Double Threshold Structure of Sensor Scheduling Policy Over a Finite-State Markov Channel

Jiang Wei and Dan Ye, *Senior Member, IEEE*

***Abstract*—In this article, we consider the optimal sensor scheduling for remote state estimation in cyber–physical systems (CPSs). Different from the existing works concerning the time-invariant channel state in the wireless communication network, our work considers the time-varying channel state modeled by a finite-state Markov channel (FSMC). We focus on the problem of how to schedule the transmission of the sensor to minimize the estimation error at the remote side with less communication cost. Using the framework of the Markov decision process (MDP), the optimal scheduling policy is shown to be deterministic stationary (DS). We further derive its double threshold structure with respect to remote estimation errors and channel states. Moreover, a necessary and sufficient condition guaranteeing the mean-square stability of the remote estimator is given based on the structured scheduling policy. Numerical simulations are provided to verify the theoretical results.**

***Index Terms*—Cyber–physical systems (CPSs), finite-state Markov channel (FSMC), Markov decision process (MDP), sensor scheduling, stability, threshold structure.**

I. Introduction

In RECENT years, cyber–physical systems (CPSs) have gained more and more attention with the rapid development of computation, control, and communication technology. These developed techniques enable devices to become "smart," for example, wireless-connected sensors with real-time sensing, calculation, and decision, which boost wide applications of CPSs in practice, including condition monitoring in smart grid and industrial CPSs, estimation/control using unmanned aerial vehicles and underwater navigation vehicles, health tracking, etc. [1], [2], [3], [4]. While CPSs bring high efficiency and huge convenience, the subsequent problems cannot be ignored, such as packet loss and delay, leading to unreliability of wireless communication and performance degradation of systems [5], [6], [7].

A lot of efforts have been devoted to enhancing reliability and further optimizing the performance of CPSs [8], [9], [10], [11], [12]. Specifically, the devices in CPSs are generally subject to limited resources, such as insufficient energy supply and finite bandwidth. Moreover, data transmission over the wireless communication channel is prone to packet loss. Hence, large efforts have been put into developing optimization strategies to address these issues. One of the interesting topics is sensor scheduling [13], [14], [15], [16], aiming at determining which and when sensors should be scheduled to transmit data packets so as to obtain accurate remote state estimation and save communication resources. Compared with the sensor scheduling in the communication community, where sensors are deployed to observe stationary source, sensor scheduling in control community aims to achieve the tradeoff between control performance for dynamic systems and limited communication resource [17], [18], [19]. Related work covers from scheduling over noiseless communication channel [20] to lossy channel [5], scheduling of data packets in the single dynamic process [14] to multiple processes [21]. With a slight difference, the power allocation problem focuses on specific transmission power of sensors [15], [22], [23].

It should be noted that the channel state is time invariant over the wireless network in [14], [15], and [22], where channel gain can be characterized by a constant. While in many practical applications, the channel state will generally be changing affected by the fading, interferences, and so on. Furthermore, the channel state reflects the quality of the current communication link and channel state information can be used to further save communication resources. This motivates research on sensor scheduling under a time-varying channel state. Different from the independent identically distributed (i.i.d.) packet dropout channel model [24], [25], the Markov channel model depicts the temporal correlation and it can reduce to the i.i.d scenario; thus, it receives a lot of attention [23], [26], [27], [28], [29]. In the work of Qi *et al.* [26], Wei and Ye [27], Leong *et al.* [30], and Wang *et al.* [31], the sensor scheduling policy was developed in the case of two channel states. For the finite-state Markov channel (FSMC), the study in [23] examined a structured optimal energy allocation for a multisensor system considering energy harvesting and sharing. A similar structural result was also obtained in [29], but the considered system was a first-order autoregressive process. A general vector system with FSMC was investigated in [32]. Among works [26], [27], [31], and [32],

Manuscript received 27 June 2022; accepted 31 July 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 62173071; in part by the Liaoning Revitalization Talents Program under Grant XLYC1907035; in part by the Liaoning BaiQianWan Talents Program under Grant 202062; and in part by the Fundamental Research Funds for the Central Universities under Grant N2204008. This article was recommended by Associate Editor R. Tagliaferri. *(Corresponding author: Dan Ye.)*

Jiang Wei is with the College of Information Science and Engineering, Northeastern University, Shenyang 110819, China (e-mail: 846489570@qq.com).

Dan Ye is with the College of Information Science and Engineering and the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China (e-mail: yedan@ise.neu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2022.3197153.

Digital Object Identifier 10.1109/TCYB.2022.3197153

threshold structures of the optimal scheduling policy were established. These special structures avoid brute force search in iterative computations and, thus, alleviate computation burden to obtain the optimal policy. Besides, the threshold structure can be used to construct event-triggered transmission schemes and the resulted scheduling policy is easy to implement. The common feature of the above-mentioned structured policy is that the optimal threshold is only involved with the estimation error and is independent of the channel state. One natural question is if the computational complexity can be further reduced by establishing the threshold structure on the time-varying channel state. This problem was explored in the work of Chakravorty and Mahajan [29], where the result was obtained over a finite-time horizon and stability issue was not discussed. Due to packet dropout or no transmission, the Kalman filter with intermittent observations occur, which could lead to instability of the remote estimator [33]. Thus, how to guarantee stability at the remote side is an important issue. This issue can be roughly categorized into the *existence* of scheduling policies that stabilize remote estimators [17], [28], [34] and the stability under *optimal* policy [16], [26], [31], [35]. Specifically, in [31], the necessary and sufficient condition (NSC) was derived in the case of a practical usage-dependent two-state Markov channel. For the FSMC, the NSC was obtained using a novel estimation-cycle-based method in [28]. However, insights were not provided in [28] for structures of the optimal sensor scheduling policy under costly transmission.

Inspired by the above-mentioned works, in this work, we investigate the double threshold structure of the optimal sensor scheduling policy under the FSMC model. Besides, the stability of the remote estimator is analyzed based on the structured optimal policy. The contributions in this work are different from that in our previous papers for considering stable dynamic processes over two-state Markov packet dropout channel with a single threshold structure in [27], for multisensor one shared Markov channel with single threshold structure over finite-horizon in [36]. As discussed previously, stabilizing remote estimator for an unstable system over infinite-horizon with an FSMC is significant and challenging. Our contributions in this work are summarized in the following.

1) The sensor scheduling problem is investigated over a time-varying wireless communication channel modeled by an FSMC. Different from the existing work involving the known real-time channel state [23], [24], [31], our work considers the channel state is available at previous time instant but is unknown at the current time.
2) Compared with the efforts in investigating existence of the scheduling policy to stabilize the remote estimator in [28], the focus of our work is the optimality of sensor scheduling policy considering the communication cost. We formulate the sensor scheduling problem as a Markov decision process (MDP) to minimize remote state estimation errors combining with communication costs. The optimal scheduling policy is shown to be deterministic stationary (DS) and is of a double threshold structure, which facilitates offline computation and online implementation of sensor scheduling policies.
3) The NSC is derived to stabilize the remote estimator under the optimal structured policy. Our stability condition coincides with that in [28] for costless transmission.

The remainder of this article is organized as follows. The mathematical model is given in Section II. Preliminary knowledge is given in Section III. In Section IV, we present threshold structures of the optimal scheduling strategy, and analyze the stability of the remote estimator. Numerical studies are shown to verify the proposed theoretical results in Section V. Finally, Section VI concludes this article.

