorem 1 is satisfied in the following Lemma. Define ϕ^K as a $(K-1)-$dimensional probability simplex, and define $\boldsymbol{p} \triangleq \{p_1,\dots,p_K\}$, where p_i is the probability of $u(n) = i \in \mathcal{U}$.

Lemma 1. *For the given cost function* $J(n,\mu)$ *defined in* (6),

$$\lim_{n\to\infty} \mathbb{P}(n \inf_{\mu} J(n,\mu) \ge Q(X) - \epsilon) \xrightarrow{\mathsf{P}(X)} 1 , \quad \forall\epsilon > 0 , \tag{20}$$

where

$$Q(X) = \inf_{\boldsymbol{p}\in\phi^K} \sum_{i=1}^{K} Q_i(X,\boldsymbol{p}) + \frac{1}{2}\sum_{i=1}^{K}\sum_{j\in B_i} Q_{ij}(X,\boldsymbol{p}) , \tag{21}$$

and

$$Q_i(X,\boldsymbol{p}) = \sum_{j=1}^{K} \frac{(a_i^j)^2}{p_j I_j(X)} , \quad \textit{and} \quad Q_{ij}(X,\boldsymbol{p}) = \sum_{l=1}^{K} \frac{(a_i^l - a_j^l)^2}{p_l I_l(X)} . \tag{22}$$

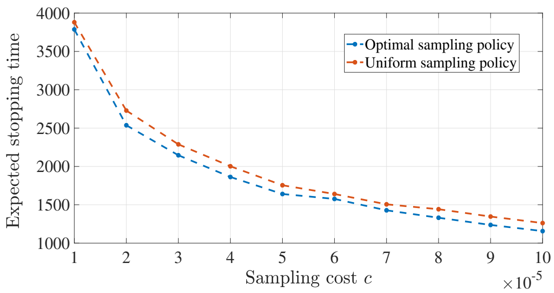

Figure 1: Expected Stopping Time vs Sampling Cost c

Proof. See [?, Appendix B]. □

Based on these, we devise the following control policy μ^*, the optimality of which is demonstrated in the ensuing theorem. At a time n, if the stopping rule is not satisfied, the $(n+1)$-th measurement is taken from a agent i with probability p_i^n, where the probabilities $\boldsymbol{p}^n = \{p_1^n, \ldots, p_K^n\}$ are determined by solving

$$\begin{aligned}\boldsymbol{p}^n = \underset{\boldsymbol{p}\in\phi^K}{\operatorname{argmin}} &\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(a_i^j)^2}{p_j I_j(X_{j,\mathsf{mle}}^n)} \\ &+ \frac{1}{2}\sum_{i=1}^{K}\sum_{j\in B_i}\left(\sum_{l=1}^{K}\frac{(a_i^l - a_j^l)^2}{p_l I_l(X_{j,\mathsf{mle}}^n)}\right) \,. \end{aligned} \tag{23}$$

Theorem 3. *The control policy μ^* applied using* (23) *satisfies the condition* (14) *in Theorem 1. Therefore, the control policy μ^* and the stopping rule $N^*(c)$ defined in Theorem 1 are weak asymptotically point-wise optimal.*

Proof. See [?, Appendix C]. □

5 Numerical Evaluation

In this section, we illustrate the application of the estimation strategy designed in Section 4.1 and the controlled sequential estimation procedure on a network consisting of 5 agents. We assume that the unknown parameter X is distributed according to the distribution $\mathcal{N}(0, 10)$ and the measurements at agent $i \in \{1, \ldots, 5\}$ are affected by i.i.d. noise distributed according to $\mathcal{N}(0, \sigma_i^2)$, where $\sigma_1^2 = 1, \sigma_2^2 = 8, \sigma_3^2 = 1, \sigma_4^2 = 8, \sigma_5^2 = 4$.

