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

Victor Wattin Håkansson, *Member, IEEE*, Naveen K. D. Venkategowda, *Member, IEEE*, Stefan Werner, *Senior Member, IEEE*, Pramod K. Varshney, *Life Fellow, IEEE*

***Abstract*—This paper proposes an optimal scheduling policy for a system where spatio-temporally dependent sensor observations are broadcast to remote estimators over a resource-limited broadcast channel. We consider a system with a measurement-blind network scheduler that transmit observations, and design scheduling schemes that minimize MSE by determining a subset of sensor observations to be broadcast based on their information freshness, as measured by their age-of-information (AoI). By modeling the problem as a finite state-space Markov decision process (MDP), we derive an optimal scheduling policy, with AoI as a state-variable, minimizing the average mean squared error for an infinite time horizon. The resulting policy has a periodic pattern that renders an efficient implementation with low data storage. We further show that for any policy that minimizes the overall AoI, the estimation accuracy depends on how the scheduling order relates to the sensor's intrinsic spatial correlation. Consequently, the estimation accuracy varies from worse than a randomized scheduling approach to near-optimal. Thus, we present an additional age-minimizing policy with optimal scheduling order. We also present alternative policies for large state spaces that are attainable with less computational effort. Numerical results validate the presented theory.**

***Index Terms*—Wireless sensor networks, age-of-information, spatio-temporal correlation, remote estimation, resource-constrained networks**

I. Introduction

Wireless sensor networks (WSN) provide the data collection infrastructure for control and estimation systems used in internet-of-things applications. In WSN and networked control systems, sensor observations are communicated to controllers or remote estimators that track physical processes by forming estimates. Sensors often share a limited number of communication channels and follow protocols to reduce interference. Measurement transmission protocols are categorized as either event- or time-triggered [1], [2]. The former refers to sensors transmitting an observation in case of an event, e.g., a measurement exceeds a predefined threshold [3]. The latter refers to allocating time slots for each sensor transmission. Time-triggered scheduling has the advantage that it can result in collision-free communication [1], [4] and is the focus of this paper.

In networked control systems, the system utility depends on the estimation accuracy of the controllers and estimators. A common objective is to design optimal sensor scheduling schemes that minimize the time-average estimation error. The processes tracked by the system are dynamic and time-dependent, and finding optimal scheduling policies involves solving sequential decision-making problems. One approach to solve this problem, is to find scheduling sequences that minimize the estimation error over shorter time horizons [5], [6]. As the time horizon grows, the number of possible scheduling trajectories rapidly increases, and short-term approximations can become sub-optimal over more extended periods. A common objective is, therefore, to find optimal policies for infinite time horizons.

Optimal scheduling schemes have been studied under various resource constraints, e.g., limited battery [7] or limited packet size for sensors monitoring sources with heterogeneous dynamics [8]. In [4], authors derive an optimal policy scheme for a system with multiple linear time-invariant sub-systems and a single communication channel. The resulting policy was to schedule the sensors in a periodic sequence. Although, addition of more communication channels improves the overall real-time accuracy, it also adds complexity in finding optimal scheduling policies as the number of possible scheduling decisions increases. In [9], authors use deep reinforcement learning to find an optimal policy for a system with multiple linear time-invariant sub-systems and multiple communication channels. The works [10], [11] propose an optimal policy for the case when a network manager, responsible for the scheduling, can observe the sensor measurements. Authors in [12] consider the system security aspect and derive an optimal scheduling policy in the presence of eavesdroppers.

All the previous works assume that sensor observations are independent. In contrast, sensor observations tend to be spatio-temporally dependent [13], [14], which can be exploited to improve the remote estimators' overall accuracy. For example, this has been done in resource-constrained WSN to achieve energy-efficient routing [15], optimal sensor location selection [16], and reducing traffic load [13]. However, only a few works have considered dependency among sensors in optimal sensor scheduling problems related to remote estimation. In [17]–[19], the authors study the transmission frequency of spatio-temporally correlated sensor measurements modeled by

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

Victor Wattin Håkansson and Stefan Werner are with Norwegian University of Science and Technology (NTNU), Norway (e-mail: victor.haakansson@ntnu.no and stefan.werner@ntnu.no).

Naveen K. D. Venkategowda is with Linköping University, Sweden (e-mail: naveen.venkategowda@liu.se).

Pramod K. Varshney is with Syracuse University, Syracuse, NY 13244, USA (e-mail: varshney@syr.edu).

a random field. Finally, in [20], the authors find an optimal scheduling policy for a system of dependent sensor processes where a network manager, responsible for the scheduling, can observe measurements before scheduling. The scheduling strategy reduces estimation error, but the setup has implications for privacy and latency.

This paper presents an optimal scheduling policy for a system where multiple spatio-temporally dependent sensor observations are broadcast to remote estimators via a network manager. However, due to limited channel capacity, only a subset of all sensor observations, which is determined by the scheduler, can be communicated by the network manager at each time instant. The estimators compensate for the lack of observations by exploiting the spatio-temporal dependencies in the received information to improve its local estimation accuracy. Our system model is similar to [7], [11], [20], [21]; however, we allow for multiple sensor observations to be broadcast as in [22] and assume that the scheduler can not view the measurements. The scheduling policy is instead based on determining the mean squared error (MSE) with respect to the timeliness of the information. Thus, the scheduler decides the scheduling decision given the *age-of-information* (AoI) [23].

The AoI refers to the freshness of information [23], i.e., the time elapsed since the information was generated. Performance metrics of the AoI, e.g., peak and average, have been studied under different system settings [2], [23]–[29]. For most remote-estimation systems, the real-time tracking accuracy depends on the AoI. However, the relationship between accuracy and the AoI is not always linear; thus, minimizing average AoI may not correspond to optimal performance [30]. The AoI can also be used as a state variable to assist in designing and evaluating scheduling policies in a variety of tasks, e.g., updating model parameters in federated learning [31], maximizing the value-of-information [30], and minimizing the time-average estimation error [5], [30], [32]. In the same way, the AoI has been utilized in several works regarding scheduling for remote state estimation and network control, see, e.g., [5], [32]–[38]. Recently, scheduling of multiple sensors that share multiple unstable communication channels that result in packet dropouts have been studied in [5], [37], [38]. In [37], the authors consider a system of multiple Markov fading communication channels and determine the system conditions in terms of LTI system parameters and channel statistics that guarantee stability, i.e., the existence of a scheduling policy that result in a bounded average estimation MSE. Similary, in [38] stability conditions for a system of multiple wireless network control systems sharing multiple imperfect communication channels for uplink and downlink transmissions is determined, and a scheduling policy is derived using deep-reinforcement-learning. However, the aforementioned works regarding AoI-based scheduling and remote estimation [5], [32]–[38] concern scheduling of independent sensor observation, whereas there exist a minority of works that consider and exploit dependency among sensor observations.

Among the works that connect AoI and remote estimation of correlated processes, ours resembles that of [17]–[19], which assumes a similar model for the spatio-temporal dependency. In contrast, [17]–[19] find optimal sensor transmission rates, where as we exploit full channel capacity and decide the subset of sensors to be scheduled during each time slot. The authors of [17], [18] primarily focus on maximizing sensor battery lifetime for the desired estimation accuracy. In [39], the average AoI is minimized for a WSN where sensors observe partial information from sources, and multiple status packet updates are required at the receiver for proper reconstruction. In [40], the overall AoI is minimized in a WSN where neighboring sensors monitor overlapping sources that produce updates according to independent Poisson processes. In contrast, system setup in this paper differs from the above as it allows for multiple communication channels. To clarify the difference to the related works regarding the scheduling of spatio-temporally dependent observations [17]–[19] for a remote estimation WSN, we: i) allow for multiple sensors to be scheduled over multiple communication channels at each time-instant; ii) do not focus on the transmission rate but regard time-discrete scheduling to exploit full channel capacity, and; iii) do not assume homogenous distributions among the processes.

The main contributions of the paper can be summarized as follows:

- We prove the existence and derive an optimal scheduling policy for a system of multiple spatio-temporally dependent observations based on the age-of-information. An optimal policy minimizes the average mean squared estimation error over an infinite time horizon.
- We show that a policy can be derived by formalizing the problem as a finite-state MDP. The finite-state MDP is possible by exploiting the property that increasing time-distance between consecutive transmissions from a single source decreases spatio-temporal correlation to observations from other sources.
- We also show that an optimal policy yields a periodic scheduling pattern, which has earlier been demonstrated for optimal single-sensor scheduling [**?**], [4], [41]. This property simplifies the practical implementations and saves data storage at the network manager.
- We show that the finite state space implies that any deterministic policy results in a periodic structure. The performance of any periodic scheduling policy can be easily calculated using the theoretical framework given in the paper.

The precursor of this work can be found in [42], where the same system was considered for two sensors. Due to the computational complexity for large sensor systems, we present low-complexity policies compared to our optimal policy. These alternative policies are respectively based on minimizing the AoI and the short-term mean squared error. We demonstrate that minimizing the AoI of a system can lead to near-optimal performance. However, the intrinsic order of the sensor scheduling significantly affects the estimation accuracy when measurements are dependent. Thus, as most works regarding AoI-based scheduling have focused on minimizing the overall AoI, the scheduling order should be accounted for when dealing with spatio-temporally dependent observations in remote estimation tasks. In all, this paper demonstrates that performance can be improved for channel constrained remote estimation systems by incorporating spatial dependencies and the AoI.

The remainder of the paper is organized as follows. In Section II, we present the system model and the scheduling problem. Next, in Section III, we formulate the problem as a finite-state Markov decision process and demonstrate how an optimal policy can be obtained and that it results in a periodic scheduling sequence. In Section IV, we show that the performance of any periodic scheduling policies is independent of the initial AoI and how this allows the scheduler to save data storage. For large WSNs, we present in Section V alternative scheduling policies that can be obtained using less computational effort. Section VI validates the theory based on numerical results, and Section VII concludes the paper.

II. Background and 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 $\theta_i[k] \in \mathbb{R}$, at time instant $k \in \mathbb{N}$ and $i = 1, ..., N$. 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. Due to limited channel capacity, the scheduler broadcasts $D \in \mathbb{N}_+$, $D \leq N$, sensor observations to the remote estimators at instant k over D orthogonal channels. We assume that the channels are reliable and packet losses are addressed by retransmission through higher layers of the communication protocol.

Below we describe the three key blocks in Fig. 1: the source processes, the scheduler, and remote estimators. Finally, we present the scheduling problem considered in this paper.

A. Source processes

Each process $\theta_i[k]$ follows a Gaussian distribution $\theta_i[k] \sim \mathcal{N}(0, \sigma_i^2)$. The processes $\{\theta_i[k]\}_{i=1}^N$ are correlated over space and time with the cross-covariance given by a positive-definite function [43], [44]

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

where $\rho_{ij} \in [-1, 1]$ represents the spatial correlation and $\varphi : \mathbb{R}_+ \to (0, 1]$ is the temporal correlation, which is a strictly decreasing function with $\varphi(0) = 1$ and $\lim_{n\to\infty} \varphi(n) = 0$. At time instant k, the ith sensor acquires measurement $x_i[k] \in \mathbb{R}$, which is modeled as

$$x_i[k] = \theta_i[k] + w_i[k], \quad k \in \mathbb{N}, i = 1, 2, ..., N, \quad (2)$$

where $w_i[k] \in \mathbb{R}$ denotes independent identically distributed (iid) measurement noise with distribution $w_i[k] \sim \mathcal{N}(0, \xi^2)$.

B. Scheduler

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 $\Delta_i[k] \in \mathbb{N}_+$, $i = 1, ..., N$,

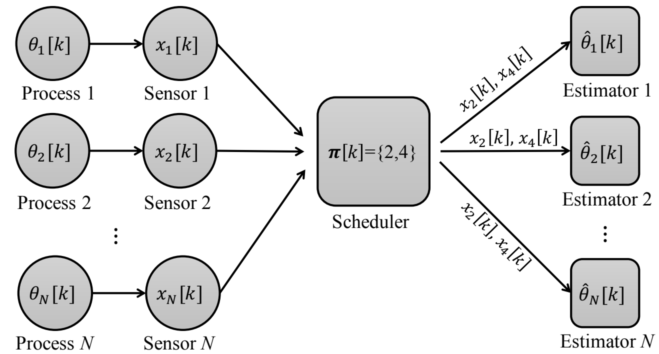

Fig. 1. Schematic of WSN scheduling problem with $D = 2$.

and defined as the time elapsed between two measurement transmissions [17], i.e.,

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

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 us define the information set $\mathcal{I}[k]$ available at the scheduler for decision at time instant k. Information set $\mathcal{I}[k]$ is the collection of the AoI for time $k = 0, 1, \ldots, k-1$ and defined as $\mathcal{I}[k] = \{\boldsymbol{\Delta}[0], \boldsymbol{\Delta}[1], \ldots, \boldsymbol{\Delta}[k-1]\}$. Let $\gamma_k : \mathcal{I}[k] \to \{1, ..., N\}^D$, denote the *scheduling strategy* at time k, i.e.,

$$\boldsymbol{\pi}[k] = \gamma_k(\mathcal{I}[k]), \quad (4)$$

which provides a mapping from $\mathcal{I}[k]$ to the scheduling decision at instant k.