Notations: The transposition of a matrix is denoted by a superscript $^{\mathrm{T}}$. Denote by $\Pr(\cdot)$ and $\mathbb{E}(\cdot)$ the probability and expectation of stochastic variable, respectively. $||\cdot||_1$ is the ℓ_1 norm of a vector.

II. Problem Setup

A. Dynamic System Model

Consider the following linear dynamic process:

$$x_{k+1} = Ax_k + w_k,$$
$$y_k = Cx_k + v_k$$

with constant system and observation matrices A and C, respectively, where $k \in \mathbb{N}$ is the time index. The variables $x_k, w_k \in \mathbb{R}^n$ are the process state and the process noise that is zero-mean Gaussian noise with covariance $Q \succeq 0$. The variables $y_k, v_k \in \mathbb{R}^m$ are the measurement and its noise which is also Gaussian with zero-mean and covariance $R \succ 0$. Besides, it is assumed that w_k and v_k are uncorrelated.

At each time k, the sensor calculates local estimates $\hat{x}_k^s = \mathbb{E}[x_k|y_1, \ldots, y_k]$. Under the assumption that the pair (A, C) is detectable and $(A, \sqrt{Q})$ is stabilizable, there exists a steady state $\bar{P}$ of estimation error covariance P_k^s, where $P_k^s \triangleq \mathbb{E}[(\hat{x}_k^s - x_k)(\hat{x}_k^s - x_k)^{\mathrm{T}}|y_1, \ldots, y_k]$. $\hat{x}_k^s$ and P_k^s are generated by the standard Kalman filter [37]. Furthermore, assume that $P_k^s = \bar{P}$ for $k \geq 0$.

Remark 1: A linear system is considered to model the dynamic process and, thus, the standard Kalman filter is the optimal estimator. For other types of systems, for example, nonlinear systems [38], designing suitable state estimators should be the first step.

In this article, we focus on the stability and performance of the remote estimator for an open-loop unstable system. It should be noted that when the open-loop system is stable, our main results are not affected. In practice, the closed-loop system may operate in an open-loop mode. In addition, the design of scheduling policies achieving some closed-loop performance can be transformed into an open-loop scheduling strategy design problem, for example, the LQG control problem [39]. Therefore, it is necessary to study the remote estimation for the open-loop scenario.

B. Communication Model

As shown in Fig. 1, the sensor schedule demand is made at the remote estimator at every time slot: activating the sensor

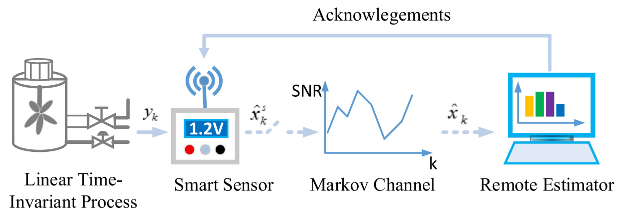


Fig. 1. System architecture.

if there is important data to transmit, otherwise leaving the sensor idle for saving energy at time k. Denote by $a_k = 1$ if the sensor is scheduled to transmit at time k; otherwise, $a_k = 0$. At every slot, acknowledgments (ACKs) are sent to the sensor through a feedback communication channel. Generally, the remote end is more powerful than the sensor; thus, the feedback channel is considered to be reliable. Benefiting from ACKs, the sensor could be dynamically scheduled depending on real-time information from the remote end. This setting is supported by many protocols, such as IEEE 802.15.4 [26]. Denote by $\gamma_k = 1$ when the remote end receives the transmitted packet at time k. Denote by I_k the available information set for the remote estimator up to time k. According to the previous discussion

$$I_k \triangleq \{\gamma_0, \gamma_1, \ldots, \gamma_{k-1}\} \cup \left\{\gamma_0\hat{x}_0^s, \gamma_1\hat{x}_1^s, \ldots, \gamma_{k-1}\hat{x}_{k-1}^s\right\}.$$

In this article, we consider the time-varying channel state. Denote by $\mathfrak{h}_k$ the channel gain at time k. To acquire the current channel state, the sensor needs to send pilot signal at the beginning of each time. The remote center calculates the channel information and then feeds it back to the sensor. A one-step time delay occurs in this process and, thus, the sensor could only know the channel state at the previous time instant. The following assumption is used in this article, which captures the temporally correlated channel states' variation in real wireless communication [40].

Assumption 1: The wireless channel between the sensor and remote estimator is an FSMC. The channel gains take values from $\mathbb{H} \triangleq \{\mathfrak{g}_1, \mathfrak{g}_2, \ldots, \mathfrak{g}_T\}$ with $\mathfrak{g}_1 \leq \mathfrak{g}_2 \leq \ldots \leq \mathfrak{g}_T$. The transition probability matrix is given by

$$\Xi = \begin{bmatrix} \xi_{11} & \cdots & \xi_{1T} \\ \vdots & \ddots & \vdots \\ \xi_{T1} & \cdots & \xi_{TT} \end{bmatrix}$$

where $\xi_{ij} = \Pr\{\mathfrak{h}_k = \mathfrak{g}_j | \mathfrak{h}_{k-1} = \mathfrak{g}_i\}, 1 \leq i, j \leq T, k \in \mathbb{N}^+$.

Based on the established model of communication theory [41], the probability of successful packet reception rate is determined by the signal-to-noise ratio at time k (SNR_k)

$$\Pr\{\gamma_k = 1\} = f(\text{SNR}_k) \tag{1}$$

with

$$\text{SNR}_k = \frac{\mathfrak{h}_k \zeta}{\sigma^2}$$

where the function $f : \mathbb{R} \mapsto [0, 1]$ is a mapping from SNR to the probability of successful packet reception rate, and the specific function form depends on the channel model and modulation method. ζ and σ^2 are the transmission power used by the sensor and the channel noise power, respectively. In this work, we mainly focus on the analysis of effect of the channel gain on the remote estimation performance, therefore, (1) is modified as

$$\Pr\{\gamma_k = 1 | \mathfrak{h}_k = \mathfrak{g}_j\} = f(\text{SNR}_{k,j}) \tag{2}$$

where $\text{SNR}_{k,j}$ is the signal-to-noise ratio at time k with channel gain $\mathfrak{g}_j$. Furthermore, since there exist finite channel gains $\mathfrak{g}_j, 1 \leq j \leq T$, we have

$$p_{k,j} \triangleq \Pr\{\gamma_k = 0 | \mathfrak{h}_k = \mathfrak{g}_j\} \tag{3}$$

where $p_{k,j}$ is the packet dropout rate at time k under channel gain $\mathfrak{g}_j$. We sometimes omit the time index k and write $p_{k,j}$ as p_j for simplicity if there is no ambiguity.

