Optimal Scheduling of Multiple Spatio-temporally Dependent Observations using Age-of-Information

Victor Wattin Håkansson, Naveen K. D. Venkategowda, Stefan Werner
Department of Electronic Systems, NTNU - Norwegian University of Science and Technology
E-mail: victor.haakansson@ntnu.no, naveen.dv@ntnu.no, stefan.werner@ntnu.no

***Abstract*—This paper proposes an optimal scheduling policy for a remote estimation problem, where spatio-temporally dependent sensor observations are broadcasted to remote estimators. At each time instant only observations from a limited number of sensors can be communicated. The system has a network scheduler that decides the set of sensor observations to be communicated. The scheduler cannot observe measurements and exploits age-of-information (AoI) to calculate the expected estimation error. The scheduling problem is modeled as a Markov decision process with the AoI representing the state and the scheduling decision representing the action. We derive an optimal scheduling policy that minimizes the average mean squared error for an infinite time horizon – the policy results in a periodic scheduling pattern. Our results show that by exploiting spatio-temporal dependencies and using optimal sensor scheduling, the overall estimation accuracy is enhanced.**

I. Introduction

In wireless sensor networks (WSN) and networked control systems, sensors communicate observations to controllers or remote estimators that track physical processes by forming estimates. Sensors often share a limited number of communication channels, and so the communication between sensors and estimators follows protocols to avoid interference. Sensor transmissions can be either event-triggered [1], [2], e.g., a measurement breaching a threshold, or time-triggered [1], i.e., scheduled time slots. The latter can result in collision-free communication [1], [3] and is the focus of this paper.

The utility of the system depends on the real-time estimation accuracy of the estimators. An important task is to design scheduling protocols that minimize the overall estimation error over time. Optimal scheduling schemes for infinite time-horizons have been studied under different resource constraints, e.g., limited packet size [4], limited battery [5], or the presence of eavesdroppers [6]. In [3], authors derive an optimal scheduling policy for a system with multiple linear time-invariant sub-systems and a single communication channel. Increasing the number of communication channels improves the estimation accuracy, and in [7], authors derive an optimal scheduling policy for a system with multiple linear time-invariant sub-systems and multiple communication channels.

Most work that regards optimal scheduling for remote estimation assumes independent processes [3], [4], [6]–[8]. However, sensor measurements tend to be spatio-temporally correlated, which can be exploited to improve estimation accuracy. In [5], [9], [10], the authors investigate the optimal transmission frequency for sensors observing spatio-temporally correlated measurements. In [11], authors consider correlated sensor measurements when a scheduler can observe measurements before scheduling. Such a scheduling strategy may reduce estimation error but has implications on the system's privacy and latency.

This paper presents an optimal scheduling policy for multiple sensors that observe spatio-temporally correlated Gaussian processes. Our system model is similar to [11], [12], where observations are communicated via a network manager to the remote estimators. In contrast, we assume spatio-temporal dependence among the sensors, that multiple observations are broadcasted, and that a system-scheduler cannot read the measurements but utilizes the *age-of-information* (AoI) [13], [14] to decide on the scheduling. Most works regarding AoI have focused on evaluating the average AoI, given the system settings [13], [15]. Recent work shows that the AoI can be used as a state-variable in a broader range of optimization tasks if the performance metric can be expressed as a function of the AoI [14], [16].

The results of this paper demonstrate that exploiting spatio-temporal dependencies, together with AoI, can improve the remote estimation accuracy in systems with communication constraints. This paper is an extension of [16], where two processes and a single communication channel were considered. We derive an optimal policy for the multiple channel case is attained by modeling the problem as a finite-state Markov decision process (MDP). Further, we show that an optimal policy results in a periodic scheduling pattern. Numerical results verify the theory and show that the policy we propose outperforms alternative policies.

II. Problem Formulation

We consider a WSN of N sensors, one scheduler, and N remote estimators as depicted in Fig. 1. Sensor i observes the stochastic process $\theta_i[k] \in \mathbb{R}$, with $\theta_i[k] \sim \mathcal{N}(0, \sigma_i^2)$, at time instant $k \in \mathbb{N}_+$ and $i = 1, ..., N$. The N processes are correlated over space and time with the cross-covariance given by a positive-definite function [17], [18]

$$\mathbb{E}[\theta_i[k]\theta_j[l]] = \sigma_i\sigma_j\rho_{ij}\rho_t(|k-l|), \quad i, j \in \{1, ..., N\}, \quad (1)$$

where $\rho_{ij} \in [-1, 1]$ represents the spatial correlation and $\rho_t : \mathbb{R}_+ \to (0, 1]$ is the temporal correlation, which is a strictly decreasing function with $\rho_t(0) = 1$ and $\lim_{n\to\infty} \rho_t(n) = 0$.

This work was partly supported by the Research Council of Norway.