C. Remote estimators

The data available at the ith remote estimator at time instant k consists of $\boldsymbol{\Delta}[k]$ and $\boldsymbol{y}[k] = [y_1[k], y_2[k], ..., y_N[k]]^{\mathrm{T}}$, where $y_i[k]$ is the most recently broadcast measurement from Sensor i, i.e.,

$$y_i[k] = x_i[k - \Delta_i[k]], \quad i = 1, ..., N. \quad (5)$$

The estimate $\hat{\boldsymbol{\theta}}[k] = [\hat{\theta}_1[k], \hat{\theta}_2[k], ..., \hat{\theta}_N[k]]^{\mathrm{T}}$ is the linear minimum mean square error (MMSE) estimate [45] given as a function of $\boldsymbol{\Delta}[k]$ and $\boldsymbol{y}[k]$ as follows,

$$\hat{\boldsymbol{\theta}}[k] = \mathbb{E}[\boldsymbol{\theta}[k]|\boldsymbol{\Delta}[k], \boldsymbol{y}[k]] = \mathbf{C}_{\theta y}[k]\mathbf{C}_{yy}^{-1}[k]\boldsymbol{y}[k], \quad (6)$$

where the elements of the cross-covariance and covariance matrices are given by

$$\begin{aligned} [\mathbf{C}_{\theta y}[k]]_{i,j} &= \sigma_i\sigma_j\rho_{ij}\varphi(\Delta_j[k]), \; i = 1, ..., N, j = 1, ..., N, \\ [\mathbf{C}_{yy}[k]]_{i,j} &= \sigma_i\sigma_j\rho_{ij}\varphi(\Delta_{ij}[k]) + \xi^2\delta(i-j), \end{aligned} \quad (7)$$

with $\Delta_{ij}[k] = |\Delta_i[k] - \Delta_j[k]| \in \mathbb{N}_+$ being the AoI differences between the two processes, and $\delta(\cdot)$ the Dirac delta function.

It should be noted that estimator (6) uses only the most recent measurement from each sensor. Even though the estimation accuracy can be improved by using previous measurements, the scheduling problem becomes intractable due to the spatio-temporal correlation of the process and the influence of

AoI on the scheduling policy. Therefore, in order to obtain insights into the properties of an optimal scheduling policy and derive efficient algorithms for scheduling, we simplify the estimator as given in (6).

D. Scheduling policy

The *scheduling policy* γ over time horizon T is defined as the collection of scheduling strategies from time instant $k=1$ to $k=T$, i.e., $\gamma=(\gamma_1,\gamma_2,\dots\gamma_T)$. As performance measure (cost), we adopt the total mean squared error (MSE) of the estimate (6) over T time slots, given by

$$J(\gamma,T)=\frac{1}{TN}\sum_{k=1}^{T}\sum_{i=1}^{N}\mathbb{E}\Big[(\theta_i[k]-\hat{\theta}_i[k])^2\Big|\gamma,\mathcal{I}[k]\Big], \tag{8}$$

where $\mathcal{I}[1]=\{\boldsymbol{\Delta}[0]\}$ is known at the scheduler.

Our objective is to find an *optimal scheduling policy* γ^* that minimizes the average cost in (8) over an infinite time horizon

$$\min_{\gamma\in\Gamma}\lim_{T\to\infty}J(\gamma,T), \tag{9}$$

where Γ is the set of all feasible policies.

III. Optimal Scheduling Policy

In this section, we first reformulate (9) as a Markov decision process (MDP). We propose an equivalent MDP with truncated states to cope with the resulting high-dimensional state-space by exploiting finite-duration temporal correlation. The two formulations are shown to give the same set of possible MSE values. Finally, we derive an optimal policy and show that it yields a periodic scheduling sequence, significantly reducing the implementation complexity.

To solve (9), we must be able to calculate the cost in (8), which depends on the process $\boldsymbol{\Delta}[k]$ during interval $k\in[1,T]$. The MSE at instant k can be expressed as a function $f:\boldsymbol{\Delta}[k]\to\mathbb{R}_+$, i.e.,

$$\begin{aligned}f(\boldsymbol{\Delta}[k])&=\sum_{i=1}^{N}\mathbb{E}\Big[(\theta_i[k]-\hat{\theta}_i[k])^2\Big|\boldsymbol{\Delta}[k]\Big]\\&=\mathrm{tr}\Big(\mathbf{C}_{\theta\theta}-\mathbf{C}_{\theta y}[k]\mathbf{C}_{yy}^{-1}[k]\mathbf{C}_{\theta y}^{\mathrm{T}}[k]\Big),\end{aligned} \tag{10}$$

where $\mathbf{C}_{\theta\theta}$ is the covariance matrix of $\boldsymbol{\theta}[k]$ and $\mathrm{tr}(\cdot)$ denotes the trace of its argument matrix. The MSE increases with respect to the AoI, i.e.,

$$\begin{aligned}&f([\Delta_1[k],...,\Delta_i[k],...,\Delta_N[k]]^{\mathrm{T}})\le\\&\quad f([\Delta_1[k],...,\Delta_i[k]+1,...,\Delta_N[k]]^{\mathrm{T}}),\quad i=1,...,N.\end{aligned} \tag{11}$$

and is upper bounded by the sum of the marginal variances, i.e.,

$$f(\boldsymbol{\Delta}[k])\le\mathrm{tr}\Big(\mathbf{C}_{\theta\theta}\Big)=\sum_{i=1}^{N}\sigma_i^2. \tag{12}$$

Proposition 1. *An optimal policy γ^* can be obtained by solving*

$$\min_{\gamma\in\Gamma}\lim_{T\to\infty}\frac{1}{TN}\sum_{k=1}^{T}\mathbb{E}\Big[f(\boldsymbol{\Delta}[k])\Big|\gamma,\boldsymbol{\Delta}[0]\Big]. \tag{13}$$

Proof. The proof is given in Appendix A. □

Proposition 1 shows that the MSE at instant k depends on $\boldsymbol{\Delta}[k]$, which in turn depends on $\boldsymbol{\Delta}[k-1]$ and $\boldsymbol{\pi}[k]$. Hence, the problem in (9) can be modeled as a Markov decision process (MDP) with the AoI as the state, the MSE as the reward and the scheduling decision as the action at instant k. In the following section, we formalize an MDP and derive an optimal scheduling policy γ^* that fulfills being time-average reward optimal.

A. Markov decision process formulation

To find γ^*, we model the system as an MDP with the AoI as the state and the MSE as the reward at instant k. We begin by deriving the set of possible AoI values, which will define the state space for our MDP. Later, we show that an optimal scheduling policy γ^* is time-average reward optimal.

The set of possible AoI values depends on the system parameters N and D. We assume that at time instant $k=0$, the system is initiated and that the AoI before initialization, $k\in\mathbb{N}_-$, is $\boldsymbol{\Delta}[k]=[\infty,\infty,...,\infty]^{\mathrm{T}}$, $\Delta_i[k]\neq\Delta_j[k]$, $i,j=1,...,N$. At each time instant $k\in\mathbb{N}_+$, D observations are scheduled, resulting in D sensors with AoI equal to zero and at most D sensors having the same AoI.

If a round-robin scheduling policy [9] is applied, there would be a maximum AoI across all sensors, denoted as $\bar{\Delta}\in\mathbb{N}_+$, i.e.,

$$\begin{aligned}\bar{\Delta}&=\min_{\gamma\in\Gamma}\lim_{k\to\infty}\sup\mathbb{E}\big[\Delta_i[k]\big|\gamma\big],\quad\forall i=1,2,...,N,\\&=\begin{cases}N/D-1, & \text{if } 0=N \mod D,\\ \lfloor N/D\rfloor, & \text{else},\end{cases}\end{aligned} \tag{14}$$

where $\lfloor\cdot\rfloor$ is the floor operator. For a round-robin scheduling policy, at each time instant k, there will be $\bar{N}$ sensors with an AoI equal to $\bar{\Delta}$. The value $\bar{N}$ is given by

$$\bar{N}=\begin{cases}D, & \text{if } 0=N \mod D,\\ N \mod D, & \text{else}.\end{cases} \tag{15}$$

Let $c:\mathbb{N}_+^N\times\mathbb{N}_+\to\mathbb{N}_+$ be an operator counting the number of elements in a vector $\boldsymbol{x}\in\mathbb{N}_+^n$ that equal to $l\in\mathbb{N}_+$, i.e.,

$$c(\boldsymbol{x},l)=\sum_{i=1}^{N}\mathbb{1}([\boldsymbol{x}]_i=l),$$

where $\mathbb{1}(\cdot)$ is an indicator function having value 1 if the condition in the argument is true and 0 otherwise.

Given N, D, (14) and (15), the set of possible AoI values $\mathcal{S}$, $\boldsymbol{\Delta}[k]\in\mathcal{S}$, $k\in\mathbb{N}_+$ generated by any policy $\gamma\in\Gamma$ becomes

$$\begin{aligned}\mathcal{S}=\Big\{\boldsymbol{x}\in\mathbb{N}_+^N\,|c(\boldsymbol{x},0)=D,&\\ c(\boldsymbol{x},l)\le D,\quad l\in\mathbb{N}_{++},&\\ c(\boldsymbol{x},\bar{\Delta})\ge\bar{N}\Big\}.&\end{aligned} \tag{16}$$

Assumption 1. *We assume $\boldsymbol{\Delta}[k]\in\mathcal{S}$, for $k\in\mathbb{N}_+$*

Definition 1. *We define the MDP $\mathcal{M}$ in the following way;*

- **State** at instant k is $\boldsymbol{\Delta}[k-1]$ and state space $\mathcal{S}$.

- **Action** at instant k is $\boldsymbol{\pi}[k]$ and the action space $\mathcal{A} = \{1, ..., N\}^D$.
- **Transition probabilities** $P(\boldsymbol{\Delta}[k] \mid \boldsymbol{\Delta}[k-1], \boldsymbol{\pi}[k])$, given state and action at instant k, can be derived using (3).
- **Reward** $r(\boldsymbol{\Delta}[k-1], \boldsymbol{\pi}[k]) \in \mathbb{R}_-$ at instant k equals $-N^{-1} f(\boldsymbol{\Delta}[k])$ in (10) and is given by the reward function $r : \{\mathcal{S}, \mathcal{A}\} \to \mathbb{R}_-$.

As can be seen above, the scheduling problem in (9) can be formulated as an MDP, where the current state, $\boldsymbol{\Delta}[k-1]$, only depends on the previous state and the scheduling decision $\boldsymbol{\pi}[k]$. Thus, we formalize the following lemma.

Lemma 1. *If the estimator uses only the most recent measurement from each sensor, then it suffices to consider only the restricted information set with only the previous AoI* $\bar{\mathcal{I}}[k] = \{\boldsymbol{\Delta}[k-1]\} \subseteq \mathcal{I}[k]$ *for the scheduling decision at time* $k = 1, ..., k-1$ *instead of* $\mathcal{I}[k] = \{\boldsymbol{\Delta}[0], \boldsymbol{\Delta}[1], \dots, \boldsymbol{\Delta}[k-1]\}$ *at the scheduler to determine an optimal policy* γ^*.

Proof. As shown, the problem in (9) can be modeled as the Markov decision process $\mathcal{M}$, where at instant k, the state is $\boldsymbol{\Delta}[k-1]$, the action is $\boldsymbol{\pi}[k]$ and the reward is $-N^{-1} f(\boldsymbol{\Delta}[k])$. The mathematical model fulfills Markovian properties, such that the state transition and reward only depend on the state and action at time k. From (3), the transition from $\boldsymbol{\Delta}[k-1]$ to $\boldsymbol{\Delta}[k]$, given $\boldsymbol{\pi}[k]$, is independent of time instant $k \in \mathbb{N}_+$. From (8) and (9), the time horizon is infinite, i.e., $T \to \infty$; hence, it is unnecessary to include the time instant $k \in \mathbb{N}_+$ in state definition. Thus, it suffices to reduce the information set $\mathcal{I}[k]$ to $\bar{\mathcal{I}}[k] = \{\boldsymbol{\Delta}[k-1]\}, \forall k \in \mathbb{N}_+$ and find a policy γ^* that minimizes the time-average reward in $\mathcal{M}$. □

Hereafter, the information set is restricted to $\bar{\mathcal{I}}[k] = \{\boldsymbol{\Delta}[k-1]\}$. Let $g_\gamma : \mathcal{S} \to \mathbb{R}_-$ be a function giving the average reward for policy γ in $\mathcal{M}$

$$g_\gamma(\boldsymbol{\Delta}[0]) = \lim_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} \mathbb{E}\Big[r(\boldsymbol{\Delta}[k-1], \boldsymbol{\pi}[k]) \Big| \gamma, \boldsymbol{\Delta}[0]\Big]. \tag{17}$$

Comparing (17) with (9), we can see that an optimal policy γ^* satisfies to maximize the average reward in $\mathcal{M}$, i.e.,

$$g_{\gamma^*}(\boldsymbol{\Delta}[0]) \geq g_\gamma(\boldsymbol{\Delta}[0]), \quad \boldsymbol{\Delta}[0] \in \mathcal{S}. \tag{18}$$

If $\boldsymbol{\Delta}[k-1]$ represents the state, $\mathcal{M}$ has an infinite countable state-space $\mathcal{S}$, for which an average reward optimal policy γ^* may not exist or is prohibitively complex to derive [46]. Therefore, we shall use another state-variable for the MDP that corresponds to a finite state-space.

In the following section, we show that if the temporal correlation in (1) is zero beyond a point, i.e., $\varphi(x) = 0,\ \forall x \geq m$, we can map $\mathcal{S}$ in (16) to an equivalent finite set and model the scheduling problem using a finite state-space to derive γ^*.