5.1 Estimation Strategy

We first illustrate the gain in the local estimation quality at each agent for the proposed estimation strategy as compared to using locally optimal estimates at each agent. For a fair comparison between the two, we assess multiple Monte Carlo realizations, in which we

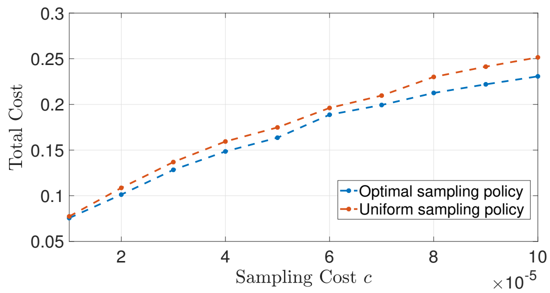

Figure 2: Total cost vs Sampling Cost c

generate X randomly and subsequently the observations at each agent according to their pdfs and evaluate the approximation of the root mean squared error metric RMSE_i at agent i. For the evaluation of the estimation strategy, we set the number of samples at each agent to n_0. We include the results for agent 2 and agent 5 in this section.

Table 1: Comparison of estimation strategy in this paper with locally optimal estimation strategy

n_0	RMSE_2 (Aggregate)	RMSE_2 (Local)	RMSE_5 (Aggregate)	RMSE_5 (Local)
1	1.1737	2.1037	1.4073	1.7059
2	0.8573	1.6871	1.00	1.2945
3	0.7063	1.4389	0.8229	1.0685
4	0.6155	1.2980	0.7275	0.9698
5	0.5426	1.1610	0.6448	0.8670

Table 1 shows the improvement in local estimation qualities at agents 2 and 5 in comparison to that from an estimation strategy based on forming optimal Bayesian estimates from the local measurements.

5.2 Controlled Sequential Estimation

The asymptotically optimal sampling policy and stopping rule are evaluated on the network. We evaluate the stopping times for the sampling strategy defined in this paper and compare the performance with another sampling strategy that chooses the agents randomly with a uniform distribution. Note that the worst case sampling strategy will consist of at least one agent not being sampled at all, which accounts for a very large stopping time as compared to the sampling policies evaluated in this section. Fig. 1 shows the expected stopping time (evaluated by averaging over multiple Monte Carlo runs) versus the sampling cost c and Fig. 2 plots the total cost at stopping time versus c. As expected, the asymptotically optimal sampling policy outperforms the uniform sampling policy in terms of stopping time as well as total cost at the stopping time. In Fig. 3, we compare the estimation performance

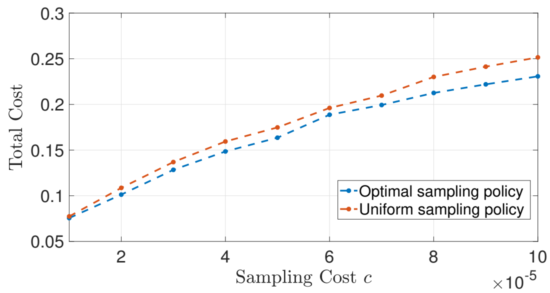

Figure 3: Estimation cost vs Number of samples for fixed sample size

at the end of a fixed number of measurements for the sequential decision rules and a policy which randomly distributes the sampling instants among the agents.

6 Conclusion

In this paper, we have considered the problem of sequential estimation in a network of interconnected agents with the objective to design data adaptive sequential decision rules that lead to sufficiently good estimation quality at all agents with minimal number of measurements. We have adopted a modified cost function to reflect the fidelity of the estimates at all agents and consistency among the estimates of neighboring agents. Under the constraint on the number of measurements that the network can collect at any instant, we have provided the design for the sequential decision rules consisting of the stopping rule to decide when to stop collecting measurements, control policy to select the subset of agents for collection of new measurements, and optimal estimator design. We have shown that these sequential decision rules admit the weak asymptotic point-wise optimality.