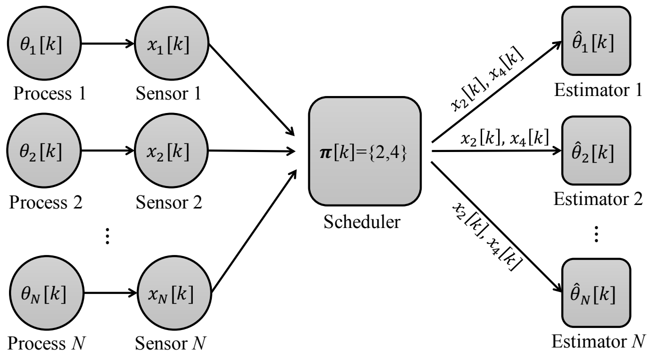

Fig. 1. Schematic of WSN scheduling problem with $D = 2$.

At time instant k, Sensor i, $i = 1, ..., N$, acquires measurement $x_i[k] \in \mathbb{R}$, which is modeled as

$$x_i[k] = \theta_i[k] + w_i[k], \quad k \in \mathbb{N}_+, \tag{2}$$

where $w_i[k] \in \mathbb{R}$ denotes independent identically distributed (iid) measurement noise with distribution $w_i[k] \sim \mathcal{N}(0, \xi^2)$. For each process $\theta_i[k]$, there is a corresponding remote estimator that tracks the process and forms an estimate $\hat{\theta}_i[k]$ based on sensor measurements communicated via the network scheduler, see Figure 1.

A. Scheduler

Due to limited channel capacity, the scheduler broadcasts $D \in \mathbb{N}_+$, $D \leq N$, sensor observations to the remote estimators. Since the processes are spatial-temporally correlated, the estimators can use every measurement to improve the local estimation accuracy. The scheduler decides the set of observations to be communicated and must maximize the total estimation accuracy over time.

Let $\boldsymbol{\pi}[k] \in \{1, ..., N\}^D$ be a scheduling variable denoting an index set of sensors to be scheduled at time k. The AoI of the ith sensor is denoted by $\Delta_i[k] \in \mathbb{N}_+$, $i = 1, ..., N$, and defined as the time elapsed between two measurement transmissions

$$\Delta_i[k] = \begin{cases} 0, & \text{if } i \in \boldsymbol{\pi}[k], \\ \Delta_i[k-1] + 1, & \text{if } i \notin \boldsymbol{\pi}[k]. \end{cases} \tag{3}$$

The scheduler is not allowed to observe the measurements, $\boldsymbol{x}[k] = [x_1[k], x_2[k], ..., x_N[k]]^{\mathrm{T}}$, but can keep track of the AoI at each sensor through vector $\boldsymbol{\Delta}[k]$, where $\boldsymbol{\Delta}[k] = [\Delta_1[k], \Delta_2[k], ..., \Delta_N[k]]^{\mathrm{T}}$. Let γ_k denote the *scheduling strategy* at time k, i.e.,

$$\boldsymbol{\pi}[k] = \gamma_k(\boldsymbol{\Delta}[k-1]), \tag{4}$$

which provides a mapping from $\boldsymbol{\Delta}[k-1]$ to the scheduling decision at instant k.

B. Remote estimators

The data available at Estimator i at time instant k contains $\boldsymbol{\Delta}[k]$ and $\boldsymbol{y}[k] = [y_1[k], y_2[k], ..., y_N[k]]^{\mathrm{T}}$, representing the most recently broadcasted measurement from each sensor, i.e.,

$$y_i[k] = x_i[k - \Delta_i[k]], \quad i = 1, ..., N. \tag{5}$$

The minimum mean square error (MMSE) estimate of $\theta_i[k]$, given $\{\boldsymbol{\Delta}[k], \boldsymbol{y}[k]\}$, is computed as

$$\hat{\theta}_i[k] = \mathbb{E}[\theta_i[k] | \boldsymbol{\Delta}[k], \boldsymbol{y}[k]], \quad i = 1, ..., N. \tag{6}$$

C. Scheduling policy

The *scheduling policy* γ is defined as the collection $\gamma = (\gamma_1, \gamma_2, \dots \gamma_T)$ where T denotes the time horizon. The performance measure (cost) is the total mean squared error (MSE) of the estimate (6) over T time slots and is given by

$$J(\gamma, T) = \sum_{k=1}^{T} \sum_{i=1}^{N} \mathbb{E}\Big[(\theta_i[k] - \hat{\theta}_i[k])^2 \Big| \boldsymbol{\Delta}^{\gamma}[k]\Big], \tag{7}$$

where $\boldsymbol{\Delta}[0] = (\infty, \infty, ..., \infty)^{\mathrm{T}}$ is the AoI when initializing the system and $\boldsymbol{\Delta}^{\gamma}[k]$ is the AoI at time k generated by γ.