B. Finite-state MDP

In this section, we will define the finite-state MDP by first introducing a state-vairble by truncating AoI values larger then m, based on the criteria that $\varphi(x) = 0,\ \forall x \geq m$. Later on, we define an optimal scheduling policy based on the truncated AoI that minimizes the time-average MSE. If such a policy is known, it can then be used to derive an optimal scheduling policy γ^*.

From (10) and (12), 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.,

$$\begin{aligned}\lim_{\Delta_i[k] \to \infty} &|f([\Delta_1[k], ..., \Delta_i[k]+1, ..., \Delta_N[k]]^{\mathrm{T}}) \\ &- f([\Delta_1[k], ..., \Delta_i[k], ..., \Delta_N[k]]^{\mathrm{T}})| = 0, \quad i = 1, ..., N.\end{aligned} \tag{19}$$

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

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

Assumption 2, together with (7), gives that the information at Estimator j, $y_j[k]$, whose AoI exceeds m, i.e., $\Delta_j[k] \geq m$, is uncorrelated with all processes at time k, i.e., $\mathrm{E}\big[\theta_i[k] y_j[[k]\big] = 0, \forall i = 1, ..., N$. As a consequence, the infinite state space $\mathcal{S}$ maps to a finite-set of MSE values, i.e., $f : \mathcal{S} \to \mathcal{Y}$ with $|\mathcal{Y}| < \infty$. This gives that any of the elements $\tilde{\Delta}_i[k], \tilde{\Delta}_{ij}[k] \in \{0, 1, ..., m\}$, $\forall i, j = 1, .., N$, belonging to the AoI vector $\boldsymbol{\Delta}[k]$ can be truncated to m while still corresponding to the same MSE value in (10).

Based on the former mentioned properties, we introduce a variable that pertains to all possible MSE values and belongs to a finite set. Let $\tilde{\boldsymbol{\Delta}}[k] \in \{0, 1, ..., m\}^{N^2}$ contain the elements $\tilde{\Delta}_i[k], \tilde{\Delta}_{ij}[k] \in \{0, 1, ..., m\}$, $\forall i, j = 1, .., N$, i.e.,

$$\begin{aligned}\tilde{\Delta}_i[k] &= \big[\Delta_i[k]\big]_+^m, \quad i = 1, ..., N, \\ \tilde{\Delta}_{ij}[k] &= \big[|\Delta_i[k] - \Delta_j[k]|\big]_+^m = \big[\Delta_{ij}[k]\big]_+^m, \quad i, j = 1, ..., N,\end{aligned} \tag{20}$$

where $m \in \mathbb{N}_+$, $\big[\,\cdot\,\big]_+^m$ is defined as the truncation operator $\big[x\big]_+^m \triangleq \min\{x, m\}$, $x \in \mathbb{R}_+$ and $\tilde{\boldsymbol{\Delta}}[k]$ denotes the truncated AoI [47].

Let $b : \mathbb{N}_+^N \to \{0, 1, ..., m\}^{N^2}$ be a mapping from $\boldsymbol{\Delta}[k]$ to $\tilde{\boldsymbol{\Delta}}[k]$, i.e., $\tilde{\boldsymbol{\Delta}}[k] = b(\boldsymbol{\Delta}[k])$. Applying b on the set of possible AoI values $\mathcal{S}$ in (16), gives the finite set of possible truncated AoI values

$$\tilde{\mathcal{S}} = \{b(\boldsymbol{\Delta}) \mid \boldsymbol{\Delta} \in \mathcal{S}\}. \tag{21}$$

We can express the MSE as a function of $\tilde{\boldsymbol{\Delta}}[k]$ as follows

$$\begin{aligned}\tilde{f}(\tilde{\boldsymbol{\Delta}}[k]) &= \sum_{i=1}^{N} \mathbb{E}\Big[(\theta_i[k] - \hat{\theta}_i[k])^2 \Big| \tilde{\boldsymbol{\Delta}}[k]\Big] \\ &= \mathrm{tr}\Big(\mathbf{C}_{\theta\theta} - \tilde{\mathbf{C}}_{\theta y}[k] (\tilde{\mathbf{C}}_{yy})^{-1}[k] (\tilde{\mathbf{C}}_{\theta y}[k])^{\mathrm{T}}\Big),\end{aligned} \tag{22}$$

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

$$\begin{aligned}[\tilde{\mathbf{C}}_{yy}[k]]_{i,j} &= \sigma_i \sigma_j \rho_{ij} \varphi(\tilde{\Delta}_{ij}[k]) + \xi^2 \delta(i-j), \\ [\tilde{\mathbf{C}}_{\theta y}[k]]_{i,j} &= \sigma_i \sigma_j \rho_{ij} \varphi(\tilde{\Delta}_j[k]), \quad i, j \in \{1, ..., N\}.\end{aligned} \tag{23}$$

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

Remark 1. Note in (20) that $\Delta_{ij}[k] = \Delta_{ji}[k]$, for $i, j = 1, .., N$. It is, therefore, sufficient to store only one of the two elements to reduce the dimension of $\tilde{\mathcal{S}}$.

Proposition 2. *Under Assumption 2, the following relationship holds*

$$f(\mathbf{\Delta}[k]) = \tilde{f}(\tilde{\mathbf{\Delta}}[k]), \quad \forall \mathbf{\Delta}[k] \in \mathbb{N}_+^N. \tag{24}$$

Proof. The proof is given in Appendix B. □

Similar to the AoI, $\mathbf{\Delta}[k]$, the truncated AoI, $\tilde{\mathbf{\Delta}}[k]$, depends on the previous value $\tilde{\mathbf{\Delta}}[k-1]$ and scheduling variable $\boldsymbol{\pi}[k]$; hence, it can be expressed as a function using (3) and (20).

Proposition 3. *The truncated AoI $\tilde{\mathbf{\Delta}}[k]$ can be expressed as a function of $\tilde{\mathbf{\Delta}}[k-1]$ and $\boldsymbol{\pi}[k]$ as*

$$\tilde{\Delta}_i[k] = \begin{cases} 0, & \text{if } i \in \boldsymbol{\pi}[k], \\ \left[\tilde{\Delta}_i[k-1]+1\right]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k], \end{cases} \tag{25}$$

$$\tilde{\Delta}_{ij}[k] = \begin{cases} 0, & \text{if } i, j \in \boldsymbol{\pi}[k], \\ \left[\tilde{\Delta}_{ij}[k-1]\right]_+^m, & \text{if } i, j \notin \boldsymbol{\pi}[k], \\ \left[\tilde{\Delta}_i[k-1]+1\right]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k],\, j \in \boldsymbol{\pi}[k], \\ \left[\tilde{\Delta}_j[k-1]+1\right]_+^m, & \text{if } i \in \boldsymbol{\pi}[k],\, j \notin \boldsymbol{\pi}[k]. \end{cases} \tag{26}$$

Proof. The proof is given in Appendix C. □

In Section III, the scheduling problem was modeled as an infinite state-space MDP with $\mathbf{\Delta}[k-1]$ as a state variable $\tilde{\mathbf{\Delta}}[k]$. Proposition 2 and Proposition 3 show that $\tilde{\mathbf{\Delta}}[k]$ and $\mathbf{\Delta}[k]$ corresponds to the same MSE and, if either $\mathbf{\Delta}[k]$, or $\tilde{\mathbf{\Delta}}[k]$, is known, any determined scheduling sequence that follows after k will result in the same sequence of MSE values.

Definition 2. *We define the finite state-space MDP, $\tilde{\mathcal{M}}$, as follows;*

- **Action** at instant k is the scheduling decision $\boldsymbol{\pi}[k]$ belonging to action-space $\mathcal{A} = \{1, ..., N\}^D$.
- **State** at instant k is the truncated AoI $\tilde{\mathbf{\Delta}}[k-1]$ belonging to state-space $\tilde{\mathcal{S}}$ in (21).
- **Transition probabilities** $P(\tilde{\mathbf{\Delta}}[k] \mid \tilde{\mathbf{\Delta}}[k-1], \boldsymbol{\pi}[k]) \in \{0, 1\}$ are binary and given by (25) and (26) in Proposition 3.
- **Reward** $\tilde{r}(\tilde{\mathbf{\Delta}}[k-1], \boldsymbol{\pi}[k]) \in \mathbb{R}$ at instant k equals $\tilde{r}[k] = -N^{-1}\tilde{f}(\tilde{\mathbf{\Delta}}[k])$ in (22) given by the reward function $\tilde{r} : \{\tilde{\mathcal{S}}, \mathcal{A}\} \to \mathbb{R}_-$.

Let $\tilde{\gamma}_k : \tilde{\mathcal{S}} \to \mathcal{A}$ be a scheduling strategy based on $\tilde{\mathbf{\Delta}}[k]$ as

$$\boldsymbol{\pi}[k] = \tilde{\gamma}_k(\tilde{\mathbf{\Delta}}[k-1]), \tag{27}$$

where $\tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, ..., \tilde{\gamma}_T)$ is a scheduling policy $\tilde{\gamma} \in \tilde{\Gamma}$. We define the average reward function in $\tilde{\mathcal{M}}$, $\tilde{g}_{\tilde{\gamma}} : \tilde{\mathcal{S}} \to \mathbb{R}_+$ as

$$\tilde{g}_{\tilde{\gamma}}(\tilde{\mathbf{\Delta}}[0]) = \lim_{T\to\infty} \frac{1}{T} \sum_{k=1}^{T} \mathbb{E}\Big[\tilde{r}(\tilde{\mathbf{\Delta}}[k-1], \boldsymbol{\pi}[k]) \Big| \tilde{\gamma}, \tilde{\mathbf{\Delta}}[0]\Big], \tag{28}$$

where an *optimal truncated scheduling policy* $\tilde{\gamma}^*$ is average reward optimal for $\tilde{\mathcal{M}}$ and fulfills

$$\tilde{g}_{\tilde{\gamma}^*}(\tilde{\mathbf{\Delta}}[0]) \geq \tilde{g}_{\tilde{\gamma}}(\tilde{\mathbf{\Delta}}[0]),\ \forall \tilde{\mathbf{\Delta}}[0]) \in \tilde{\mathcal{S}}. \tag{29}$$

Theorem 1. *Under Assumption 1 and Assumption 2, if an optimal truncated scheduling policy $\tilde{\gamma}^* = (\tilde{\gamma}_1^*, \tilde{\gamma}_2^*, ..., \tilde{\gamma}_T^*)$ exists, we can obtain $\gamma^* = (\gamma_1^*, \gamma_2^*, ..., \gamma_T^*)$ as*

$$\gamma_k^* := \tilde{\gamma}_k^* \circ b, \quad k \in \mathbb{N}_+, \tag{30}$$

where $\circ$ is the function composition operator.

Proof. The proof is given in Appendix D. □

Theorem 1 states the relationship between an optimal truncated scheduling policy $\tilde{\gamma}^*$ and an optimal scheduling policy γ^*. In the following section, we first prove the existence of $\tilde{\gamma}^*$ and how to derive it. Later on, we use $\tilde{\gamma}^*$ to derive γ^*

C. Optimal scheduling policy

We begin this section by presenting some important definitions and mathematical properties of $\tilde{\mathcal{M}}$, to be used to prove the existence of $\tilde{\gamma}^*$, and how to derive it. We then obtain γ^* using (30). The section ends by demonstrating that an optimal scheduling policy results in a periodic scheduling pattern.

A policy $\tilde{\gamma} = (\tilde{\gamma}_1, \tilde{\gamma}_2, ..., \tilde{\gamma}_T)$ is said to be deterministic if all the scheduling strategies $\tilde{\gamma}_k$ are deterministic functions, $\tilde{\gamma}_k : \tilde{\mathcal{S}} \to \mathcal{A}$, $\forall k \in \mathbb{N}_+$, which is the case in (27). A policy $\tilde{\gamma}$ is said to be stationary if the decision rules $\tilde{\gamma}_k$ are independent of time k, i.e., $\tilde{\gamma}_k := \tilde{\gamma}_0, \forall k \in \mathbb{N}_+$. Let $\tilde{\Gamma}^S \subset \tilde{\Gamma}$ be the set of all stationary policies $\tilde{\gamma}$.

Let $P_{\tilde{\gamma}}(\tilde{\mathbf{\Delta}}[k+l] \mid \tilde{\mathbf{\Delta}}[k])$, $l \in \mathbb{N}_+$, be the probability that the Markov chain transitions from $\tilde{\mathbf{\Delta}}[k]$ to $\tilde{\mathbf{\Delta}}[k+l]$ in l-time instances, given policy $\tilde{\gamma}$. A state $\tilde{\mathbf{\Delta}} \in \tilde{\mathcal{S}}$ is recurrent if, once reached, the process will return to that state within a finite time horizon, i.e., $\exists l < \infty$, $P_{\tilde{\gamma}}(\tilde{\mathbf{\Delta}}[k+l] = \tilde{\mathbf{\Delta}} \mid \tilde{\mathbf{\Delta}}[k] = \tilde{\mathbf{\Delta}}) = 1$.