Now, we are ready to introduce the remote state estimation under the aforementioned communication model. We use a stochastic variable τ_k to represent for the number of successive packet dropout up to time k, that is

$$\tau_k = k - \max\{t | \gamma_t = 1, 0 \leq t \leq k\}. \tag{4}$$

It can be observed that

$$\tau_k = \begin{cases} 0, & \text{if } \gamma_k = 1 \\ \tau_{k-1} + 1, & \text{if } \gamma_k = 0. \end{cases} \tag{5}$$

The remote state estimate is updated as follows:

$$\hat{x}_k = \gamma_k \hat{x}_k^s + (1 - \gamma_k) A \hat{x}_{k-1}.$$

Since we assume that $P_k^s = \bar{P}$ for $k \geq 0$, the estimation error covariance at the remote side is given by

$$P_k = \gamma_k \bar{P} + (1 - \gamma_k) h(P_{k-1}) \tag{6}$$

where $h(X) = AXA^{\mathrm{T}} + Q$ and $h^n(X) = h(h^{n-1}(X))$ with $h^0(X) = \bar{P}$. By (4) and (6), we have $P_k = h^{\tau_k}(\bar{P})$.

C. Problem of Interest

When the sensor is allowed to transmit its packet at time k, that is, $a_k = 1$, a fixed energy consumption ζ (including the circuit consumption of the sensor and the transmission power) will occur, where $0 \leq \zeta \leq \bar{\zeta}$ and $\bar{\zeta}$ is the upper bound for the consumed power.

Then, the one-step cost is defined as the weighted sum of the state estimation accuracy at the remote side and energy consumption for the sensor at time k, which is given by

$$r_k(\tau_{k-1}, \mathfrak{h}_{k-1}, a_k) = \text{Tr} h^{\tau_{k-1}}\left(\bar{P}\right) + a_k \kappa \zeta \tag{7}$$

where κ is the weighting factor. By designing an appropriate scheduling strategy to balance the two, the long-term cumulative performance index can be optimized.

In this article, we seek a transmission policy π which only depends on the estimation error (or equivalently packet loss duration time) and the channel gain, that is, $a_k = a_k(\tau_{k-1}, \mathfrak{h}_{k-1})$. Denote by Π all admissible transmission policies with $\Pi = (a_1(s_1), a_2(s_2), a_3(s_3), \ldots)$ and $\Pi_{d,s}$ all admissible DS transmission policies. The scheduler exploits the channel state $\mathfrak{h}$ and holding time τ at time $k-1$ to minimize the estimation error covariance while reduces communication consumption. We optimize the remote estimation performance over the infinite-horizon

$$R_\alpha(\pi) = \mathbb{E}^\pi\left[\sum_{k=1}^{\infty}\alpha^{k-1} r_k | \tau_0, \mathfrak{h}_0\right] \tag{8}$$

where r_k, α, τ_0, $\mathfrak{h}_0$, and π are the one-step cost at time k, discount factor, duration time, channel gain, and scheduling policy, respectively. Specifically, we aim to find an optimal DS policy π^*.

III. Preliminaries

First, the MDP framework is established using a tuple $\{\mathbb{S}, \mathbb{A}, \mathbb{P}\{\cdot|\cdot,\cdot\}, r(\cdot,\cdot)\}$.

1) The state space $\mathbb{S} \triangleq \mathbb{N} \times \mathbb{H}$. The state consists of two substates $s_k = (\tau_{k-1}, \mathfrak{h}_{k-1})$: a) the packet loss duration time and b) channel gain.
2) The action space $\mathbb{A} \triangleq \{0, 1\}$, and decision action $a_k \in \mathbb{A}$.
3) The state transition probability $\mathbb{P}\{\cdot|\cdot,\cdot\}$: Denote by $\Pr\{s_{k+1}|s_k, a_k\}$ the transition probability from state $s_k = (\tau_{k-1}, \mathfrak{h}_{k-1})$ to $s_{k+1} = (\tau_k, \mathfrak{h}_k)$ when action a_k is taken at time k. According to (5) and (6), the state transition probability is given by

$$\begin{aligned}&\Pr\{\tau_k, \mathfrak{h}_k | \tau_{k-1}, \mathfrak{h}_{k-1}, a_k\}\\ &= \begin{cases} \xi_{ij}(1-p_j), & \text{if } \tau_k = 0, \mathfrak{h}_k = \mathfrak{g}_j, \mathfrak{h}_{k-1} = \mathfrak{g}_i, a_k = 1\\ \xi_{ij}p_j, & \text{if } \tau_k = \tau_{k-1}+1, \mathfrak{h}_k = \mathfrak{g}_j, \mathfrak{h}_{k-1} = \mathfrak{g}_i, a_k = 1\\ \xi_{ij}, & \text{if } \tau_k = \tau_{k-1}+1, \mathfrak{h}_k = \mathfrak{g}_j, \mathfrak{h}_{k-1} = \mathfrak{g}_i, a_k = 0\\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{9}$$

4) The immediate cost r is defined as

$$r(\tau_{k-1}, \mathfrak{h}_{k-1}, a_k) \triangleq \mathrm{Tr} h^{\tau_{k-1}}(\bar{P}) + a_k \kappa \zeta. \tag{10}$$

Under the previously introduced MDP framework, (8) can be written as

$$\mathcal{R}_\alpha(\pi, s_0) \triangleq \mathbb{E}^\pi_{s_0}\left[\sum_{k=1}^{\infty}\alpha^{k-1} r(s_k, a_k)\right]. \tag{11}$$

Accordingly, the problem of interest is formulated as follows:

Problem 1:

$$\inf_{\pi\in\Pi_{d,s}} \mathcal{R}_\alpha(\pi, s_0).$$

Denote by $V_\alpha(s_0)$ the optimal value function, that is

$$V_\alpha(s_0) \triangleq \mathbb{E}^{\pi^*}_{s_0}\left[\sum_{k=1}^{\infty}\alpha^{k-1} r(s_k, a_k)\right]. \tag{12}$$

If there exists an optimal DS policy, the following Bellman optimality equation holds $\forall s \in \mathbb{S}$:

$$V_\alpha(s) = \min_{a\in\mathbb{A}}\left[r(s,a) + \alpha\sum_{s^+\in\mathbb{S}}\Pr(s^+|s,a)V_\alpha(s^+)\right]. \tag{13}$$

Denote by $Q_\alpha(s, a)$ the action-value function, where

$$Q_\alpha(s,a) = r(s,a) + \alpha\sum_{s^+}\Pr(s^+|s,a)V_\alpha(s^+). \tag{14}$$

It can be seen that $V_\alpha(s) = \min_{a\in\mathbb{A}} Q_\alpha(s, a)$.

Two essential definitions are introduced in the following.

Definition 1 [42]: Let W be an arbitrary function: $\mathbb{S} \mapsto \mathbb{R}^+$ satisfying $\inf_{s\in S} W(s) > 0$. The weighed supremum norm $||\cdot||$ for real-valued functions φ on $\mathbb{S}$ is given by

$$||\varphi||_W = \sup_{s\in\mathbb{S}} W(s)^{-1}|\varphi(s)|.$$

Definition 2 [42]: For two partially ordered sets S_1 and S_2, and a real-valued function ϕ on $S_1 \times S_2$, we say ϕ is subadditive function if for any $s_1 \le s_1'$ in S_1 and $s_2 \le s_2'$ in S_2, there holds

$$\phi(s_1', s_2') + \phi(s_1, s_2) \le \phi(s_1', s_2) + \phi(s_1, s_2'). \tag{15}$$

IV. Main Results

In this section, the double threshold structure of the optimal scheduling policy and the associated stability of the remote estimator are presented.