Our objective is to find an *optimal scheduling policy* γ^* that minimizes the average cost in (7) over an infinite time horizon

$$\min_{\gamma \in \Gamma} \limsup_{T \to \infty} \frac{1}{TN} J(\gamma, T), \tag{8}$$

where Γ is the set of all feasible policies.

III. Optimal Scheduling Policy

To solve (8), we need to calculate the cost (7), which depends on the process $\boldsymbol{\Delta}^{\gamma}[k]$ during interval $k \in [1, T]$. To do so, we derive a closed-form expression for the MSE at instant k given $\boldsymbol{\Delta}[k]$. The process vector $\boldsymbol{\theta}[k] = [\theta_1[k], \theta_2[k], ..., \theta_N[k]]^{\mathrm{T}}$ follows a zero-mean Gaussian distribution with covariance matrix $\mathbf{C}_{\theta\theta}$. Substituting (2) in (5), we obtain $\boldsymbol{y}[k] \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{yy}[k])$. The covariance matrix $\mathbf{C}_{yy}[k]$ can be obtained from (1)–(5) as

$$[\mathbf{C}_{yy}[k]]_{i,j} = \sigma_i \sigma_j \rho_{ij} \rho_t(\Delta_{ij}[k]) + \xi^2 \delta(i-j), \; i, j \in \{1, ..., N\} \tag{9}$$

where $\Delta_{ij}[k] = |\Delta_i[k] - \Delta_j[k]|$ is the absolute difference of AoI between processes i and j, and $\delta(\cdot)$ denotes the Dirac delta function.

The vector estimate $\hat{\boldsymbol{\theta}}[k] = [\hat{\theta}_1[k], \hat{\theta}_2[k], ..., \hat{\theta}_N[k]]^{\mathrm{T}}$ becomes [19]

$$\hat{\boldsymbol{\theta}}[k] = \mathbb{E}[\boldsymbol{\theta}[k] | \boldsymbol{\Delta}[k], \boldsymbol{y}[k]] = \mathbf{C}_{\theta y}[k] \mathbf{C}_{yy}^{-1}[k] \boldsymbol{y}[k], \tag{10}$$

where $\mathbf{C}_{\theta y}[k] \in \mathbb{R}^{N \times N}$ is the cross-covariance between $\boldsymbol{y}[k]$ and $\boldsymbol{\theta}[k]$ given by

$$[\mathbf{C}_{\theta y}[k]]_{i,j} = \sigma_i \sigma_j \rho_{ij} \rho_t(\Delta_j[k]), \quad i, j \in \{1, ..., N\}. \tag{11}$$

The MSE at instant k can be expressed as a function of $\boldsymbol{\Delta}[k]$

$$\begin{aligned} E(\boldsymbol{\Delta}[k]) &= \sum_{i=1}^{N} \mathbb{E}\Big[(\theta_i[k] - \hat{\theta}_i[k])^2 \Big| \boldsymbol{\Delta}[k]\Big] \\ &= \mathrm{tr}\Big(\mathbf{C}_{\theta\theta} - \mathbf{C}_{\theta y}[k] \mathbf{C}_{yy}^{-1}[k] \mathbf{C}_{\theta y}^{\mathrm{T}}[k]\Big), \end{aligned} \tag{12}$$

where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. The MSE is upper bounded by the sum of all the marginal variances, i.e.,

$$E(\boldsymbol{\Delta}[k]) \leq \mathrm{tr}\Big(\mathbf{C}_{\theta\theta}\Big) = \sum_{i=1}^{N} \sigma_i^2. \tag{13}$$

It can be seen from (3) that $\mathbf{\Delta}[k]$ only depends on $\mathbf{\Delta}[k-1]$ and $\boldsymbol{\pi}[k]$. We can, therefore, model the system as a Markov decision process (MDP) [20], where at instant k, $\mathbf{\Delta}[k-1]$ is the state, $\boldsymbol{\pi}[k]$ is the action and $E(\mathbf{\Delta}[k])$ is the reward. Using dynamic programming, we can then derive a state-action policy that minimizes the average reward and corresponds to an optimal scheduling policy.

Using $\mathbf{\Delta}[k]$ as a state-variable leads to an infinite countable state space, and an optimal state-action policy may not exist or be prohibitive to derive. From (12) and (13), we see that as the AoI grows, the temporal correlation becomes negligible, and the MSE does not increase with respect to the marginal AoI, i.e.,

$$\lim_{\Delta_i[k]\to\infty} \frac{\partial E(\mathbf{\Delta}[k])}{\partial \Delta_i[k]} = 0, \quad i = 1, ..., N. \tag{14}$$

Therefore, we can reduce the state-space in our MDP to only AoI values that correspond to distinct MSE values. Since ρ_t in (2) is continuous, we restrict the set of possible correlation functions ρ_t to the type given in Assumption 1.

Assumption 1. *The temporal correlation function $\rho_t : \mathbb{R}_+ \to [0,1]$ in (2), satisfies $\rho_t(\Delta) = 0$, for all $\Delta \geq m$, $m \in \mathbb{N}_+$.*

If Assumption 1 holds, we can try to find a state-variable that corresponds to all possible MSE values, and results in a finite state MDP.