Proposition 4. *If $\tilde{\mathbf{\Delta}}[0] \in \tilde{\mathcal{S}}$ and $\tilde{\gamma} \in \tilde{\Gamma}^S$ are applied, the process $\tilde{\mathbf{\Delta}}[k]$ evolves to a set of recurrent states $\tilde{\mathcal{S}}_{\tilde{\gamma}} \subseteq \tilde{\mathcal{S}}$, i.e.,*

$$\tilde{\mathbf{\Delta}}[k] \in \tilde{\mathcal{S}}_{\tilde{\gamma}}, \quad \forall k \geq |\mathcal{S}| \tag{31}$$

Proof. Assume $\tilde{\gamma}$ is a stationary deterministic policy $\tilde{\gamma} \in \tilde{\Gamma}^S$. From the definition of $\tilde{\mathcal{M}}$, the transition probabilities are binary, i.e., $P(\tilde{\mathbf{\Delta}}[k+1] \mid \tilde{\mathbf{\Delta}}[k], \boldsymbol{\pi}[k]) \in \{0, 1\}$. Hence, given the policy $\tilde{\gamma}$ and state at instant k, $\tilde{\mathbf{\Delta}}[k]$, any future state $\tilde{\mathbf{\Delta}}[k+l]$ is perfectly known. Since, the number of states is finite, i.e., $|\tilde{\mathcal{S}}| < \infty$, there can be a maximum number of $|\tilde{\mathcal{S}}|$ actions before any state is re-visited for any stationary deterministic policy $\tilde{\gamma} \in \tilde{\Gamma}^S$. If any state is re-visited, the state-action process will keep repeating itself. Hence, $\tilde{\mathcal{S}}_{\tilde{\gamma}} \subseteq \tilde{\mathcal{S}}$ represents a set of recurrent states $\tilde{\mathbf{\Delta}}[k] \in \tilde{\mathcal{S}}_{\tilde{\gamma}}$, $\forall k \geq |\mathcal{S}|$. □

From Proposition 4 and [46], we find that if an average optimal policy $\tilde{\gamma}^*$ for $\tilde{\mathcal{M}}$ exist, it results in a constant average reward $g_{\gamma^*}(\tilde{\mathbf{\Delta}}[0]) = g^* \in \mathbb{R}$, $g^* \in \mathbb{R}$, $\forall \tilde{\mathbf{\Delta}}[0] \in \tilde{\mathcal{S}}$. The scalar g^* must then satisfy the optimality equations

$$\max_{\boldsymbol{\pi}\in\mathcal{A}} \Big\{ r(\tilde{\mathbf{\Delta}}, \boldsymbol{\pi}) - g^* + \sum_{\tilde{\mathbf{\Delta}}' \in \mathcal{S}} P(\tilde{\mathbf{\Delta}}' | \tilde{\mathbf{\Delta}}, \boldsymbol{a}) h(\tilde{\mathbf{\Delta}}') - h(\tilde{\mathbf{\Delta}}) \Big\} = 0, \tag{32}$$

where $h : \tilde{\mathcal{S}} \to \mathbb{R}$, and $h \in V$, where V is the set of bounded functions on $\tilde{\mathcal{S}}$. A solution for (32) is given in the following lemma.

Algorithm 1 Finding scheduling policy γ^*

1: Define $\tilde{\mathcal{M}} = \{\mathcal{A}, \tilde{\mathcal{S}}, \tilde{r}, P(\cdot \mid \cdot)\}$ as in Section III-B, given N, D, $\mathbf{C}_{\theta\theta}$, ξ, φ and m
2: Set $n = 0$ and select arbitrary policy $\tilde{\gamma}^n \in \tilde{\Gamma}^S$
3: Obtain $g \in \mathbb{R}$ and $\boldsymbol{h} \in \mathbb{R}^{|\tilde{\mathcal{S}}|}$, $[\boldsymbol{h}]_i = h(\tilde{\boldsymbol{\Delta}}_i)$, $\forall \tilde{\boldsymbol{\Delta}}_i \in \tilde{\mathcal{S}}$, by solving (32), below, represented in vector form

$$\tilde{\boldsymbol{r}}_{\tilde{\gamma}^n} + [-\mathbf{1} \mid (\mathbf{P}_{\tilde{\gamma}} - \mathbf{I})] \begin{bmatrix} g \\ \boldsymbol{h} \end{bmatrix} = \mathbf{0},$$

where $\tilde{\boldsymbol{r}}_{\tilde{\gamma}} \in \mathbb{R}^{|\tilde{\mathcal{S}}|}$, $[\tilde{\boldsymbol{r}}]_i = r(\tilde{\boldsymbol{\Delta}}_i, \tilde{\gamma}(\tilde{\boldsymbol{\Delta}}_i))$, $\forall \tilde{\boldsymbol{\Delta}}_i \in \tilde{\mathcal{S}}$, is a reward vector, $\mathbf{1} = (1, 1, ..., 1)^{\mathrm{T}}$, $\mathbf{1} \in \mathbb{R}^{|\tilde{\mathcal{S}}|}$ is a vector of ones, $\mathbf{P}_{\tilde{\gamma}} \in \mathbb{R}^{|\tilde{\mathcal{S}}| \times |\tilde{\mathcal{S}}|}$, $[\mathbf{P}_{\tilde{\gamma}}]_{i,j} = P(\tilde{\boldsymbol{\Delta}}_j | \tilde{\boldsymbol{\Delta}}_i, \tilde{\gamma}(\tilde{\boldsymbol{\Delta}}_i))$ is the transition matrix..
4: Get policy $\tilde{\gamma}^{n+1} = (\tilde{\gamma}_0^{n+1}, \tilde{\gamma}_0^{n+1}, ..., \tilde{\gamma}_0^{n+1})$, $\forall \tilde{\boldsymbol{\Delta}} \in \mathcal{S}$, by solving (33)

$$\tilde{\gamma}_0^{n+1}(\tilde{\boldsymbol{\Delta}}) = \underset{\boldsymbol{\pi} \in \mathcal{A}}{\arg\max} \{ r(\tilde{\boldsymbol{\Delta}}, \boldsymbol{a}) + \sum_{\tilde{\boldsymbol{\Delta}}' \in \tilde{\mathcal{S}}} P(\tilde{\boldsymbol{\Delta}}' | \tilde{\boldsymbol{\Delta}}, \boldsymbol{a}) h(\tilde{\boldsymbol{\Delta}}') \}$$

5: **if** $\tilde{\gamma}^{n+1} = \tilde{\gamma}^n$ **then**
6: Stop and set $\tilde{\gamma}^* = \tilde{\gamma}^{n+1}$
7: **else**
8: Return to Step 2 using $\tilde{\gamma}^{n+1}$
9: **end if**
10: Obtain $\gamma^* = (\gamma_0^*, ..., \gamma_0^*)$ as $\gamma_0^* := \tilde{\gamma}_0^* \circ b$ in (30)

Lemma 2. *For the finite state MDP $\tilde{\mathcal{M}}$, there exists an optimal truncated scheduling policy $\tilde{\gamma}^* = (\tilde{\gamma}_0^*, \tilde{\gamma}_0^*, ..., \tilde{\gamma}_0^*)$ with $\tilde{\gamma}^* \in \tilde{\Gamma}^S$ corresponding to reward $g_{\tilde{\gamma}^*}(\tilde{\boldsymbol{\Delta}}) = g^*$, $g^* \in \mathbb{R}_-$, $\forall \tilde{\boldsymbol{\Delta}} \in \tilde{\mathcal{S}}$, given by*

$$\tilde{\gamma}_0^*(\tilde{\boldsymbol{\Delta}}) = \underset{\boldsymbol{\pi} \in \mathcal{A}}{\arg\max} \{ r(\tilde{\boldsymbol{\Delta}}, \boldsymbol{a}) + \sum_{\tilde{\boldsymbol{\Delta}}' \in \mathcal{S}} P(\tilde{\boldsymbol{\Delta}}' | \tilde{\boldsymbol{\Delta}}, \boldsymbol{a}) h^*(\tilde{\boldsymbol{\Delta}}') \}, \tag{33}$$

where $h^* \in V$ and g^* satisfy (32). The policy $\tilde{\gamma}^*$ can be obtained in a finite number of iterations using algorithm policy iteration.

Proof. The proof is given in Appendix E. □

Given the existence and possibility to derive $\tilde{\gamma}^*$, we summarize the theoretical findings and formulate the following theorem.

Theorem 2. *Under Assumption 1 and Assumption 2, there exists an optimal stationary scheduling policy $\gamma^* = (\gamma_0^*, ..., \gamma_0^*)$, where $\gamma_0^* = \tilde{\gamma}_0^* \circ b$ and $\tilde{\gamma}^* = (\tilde{\gamma}_0^*, \tilde{\gamma}_0^*, ..., \tilde{\gamma}_0^*)$, which can be derived in a finite number of iterations using policy iteration. The policy results in a periodic scheduling sequence.*

Proof. The proof follows from Theorem 1 and Lemma 2. □

Based on Lemma 2 and Theorem 2, we show in Algorithm 1 how an optimal scheduling policy γ^* can be derived, by first deriving an optimal truncated scheduling policy $\tilde{\gamma}^*$ using policy iteration, to later obtain γ^* using (30).

Theorem 2 states that γ^* results in a periodic scheduling sequence. The following section will demonstrate that applying a periodic scheduling sequence results in the same periodic sequence of truncated AoI states, regardless of the initial truncated AoI. We show how this result can be utilized to save data storage at the scheduler, as it only needs to store the periodic scheduling pattern and not the entire policy, mapping every truncated AoI value to a scheduling decision.

Recall from (6), that the estimator is based on solely the most recent measurement from each sensor. If there is no temporal correlation, the estimator in (6) is optimal, since previous measurements are uncorrelated. In that case, the estimator together with the optimal scheduling policy results in a joint optimal scheduling-estimator pair. On the other hand, if the spatial correlation is weak and the temporal correlation is strong, the estimator is suboptimal and can be improved by utilizing the measurements received in prevision time instants. However, there is a trade-off between the number of previous measurements utilized in the estimator and the numerical complexity in deriving an optimal scheduling policy.

IV. Properties of Periodic Scheduling Sequences

As stated in the previous section, an optimal scheduling policy results in a periodic scheduling pattern. The scheduler could store and execute the periodic scheduling pattern instead of the complete state-action policy to save data storage. The requirement for this is that both approaches result in the same performance. In this section, we demonstrate that any periodic scheduling decision results in the same performance regardless of the initial AoI value.

Let $\gamma \in \Gamma$ be a scheduling policy that results in a periodic scheduling sequence with period $n \in \mathbb{N}_+$, such that

$$\boldsymbol{\pi}[k+n] = \boldsymbol{\pi}[k], \quad \forall k \in \mathbb{N}_+$$

regardless of the previous AoI $\boldsymbol{\Delta}[k-1] \in \mathcal{S}$. Let $\boldsymbol{\Pi} \in \{1, 2, ..., N\}^{D \times n}$, $n \in \mathbb{R}_+$, $n < \infty$, be a matrix that represents a defined periodic scheduling-sequence that the policy γ results in, where each of the n columns represents a scheduling decision in the sequence, i.e., $[\boldsymbol{\Pi}]_{*,j} \in \{1, 2, ..., N\}^D$, $j = 1, ..., n$, is the jth column of matrix $\boldsymbol{\Pi}$. Based on $\boldsymbol{\Pi}$, the scheduling strategy at instant k is defined as

$$[\boldsymbol{\Pi}]_{*,\phi(k)} = \gamma_k(\boldsymbol{\Delta}[k-1]), \tag{34}$$

where $\phi : \mathbb{N}_+ \to \{1, 2, .., n\}$ is defined as

$$\phi(k) = \begin{cases} k \mod n, & \text{if } k \mod n \neq 0, \\ n, & \text{if } k \mod n = 0. \end{cases}$$

As seen in (34), the scheduling decision only depends on the time instant k and results in a periodic scheduling sequence

$$\gamma_{k+nl}(\boldsymbol{\Delta}[k-1+nl]) = \gamma_k(\boldsymbol{\Delta}[k-1]), \quad \forall k, l \in \mathbb{N}_+. \tag{35}$$

We now want to determine how the AoI sequence, generated by the policy in (34), evolves over time. If γ is applied and sensor i is never scheduled, the AoI goes to infinity; otherwise, it becomes periodic, i.e.,

$$\mathbb{E}\big[\Delta_i[k] \big| \gamma\big] = \begin{cases} n - \sup\limits_{j \in \{1, ..., n\}} \{ i \in [\boldsymbol{\Pi}_k]_{*,j} \}, & \text{if } i \in \boldsymbol{\Pi} \\ \infty, & \text{if } i \notin \boldsymbol{\Pi}, \end{cases} \tag{36}$$

where $\mathbf{\Pi}_k = \big[[\mathbf{\Pi}]_{*,\phi(k)+1:n}, [\mathbf{\Pi}]_{*,1:\phi(k)}\big]$, $\mathbf{\Pi}_k \in \{1,2,...,N\}^{D\times n}$, and is derived by rearranging the n columns of matrix $\mathbf{\Pi}$.