A Proof of Theorem 2

Under the constraint that the joint distribution of the measurements cannot be formed by any agent, the partial derivative of $J(n, \mu)$ with respect to the estimator X_n^i is given by

$$\begin{aligned}\frac{\partial J(n,\mu)}{\partial X_n^i} &= 2X_n^i - 2\mathbb{E}[X \mid \mathbf{Y}_i^n, \mathcal{U}^n] + 2|B_i|X_n^i - 2\sum_{j\in B_i} X_n^j \,, \\ &= 2(|B_i|+1)X_n^i - 2\sum_{j\in B_i} X_n^j - 2\mathbb{E}[X \mid \mathbf{Y}_n^i, \mathcal{U}^n] \,. \end{aligned} \tag{24}$$

Setting the gradient of $J(n, \mu)$ with respect to the estimators $\mathcal{X}_n = [X_n^1, \dots, X_n^K]$ to 0 leads to the system of K linearly independent equations, given by

$$(|B_i|+1)X_n^i - \sum_{j\in B_i} X_n^j = \mathbb{E}[X \mid \mathbf{Y}_n^i, \mathcal{U}^n] \,, \quad i \in \{1, \dots, K\} \tag{25}$$

The system of equations in (25) can be represented in the matrix form as

$$A[X_n^i] = [\mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n]] \ , \tag{26}$$

where A is a $K \times K$ matrix, such that $A_{kk} = (|B_k|+1), \forall k \in \mathcal{U}$, $A_{kl} = -1$ iff $l \in B_k, \forall k, l \in \mathcal{U}$ and $A_{kl} = 0$, otherwise. Also,

$$[X_n^i] = [X_n^1, \ldots, X_n^K]^\mathsf{T} \ , \tag{27}$$

and

$$[\mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n]] = \left[\mathbb{E}[X \mid \boldsymbol{Y}_n^1, \mathcal{U}^n], \ldots, \mathbb{E}[X \mid \boldsymbol{Y}_n^K, \mathcal{U}^n]\right]^\mathsf{T} \ . \tag{28}$$

Note that the matrix A satisfies the following equation

$$A\mathbb{1}_{K\times 1} = \mathbb{1}_{K\times 1} \ , \tag{29}$$

where $\mathbb{1}_{K\times 1}$ is a $K \times 1$ matrix of 1's. Therefore,

$$\mathbb{1}_{K\times 1} = A^{-1}\mathbb{1}_{K\times 1} \ , \tag{30}$$

which implies that the sum of the elements of rows of A^{-1} is 1. Since,

$$[X_i^n] = A^{-1}[\mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n]] \ , \tag{31}$$

it can be concluded that the estimators $[X_n^i]$ that satisfy (26) are of the form

$$X_n^i = \sum_{j=1}^{K} a_i^j \mathbb{E}[X \mid \boldsymbol{Y}_n^j, \mathcal{U}^n] \ , \tag{32}$$

where a_i^j is the (i,j)-th element of A^{-1}. Furthermore, since the Hessian of $J(n,\mu)$ is positive semi-definite, the estimators obtained from (31) are optimal. Note that at any instant, only one agent collects a new measurement. Therefore,

$$\mathbb{E}[X \mid \boldsymbol{Y}_n^j, \mathcal{U}^n] = \mathbb{E}[X \mid \boldsymbol{Y}_{n-1}^j, \mathcal{U}^{n-1}], \quad \forall j \neq u(n) \ , \tag{33}$$

which allows for forming the update rules for $\mathcal{X}_n$ in terms of $\mathcal{X}_{n-1}$ and the set of observations $\boldsymbol{Y}_n^i$, where $i = u(n)$. For $i = u(n)$,

$$X_n^i = \sum_{\substack{m=1 \\ m\neq i}}^{K} a_i^m \mathbb{E}[X \mid \boldsymbol{Y}_{n-1}^m, \mathcal{U}^{n-1}] + a_i^i \mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n] \ . \tag{34}$$