A. Truncated AoI

Let $\mathbf{\Delta}^m[k] \in \{0,1,...,m\}^{N^2}$ denote the truncated AoI [15] that contains the elements $\Delta_i^m[k], \Delta_{ij}^m[k] \in \{0,1,...,m\}$, $\forall i,j = 1,..,N$, i.e.,

$$\Delta_i^m[k] = \big[\Delta_i[k]\big]_+^m, \quad i = 1,...,N, \tag{15}$$
$$\Delta_{ij}^m[k] = \big[|\Delta_i[k] - \Delta_j[k]|\big]_+^m = \big[\Delta_{ij}[k]\big]_+^m, \quad i,j = 1,...,N,$$

where $m \in \mathbb{N}_+$ and $\big[\,\cdot\,\big]_+^m$ is defined as the truncation operator $\big[x\big]_+^m \triangleq \min\{x, m\}, x \in \mathbb{R}_+$. Let $f^m : \mathbb{N}_+^N \to \{0,1,...,m\}^{N^2}$ be a mapping from $\mathbf{\Delta}[k]$ to $\mathbf{\Delta}^m[k]$, i.e., $\mathbf{\Delta}^m[k] = f^m(\mathbf{\Delta}[k])$.

We can express the MSE as a function of $\mathbf{\Delta}^m[k]$, i.e.,

$$E^m(\mathbf{\Delta}^m[k]) = \sum_{i=1}^{N} \mathbb{E}\Big[(\theta_i[k] - \hat{\theta}_i[k])^2 \Big| \mathbf{\Delta}^m[k]\Big]. \tag{16}$$

The function $E^m(\mathbf{\Delta}^m[k])$ is obtained in a similar fashion as $E(\mathbf{\Delta}[k])$ in (12), i.e.,

$$E^m(\mathbf{\Delta}[k]) = \mathrm{tr}\Big(\mathbf{C}_{\theta\theta} - \mathbf{C}_{\theta y}^m[k](\mathbf{C}_{yy}^m)^{-1}[k](\mathbf{C}_{\theta y}^m[k])^{\mathrm{T}}\Big), \tag{17}$$

with $\mathbf{C}_{yy}^m[k]$ and $\mathbf{C}_{\theta y}^m[k]$ calculated using $\mathbf{\Delta}^m[k]$ as

$$[\mathbf{C}_{yy}^m[k]]_{i,j} = \sigma_i\sigma_j\rho_{ij}\rho_t(\Delta_{ij}^m[k]) + \xi^2\delta(i-j),$$
$$[\mathbf{C}_{\theta y}^m[k]]_{i,j} = \sigma_i\sigma_j\rho_{ij}\rho_t(\Delta_j^m[k]), \quad i,j \in \{1,...,N\}. \tag{18}$$

In the following propositions, we show that $\mathbf{\Delta}^m[k]$ can be used as a state-variable for modeling the system as an MDP.

Proposition 1. *Under Assumption 1 the following relationship holds*

$$E(\mathbf{\Delta}[k]) = E^m(\mathbf{\Delta}^m[k]), \quad \forall \mathbf{\Delta}[k] \in \mathbb{N}_+^N. \tag{19}$$

Proof. If Assumption 1 holds, we can see from expressions (12) and (17) that $E(\mathbf{\Delta}[k]) - E^m(f^m(\mathbf{\Delta}[k])) = 0$, since $\rho_t(\Delta_i[k]) = \rho_t(\Delta_i^m[k])$ and $\rho_t(\Delta_{ij}[k]) = \rho_t(\Delta_{ij}^m[k])$, $\forall \Delta_i[k], \Delta_{ij}[k] \in \mathbb{N}_+$. This gives that $E(\mathbf{\Delta}[k]) = E^m(\mathbf{\Delta}^m[k])$, $\forall \mathbf{\Delta}[k] \in \mathbb{N}_+^N$. ☐

Proposition 1 states that if $\mathbf{\Delta}[k]$ or $\mathbf{\Delta}^m[k]$ represents the state, either can be used to calculate the MSE in (12). For $\mathbf{\Delta}^m[k]$ to represent the state in the MDP, we need to express it as a function of $\mathbf{\Delta}^m[k-1]$ and $\boldsymbol{\pi}[k]$.