Applying (20) to (36), yields the corresponding truncated AoI sequence. Due to the truncation operator, the truncated AoI sequence becomes periodic, regardless of whether every sensor is scheduled or not. Let $l_\gamma \in \mathbb{N}_+$ $l_\gamma \le n$, represent the fundamental period of the truncated AoI sequence generated by γ

$$l_\gamma = \inf_{l\le n} \lim_{k\to\infty} \Big\{\mathbb{E}\big[\tilde{\mathbf{\Delta}}[k+l]\big|\gamma\big] = \mathbb{E}\big[\tilde{\mathbf{\Delta}}[k]\big|\gamma\big]\Big\}. \tag{37}$$

Let $\tilde{\mathcal{S}}_\gamma$ be the periodic sequence of $\tilde{\mathbf{\Delta}}[k]$, i.e.,

$$\tilde{\mathcal{S}}_\gamma = \lim_{k\to\infty} \Big\{\mathbb{E}\big[\tilde{\mathbf{\Delta}}[k+1]\big|\gamma\big], \mathbb{E}\big[\tilde{\mathbf{\Delta}}[k+2]\big|\gamma\big], ..., \\ \mathbb{E}\big[\tilde{\mathbf{\Delta}}[k+l_\gamma]\big|\gamma\big] \,\Big|\, \forall \tilde{\mathbf{\Delta}}[k] \in \tilde{\mathcal{S}}\Big\}, \tag{38}$$

where $l_\gamma = |\tilde{\mathcal{S}}_\gamma|$. The set $\tilde{\mathcal{S}}_\gamma$ is independent of the initial state $\mathbf{\Delta}[0]$. Thus, based on (37) and (38) the cost for any policy that result in a periodic scheduling sequence $\mathbf{\Pi}$ can be calculated as

$$\lim_{T\to\infty} J(\gamma,T) = \frac{1}{Nl_\gamma} \sum_{\tilde{\mathbf{\Delta}}\in\tilde{\mathcal{S}}_\gamma} \tilde{f}(\tilde{\mathbf{\Delta}}). \tag{39}$$

From Theorem 2, we know that an optimal stationary policy results in a periodic scheduling sequence. The results in (36)-(39) show that if the particular periodic scheduling sequence is known, we can derive an optimal scheduling policy γ^* as in (34) based on the periodic scheduling sequence that γ^* results. This allows the scheduler to save data storage as it does not need to store a look-up table mapping each possible truncated AoI value to a scheduling decision. For an optimal stationary scheduling policy, the period of the resulting periodic scheduling sequence, l_{γ^*}, is bounded by the cardinality of the state-space $|\tilde{\mathcal{S}}|$ as $l_{\gamma^*} \le |\tilde{\mathcal{S}}|$.

The computational complexity of deriving γ^* becomes challenging for large values of N or m. As seen in (16) and (21), the set of truncated AoI values $\tilde{\mathbf{\Delta}}$, hence the state-space $\tilde{\mathcal{S}}$, grows exponentially with N and m. In contrast, it decreases with D. The system parameters N, D and m affect the number of possible states, the number of possible actions, and the computational complexity to calculate the reward for each state and action using expression (22). In the following section, we present periodic suboptimal scheduling policies that allow for a reduced computational complexity.

V. Construction of Suboptimal Policies

This section considers suboptimal policies to avoid the exponential growth of the state space associated with the optimal policies in previous sections. One approach to overcome this problem is to derive scheduling policies using other less-computationally heavy methods than the one presented in Section III-C. These are based on optimizing other objective functions than the infinity time-average MSE.

In this section we present three suboptimal scheduling policies; the first two are based on; minimizing the AoI across the sensors, while the third; approximating the truncated AoI state space to a smaller finite state space.

A. Minimizing the AoI

For most works regarding AoI, the objective is to minimize the maximum or average AoI [23]–[25]. To achieve this for our system, one can schedule the D sensors with the highest AoI in a round-robin fashion, which results in a periodic scheduling pattern. To calculate the performance of a round-robin policy, one can make use of expressions (36)-(39).

The performance of a round-robin policy depends on the spatial dependencies ρ_{ij} and the marginal variances σ_i, $i = 1,...,N$ in (2). The performance can be enhanced if the order of the periodic scheduling sequence is re-organized in an optimal fashion. Thus, we present an optimal order round-robin scheduling policy as

$$\min_{\gamma} \lim_{T\to\infty} J(\gamma,T), \\ \text{s.t.} \lim_{k\to\infty} \sup \mathbb{E}\big[\Delta_i[k]\big|\gamma\big] = \bar{\Delta}, \quad \forall i = 1,2,...,N, \tag{40}$$

where the AoI is limited to fairness constraint $\bar{\Delta}$, presented in (14). The way to solve (40) is to find which sensor index order of a round-robin policy minimizes the cost in (39).

Proposition 5. *For $\sigma_i = \sigma_j$ and $\rho_{ij} = \rho_0, \rho_0 \in [-1,1]$, $\forall i,j = 1,2,...,N$, a policy γ that satisfies (40) is an optimal policy γ^*.*

Proof. From (11) we know that the MSE either increases, or is the same, with respect to an increase of $\Delta_i[k]$, $i = 1,...,N$. Assume all process parameters are equal, i.e., $\sigma_i = \sigma_j$ and $\rho_{ij} = \rho_0$, $\rho_0 \in [-1,1]$, $\forall i,j = 1,2,...,N$, then if $\Delta_i[k] = \Delta_j[k]$, $\forall i,j = 1,2,...,N$, the change in MSE is equivalent with respect to $\Delta_i[k]$ and $\Delta_j[k]$, i.e.,

$$f([\Delta_1[k],...,\Delta_i[k]+1,...,\Delta_N[k]]^{\mathrm{T}}) = \\ f([\Delta_1[k],...,\Delta_j[k]+1,...,\Delta_N[k]]^{\mathrm{T}}). \tag{41}$$

Furthermore, all permutations of an AoI vector results in the same MSE. Given (11), (41) and (16), the smallest possible value of $f(\mathbf{\Delta}[k])$ corresponds to the one minimizing the AoI across all sensors. To achieve this, a round-robin scheduling policy is applied. □

B. Finite-horizon minimization

Similar to the works in [4], we propose a scheduling policy where the scheduler minimizes the time-average MSE over L-time steps given the current truncated AoI. The proposed scheduling policy is referred to as FHM-L, referring to an L-step finite-horizon minimization. The policy is constructed such that the scheduler decides the same scheduling sequence every time instant it returns to a previous truncated AoI value. Since the number of truncated AoI values are finite, the policy will result in a stationary deterministic policy γ defined by a periodic scheduling sequence $\mathbf{\Pi}$. An FHM-L scheduling policy γ can, therefore, easily be calculated offline using Algorithm 2.

If there is no temporal dependency, i.e., $m = 1$ in Assumption 2, the scheduling problem becomes a one-shot decision problem, independent of previous scheduling decisions. Thus, the scheduler will choose the same D sensors, which minimize (10), at every time instant. An FHM-L scheduling policy is in that case an optimal policy γ^*.

Algorithm 2 Finite-horizon minimization over L-time steps

1: Set $b(\mathbf{\Delta}[0]) = \tilde{\mathbf{\Delta}}[0]$, $n = 1$, $\tilde{\mathbf{\Delta}}_n = \mathbf{\Delta}[0]$ and $k = 1$

2: Obtain and store $\boldsymbol{\pi}_n = \boldsymbol{\pi}^*_{k:k+L-1}$ as

$$\boldsymbol{\pi}^*_{k:k+L-1} = \underset{\boldsymbol{\pi}_{k:k+L-1}}{\arg\min} \frac{1}{NL} \sum_{j=k}^{k+L-1} \tilde{f}(\tilde{\mathbf{\Delta}}[j]),$$

where $\boldsymbol{\pi}_{k:k+L-1} = (\boldsymbol{\pi}[k], \boldsymbol{\pi}[k+1], ..., \boldsymbol{\pi}[k+L-1])$

3: Apply $\boldsymbol{\pi}^*_{k:k+L-1}$ to obtain $\tilde{\mathbf{\Delta}}[k+L-1]$

4: Store $\tilde{\mathbf{\Delta}}_n = \tilde{\mathbf{\Delta}}[k+L-1]$

5: **if** $\tilde{\mathbf{\Delta}}_n \in \{\tilde{\mathbf{\Delta}}_0, \tilde{\mathbf{\Delta}}_1 ..., \tilde{\mathbf{\Delta}}_{n-1}\}$ **then**

6: Find $n_0 = \sup_{k \leq n}\{\tilde{\mathbf{\Delta}}_k = \tilde{\mathbf{\Delta}}_n\}$

7: Set $\mathbf{\Pi} = [\boldsymbol{\pi}_{n_0}, \boldsymbol{\pi}_{n_0+1}, ..., \boldsymbol{\pi}_{n-1}]$

8: Stop and set γ from $\mathbf{\Pi}$

9: **else**

10: Return to Step 2

11: **end if**

C. Finite state-space approximation

To handle the computational complexity of using dynamic programming for large state spaces $\tilde{\mathcal{S}}$, one can reduce the state-space by introducing an approximation for elements $\tilde{\mathbf{\Delta}}$ in $\tilde{\mathcal{S}}$ that correspond to similar MSE values. The approximation is done such that if the temporal correlation between process i and the most recent measurement of a process j is zero, i.e., $\mathrm{E}[\theta_i[k]x_j[k-\Delta_j[k]] = 0$ due to $\varphi(\Delta_j[k]) = 0$, but the AoI difference between the recent measurements $\Delta_{ij}[k]$ is not greater then m, i.e., $\Delta_{ij}[k] < m$, the AoI difference $\Delta_{ij}[k]$ has little contribution to the MSE. In that case, the AoI difference $\Delta_{ij}[k]$ can be approximated as m, without changing the corresponding MSE value much. The approximation $\hat{\mathbf{\Delta}}[k]$ is defined as

$$\hat{\Delta}_i[k] = \tilde{\Delta}_i[k]$$
$$\hat{\Delta}_{ij}[k] = \begin{cases} m, & \text{if } m = \max\{\tilde{\Delta}_i[k], \tilde{\Delta}_j[k]\} \\ \tilde{\Delta}_{ij}, & \text{else,} \end{cases} \tag{42}$$

for all $i, j = 1, 2, ..., N$.

Based on the definition in (42), we then model an MDP similar to $\mathcal{M}$, but use $\hat{\mathbf{\Delta}}[k]$ to represent the state-variable instead of $\tilde{\mathbf{\Delta}}[k]$. The transition probabilities can then be derived using (25) and (26) in Proposition 3. The action set and reward function is the same as in $\mathcal{M}$. Further, we use the same methodology as in Section III to first derive an optimal policy $\hat{\gamma}^*$ based on $\hat{\mathbf{\Delta}}[k-1]$ and then map it to a scheduling policy γ, similarly as in Theorem 1.

VI. Numerical Examples

In this section, we perform numerical simulations to evaluate the performance of an optimal policy and the presented suboptimal policies. To begin, we define a system with equal marginal variance across the sensors and a spatio-temporal correlation model that becomes zero beyond AoI m. We then investigate the performance discrepancy for a policy that is derived using the methods in Section III-C, while applying a smaller truncation time than m to derive the state space. Later on, we investigate the performance of an optimal policy

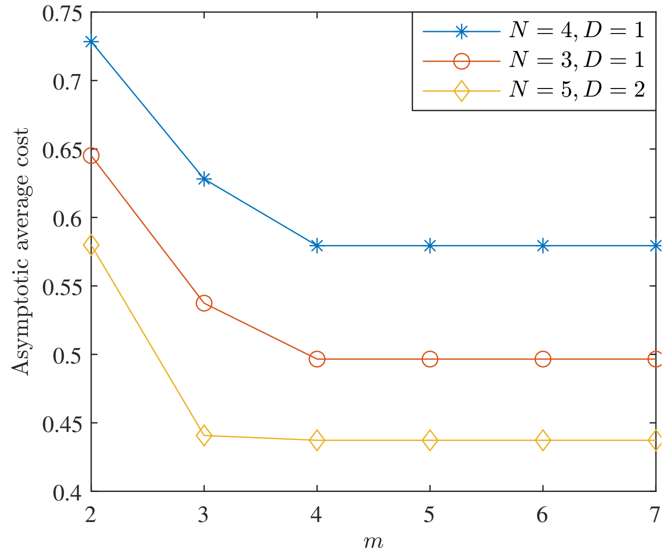

Fig. 2. Asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, vs m with $\sigma_i = 1$, $\forall i = 1, 2, ..., N$, $\xi = 0.5$, $\lambda = 0.35$ and $r_0 = 0.9$.

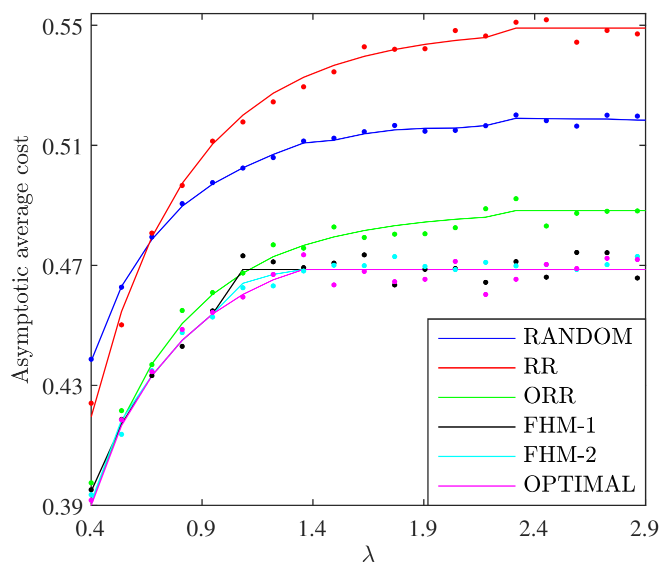

Fig. 3. Asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, vs λ with $N = 5$, $D = 2$, $\sigma_i = 1$, $\forall i = 1, 2, ..., N$, $\xi = 0.5$, and $r_0 = 0.5$.

and the suboptimal policies, firstly, given the degree of spatio-temporal, and secondly, the system size N and scheduling capacity D.