A. Double Threshold Structure of the Optimal Policy

Since the boundness of the value function depends on the specific sensor scheduling policy, the focus of this section is to explore the structure of the optimal policy under the assumption that there exists a π^* such that $\mathcal{R}_\alpha(\pi^*)$ is bounded.

The following lemma shows the monotonicity of the optimal discounted value function $V_\alpha(s)$.

Lemma 1: For the optimal discounted value function in (13), there holds $V_\alpha(s) \le V_\alpha(s')$, where $s = (\tau, \mathfrak{h}), s' = (\tau', \mathfrak{h}) \in \mathbb{S}$ with $\tau \le \tau' \in \mathbb{N}$, $\mathfrak{h} \in \mathbb{H}$.

Proof: The proof of Lemma 1 is completed by the induction method. Define operator $T_\alpha v(s) \triangleq \min_a[r(s,a) + \alpha\sum_{s^+} v(s^+)\Pr(s^+|s,a)]$. Then, it suffices to prove the monotonicity of $v(s)$ can be preserved by the operator T_α. Assume that $\forall s, s' \in \mathbb{S}, \tau \le \tau', v_0(\cdot) = 0$

$$v_k(s = (\tau, \mathfrak{h})) \le v_k(s' = (\tau', \mathfrak{h})). \tag{16}$$

For notational simplicity, we sometimes omit the time index k if there is no ambiguity in the following. Then, it suffices to prove that, $v_{k+1}(s) \le v_{k+1}(s')$, that is

$$\begin{aligned} &r(s,a) + \alpha\sum_{s^+} v(s^+)\Pr(s^+|s,a)\\ &\quad\le r(s',a) + \alpha\sum_{s'^+} v(s'^+)\Pr(s'^+|s',a)\end{aligned} \tag{17}$$

$\forall a \in \mathbb{A}$. By (10) and the fact that $h^\tau(\bar{P}) \preceq h^{\tau'}(\bar{P})$ $\forall \tau \le \tau'$, it can be seen that the immediate cost $r(\tau, \mathfrak{h}, a)$ is nondecreasing in τ. Let $\mathfrak{g}_i = \mathfrak{h} \le \mathfrak{h}' = \mathfrak{g}_{i'} \in \mathbb{H}$. In the following, we will prove (17) by the following cases.

1) Case $a(s) = 0$, that is, not to transmit data packet: Equation (17) holds if $\sum_{s^+} v(s^+)\Pr(s^+|s,0) \le \sum_{s'^+} v(s'^+)\Pr(s'^+|s',0)$, which is equivalent to

$$\sum_{j=1}^{\mathrm{T}} v_k(s^+ = (\tau+1, \mathfrak{g}_j))\xi_{ij} \le \sum_{j=1}^{\mathrm{T}} v_k(s'^+ = (\tau'+1, \mathfrak{g}_j))\xi_{ij}. \tag{18}$$

Equation (18) is true due to (16), which implies that (17) holds when $a = 0$.

2) Case $a(s)=1$: Similar to case 1), (17) holds if

$$\sum_{j=1}^{\mathrm{T}} v(0,\mathfrak{g}_j)(1-p_j)\xi_{ij}+\sum_{j=1}^{\mathrm{T}} v\big(\tau+1,\mathfrak{g}_j\big)p_j\xi_{ij} \leq \sum_{j=1}^{\mathrm{T}} v(0,\mathfrak{g}_j)(1-p_j)\xi_{ij}+\sum_{j=1}^{\mathrm{T}} v\big(\tau'+1,\mathfrak{g}_j\big)p_j\xi_{ij}. \quad (19)$$

Equation (19) is true due to (16), which implies that (17) holds in the case of $a=1$.

Combining (18) and (19), we know that $T_\alpha v(s=(\tau,\mathfrak{h}))\leq T_\alpha v(s'=(\tau',\mathfrak{h}))$, which means the monotonicity is preserved. Using the argument of induction method, we can obtain $V_\alpha(s=(\tau,\mathfrak{h}))\leq V_\alpha(s'=(\tau',\mathfrak{h}))$. This completes the proof. ■

Next, we show the optimal policy is nondecreasing in the remote estimation error.

Theorem 1: For any fixed channel gain $\mathfrak{h}\in\mathbb{H}$, the optimal sensor scheduling policy $a^*(\tau,\mathfrak{h})$ is nondecreasing in packet duration time $\tau\in\mathbb{N}$.

Proof: To prove the result in Theorem 1, we show that if $a^*(\tau,\mathfrak{h})=1$, $a^*(\tau',\mathfrak{h})=1$ with $\tau'\geq\tau\in\mathbb{N},\mathfrak{h}\in\mathbb{H}$. Define $\Delta Q_\alpha(s)\triangleq Q_\alpha(s,0)-Q_\alpha(s,1)$. From (14), one can see that $a^*(s)=1$ if $\Delta Q_\alpha(s)\geq 0$ and $a^*(s)=0$ otherwise. Then, we have

$$\begin{aligned}\Delta Q_\alpha(s) &= r(s,0)+\alpha\sum_{s^+\in\mathbb{S}}V_\alpha(s^+)\Pr\big(s^+|s,0\big)\\ &\quad-\left[r(s,1)+\alpha\sum_{s^+\in\mathbb{S}}V_\alpha\big(s^+\big)\Pr(s^+|s,1)\right]\\ &= \mathrm{Tr}h^\tau\big(\bar P\big)+\alpha\sum_{j=1}^{\mathrm{T}}\xi_{ij}V_\alpha\big(\tau+1,\mathfrak{g}_j\big)-\Bigg[\mathrm{Tr}h^\tau\big(\bar P\big)+\kappa\zeta\\ &\quad+\alpha\left(\sum_{j=1}^{\mathrm{T}}\xi_{ij}p_jV_\alpha\big(\tau+1,\mathfrak{g}_j\big)+\sum_{j=1}^{\mathrm{T}}\xi_{ij}(1-p_j)V_\alpha(0,\mathfrak{g}_j)\right)\Bigg]\\ &= -\kappa\zeta+\alpha\sum_{j=1}^{\mathrm{T}}\xi_{ij}(1-p_j)\big[V_\alpha(\tau+1,\mathfrak{g}_j)-V_\alpha(0,\mathfrak{g}_j)\big]. \quad (20)\end{aligned}$$

For a fixed channel gain $\mathfrak{h}=\mathfrak{g}_i$, suppose that there exists a state $s^*=(\tau^*,\mathfrak{h})$ such that $a(s^*)=1$, we have $\Delta Q_\alpha(s^*)\geq 0$. Let $s'^*=(\tau'^*,\mathfrak{h})\in\mathbb{S}$ with $\tau'^*\geq\tau^*$, due to (20) and the monotonicity of $V_\alpha(s)$ given in Lemma 1, we obtain $\Delta Q_\alpha(s'^*)\geq\Delta Q_\alpha(s^*)\geq 0$, which implies that $a(s'^*)=1$. This completes the proof. ■

Theorem 1 implies $a^*(s)=1$ if $\tau\geq\tau_j^*$ given a channel state $\mathfrak{g}_j\in\mathbb{H}$, where $\tau_j^*\in\mathbb{N}$. According to the proof of Theorem 1, it can be seen that τ_j^* is the minimum τ such that $\Delta Q_\alpha(s)\geq 0$. From the result in Lemma 1, we know that given $\mathfrak{g}_j\in\mathbb{H}$, $V_\alpha(\tau+1,\mathfrak{g}_j)$ is increasing in τ. Then, τ_j^* for each $\mathfrak{g}_j\in\mathbb{H}$ such that $\Delta Q_\alpha(s)\geq 0$ is bounded.