Proposition 2. *The truncated AoI $\mathbf{\Delta}^m[k]$ can be expressed as a function of $\mathbf{\Delta}^m[k-1]$ and $\boldsymbol{\pi}[k]$ as*

$$\Delta_i^m[k] = \begin{cases} 0, & \text{if } i \in \boldsymbol{\pi}[k], \\ \big[\Delta_i^m[k-1]+1\big]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k], \end{cases} \tag{20}$$

$$\Delta_{ij}^m[k] = \begin{cases} 0, & \text{if } i,j \in \boldsymbol{\pi}[k], \\ \big[\Delta_{ij}^m[k-1]\big]_+^m, & \text{if } i,j \notin \boldsymbol{\pi}[k], \\ \big[\Delta_i^m[k-1]+1\big]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k],\ j \in \boldsymbol{\pi}[k], \\ \big[\Delta_j^m[k-1]+1\big]_+^m, & \text{if } i \in \boldsymbol{\pi}[k],\ j \notin \boldsymbol{\pi}[k]. \end{cases} \tag{21}$$

Proof. Applying the truncation operator on the expression (3) for $\Delta_i[k]$, we obtain

$$\Delta_i^m[k] = \begin{cases} 0, & \text{if } i \in \boldsymbol{\pi}[k], \\ \big[\Delta_i[k-1]+1\big]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k]. \end{cases} \tag{22}$$

Further, we know that the following relationship holds

$$\big[\Delta_i[k-1]+1\big]_+^m = \Big[\big[\Delta_i[k-1]\big]_+^m + 1\Big]_+^m = \big[\Delta_i^m[k-1]+1\big]_+^m. \tag{23}$$

Now, substituting (23) in (22), we obtain the relationship in (20). Similarly, substituting (3) in $\Delta_{ij}[k] = |\Delta_i[k] - \Delta_j[k]|$, we can express $\Delta_{ij}[k]$ as

$$\Delta_{ij}[k] = \begin{cases} 0, & \text{if } i,j \in \boldsymbol{\pi}[k], \\ \Delta_{ij}[k-1], & \text{if } i,j \notin \boldsymbol{\pi}[k], \\ \Delta_i[k-1]+1, & \text{if } i \notin \boldsymbol{\pi}[k],\ j \in \boldsymbol{\pi}[k], \\ \Delta_j[k-1]+1, & \text{if } i \in \boldsymbol{\pi}[k],\ j \notin \boldsymbol{\pi}[k]. \end{cases} \tag{24}$$

Next, we use the truncation operator on (24) and employ the relationship in (23) to derive expression (21). ☐

Proposition 1 and 2 state that $\mathbf{\Delta}[k]$ or $\mathbf{\Delta}^m[k]$ can be used as state-variable to model the system as an MDP. Employing the latter leads to a finite state-space, which is preferable when trying to derive γ^*.

B. Finite-state Markov decision process

To find γ^*, we model our scheduling problem as a finite-state MDP [20]. We define the MDP as the following;

- **Action** $\boldsymbol{a}[k]$ at instant k is the scheduling decision $\boldsymbol{\pi}[k]$ belonging to action-space $\mathcal{A} = \{1,...,N\}^D$, $\boldsymbol{a}[k] \in \mathcal{A}$.
- **State** $\boldsymbol{s}[k]$ at instant k is the truncated AoI $\mathbf{\Delta}^m[k-1]$ belonging to state-space $\mathcal{S} = \{f^m(\boldsymbol{\delta}) \mid \boldsymbol{\delta} \in \mathbf{\Delta}\}$, $\boldsymbol{s}[k] \in \mathcal{S}$,

where $\boldsymbol{\Delta}$ is the set of possible AoI values, i.e., $\boldsymbol{\Delta}[k] \in \boldsymbol{\Delta}$, which depends on N and D.

- **Transition probabilities** $P(\boldsymbol{s}[k+1] \mid \boldsymbol{s}[k], \boldsymbol{a}[k]) \in \{0,1\}$ are binary and given by (20) and (21) in Proposition 2, where $\boldsymbol{s}[k]$ corresponds to $\boldsymbol{\Delta}^m[k-1]$ and $\boldsymbol{a}[k]$ corresponds to $\boldsymbol{\pi}[k]$.
- **Reward** at instant k corresponds to $E^m(\boldsymbol{\Delta}^m[k])$ in (17) and is given by the reward function

$$r(\boldsymbol{s}[k], \boldsymbol{a}[k]) = -\sum_{\boldsymbol{s}[k+1] \in \mathcal{S}} E^m(\boldsymbol{s}[k+1]) P(\boldsymbol{s}[k+1] \mid \boldsymbol{s}[k], \boldsymbol{a}[k]). \tag{25}$$

A policy $\mu = (\mu_1, ..., \mu_T)$ maps action $\boldsymbol{a}[k]$ to state $\boldsymbol{s}[k]$, i.e., $\boldsymbol{a}[k] = \mu_k(\boldsymbol{s}[k])$. The *average expected reward* g_μ is defined as

$$g_\mu(\boldsymbol{s}) = \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\Big[\sum_{k=0}^{T} r(\boldsymbol{s}[k], \boldsymbol{a}[k]) \Big| \boldsymbol{s}[0] = \boldsymbol{s}, \boldsymbol{a}[k] = \mu(\boldsymbol{s}[k])\Big], \tag{26}$$

where a policy μ^* is optimal in average-sense if it fulfills

$$g_{\mu^*}(\boldsymbol{s}) \geq g_\mu(\boldsymbol{s}), \ \vee \boldsymbol{s} \in \mathcal{S}. \tag{27}$$

We state the following property of the MDP that will be useful to derive μ^*.