We assume a system where N sensors observe dependent processes with equal marginal variances, i.e., $\sigma_i = 1$, $i = 1, ..., N$, measurement noise $\xi = 0.5$ and the spatio-temporal dependency components in (1) are given by [43], [44]

$$\rho_{ij} = e^{-r_0|i-j|}, \quad \varphi(x) = e^{-\lambda x}\mathbb{1}(e^{-\lambda x} \geq 0.1),\ x \in \mathbb{R}_+, \tag{43}$$

where $\lambda \in \mathbb{R}_+$, is the temporal correlation decay factor, $r_0 \in \mathbb{R}_+$, is the spatial correlation decay factor and, $|i-j|$, represents the Euclidean distance between sensors i and j.

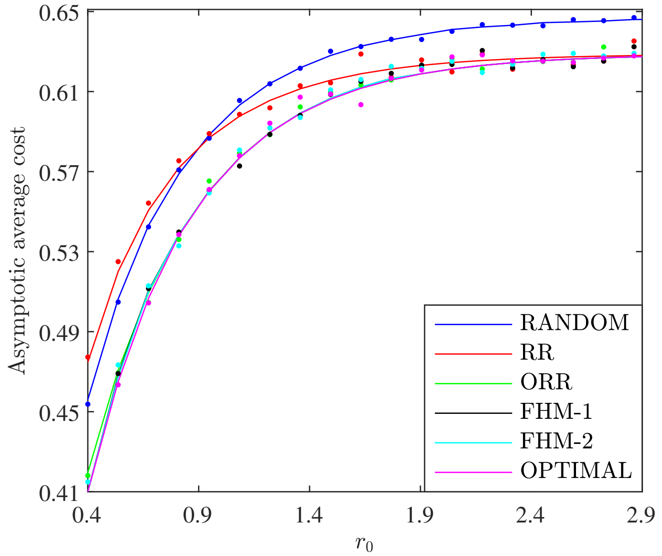

Fig. 4. Asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, vs r_0 with $N = 5$, $D = 2$, $\xi = 0.5$, $\sigma_i = 1$, $\forall i = 1, 2, ..., N$, and $\lambda = 0.8$.

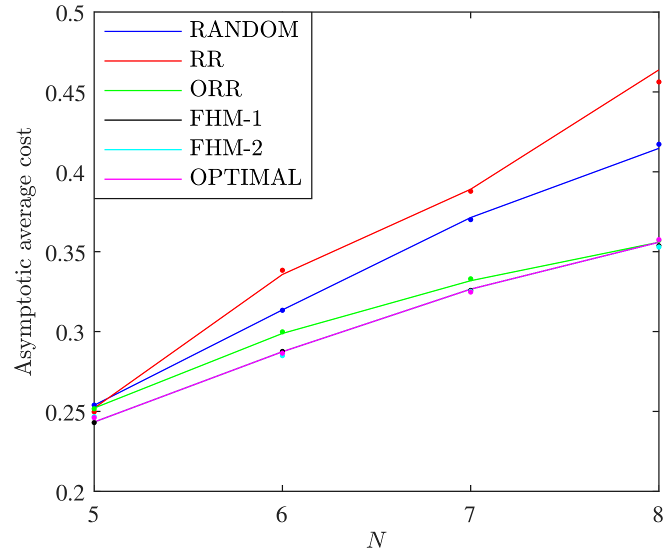

Fig. 5. Asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, vs N with $D = 4$, $\sigma_i = 1$, $\forall i = 1, 2, ..., N$, $\xi = 0.5$, $\lambda = 0.8$, , and $r_0 = 0.5$.

A. *Performance vs truncation time m*

The temporal correlation in (43) is zero for an AoI equal to and larger than 7, i.e., $\varphi(x) = 0$ for $x \geq 7$. If a scheduling policy is derived using the methods in Section III-C with $\tilde{\mathcal{S}}$ in (21) based on $m = 7$, an optimal policy γ^* is obtained. We investigate the performance difference if the state space $\tilde{\mathcal{S}}$ in (21) is calculated using a truncation value m smaller than 7, which results in a smaller state-space for the MDP and reduces the computational workload.

Figure 2 shows the asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, versus the truncation time m, when deriving the policy γ using Algorithm 1, with $\xi = 0.5$, $\lambda = 0.35$ and $r_0 = 0.9$. We see that a policy derived using $m \geq 4$, performs near-to-optimal, for all three combinations of N and D. This shows that it is possible to derive an near-to-optimal policy using a truncation time m smaller than $\inf_{x\in\mathbb{N}_+}\{\varphi(x) = 0\}$.

In the following sections, we examine the time-averaged MSE per sensor, i.e., cost in (8), for an optimal policy and also compare it to alternative scheduling policies under different combinations of N, D, λ and r_0. The truncation time m in (20) will then be strictly set to $m = \inf_{x\in\mathbb{N}_+}\{e^{-\lambda x} \leq 0.1\}$.

B. *Performance vs spatio-temporal dependency*

The periodic scheduling pattern and performance of an optimal policy γ^* depends on how fast the temporal correlation decays over time, and the spatial dependencies of the sensors. To determine the performance gain of deriving an optimal policy compared to the policies in Section V, we investigate the performance across all policies given the spatio-temporal decay factors λ and r_0 in (43).

Figure 3 shows the asymptotic average cost versus λ for different scheduling policies, where $N = 5$, $D = 2$, and $r_0 = 0.5$. We compare an optimal policy (OPTIMAL) with randomized scheduling (RANDOM), round-robin (RR), optimal round-robin (ORR) and two FHM policies with $L = 1$ and $L = 2$, respectively, i.e., (FHM-1) and (FHM-2). Solid lines show theoretical results, whereas dots represent simulation results, obtained by averaging 200 ensembles for a time horizon of $T = 100$. We see that simulations are in close agreement with the theoretical predictions.

For $\lambda \to 0$, the temporal correlation increases, and so, the asymptotic average cost decreases for all policies. Both FHM policies perform very close to optimal, which is due to λ resulting in small values of m. The performance of the ORR policy becomes near-optimal, since the temporal correlation exceeds the spatial correlation and more information is inhibited in recent measurements than in dependent processes. As $\lambda \to \infty$ the temporal correlation vanishes, FHM-1 becomes optimal and the performance difference to ORR increases. This is because the scheduling problem becomes a one-shot optimization problem, where an optimal policy results in the same set of sensors being chosen at every time instant. The ORR consistently outperforms RR since it exploits the spatial dependency more efficiently. For $\lambda > 0.75$, RANDOM even performs better than RR. FHM-2 performs better than FHM-1 as it can plan over a longer time horizon.

Figure 4 shows the asymptotic average cost versus the distance between neighboring sensors for $N = 5$, $D = 2$, and $T = 0.8$. We see that all policies result in lower MSE as the spatial dependency increases, i.e., $r_0 \to 0$. FHM-1 and FHM-2 are optimal for large parts of r_0. Note that for $r_0 < 0.8$, RR performs worse than RANDOM since it does not utilize the spatial dependency efficiently. As the distance r_0 increases, RR and ORR become optimal policies. This is because the spatial correlation becomes zero, leading to a symmetric covariance matrix $\mathbf{C}_{\theta\theta}$, which matches with the theory in Proposition 5.

As seen in Figures 3 and 4, the performance of an OPTIMAL policy becomes bounded if either the temporal or the spatial correlation vanishes, and is lower than if compared to a system where no spatial or temporal correlation would be exploited. This demonstrates the benefit of exploiting both

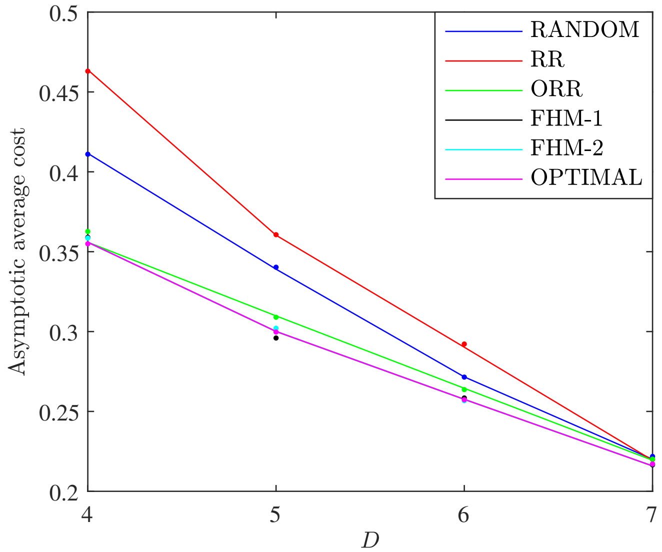

Fig. 6. Asymptotic average cost, $\lim_{T\to\infty} J(\gamma, T)$, vs D with $N = 8$, $\sigma_i = 1$, $\forall i = 1, 2, ..., N$, $\xi = 0.5$, $\lambda = 0.8$, , and $r_0 = 0.5$.

temporal ans spatial correlation, since, a weak correlation component can be counterbalanced by the other.

C. Performance vs system size

A higher ratio between the number of sensors N in relation to the scheduling capacity D is expected to reduce performance. In this section, we analyze the performance given N or D to see how it changes across the policies.

Figure 5 shows the asymptotic average cost versus the number of sensors N with $D = 4$, $\lambda = 0.8$, and $r_0 = 0.5$. For $N = 8$, we used the approximated FS-MDP presented in Section V-C to derive a policy represented by the OPTIMAL curve. We see that the cost increases as the number of sensors N grows. The curve eventually flattens with N, as it can never be higher than the marginal variances $\sigma_i = 1$, as shown in (12). We see that RR performs the worse, which again demonstrates the importance for the intrinsic scheduling order. The FHM-1 and FHM-2 is optimal for most parts of the figure.

Figure 6 shows the time-average MSE versus the number of sensors D receiving broadcasts from the scheduler, with $N = 8$, $\lambda = 0.8$, and $r_0 = 0.5$. For $D = 4$, we used the approximated FS-MDP in Section V-C to derive a policy represented by the OPTIMAL curve. We see a similar, but inverted result comparing to Figure 5, as the cost decrease with the number of sensors receiving broadcasts D.

VII. Conclusion

This paper proposed optimal scheduling policies for transmitting observations of spatio-temporally dependent processes from multiple sensors to remote estimators over a limited number of communication channels. The problem was modeled as a finite state Markov decision process with the AoI as state-variable. An optimal scheduling policy was derived that minimizes the time-average mean squared error (MSE), resulting in a periodic scheduling sequence. Due to increased computational complexity for large systems, we also considered computationally less demanding scheduling policies, minimizing the MSE over short-time horizons or the AoI across the sensors, which both performed well.

Our paper expands the work regarding utilizing AoI for remote estimation scheduling of dependent observations. We showed that if the main objective is minimizing the average AoI, regarding the intrinsic scheduling order can enhance the estimation accuracy if the sensor observations are spatio-temporally dependent.

For future work, a sequential estimator incorporating past measurements can be considered to derive the optimal scheduling policies for estimating spatio-temporal processes. Another extension of our work, could be to model the measurements using a different dynamic spatio-temporally process, e.g., a Gauss-Markov model [9]. Similarly, the research would involve proposing how it could be modeled, determining if an optimal scheduling policy exists and, if so, how it can be derived. Finally, another extension of our work could be to design AoI-based scheduling policies for multiple-access channels considering queuing processes and random arrival times that would influence the AoI. In that case, the MDP would need to be re-designed for the considered system model.