The following stochastic monotonicity of transition probability matrix Ξ is introduced to establish the threshold structure on the channel state.

Assumption 2: It is assumed that the transition probability matrix of the channel gain Ξ is stochastic monotone [29], that is $\forall\ 1\leq i<i'\leq T$ and $1\leq t\leq T$, there holds

$$\sum_{j=t}^{\mathrm{T}}\xi_{i'j}\geq\sum_{j=t}^{\mathrm{T}}\xi_{ij}. \quad (21)$$

The stochastic monotonicity of transition matrix Ξ allows an intuitive interpretation: a larger channel gain at time $k-1$ leads to a larger channel gain at time k with bigger probability [26]. This implies that the better the channel quality at the previous time instant, the more likely it is to have a better channel quality at the current moment. Based on the stochastic monotonicity, we have the following result.

Theorem 2: If Assumption 2 holds, the optimal policy $a^*(\tau,\mathfrak{h})$ is nonincreasing with respect to the channel state $\mathfrak{h}$ given packet loss duration time τ.

Proof: In what follows, the time index k is omitted sometimes for notational simplicity. According to [42, Th. 6.11.6], to show the monotonicity of the optimal policy in $\mathfrak{h}$, it suffices to verify the following conditions.

1) $r(s,a)$ is nondecreasing in $\mathfrak{h}$ $\forall a\in\mathbb{A}$, with fixed $\tau\in\mathbb{Z}_+$.
2) $\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s,a\}$ is nondecreasing in $\mathfrak{h}$ $\forall a\in\mathbb{A}$ with fixed $\tau\in\mathbb{Z}_+$, where $\mathfrak{h}'$ is the successive channel state of s, $t\in\mathbb{H}$.
3) $r(s,a)$ is a subadditive function of $\mathfrak{h}$ and a on $\mathbb{H}\times\mathbb{A}$ for fixed $\tau\in\mathbb{Z}_+$.
4) $\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s,a\}$ is a subadditive function of $\mathfrak{h}$ and a on $\mathbb{H}\times\mathbb{A}$, with fixed $\tau\in\mathbb{Z}_+$.

Suppose that $s=(\tau,\mathfrak{h})$, $s^+=(\tau,\mathfrak{h}^+)$, $s^-=(\tau,\mathfrak{h}^-)$ with $\mathfrak{h}^+=\mathfrak{g}_{i^+}\geq\mathfrak{g}_{i^-}=\mathfrak{h}^-$.

It is easy to see from (7) that $r(\tau,\mathfrak{h}^+,a)=r(\tau,\mathfrak{h}^-,a)$ $\forall a\in\mathbb{A}$, $\mathfrak{h}^+,\mathfrak{h}^-\in\mathbb{H}$. Thus, conditions 1) and 3) are true.

Suppose that the index of channel state t is l, that is, $\mathfrak{g}_l=t$. Then, for fixed $s'=(\tau',\mathfrak{h}')$, there holds the following.

1) For $a=1$, the following is true:

$$\begin{aligned}&\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'|s^+=(\tau,\mathfrak{h}^+),a\big\}\\ &=\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'=\big(0,\mathfrak{h}'\big)|s^+=\big(\tau,\mathfrak{h}^+\big),a\big\}\\ &\quad+\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'=\big(\tau+1,\mathfrak{h}'\big)|s^+=\big(\tau,\mathfrak{h}^+\big),a\big\}\\ &=\sum_{j=l}^{\mathrm{T}}\xi_{i^+j}(1-p_j)+\sum_{j=l}^{\mathrm{T}}\xi_{i^+j}p_j\\ &=\sum_{j=l}^{\mathrm{T}}\xi_{i^+j}\geq\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'|s^-=\big(\tau,\mathfrak{h}^-\big),a\big\}=\sum_{j=l}^{\mathrm{T}}\xi_{i^-j}\end{aligned}$$

where the inequality is due to Assumption 2.

2) For $a=0$, $\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s^+,a\}=\sum_{j=l}^{\mathrm{T}}\xi_{i^+j}\geq\sum_{j=l}^{\mathrm{T}}\xi_{i^-j}=\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s^-,a\}$.

Hence, we have verified condition 2). Now, we are ready to check condition 4). For $a^+\geq a^-$ in $\mathbb{A}$, the following equations are true:

$$\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'|s^+,a^+\big\}=\sum_{\mathfrak{h}'\geq t}\Pr\big\{s'|s^+,a^-\big\}=\sum_{j=l}^{\mathrm{T}}\xi_{i^+j}$$

$$\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s^-,a^+\}=\sum_{\mathfrak{h}'\geq t}\Pr\{s'|s^-,a^-\}=\sum_{j=l}^{\mathrm{T}}\xi_{i-j}.$$

Therefore, condition 4) is true. By [42, Th. 6.11.6], we obtain the optimal scheduling policy $a^*(\tau,\mathfrak{h})$ is nondecreasing in $\mathfrak{h}\in\mathbb{H}$ with fixed $\tau\in\mathbb{Z}_+$. ■

Remark 2: Compared with the structural results on packet loss duration time τ over the i.i.d packet dropout channel [25], the two-state Markov packet dropout channel [27], [36], or two-state usage-dependent Markov channel [31], Theorems 1 and 2 indicate that $a^*(s)$ is of threshold structure for both τ and $\mathfrak{h}$ over an FSMC.

An available implementation method of the studied scheduling policy is illustrated as follows. *Offline Calculation Phase*: The critical thresholds of the optimal policy is calculated according to the system parameters and stored in the scheduler. *Online Implementation Phase:* At each time slot, the remote center estimates the current channel state using pilot signal and obtains the optimal scheduling action by comparing real-time packet loss duration time τ and the channel state $\mathfrak{h}$ with the critical thresholds. Finally, the scheduling action is fed back to the sensor through a reliable channel. This setting and mechanism are considered for voltage control in microgrid [3] and for power management in networked control systems.

B. Existence of DS Policy and Stability of the Remote Estimator

In this section, the existence of a DS policy is shown. Then, an NSC guaranteeing mean-square stability of the remote estimator is derived.

Assumption 3: Assume that spectral radius of the system matrix, packet loss rate, and channel state transition probability satisfy

$$\rho^2(A)\max_{i,j}\xi_{i,j}\leq 1. \tag{22}$$

By Assumption 3, we can restrict the optimal policy of Problem 1 to a DS class.

Theorem 3: Under Assumption 3, the optimal sensor scheduling policy is DS.

Proof: According to [42, Th. 6.10.4], we shall examine the following two conditions (see [42, Assumptions 6.10.1 and 6.10.2, and Proposition 6.10.5]):

1) There exists a constant $\varrho<\infty$ such that

$$\sup_{a\in\mathbb{A}}|r(s,a)|\leq\varrho W(s). \tag{23}$$

2) There exists a constant $\mathcal{C}>0$ such that

$$\sum_{s^+\in\mathbb{S}}\Pr\big(s^+|s,a\big)W\big(s^+\big)\leq W(s)+\mathcal{C} \tag{24}$$

for $a\in\mathbb{A}$ and $s\in\mathbb{S}$, where s^+ is the next state of s.