Lemma 1. *For any stationary policy $\mu = (\mu_0, ..., \mu_0)$, $\mu_0 : \mathcal{S} \to \mathcal{A}$, the MDP results in a periodic state-action sequence. In other words, for $k \to \infty$, we have $\mu_0(\boldsymbol{s}[k]) = \mu_0(\boldsymbol{s}[k+L])$ and $\boldsymbol{s}[k] = \boldsymbol{s}[k+L]$, $L \in \mathbb{R}_+, L \leq |\mathcal{S}|$.*

Proof. From (20) and (21), we know that the transition probabilities are binary, i.e., $P(\boldsymbol{s}[k+1] \mid \boldsymbol{s}[k], \boldsymbol{a}[k]) \in \{0,1\}$. Hence, given the state at instant k, $\boldsymbol{s}[k]$, the next state $\boldsymbol{s}[k+1]$ is perfectly known. Similarly, the states that follow after $\boldsymbol{s}[k+1]$ are also perfectly known. If a state is revisited and μ is stationary, the state-sequence that occurred in-between the state was last visited will repeat itself. Since the state-space is finite, i.e., $|\mathcal{S}| < \infty$, at least one state will be revisited for $k \to \infty$. Hence, we conclude that for $k \to \infty$ the sequence of states and actions is periodic, i.e., $\mu_0(\boldsymbol{s}[k]) = \mu_0(\boldsymbol{s}[k+L])$ and $\boldsymbol{s}[k] = \boldsymbol{s}[k+L]$, $L \in \mathbb{R}_+, L \leq |\mathcal{S}|$. □

From Lemma 1 and [20], if an optimal policy μ^* exist, it results in a constant average reward $g_{\mu^*}(\boldsymbol{s}) = g^* \in \mathbb{R}$, $g^* \in \mathbb{R}$. The scalar g^* must then satisfy the optimality equations

$$\max_{\boldsymbol{a}\in\mathcal{A}} \{ r(\boldsymbol{s}, \boldsymbol{a}) - g^* + \sum_{\boldsymbol{s}' \in \mathcal{S}} P(\boldsymbol{s}' | \boldsymbol{s}, \boldsymbol{a}) h(\boldsymbol{s}') - h(\boldsymbol{s}) \} = 0, \tag{28}$$

where h is a function, $h : \mathcal{S} \to \mathbb{R}$, from the set V of bounded functions on $\mathcal{S}$, $h \in V$. The existence of an optimal policy μ^* is stated in the lemma below.

Lemma 2. *There exists a stationary average reward optimal policy $\mu^* = (\mu_0^*, ..., \mu_0^*)$, corresponding to an optimal constant reward $g_{\mu^*} = g^*$, $g^* \in \mathbb{R}$. The policy μ^* is given by*

$$\mu_0^*(\boldsymbol{s}) = \arg\max_{\boldsymbol{a}\in\mathcal{A}} \{ r(\boldsymbol{s}, \boldsymbol{a}) + \sum_{\boldsymbol{s}' \in \mathcal{S}} P(\boldsymbol{s}' | \boldsymbol{s}, \boldsymbol{a}) h^*(\boldsymbol{s}') \}, \tag{29}$$

where $h^* \in V$ and g^* satisfy (28), which can be attained in a finite number of iterations using policy iteration.

Proof. The MDP has a finite action set $|\mathcal{A}| < \infty$, a finite state-space $|\mathcal{S}| < \infty$, stationary bounded rewards $|r(\boldsymbol{s}[k], \boldsymbol{a}[k]| < \infty$ and stationary binary transition probabilities. Given the aforementioned properties of the MDP and Lemma 1, [20, Th. 8.4.5] states that there exist a stationary optimal policy μ^* and a pair (g^*, h^*) that satisfy (28). It also states the relationship between μ^* and (g^*, h^*) presented in (29). Furthermore, [20, Th. 8.6.6], states that (g^*, h^*) can be derived in a finite number of iterations using policy iteration. The full proofs are given in [20]. □

Based on Lemma 1 and Lemma 2, we formalize the following theorem that states the existence of an optimal scheduling policy γ^* and how it can be derived.