Appendix A
Proof of Proposition 1

Proof. If $\mathbf{\Delta}[k]$ is given, the MSE at instant k is obtained from (10). If $\mathbf{\Delta}[0]$ and γ are known, it is possible to determine $\mathbf{\Delta}[k]$ using (3) and (4). Thus, the cost function in (8) can be expressed as

$$\begin{aligned} J(\gamma, T) &= \frac{1}{TN}\sum_{k=1}^{T}\sum_{i=1}^{N}\mathbb{E}\Big[(\theta_i[k]-\hat{\theta}_i[k])^2\Big|\gamma, \mathbf{\Delta}[0]\Big] \\ &= \frac{1}{TN}\sum_{k=1}^{T}\mathbb{E}\Big[f(\mathbf{\Delta}[k])\Big|\gamma, \mathbf{\Delta}[0]\Big]. \end{aligned} \tag{44}$$

Substituting (44) in (9), we see that an optimal policy γ^* satisfies (13). □

Appendix B
Proof of Proposition 2

Proof. Let $\mathbf{\Delta}[k]$ and be known and have $\tilde{\mathbf{\Delta}}[k] = b(\mathbf{\Delta}[k])$ given by (20). The value $f(\mathbf{\Delta}[k])$ in (10) depends on the covariance terms $\mathbf{C}_{\theta y}[k]]$ and $\mathbf{C}_{yy}[k]$, given as functions of $\mathbf{\Delta}[k]$ in (7). Similarly, $\tilde{f}(\tilde{\mathbf{\Delta}}[k])$ in (10) is given by the terms $\tilde{\mathbf{C}}_{\theta y}[k]]$ and $\tilde{\mathbf{C}}_{yy}[k]$, which depends on $\tilde{\mathbf{\Delta}}[k]$ in (23). By substituting (20) in (23) we get

$$\begin{aligned} &[\tilde{\mathbf{C}}_{yy}[k]]_{i,j} = \sigma_i\sigma_j\rho_{ij}\varphi\big(\big[|\Delta_i[k]-\Delta_j[k]|\big]_+^m\big) + \xi^2\delta(i-j), \\ &[\tilde{\mathbf{C}}_{\theta y}[k]]_{i,j} = \sigma_i\sigma_j\rho_{ij}\varphi\big(\big[\Delta_i[k]\big]_+^m\big), \quad i,j \in \{1,...,N\}. \end{aligned} \tag{45}$$

If Assumption 2 holds, the temporal correlation in (1) satisfy $\varphi(x) = \varphi(\big[x\big]_+^m)$ for all $x \geq m$, which leads to (7) and (45) being equal, i.e.,

$$\begin{aligned} &[\tilde{\mathbf{C}}_{yy}[k]]_{i,j} = [\mathbf{C}_{yy}[k]]_{i,j}, \\ &[\tilde{\mathbf{C}}_{\theta y}[k]]_{i,j} = [\mathbf{C}_{\theta y}[k]]_{i,j}, \quad i,j \in \{1,...,N\}. \end{aligned} \tag{46}$$

From (46) we substitue $\mathbf{C}_{\theta y}[k]]$ with $\tilde{\mathbf{C}}_{\theta y}[k]]$ and $\mathbf{C}_{yy}[k]$ with $\tilde{\mathbf{C}}_{yy}[k]$ in (20) and see that

$$f(\boldsymbol{\Delta}[k]) = \tilde{f}(\tilde{\boldsymbol{\Delta}}[k]), \quad \forall \boldsymbol{\Delta}[k] \in \mathbb{N}_+^N.$$

□

APPENDIX C
PROOF OF PROPOSITION 3

Proof. Applying the truncation on $\Delta_i[k]$ in (3), yields

$$\tilde{\Delta}_i[k] = \begin{cases} 0, & \text{if } i \in \boldsymbol{\pi}[k], \\ \left[\Delta_i[k-1]+1\right]_+^m, & \text{if } i \notin \boldsymbol{\pi}[k]. \end{cases} \tag{47}$$

Further, the following relationship holds true

$$\left[\Delta_i[k-1]+1\right]_+^m = \left[\left[\Delta_i[k-1]\right]_+^m + 1\right]_+^m = \left[\tilde{\Delta}_i[k-1]+1\right]_+^m. \tag{48}$$

After substituting (48) in (47), we obtain (25). Similarly, substituting (3) in $\Delta_{ij}[k] = |\Delta_i[k] - \Delta_j[k]|$, gives

$$\Delta_{ij}[k] = \begin{cases} 0, & \text{if } i,j \in \boldsymbol{\pi}[k], \\ \Delta_{ij}[k-1], & \text{if } i,j \notin \boldsymbol{\pi}[k], \\ \Delta_i[k-1]+1, & \text{if } i \notin \boldsymbol{\pi}[k],\, j \in \boldsymbol{\pi}[k], \\ \Delta_j[k-1]+1, & \text{if } i \in \boldsymbol{\pi}[k],\, j \notin \boldsymbol{\pi}[k]. \end{cases} \tag{49}$$

Finally, (26) is obtained by applying the truncation operator in (49) and employing the relationship in (48). □

APPENDIX D
PROOF OF THEOREM 1

Proof. Throughout this proof, we assume that Assumption 1 and Assumption 2 holds. To prove the relationship between γ^* and $\tilde{\gamma}^*$ in (30), we present four important properties showing how $\mathcal{M}$ and $\tilde{\mathcal{M}}$ relate.

Firstly, as seen in (21), every state in $\mathcal{S}$ maps to a state in $\tilde{\mathcal{S}}$. Secondly, based on Proposition 3, for any two states $\boldsymbol{\Delta}' \in \mathcal{S}$ and $\boldsymbol{\Delta}'' \in \mathcal{S}$, the following relationship holds

$$P\Big(\boldsymbol{\Delta}[k] = \boldsymbol{\Delta}'' \Big| \boldsymbol{\Delta}[k-1] = \boldsymbol{\Delta}', \boldsymbol{\pi}[k]\Big) = P\Big(\tilde{\boldsymbol{\Delta}}[k] = b(\boldsymbol{\Delta}'') \Big| \tilde{\boldsymbol{\Delta}}[k-1] = b(\boldsymbol{\Delta}'), \boldsymbol{\pi}[k]\Big), \forall \boldsymbol{\Delta}', \boldsymbol{\Delta}'' \in \mathcal{S},$$

showing that the mapping between $\mathcal{S}$ and $\tilde{\mathcal{S}}$ is consistent after any action is taken. Thirdly, from Proposition 2 we have that

$$f(\boldsymbol{\Delta}[k]) = \tilde{f} \circ b(\boldsymbol{\Delta}[k]) = \tilde{f}(\tilde{\boldsymbol{\Delta}}[k]), \quad \forall \boldsymbol{\Delta}[k] \in \mathcal{S},$$

showing that for a given state $\boldsymbol{\Delta}[k]$ in $\mathcal{S}$, and the mapping state $\tilde{\boldsymbol{\Delta}}[k] = b(\boldsymbol{\Delta}[k])$ in $\tilde{\mathcal{S}}$, gives the same MSE value. Fourthly, the two images $f : \mathcal{S} \to \mathcal{Y}$ and $\tilde{f} : \tilde{\mathcal{S}} \to \tilde{\mathcal{Y}}$ are equal, i.e., $\mathcal{Y} = \tilde{\mathcal{Y}}$. Showing that every possible MSE value in $\mathcal{M}$ exists in $\tilde{\mathcal{M}}$.

Given the aforementioned relating properties between $\mathcal{M}$ and $\tilde{\mathcal{M}}$, if the initial values $\boldsymbol{\Delta}[0]$ and $\tilde{\boldsymbol{\Delta}}[0] = b(\boldsymbol{\Delta}[0])$ are known and two policies $\gamma = (\gamma_1, ..., \gamma_T)$ and $\tilde{\gamma} = (\tilde{\gamma}_1, ..., \tilde{\gamma}_T)$, relate as $\gamma_k = \tilde{\gamma}_k \circ b$, $\forall k \in \mathbb{N}_+$, we have that

$$\mathbb{E}\left[r(\boldsymbol{\Delta}[k-1], \boldsymbol{\pi}[k]) | \gamma, \boldsymbol{\Delta}[0]\right] = \mathbb{E}[\tilde{r}(\tilde{\boldsymbol{\Delta}}[k-1], \boldsymbol{\pi}[k]) | \tilde{\gamma}, \tilde{\boldsymbol{\Delta}}[0]], \quad \forall \boldsymbol{\Delta}[0] \in \mathcal{S}, k \in \mathbb{N}_+. \tag{50}$$

Let two policies γ and $\tilde{\gamma}$ that satisfy (50) be referred to as *replicable*. From (17) and (28) we see that two replicable policies result in

$$g_\gamma(\boldsymbol{\Delta}[0]) = \tilde{g}_{\tilde{\gamma}} \circ b(\boldsymbol{\Delta}[0]), \quad \forall \boldsymbol{\Delta}[0] \in \mathcal{S}.$$

Thus, if an optimal policy γ^* is replicable, we have that

$$g_{\gamma^*}(\boldsymbol{\Delta}[0]) = \tilde{g}_{\tilde{\gamma}^*} \circ b(\boldsymbol{\Delta}[0]), \quad \forall \boldsymbol{\Delta}[0] \in \mathcal{S}. \tag{51}$$

Since we can always construct a policy $\gamma_k = \tilde{\gamma}_k \circ b$, there always exists a replicable policy $\gamma \in \Gamma$ for every policy $\tilde{\gamma} \in \tilde{\Gamma}$. If γ^* is not replicable, the following inequality holds

$$g_{\gamma^*}(\boldsymbol{\Delta}[0]) \geq \tilde{g}_{\tilde{\gamma}^*} \circ b(\boldsymbol{\Delta}[0]), \quad \forall \boldsymbol{\Delta}[0] \in \mathcal{S}. \tag{52}$$

To prove (51), we will followingly prove by contradiction that $g_{\gamma^*}(\boldsymbol{\Delta}[0]) > \tilde{g}_{\tilde{\gamma}^*} \circ b(\boldsymbol{\Delta}[0])$ does not hold.

As seen in (21), some states in $\mathcal{S}$ maps to the same state in $\tilde{\mathcal{S}}$ and we have that $|\tilde{\mathcal{S}}| \leq |\mathcal{S}|$. From (3), in every state, an action $\boldsymbol{\pi}[k] \in \mathcal{A}$ in $\mathcal{M}$ results in a state transition. Whereas for $\tilde{\mathcal{M}}$, in (20), the states in $\tilde{\mathcal{S}}$ can be grouped into two types, where; i) every action results in a state transition, or; ii) a particular action results in returning to the same state, while all other actions result in a transition. The state of type ii) can only be reached in $\mathcal{M}$ if an action is repeated over a consecutive number of instants. Given $|\tilde{\mathcal{S}}| \leq |\mathcal{S}|$, there exists policies $\gamma \in \Gamma$ that are not replicable in $\tilde{\gamma} \in \tilde{\Gamma}$. We must, therefore, prove that the non-replicable policies in Γ are not exclusively optimal.

To begin, we present some important properties of $\mathcal{M}$ and $\tilde{\mathcal{M}}$. In both MDPs, every state is reachable from every other state. Every state in $\mathcal{M}$ and $\tilde{\mathcal{M}}$ can only be reached from a unique state. Also, the rewards and transition probabilities are stationary and the rewards are upper-bounded in (12). From (11), the MSE increases, but is upper-bounded, with respect to the marginal AoI. Thus, the MSE increases or is equal every time an action is consecutively repeated. Given the aforementioned properties, the average reward function in $\mathcal{M}$ does not increase with the marginal AoI, i.e.,

$$\max_{\gamma \in \Gamma} g_\gamma\big([\Delta_1[k], ..., \Delta_i[k], ..., \Delta_N[k]]^{\mathrm{T}}\big) \geq \max_{\gamma \in \Gamma} g_\gamma\big([\Delta_1[k], ..., \Delta_i[k]+1, ..., \Delta_N[k]]^{\mathrm{T}}\big), \quad i = 1, ..., N. \tag{53}$$

Let $\boldsymbol{\Delta}' \in \mathcal{S}$, $\tilde{\boldsymbol{\Delta}}' = b(\boldsymbol{\Delta}') \in \tilde{\mathcal{S}}$ and $\boldsymbol{\pi}[k] = \boldsymbol{\pi}' \in \mathcal{A}$ be a state-action combo that satisfy

$$P\Big(\boldsymbol{\Delta}[k] \neq \boldsymbol{\Delta}' \Big| \boldsymbol{\Delta}[k-1] = \boldsymbol{\Delta}', \boldsymbol{\pi}[k] = \boldsymbol{\pi}'\Big) = 1,$$
$$P\Big(\tilde{\boldsymbol{\Delta}}[k] = \tilde{\boldsymbol{\Delta}}' \Big| \tilde{\boldsymbol{\Delta}}[k-1] = \tilde{\boldsymbol{\Delta}}', \boldsymbol{\pi}[k] = \boldsymbol{\pi}'\Big) = 1.$$

Assume $\boldsymbol{\Delta}[k-1] \in \mathcal{S}$ and that action $\boldsymbol{\pi}[k] = \boldsymbol{\pi}'$ results in a transition to $\boldsymbol{\Delta}[k] = \boldsymbol{\Delta}'$. Let $\gamma \in \Gamma$ be a policy that, regardless of initial state $\boldsymbol{\Delta}[0] \in \mathcal{S}$, if $\boldsymbol{\Delta}[k] = \boldsymbol{\Delta}'$ for any $k \in \mathbb{N}_+$, results in action $\boldsymbol{\pi}[k+l] = \boldsymbol{\pi}'$ for $l = 1, ..., L$, $L \in \mathbb{N}_{++}$, consecutive instances, until instant $k+L+1$, i.e., $\boldsymbol{\pi}[k+L+1] \neq \boldsymbol{\pi}'$. Such a policy γ is not replicable in $\tilde{\gamma} \in \tilde{\Gamma}$. However, from (53), γ is not optimal, since the average reward corresponding to transitioning states $\mathbf{\Delta}[k]$ is non-increasing, and the expected reward is the same over the L consecutive time instants, i.e.,

$$\mathbb{E}[r(\mathbf{\Delta}[k+l-1], \boldsymbol{\pi}[k+l])]|\gamma, \mathbf{\Delta}[k] = \mathbf{\Delta}'] = \mathbb{E}[r(\mathbf{\Delta}[k-1], \boldsymbol{\pi}[k])|\gamma, \mathbf{\Delta}[k-1]], \quad l = 2, .., L.$$

For this reason, it would not exist a finite value $L \geq 1$ for an optimal policy γ^*. If L is either $L = 0$ or $L \to \infty$, then γ^* is again replicable. Thus, the inequality in (52) becomes (51), which proves (30). □

Appendix E
Proof of Lemma 2

Proof. A finite-state MDP is classified as *unichain* if every stationary deterministic policy correspond to a Markov chain with a single set of recurrent states and a set of transient states, i.e., $\exists n < \infty$, $P(\tilde{\mathbf{\Delta}}[k+n] = \tilde{\mathbf{\Delta}} \mid \tilde{\mathbf{\Delta}}[k] = \tilde{\mathbf{\Delta}}) < 1$. From Proposition 4, we see that $\tilde{\mathcal{S}}_{\tilde{\gamma}}$ is a set of recurrent states and that the complement, i.e., $\tilde{\mathcal{S}}^c = \tilde{\mathcal{S}} \setminus \tilde{\mathcal{S}}_{\tilde{\gamma}}$, is a set of transient states. We conclude that $\tilde{\mathcal{M}}$ is unichain.