The verification of condition (23) follows from the argument in [14] and [25]. By letting $W(\tau_{k-1},\mathfrak{h}_{k-1})=\rho^{2\tau_{k-1}}(A)$, it can be shown $||r(s,a)||_W<\varrho$. More details can be found in [14]. To check condition (24), consider the structured policy in Theorem 1, $a^*(s)=1$ if τ exceeds some threshold given a channel state, $a^*(s)=0$ otherwise. Then, we have the following.

1) For state $s_k=(\tau_{k-1},\mathfrak{g}_i)$ with $\tau_{k-1}\geq\tau_i^*$, it follows $a^*(s_k)=1$ and:

$$\Pr\big(s^+|s,1\big)W\big(s^+\big)=\begin{cases}(1-p_j)\xi_{ij}, & \text{if } s^+=\big(0,\mathfrak{h}_k=\mathfrak{g}_j\big),\\ p_j\xi_{ij}\rho^2(A)W(s), & \text{if } s^+=\big(\tau_{k-1}+1,\mathfrak{h}_k=\mathfrak{g}_j\big)\\ 0, & \text{otherwise.}\end{cases}$$

That is

$$\sum_{s^+\in\mathbb{S}}\Pr\big(s^+|s,1\big)W\big(s^+\big)=p_j\xi_{ij}\rho^2(A)W(s)+(1-p_j)\xi_{ij}.$$

If $\max_{i,j}p_j\xi_{ij}\rho^2(A)\leq 1$, one obtains $\sum_{s^+\in\mathbb{S}}\Pr(s^+|s,a)$ $W(s^+)\leq W(s)+1$.

2) For state $s_k=(\tau_{k-1},\mathfrak{g}_i)$ with $\tau_{k-1}<\tau_i^*$, $a^*(s_k)=0$, we obtain

$$\Pr\big(s^+|s,0\big)W\big(s^+\big)=\xi_{ij}\rho^2(A)W(s).$$

If $\max_{i,j}\xi_{ij}\rho^2(A)\leq 1$, it follows $\sum_{s^+\in\mathbb{S}}\Pr(s^+|s,a)$ $W(s^+)\leq W(s)$.

Therefore, combining the above two cases, let $\max_{i,j}\xi_{ij}\rho^2(A)\leq 1$ and $\mathcal{C}=1$, condition (24) holds. This completes the proof. ■

The derived condition (22) implies that there exists an optimal policy π^* bounding the discount value function, that is, $||V_\alpha||_W<\infty$. Inspired by [31] and [28], in the following, an NSC is given to guarantee mean-square stability of the remote estimator.

Theorem 4: The remote state estimation is mean-square stable under the optimal sensor scheduling policy if and only if

$$\rho(\Xi\Gamma)\rho^2(A)<1 \tag{25}$$

where $\Gamma=\mathrm{diag}\{p_1,p_2,\ldots,p_T\}$.

Proof: Denote by ω_{ij} the limiting distribution of states $s=(i,j)\triangleq(\tau,\mathfrak{g}_j)$, with a shorthand s_{ij}, induced by a scheduling policy. Recall the structured scheduling policy, denoted by π^*, in Theorem 1, given a channel state $\mathfrak{g}_j\in\mathbb{H}$, $a^*(s)=1$ if $i\geq\tau_j^*$. Following a similar argument of [16, Th. 1], it can be shown that, under π^*, the Markov chain $\{s_{ij}\}$ is irreducible, aperiodic and recurrent. The first two properties are easy to verify since for any state $(\tau,\mathfrak{h})$, there is a positive probability to reach one of the states $(0,\mathfrak{h}')$. For the recurrent property, denote by $\mathbb{E}(T_1)$ the expectation of the first time to revisit the state $(0,\mathfrak{g}_1)$, there holds $\mathbb{E}(T_1)<\sum_{r=0}^{\infty}[1-\underline{\xi}(1-p_1)]^r<\infty$, where $\underline{\xi}\triangleq\min_{i,\xi_{i1}\neq 0}\xi_{i1}$. Then, the unique limiting distribution of $\{s_{ij}\}$ exists. Denote by $\bar{\tau}^*\triangleq\max\{\tau_j^*|1\leq j\leq T\}$. Then, for the limiting distribution at state s_{ij} under the optimal action $\pi^*(s_{ij})$, there hold the following.

1) If $i=0$, $\omega_{0j}=\sum_{i'\in\mathbb{N}}\sum_{1\leq j'\leq T}\omega_{i'j'}\Pr(s_{0j}|s_{i'j'},\pi^*(s_{i'j'}))$ with $\Pr(s_{0j}|s_{i'j'},\pi^*(s_{i'j'}))=(1-p_j)\xi_{j'j}$.
2) If $0<i\leq\bar{\tau}^*$, $\omega_{ij}=\sum_{1\leq j'\leq T}\omega_{(i-1)j'}\Pr\ (s_{ij}\ |s_{(i-1)j'},\pi^*(s_{(i-1)j'}))$, where $\Pr\ (s_{ij}\ |s_{(i-1)j'},\pi^*(s_{(i-1)j'}))=\xi_{j'j}$ for $i-1<\tau_{j'}^*$ and $\Pr\ (s_{ij}\ |s_{(i-1)j'},\pi^*(s_{(i-1)j'}))=p_j\xi_{j'j}$ for $i-1\geq\tau_{j'}^*$.
3) If $i>\bar{\tau}^*$, $\omega_{ij}=\sum_{1\leq j'\leq T}\omega_{(i-1)j'}\Pr\ (s_{ij}\ |s_{(i-1)j'},\pi^*(s_{i'j'}))$ with $\Pr\ (s_{ij}|s_{(i-1)j'},\pi^*(s_{(i-1)j'}))=p_j\xi_{j'j}$.

It can be seen that the states satisfying $i = 0$ and $0 < i \leq \bar{\tau}^*$ induce bounded estimation errors, while estimation errors corresponding to states $i > \bar{\tau}^*$ may grow unbounded. According to case 3) above, we have the following:

$$\begin{cases} \omega_{(\bar{\tau}^*+1)1} = \omega_{\bar{\tau}^*1} p_1 \xi_{11} + \omega_{\bar{\tau}^*2} p_1 \xi_{21} + \cdots + \omega_{\bar{\tau}^*T} p_1 \xi_{T1} \\ \vdots \\ \omega_{(\bar{\tau}^*+1)T} = \omega_{\bar{\tau}^*1} p_T \xi_{1T} + \omega_{\bar{\tau}^*2} p_T \xi_{2T} + \cdots + \omega_{\bar{\tau}^*T} p_1 \xi_{TT} \end{cases}$$

which is compactly written as $\boldsymbol{\omega}_{\bar{\tau}^*+1} = \boldsymbol{\omega}_{\bar{\tau}^*}(\Xi\Gamma)$, where $\boldsymbol{\omega}_i \triangleq (\omega_{(\bar{\tau}^*+1)1}, \ldots, \omega_{(\bar{\tau}^*+1)T})$. By repeatedly using the property in case 3), one can derive that

$$\boldsymbol{\omega}_{\bar{\tau}^*+m} = \boldsymbol{\omega}_{\bar{\tau}^*}(\Xi\Gamma)^m, m \in \mathbb{N}. \tag{26}$$

Denote by $\tilde{\boldsymbol{\omega}}_i \triangleq \sum_{1 \leq j \leq T} \omega_{ij} = ||\boldsymbol{\omega}_i||_1$. The remote estimator is mean-square stable iff the following average remote estimation error is bounded [28]:

$$\lim_{K\to\infty} \frac{1}{K+1} \sum_{k=0}^{K} \mathbb{E}[P_k] = \sum_{i=0}^{\infty} \tilde{\boldsymbol{\omega}}_i h^i(\bar{P}).$$

Due to $\sum_{i=0}^{\infty} \tilde{\boldsymbol{\omega}}_i = 1$ and the nondecreasing property of $h^i(\bar{P})$, one concludes that the remote estimator is stable iff $\sum_{i=\bar{\tau}^*}^{\infty} \tilde{\boldsymbol{\omega}}_i h^i(\bar{P}) < \infty$.