Theorem 1. *There exists a stationary optimal scheduling policy $\gamma^* = (\gamma_0^*, ..., \gamma_0^*)$, where $\gamma_0^* = \mu_0^* \circ f^m$ and μ_0^* is found using policy iteration. The policy γ^* results in a periodic scheduling pattern, i.e., $\gamma_0^*(\boldsymbol{\Delta}^{\gamma^*}[k-1]) = \gamma_0^*(\boldsymbol{\Delta}^{\gamma^*}[k-1+L])$, $L \in \mathbb{R}_+, L < \infty$.*

Proof. Consider the MDP defined in Section III-B, where at instant k, $\boldsymbol{\Delta}^m[k-1]$ represents the state-variable $\boldsymbol{s}[k]$, $-E^m(\boldsymbol{\Delta}^m[k])$ represent the reward $r(\boldsymbol{s}[k], \boldsymbol{a}[k])$, $\boldsymbol{\pi}[k]$ represents the action $\boldsymbol{a}[k]$ and the policy μ represents the mapping between $\boldsymbol{\pi}[k]$ and $\boldsymbol{\Delta}^m[k-1]$. Lemma 2 together with (27) states that a state-action policy $\mu^* = (\mu_0^*, ..., \mu_0^*)$ exists that maximizes the average expected reward in (26). If we compare (26) with (8), this implies, that if μ^* and the mapping between $\boldsymbol{\Delta}[k-1]$ and $\boldsymbol{\Delta}^m[k-1]$ in (15) is known, we can derive an optimal scheduling policy γ^* that minimizes the average cost in (8). Hence, an stationary optimal scheduling policy is given by $\gamma^* = (\gamma_0^*, ..., \gamma_0^*)$, where $\gamma_0^* = \mu_0^* \circ f^m$. Lemma 2 states that μ_0^* can be derived using policy iteration to solve (29). Lemma 1 states that μ_0^* results in periodic action sequence. Hence, γ_0^* results in a periodic action sequence. □

IV. Simulation results

We assume a system where $N = 5$ sensors observe dependent processes with equal marginal variances, i.e., $\sigma_i = 1, \forall i = 1, ..., N$, where the scheduler can broadcast $D = 2$ sensors at each time instant k. The spatio-temporal dependency components in (2) are given by [17], [18]

$$\rho_{ij} = e^{-r_0|i-j|}, \quad \rho_t(\Delta) = e^{-T_0\Delta} \mathbb{1}(e^{-T_0\Delta} \geq 0.1) \tag{30}$$

where $T_0 \in \mathbb{R}_+$, represents the time interval between two broadcasting sessions, $r_0 \in \mathbb{R}_+$ represents the Euclidean distance between two neighboring sensors, $\mathbb{1}(\cdot)$ is an indicator function having value 1 if the condition in the argument is true and 0 otherwise. The truncation time m in (15) is set to $m = \inf_{\Delta \in \mathbb{N}_+} \{e^{-T\Delta} \leq 0.1\}$.

Figure 2 shows the average cost versus T_0 with $\xi = 0.5$, $r_0 = 0.5$, for an optimal policy γ^*, round-robin [13], random scheduling, and a greedy policy, i.e., choosing the set of sensors at time k that minimizes the MSE $E(\boldsymbol{\Delta}[k])$ in (12).

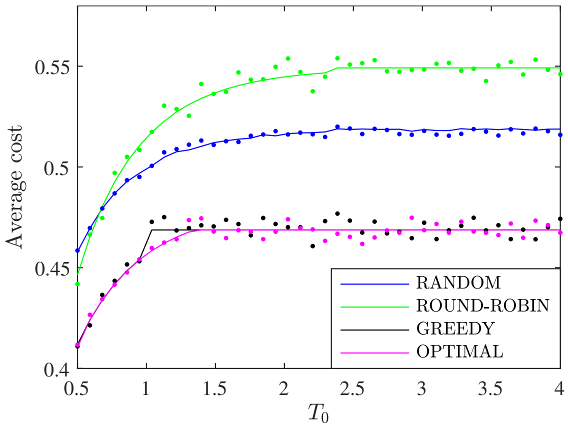

Fig. 2. Average cost versus T_0 for different policies with system parameters $N = 5$, $D = 2$, $\xi = 0.5$, $r_0 = 0.5$ and $\sigma_i = 1$, $\forall i = 1, ..., N$.

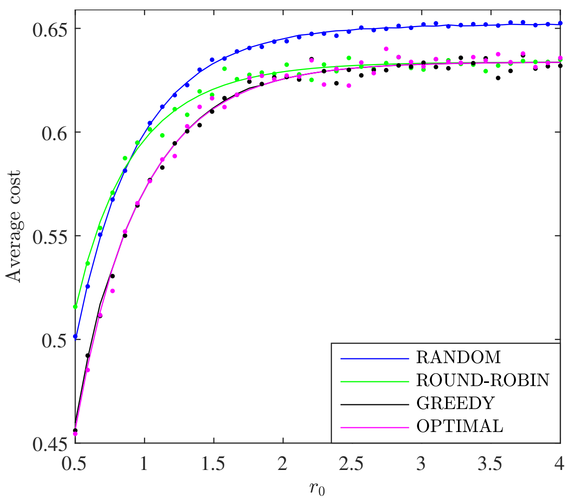

Fig. 3. Average cost versus r_0 for different policies with system parameters $N = 5$, $D = 2$, $\xi = 0.5$, $T_0 = 1$ and $\sigma_i = 1$, $\forall i = 1, ..., N$.