Thus, $\tilde{\mathcal{M}}$, is unichain, has a finite action set $|\mathcal{A}| < \infty$, a finite state-space $|\tilde{\mathcal{S}}| < \infty$, bounded rewards $|r(\tilde{\mathbf{\Delta}}[k-1], \boldsymbol{a}[k]| < \infty$ and stationary rewards and transition probabilities. Given the aforementioned properties of $\mathcal{M}$, [46, Th. 8.4.5] states that there exist a stationary optimal policy $\tilde{\gamma}^*$ and a pair (g^*, h^*) that satisfy (32). The theorem also states the relationship between $\tilde{\gamma}^*$ and (g^*, h^*) in (33). Based on [46, Th. 8.6.6], the pair (g^*, h^*) can be derived in a finite number of iterations using policy iteration. The full proofs of the referred theorems are presented in [46]. □

References

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

[2] M. Emara, H. Elsawy, and G. Bauch, "A spatiotemporal model for peak AoI in uplink IoT networks: Time versus event-triggered traffic," *IEEE Internet Things J.*, vol. 7, no. 8, pp. 6762–6777, Aug. 2020.

[3] 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 Conf.*, pp. 1–5, 2019.

[4] 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.

[5] O. Ayan, M. Vilgelm, M. Klügel, S. Hirche, and W. Kellerer, "Age-of-information vs. value-of-information scheduling for cellular networked control systems," *Proc. 10th ACM/IEEE Int. Conf. Cyber-Physical Syst.*, pp. 109–117, 2019.

[6] T. H. Chung, V. Gupta, B. Hassibi, J. Burdick, and R. M. Murray, "Scheduling for distributed sensor networks with single sensor measurement per time step," *IEEE Int. Conf. Robot. Autom.*, vol. 1, pp. 187–192 Vol.1, 2004.

[7] M. M. Vasconcelos, M. Gagrani, A. Nayyar, and U. Mitra, "Optimal scheduling strategy for networked estimation with energy harvesting," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 4, pp. 1723–1735, Dec. 2020.

[8] 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.

[9] 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.

[10] M. Gagrani, M. M. Vasconcelos, and A. Nayyar, "Scheduling and estimation strategies in a sequential networked estimation problem," in *56th Annu. Allert. Conf. Commun. Control Comput.*, 2018, pp. 871–878.

[11] M. M. Vasconcelos, A. Nayyar, and U. Mitra, "Optimal sensor scheduling strategies in networked estimation," in *IEEE 56th Conf. Decision and Control*, 2017, pp. 5378–5384.

[12] 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 Trans. Automat. Control*, vol. 64, no. 9, pp. 3732–3739, Sep. 2019.

[13] H. Yetgin, K. T. K. Cheung, M. El-Hajjar, and L. Hanzo, "A survey of network lifetime maximization techniques in wireless sensor networks," *IEEE Commun. Surveys Tut.*, vol. 19, no. 2, pp. 828–854, May 2017.

[14] M. C. Vuran, Ö. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Comput. Netw.*, vol. 45, no. 3, pp. 245–259, Jun. 2004.

[15] L. A. Villas, A. Boukerche, H. A. De Oliveira, R. B. De Araujo, and A. A. Loureiro, "A spatial correlation aware algorithm to perform efficient data collection in wireless sensor networks," *Ad Hoc Netw.*, vol. 12, no. 1, pp. 69–85, Jan. 2014.

[16] V. Roy, A. Simonetto, and G. Leus, "Spatio-temporal sensor management for environmental field estimation," *Signal Process.*, vol. 128, pp. 369–381, Nov. 2016.

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

[18] 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 Int. Conf. Commun. Workshops*, 2019, pp. 1–6.

[19] Z. Jiang and S. Zhou, "Status from a random field: How densely should one update?" *IEEE Int. Symp. Inform. Theory*, pp. 1037–1041, 2019.

[20] M. M. Vasconcelos and U. Mitra, "Observation-driven scheduling for remote estimation of two Gaussian random variables," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 1, pp. 232–244, Mar. 2020.

[21] Y. P. Hsu, E. Modiano, and L. Duan, "Scheduling algorithms for minimizing age of information in wireless broadcast networks with random arrivals," *IEEE Trans. Mobile Comput.*, vol. 19, no. 12, pp. 2903–2915, Dec. 2020.

[22] B. Zhou, Y. Cui, and M. Tao, "Optimal dynamic multicast scheduling for cache-enabled content-centric wireless networks," *IEEE Trans. Commun.*, vol. 65, no. 7, pp. 2956–2970, Jul. 2017.

[23] A. Kosta, N. Pappas, and V. Angelakis, "Age of information: A new concept, metric, and tool," *Foundations and Trends® in Networking*, vol. 12, no. 3, pp. 162–259, Nov. 2017.

[24] M. A. Abd-Elmagid, N. Pappas, and H. S. Dhillon, "On the role of age of information in the internet of things," *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 72–77, Dec. 2019.

[25] H. Tang, J. Wang, L. Song, and J. Song, "Minimizing age of information with power constraints: Multi-user opportunistic scheduling in multi-state time-varying channels," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 5, pp. 854–868, Mar. 2020.

[26] T. Park, W. Saad, and B. Zhou, "Centralized and distributed age of information minimization with nonlinear aging functions in the internet of things," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 8437–8455, May 2021.

[27] B. Zhou and W. Saad, "Joint status sampling and updating for minimizing age of information in the internet of things," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7468–7482, Nov. 2019.

[28] L. Guo, Z. Chen, D. Zhang, K. Liu, and J. Pan, "AoI inspired collaborative information collection for auv assisted internet of underwater things," *IEEE Internet Things J.*, vol. 8, no. 5, pp. 3851–3863, Mar. 2021.

[29] L. Hu, Z. Chen, Y. Dong, Y. Jia, L. Liang, and M. Wang, "Status update in IoT networks: Age of information violation probability and optimal update rate," *IEEE Internet Things J.*, no. early acces, 2021.

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

[31] H. H. Yang, A. Arafa, T. Q. S. Quek, and H. Vincent Poor, "Age-based scheduling policy for federated learning in mobile edge networks," *IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 8743–8747, 2020.

[32] M. Klügel, M. H. Mamduhi, S. Hirche, and W. Kellerer, "AoI-penalty minimization for networked control systems with packet loss," *IEEE Conf. Computer Commun. Workshops*, pp. 189–196, 2019.

[33] Y. Sun, Y. Polyanskiy, and E. Uysal, "Sampling of the wiener process for remote estimation over a channel with random delay," *IEEE Trans. Inform. Theory*, vol. 66, no. 2, pp. 1118–1135, Feb. 2020.

[34] K. Huang, W. Liu, Y. Li, B. Vucetic, and A. Savkin, "Optimal downlink-uplink scheduling of wireless networked control for industrial IoT," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1756–1772, Mar. 2020.

[35] K. Huang, W. Liu, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Real-time remote estimation with hybrid ARQ in wireless networked control," *IEEE Trans. Wirel. Commun.*, vol. 19, no. 5, pp. 3490–3504, May 2020.

[36] K. Huang, W. Liu, Y. Li, A. Savkin, and B. Vucetic, "Wireless feedback control with variable packet length for industrial IoT," *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1586–1590, 2020.

[37] W. Liu, D. E. Quevedo, K. H. Johansson, B. Vucetic, and Y. Li, "Remote state estimation of multiple systems over multiple markov fading channels," 2021, arXiv Preprint. [Online]. Available: https://arxiv.org/abs/2104.04181

[38] W. Liu, K. Huang, D. E. Quevedo, B. Vucetic, and Y. Li, "Deep reinforcement learning for wireless scheduling in distributed networked control," 2021, arXiv Preprint. [Online]. Available: https://arxiv.org/abs/2109.12562

[39] B. Zhou and W. Saad, "On the age of information in internet of things systems with correlated devices," *IEEE Glob. Commun. Conf.*, pp. 1–6, 2020.

[40] A. E. Kalor and P. Popovski, "Minimizing the age of information from sensors with common observations," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1390–1393, May 2019.

[41] M. P. Vitus, W. Zhang, A. Abate, J. Hu, and C. J. Tomlin, "On efficient sensor scheduling for linear dynamical systems," *Automatica*, vol. 48, no. 10, pp. 2482–2493, Oct. 2012.

[42] V. W. Håkansson, N. K. D. Venkategowda, and S. Werner, "Optimal scheduling policy for spatio-temporally dependent observations using age-of-information," *IEEE 23rd Int. Conf. Inform. Fusion*, pp. 1–6, 2020, doi: 10.23919/FUSION45008.2020.9190323.

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

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

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

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

[47] Y. Hsu, E. Modiano, and L. Duan, "Age of information: Design and analysis of optimal scheduling algorithms," *IEEE Int. Symp. Inform. Theory*, pp. 561–565, 2017.

Victor Wattin Håkansson received the M.Sc. degree in Engineering Physics from Lund University, Sweden, in 2014. He is currently pursuing a Ph.D. degree at the Norwegian University of Science and Technology (NTNU). He is presently with the Signal Processing Group in the Department of Electronic Systems at NTNU. His experience and research interests include machine learning, age-of-information, wireless sensor networks, and financial modeling.

Naveen K. D. Venkategowda (S'12–M'17) received the B.E. degree in electronics and communication engineering from Bangalore University, Bengaluru, India, in 2008, and the Ph.D. degree in electrical engineering from Indian Institute of Technology, Kanpur, India, in 2016. He is currently an Universitetslektor at the Department of Science and Technology, Linköping University, Sweden. From Oct. 2017 to Feb. 2021, he was postdoctoral researcher at the Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway. He was a Research Professor at the School of Electrical Engineering, Korea University, South Korea from Aug. 2016 to Sep. 2017. He was a recipient of the TCS Research Fellowship (2011-15) from TCS for graduate studies in computing sciences and the ERCIM Alain Bensoussan Fellowship in 2017.

Stefan Werner (SM'07) received the M.Sc. degree in electrical engineering from the Royal Institute of Technology, Stockholm, Sweden, in 1998, and the D.Sc. degree (Hons.) in electrical engineering from the Signal Processing Laboratory, Helsinki University of Technology, Espoo, Finland, in 2002. He is currently a Professor at the Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), Director of IoT@NTNU, and Adjunct Professor with Aalto University in Finland. He was a visiting Melchor Professor with the University of Notre Dame during the summer of 2019 and an Adjunct Senior Research Fellow with the Institute for Telecommunications Research, University of South Australia, from 2014 to 2020. He held an Academy Research Fellowship, funded by the Academy of Finland, from 2009 to 2014. His research interests include adaptive and statistical signal processing, wireless communications, and security and privacy in cyber-physical systems. He is a member of the editorial boards for the EURASIP Journal of Signal Processing and the IEEE Transactions on Signal and Information Processing over Networks.

Pramod K. Varshney (S'72–M'77–SM'82–F'97–LF'18) was born in Allahabad, India, in 1952. He received the B.S. degree (Highest Honors) in electrical engineering and computer science and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Urbana– Champaign, USA, in 1972, 1974, and 1976, respectively. From 1972 to 1976, he held teaching and research assistantships with the University of Illinois. Since 1976, he has been with Syracuse University, Syracuse, NY, USA, where he is currently a Distinguished Professor of electrical engineering and computer science and the Director of the Center for Advanced Systems and Engineering. As the Director of CASE, he is responsible for technology transition of university expertise to make economic impact in the high-tech economy of New York. He served as an Associate Chair of the department from 1993 to 1996. He is also an Adjunct Professor of radiology with Upstate Medical University, Syracuse, NY, USA. He has published extensively. He is the author of Distributed Detection and Data Fusion (New York, NY, USA: Springer-Verlag, 1997). His current research interests include distributed sensor networks and data fusion, detection and estimation theory, wireless communications, image processing, radar signal processing, physical layer security, and machine learning. Dr. Varshney was a James Scholar, a Bronze Tablet Senior, and a fellow while at the University of Illinois. He is a member of Tau Beta Pi. He received the 1981 ASEE Dow Outstanding Young Faculty Award. In 1997, he was elected to the grade of fellow of the IEEE for his contributions in the area of distributed detection and data fusion. In 2000, he received the Third Millennium Medal from the IEEE and the Chancellor's Citation for exceptional academic achievement at Syracuse University. He also received the IEEE 2012 Judith A. Resnik Award, Doctor of Engineering Honoris Causa from Drexel University in 2014, the ECE Distinguished Alumni Award from the University of Illinois in 2015, the ISIF's Yaakov Bar-Shalom Award for a Lifetime of Excellence in Information Fusion in 2018, the Claude Shannon-Harry Nyquist Technical Achievement Award from the IEEE Signal Processing Society and the Pioneer Award from the IEEE Aerospace and Electronic Society in 2021. He was the Guest Editor of the special issue on data fusion of the IEEE PROCEEDINGS in 1997. He has served on the Editorial Boards of the Journal on Advances in Information Fusion, IEEE Transactions of Signal Processing as well as IEEE Signal Processing Magazine among others. He was the President of International Society of Information Fusion in 2001.