From (25),

$$X_n^i = \sum_{\substack{m=1 \\ m\neq i}}^{K} a_i^m ((|B_m|+1)X_{n-1}^m - \sum_{l\in B_m} X_{n-1}^l) + a_i^i \mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n] \ . \tag{35}$$

The function $g_i(\mathcal{X}_{n-1}, \boldsymbol{Y}_n^i)$ is determined by (35). For $j \neq u(n)$,

$$X_n^j = \sum_{\substack{l=1 \\ l \neq i}}^{K} a_j^l \mathbb{E}[X \mid \boldsymbol{Y}_{n-1}^l, \mathcal{U}^{n-1}] + a_j^i \mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n] \ . \tag{36}$$

From (25) and (35),

$$\begin{aligned} X_n^j &= \sum_{\substack{l=1 \\ l \neq i}}^{K} a_j^l ((|B_l| + 1) X_{n-1}^l - \sum_{m \in B_l} X_{n-1}^m) \\ &+ \frac{a_j^i}{a_i^i} X_n^i - \frac{a_j^i}{a_i^i} \sum_{\substack{l=1 \\ l \neq i}}^{K} a_i^l ((|B_l| + 1) X_{n-1}^l - \sum_{m \in B_l} X_{n-1}^m) \ , \\ &= \sum_{\substack{l=1 \\ l \neq i}}^{K} \left(a_j^l - \frac{a_j^i}{a_i^i} a_i^l \right) ((|B_l| + 1) X_{n-1}^l - \sum_{m \in B_l} X_{n-1}^m) + \frac{a_j^i}{a_i^i} X_n^i \ . \end{aligned} \tag{37}$$

Therefore, the function $h_j(\mathcal{X}_{n-1}, X_n^i)$ is determined by (37).

B Proof of Lemma 1

Note that

$$\begin{aligned} \sqrt{n}(X_n^i - X) &= \sqrt{n}(\sum_{j=1}^{K} a_i^j (\mathbb{E}[X \mid \boldsymbol{Y}_n^j, \mathcal{U}^n] - X) \ , & (38) \\ &= \sum_{j=1}^{K} a_i^j \left(\sqrt{n} (\mathbb{E}[X \mid \boldsymbol{Y}_n^j, \mathcal{U}^n] - X) \right) \ , & (39) \\ &= \sum_{j=1}^{K} a_i^j \left(\frac{\sqrt{n_i}}{\sqrt{n_i / n}} (\mathbb{E}[X \mid \boldsymbol{Y}_n^j, \mathcal{U}^n] - X) \right) \ . & (40) \end{aligned}$$

where $n_i = \sum_{m=1}^{n} \mathbb{1}_{u(m)=i}$. For any measurable control policy μ, $n_i \to \infty$ as $n \to \infty$, such that

$$\lim_{\substack{n_i \to \infty \\ n \to \infty}} \frac{n_i}{n} = q_i^\mu \ , \tag{41}$$

and $\sum_{i=1}^{K} q_i^\mu = 1$. Define $\boldsymbol{q}^\mu = \{q_1^\mu, \ldots, q_K^\mu\}$. From asymptotic efficiency of Bayesian estimators,

$$\lim_{n \to \infty} \sqrt{n} (\mathbb{E}[X \mid \boldsymbol{Y}_n^i, \mathcal{U}^n] - X) \xrightarrow{\mathsf{P}(X)} \mathcal{N} \left(0, \frac{1}{q_i^\mu I_i(X)} \right) \ . \tag{42}$$

By using (42) and (38), it can be easily concluded that

$$\lim_{n\to\infty} \sqrt{n}(X_n^i - X) \xrightarrow{\mathsf{P}(X)} \mathcal{N}\left(0, \sum_{j=1}^{K} \frac{(a_i^j)^2}{q_j^\mu I_j(X)}\right) . \tag{43}$$