According to [43, Th. 4.2], $\rho(\Xi\Gamma)$ is an eigenvalue of nonnegative matrix $\Xi\Gamma$. Denote by μ^* the eigenvector with respect to $\rho(\Xi\Gamma)$. By repeatedly using $\rho(\Xi\Gamma)\mu^* = \Xi\Gamma\mu^*$, there holds $\rho^m(\Xi\Gamma)\mu^* = (\Xi\Gamma)^m\mu^*$. In light of (26), one obtains

$$\boldsymbol{\omega}_{\bar{\tau}^*+m}\mu^* = \rho^m(\Xi\Gamma)\boldsymbol{\omega}_{\bar{\tau}^*}\mu^*, m \in \mathbb{N}. \tag{27}$$

Then, there exists a constant c_m^* for each $m \in \mathbb{N}$ such that

$$||\boldsymbol{\omega}_{\bar{\tau}^*+m}||_1 = c_m^* \rho^m(\Xi\Gamma)||\boldsymbol{\omega}_{\bar{\tau}^*}||_1. \tag{28}$$

Let $\bar{c}^* = \max\{c_m^* | m \in \mathbb{N}\}$, $\underline{c}^* = \min\{c_m^* | m \in \mathbb{N}\}$. It follows that:

$$\begin{aligned} \underline{c}^* \sum_{i=\bar{\tau}^*}^{\infty} \rho^{i-\bar{\tau}^*}(\Xi\Gamma)||\boldsymbol{\omega}_{\bar{\tau}^*}||_1 h^i(\bar{P}) &\leq \sum_{i=\bar{\tau}^*}^{\infty} \tilde{\boldsymbol{\omega}}_i h^i(\bar{P}) \\ &\leq \bar{c}^* \sum_{i=\bar{\tau}^*}^{\infty} \rho^{i-\bar{\tau}^*}(\Xi\Gamma)||\boldsymbol{\omega}_{\bar{\tau}^*}||_1 h^i(\bar{P}). \end{aligned} \tag{29}$$

Due to $h^i(\bar{P}) = A^i \bar{P} A^{\mathrm{T}^i} + \sum_{n=0}^{i-1} A^n Q A^{\mathrm{T}^n}$, for the LHS of inequality (29), we derive

$$\begin{aligned} &\underline{c}^* \sum_{i=\bar{\tau}^*}^{\infty} \rho^{i-\bar{\tau}^*}(\Xi\Gamma)||\boldsymbol{\omega}_{\bar{\tau}^*}||_1 h^i(\bar{P}) \\ &= \underline{c}^* ||\tilde{\boldsymbol{\omega}}_{\bar{\tau}^*}||_1 A^{\bar{\tau}^*} \left(\sum_{i=0}^{\infty} \rho^i(\Xi\Gamma) A^i \bar{P} A^{\mathrm{T}^i} \right) A^{\mathrm{T}^{\bar{\tau}^*}} \\ &\quad + \underline{c}^* ||\tilde{\boldsymbol{\omega}}_{\bar{\tau}^*}||_1 \frac{1}{\rho^{\bar{\tau}^*-1}(\Gamma\Xi)} \sum_{i=\bar{\tau}^*-1}^{\infty} \rho^i(\Xi\Gamma) A^i Q A^{\mathrm{T}^i}. \end{aligned}$$

According to [35, Lemma 3], $\sum_{i=0}^{\infty} A^i X A^{\mathrm{T}^i}$ converges iff $\rho(A) < 1$. We conclude that the series in the LHS of (29) is bounded iff $\rho(\Xi\Gamma)\rho^2(A) < 1$. Similarly, the RHS in (29) is bounded under the same condition. Hence, the remote estimator is mean-square stable if and only if $\rho(\Xi\Gamma)\rho^2(A) < 1$. ■

Remark 3: In the case of a usage-dependent two-state Markov channel model [31], the stability condition was obtained using the stability theory of linear systems. An estimation-cycle-based method was proposed in [28] to show that condition (25) ensures the existence of scheduling policies that stabilize the remote estimator. Our technical line differs from [28] and [31] in analyzing limiting distribution of the states under the optimal transmission scheme.

V. Simulation Studies

In this section, numerical simulations, including a practical system, that is, involved in vehicle tracking in a 2-D space, are provided to illustrate our theoretical results. For calculating the optimal policy, the infinite state space $\mathbb{S}$ is truncated as a finite one $\bar{\mathbb{S}}$ by letting $\tau = \bar{\tau} = 9\ \forall\tau \geq 9$. Besides, the function $f(\cdot)$ in (1) is selected as $f(\cdot) = [1 - 2\mathcal{F}(\sqrt{\cdot})]^L$, where L is the length of the transmitted data packet and $\mathcal{F}(x) = (1/\sqrt{2\pi}) \int_x^{+\infty} e^{(-t^2/2)} dt$ [41]. We set the communication parameters to be: noise power $\sigma^2 = 8$ and packet length $L = 10$.

A. Performance and Threshold Structure of the Optimal Policy

In this section, an unstable system matrix is considered, that is, $\rho(A) > 1$. The system parameters are given by $A = 1.2, Q = 1, C = 1, R = 1$ with $\bar{P} = 0.6613$. We assume that there are ten channel states with $\mathbb{H} = \{0.2, 0.4, \ldots, 2\}$ and the transition matrix is given by $\Xi =$

$$\begin{bmatrix} 0.91 & 0.09 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.9 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.8 & 0.12 & 0.08 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.7 & 0.11 & 0.11 & 0.08 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.6 & 0.11 & 0.1 & 0.1 & 0.09 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0.18 & 0.1 & 0.08 & 0.07 & 0.07 & 0 & 0 & 0 & 0 \\ 0.4 & 0.28 & 0.1 & 0.06 & 0.05 & 0.06 & 0.05 & 0 & 0 & 0 \\ 0.3 & 0.1 & 0.1 & 0.1 & 0.16 & 0.14 & 0.09 & 0.01 & 0 & 0 \\ 0.2 & 0.11 & 0.1 & 0.1 & 0.1 & 0.05 & 0.09 & 0.13 & 0.12 & 0 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.1 \end{bmatrix}.$$

One can verify that Assumption 2 is satisfied. The following four scheduling policies are considered: the policy solved by value iteration algorithm P_{MDP}, two different greedy policies P_{Greedy}^1 and P_{Greedy}^2, and random policy P_{Random}. The last three policies are conducted as follows.