Solid lines depicts theoretical values and markers show Monte Carlo simulated estimates, based on simulating 200 sequences with $T = 100$ per T_0. We see that the simulations matches the theory. An optimal policy performs the best and the greedy policy performs close to optimal for most regions of T_0.

Figure 3 shows the average cost versus r_0 with $\xi = 0.5$ and $T_0 = 1$. Again, an optimal policy performs best followed by the greedy policy.

V. Conclusion

This paper studied a scheduling problem for sensors observing multiple spatio-temporally dependent processes to be communicated to remote estimators. At each time instant, the scheduler broadcasts a limited number of sensor measurements to the estimators. The scheduler cannot view the measurements but decides the set of sensors based on the age-of-information. We derived an optimal scheduling policy that achieves the minimum average MSE over time by modeling the problem as a finite state-MDP, with the AoI as a state variable. The optimal scheduling policy results in a periodic scheduling pattern.

References

[1] M. Xia, V. Gupta, and P. J. Antsaklis, "Networked state estimation over a shared communication medium," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1729–1741, Apr. 2017.

[2] V. W. Håkansson, N. K. D. Venkategowda, F. A. Kraemer, and S. Werner, "Cost-aware dual prediction scheme for reducing transmissions at IoT sensor nodes," *27th European Signal Processing Conference (EUSIPCO)*, pp. 1–5, 2019.

[3] D. Han, J. Wu, H. Zhang, and L. Shi, "Optimal sensor scheduling for multiple linear dynamical systems," *Automatica*, vol. 75, pp. 260 – 270, Jan. 2017.

[4] S. Wu, X. Ren, S. Dey, and L. Shi, "Optimal scheduling of multiple sensors over shared channels with packet transmission constraint," *Automatica*, vol. 96, pp. 22 – 31, Oct. 2018.

[5] J. Hribar, A. Marinescu, G. A. Ropokis, and L. A. DaSilva, "Using deep Q-learning to prolong the lifetime of correlated internet of things devices," in *IEEE International Conference on Communications Workshops*, 2019, pp. 1–6.

[6] A. S. Leong, D. E. Quevedo, D. Dolz, and S. Dey, "Transmission scheduling for remote state estimation over packet dropping links in the presence of an eavesdropper," *IEEE Transactions on Automatic Control*, vol. 64, no. 9, pp. 3732–3739, Sep. 2019.

[7] A. Leong, A. Ramaswamy, D. Quevedo, H. Karl, and L. Shi, "Deep reinforcement learning for wireless sensor scheduling in cyber–physical systems," *Automatica*, vol. 113, pp. 1–8, Mar. 2020.

[8] M. Gagrani, M. M. Vasconcelos, and A. Nayyar, "Scheduling and estimation strategies in a sequential networked estimation problem," in *56th Annual Allerton Conference on Communication, Control, and Computing*, 2018, pp. 871–878.

[9] J. Hribar, M. Costa, N. Kaminski, and L. A. DaSilva, "Using correlated information to extend device lifetime," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2439–2448, Apr. 2019.

[10] Z. Jiang and S. Zhou, "Status from a random field: How densely should one update?" *IEEE International Symposium on Information Theory*, vol. 2019-July, pp. 1037–1041, 2019.

[11] M. M. Vasconcelos and U. Mitra, "Observation-driven scheduling for remote estimation of two Gaussian random variables," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 1, pp. 232–244, Mar. 2020.

[12] M. M. Vasconcelos, A. Nayyar, and U. Mitra, "Optimal sensor scheduling strategies in networked estimation," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017, pp. 5378–5384.

[13] A. Kosta, N. Pappas, A. Ephremides, and V. Agelakis, "Age of information performance of multiaccess strategies with packet management," *Journal of Communications and Networks*, vol. 21, no. 3, pp. 244–255, Jun. 2019.

[14] A. Kosta, N. Pappas, A. Ephremides, and V. Angelakis, "The cost of delay in status updates and their value: non-linear ageing," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 1–1, Aug. 2020.

[15] Y. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," *IEEE International Symposium on Information Theory (ISIT)*, pp. 561–565, 2017.

[16] V. W. Håkansson, N. K. D. Venkategowda, and S. Werner, "Optimal scheduling policy for spatio-temporally dependent observations using age-of-information," *IEEE 23rd International Conference on Information Fusion (FUSION)*, pp. 1–6, 2020.

[17] N. Cressie and C. Wikle, *Statistics for Spatio-Temporal Data*. Wiley, 2011.

[18] R. Furrer, M. G. Genton, and D. Nychka, "Covariance tapering for interpolation of large spatial datasets," *Journal of Computational and Graphical Statistics*, vol. 15, no. 3, pp. 502–523, 2006.

[19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1997.

[20] M. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, 1994.