Therefore, (43) provides an equivalent of a Cramer Rao Lower bound for the estimator X_i^n. Define $R_i(n,\mu)$ as $\mathbb{E}[(X_n^i - X)^2 \mid \boldsymbol{Y}_n^i, \mathcal{U}^n]$ for $i \in \mathcal{U}$. It can be concluded that

$$\mathbb{P}\left(nR_i(n,\mu) \geq \sum_{j=1}^{K} \frac{(a_i^j)^2}{q_j^\mu I_j(X)}\right) = 1 . \tag{44}$$

Therefore, as $n \to \infty$,

$$\mathbb{P}\left(nR_i(n,\mu) \geq \sum_{j=1}^{K} \frac{(a_i^j)^2}{q_j^\mu I_j(X)} - \epsilon\right) = 1 , \quad \forall \epsilon > 0 . \tag{45}$$

Define $Q_i(X, \boldsymbol{q}^\mu) = \sum_{j=1}^{K} \frac{(a_i^j)^2}{(q_j^\mu) I_j(X)}$, for $i \in \mathcal{U}$. We follow same steps to evaluate the asymptotic behavior of the terms $R_{ij} = ||X_n^i - X_n^j||^2$. Note that

$$\sqrt{n}(X_n^i - X_n^j) = \sqrt{n}(X_n^i - X) - \sqrt{n}(X_n^j - X) , \tag{46}$$

$$= \sqrt{n} \sum_{l=1}^{K} (a_i^l - a_j^l)(\mathbb{E}[X \mid \boldsymbol{Y}_n^l] - X) . \tag{47}$$

Using (42), it can be concluded that

$$\lim_{n\to\infty} \sqrt{n}(X_n^i - X_n^j) \xrightarrow{\mathsf{P}(X)} \mathcal{N}\left(0, \sum_{l=1}^{K} \frac{(a_i^l - a_j^l)^2}{q_l^\mu I_l(X)}\right) . \tag{48}$$

Also, define

$$Q_{ij}(X, \boldsymbol{q}^\mu) = \sum_{l=1}^{K} \frac{(a_i^l - a_j^l)^2}{q_l^\mu I_l(X)} . \tag{49}$$

Therefore, using the similar arguments as before,

$$\mathbb{P}(nR_{ij} \geq Q_{ij}(X) - \epsilon) = 1 , \quad \forall \epsilon > 0 . \tag{50}$$

On combining the results from (45) and (50),

$$\mathbb{P}\left(nJ(n,\mu) \geq \sum_{i=1}^{K} Q_i(X, \boldsymbol{q}^\mu) + \frac{1}{2} \sum_{i=1}^{K} \sum_{j \in B_i} Q_{ij}(X, \boldsymbol{q}^\mu) - \epsilon\right) = 1 , \quad \forall \epsilon > 0 . \tag{51}$$

Therefore,

$$\mathbb{P}\left(n \inf_{\mu} J(n,\mu) \geq Q(X) - \epsilon\right) = 1, \quad \forall \epsilon > 0\ , \tag{52}$$

where

$$Q(X) = \inf_{\boldsymbol{p}\in\phi^K} \sum_{i=1}^{K} Q_i(X,\boldsymbol{p}) + \frac{1}{2}\sum_{i=1}^{K}\sum_{j\in B_i} Q_{ij}(X,\boldsymbol{p})\ . \tag{53}$$

C Proof of Theorem 3

From Theorem 1 and Lemma 1, the control policy μ^* based on the sampling rule in (23) is weak asymptotically point-wise optimal for the given stopping rule $N^*(c)$ if

$$nJ(n,\mu^*) \xrightarrow{\mathsf{P}(X)} Q(X)\ . \tag{54}$$

Note that, all the maximum likelihood estimates $X_{i,\mathsf{mle}}$ converge to the true value X almost surely. Corresponding to the control policy μ^*, define $\tilde{q}_j^{\mu^*} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{u^i=j}$ and $\tilde{\boldsymbol{q}}^{\mu^*} = \{\tilde{q}_1^{\mu^*},\dots,\tilde{q}_K^{\mu^*}\}$. From the asymptotic optimality of Maximum Likelihood estimators and using Lemma 2.6 in [1], we can conclude that $\tilde{\boldsymbol{q}}^{\mu^*}$ converges to $\boldsymbol{p}^*$, where

$$\boldsymbol{p}^* = \arg\inf_{\boldsymbol{p}\in\phi^K} \sum_{i=1}^{K} Q_i(X,\boldsymbol{p}) + \frac{1}{2}\sum_{i=1}^{K}\sum_{j=1}^{K} Q_{ij}(X,\boldsymbol{p})\ . \tag{55}$$

Note that, from (50) and the fact that $\tilde{\boldsymbol{q}}^{\mu^*}$ converges to $\boldsymbol{p}^* = \{p_1^*,\dots,p_K^*\}$,

$$n||X_n^i - X_n^j||^2 \xrightarrow{\mathbb{P}(X)} \sum_{l=1}^{K} \frac{(a_i^l - a_j^l)^2}{p_l^* I_l(X)}\ . \tag{56}$$

Next, to show that the terms $n\mathbb{E}[(X - X_n^I)^2]$ converge to a positive random variable in probability, we need to show that the posterior distribution $\phi(\sqrt{n}(X_n^i - X) \mid \boldsymbol{Y}_n^i, \mathcal{U}^n)$ is asymptotically normal, i.e. it converges to $\mathcal{N}\left(0, \sum_{j=1}^{K} \frac{(a_i^j)^2}{p_j^* I_j(X)}\right)$ in distribution. We use the expansion of X_n^i again to establish the asymptotic normality. Note that under probability measure $\mathsf{P}(X)$, all observations in the set $\mathcal{Y}^n$ are independent of each other. Therefore, we only need to establish the asymptotic normality of the distribution $\phi(\sqrt{n}(\mathbb{E}[X \mid \boldsymbol{Y}_n^i] - X) \mid \boldsymbol{Y}_n^i, \mathcal{U}^n)$, which follows directly from the results in [1] and [2]. Therefore, $nJ(n,\mu^*)$ converges to $Q(X)$ in probability $\mathsf{P}(X)$. Hence, the proof is completed.

References

[1] P. J. Bickel and J. A. Yahav, "Some contributions to the asymptotic theory of Bayes solutions," in *Proc. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 11, no. 4, 1969, pp. 257–276.

[2] ——, "Asymptotically pointwise optimal procedures in sequential analysis," in *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 1967.

[3] G. Fellouris, "Asymptotically optimal parameter estimation under communication constraints," *The Annals of Statistics*, vol. 40, no. 4, pp. 2239–2265, Aug. 2012.

[4] G. V. Moustakides, T. Yaacoub, and Y. Mei, "Sequential estimation based on conditional cost," in *Proc. IEEE International Symposium on Information Theory*, June 2017, pp. 436–440.

[5] P. Grambsch, "Sequential sampling based on the observed fisher information to guarantee the accuracy of the maximum likelihood estimator," *The Annals of Statistics*, vol. 11, no. 1, pp. 68–77, Mar. 1983.

[6] V. J. Yohai, "Asymptotically optimal bayes sequential design of experiments for estimation," *The Annals of Statistics*, vol. 1, no. 5, pp. 822–837, Sep. 1973.

[7] G. Atia and S. Aeron, "Asymptotic optimality results for controlled sequential estimation," in *Proc. 51st Annual Allerton Conference Communication, Control, and Computing*, Oct. 2013.

[8] Y. Yilmaz, S. Li, and X. Wang, "Sequential joint detection and estimation: Optimum tests and applications," *arXiv:1411.1440v1*, 2014.

[9] Y. Yilmaz, G. V. Moustakides, and X. Wang, "Sequential joint detection and estimation," *Theory of Probability & Its Applications*, vol. 59, no. 3, pp. 452–465, 2015.