1) P_{Greedy}^1: $a(\tau, \mathfrak{h}) = 1$ if $\tau \geq 5$ and $\mathfrak{h} \geq \mathfrak{g}_5 = 1.0$. One can see that states in the set $\{s | a(s) = 1\}$ under policy P_{Greedy}^1 occupy the upper right quarter of the finite state region.
2) P_{Greedy}^2: $a(\tau, \mathfrak{h}) = 1$ if $\tau + i \geq 9$ with $0 \leq i \leq 9$, it can be seen that states in the set $\{s | a(s) = 1\}$ under policy P_{Greedy}^2 occupy the upper right half of the finite state region.
3) P_{Random}: $a(s) = 1$ with probability 0.5 $\forall s \in \bar{\mathbb{S}}$.

We run the above three policies 10 000 times and take the average values to evaluate their performance. The comparison results for different discount factors are reported in Table I

TABLE I
PERFORMANCE COMPARISONS FOR DIFFERENT POLICIES UNDER DIFFERENT DISCOUNT FACTORS

α \ Policy	P_{MDP}	P^1_{Greedy}	P^2_{Greedy}	P_{Random}
0.6	4.8287	17.2198	18.0691	29.1356
0.7	8.9531	18.7125	19.8089	30.8694
0.8	15.5494	24.3962	23.1164	33.5825
0.9	24.2600	40.0275	28.8287	39.3305

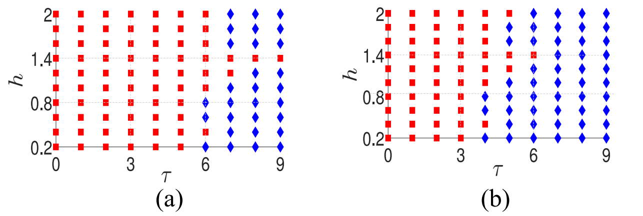


Fig. 2. Optimal policies without Assumption 2 under different parameters: Red squares and blue diamonds indicate actions $a^*(s) = 0$ and $a^*(s) = 1$, respectively. (a) $\alpha = 0.6$. (b) $\alpha = 0.9$.

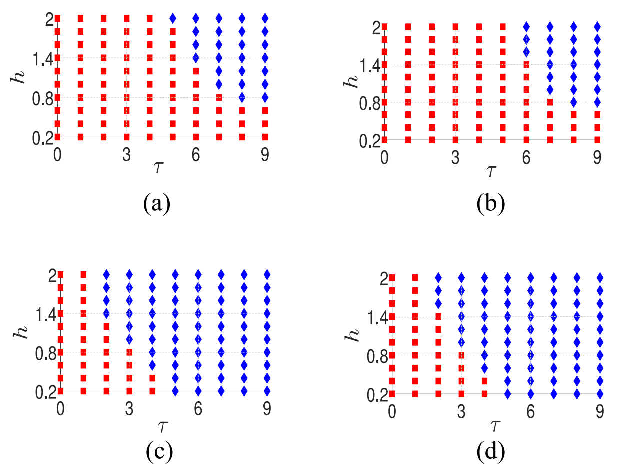


Fig. 3. Optimal policies satisfying Assumption 2 with different parameters. (a) $\zeta = 45, \alpha = 0.6$. (b) $\zeta = 50, \alpha = 0.6$. (c) $\zeta = 45, \alpha = 0.9$. (d) $\zeta = 50, \alpha = 0.9$.

after normalizing by $(1 - \alpha)$. The transmission power and weighting factor are set to be $\zeta = 50$ and $\kappa = 1$, respectively.

We can see that policy P_{MDP} outperforms the other policies and the two greedy policies are better than the random policy except that when $\alpha = 0.9$. It also can be found that P^1_{Greedy} is better than P^2_{Greedy} when $\alpha = 0.6, 0.7$ but worse than P^2_{Greedy} under $\alpha = 0.8, 0.9$ in Table I. This is because thresholds τ_j^* changes with the discount factor α. Specifically, these critical thresholds are also affected by the transition law between different states (the channel state transition probability and packet loss rate) and transmission cost. More details for this aspect can be seen in Fig. 3. Besides, as we expect that the total discounted cost for all policies are increasing in discount factor α.

Fig. 2 plots the monotonicity of the optimal scheduling policy as stated in Theorem 1, where $\zeta = 50$ and the transition matrix Ξ is renewedly selected such that Assumption 2 is not satisfied. We can see that a larger α could result $a^*(s) = 0$ changes to $a^*(s) = 1$ for some states $s \in \bar{\mathbb{S}}$. In Fig. 3, when

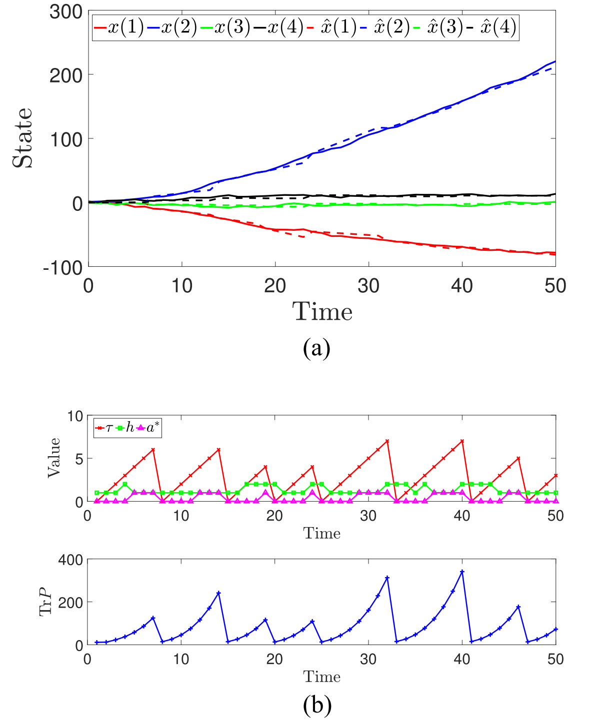


Fig. 4. Optimal scheduling policy with application in a practical system. (a) System states. (b) MDP states and the optimal actions.

Assumption 2 is satisfied with the given Ξ at the beginning of this section, the optimal actions exhibit double thresholds. It can be observed in Fig. 3(a) and (b) that the optimal policy $a^*(s)$ for some states change when the power ζ becomes larger due to the tradeoff between estimation error and power consumption. We shall note that more transmission power ζ means the decrease of the packet loss rate under increasing communication cost. Thus, it can be seen that for some states with high channel gain, $a^*(s) = 1$ changes to be $a^*(s) = 0$. Besides, one can see in Fig. 3(a) and (c) that the optimal policy $a^*(s) = 0$ becomes $a^*(s) = 1$ for some states under larger α.

B. Performance Evaluation by Practical System

In this section, we evaluate the performance of the optimal sensor scheduling policy in a practical system model, which is involved in vehicle tracking in a 2-D space. This example was studied in [44]. The system parameters are given by

$$A = \begin{bmatrix} 1 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0.5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

as well as $C = \begin{bmatrix} 1.2 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 \end{bmatrix}$ and $R = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix}$. This system matrix is critically stable since $\rho(A) = 1$. The channel gains are reset to be $\mathbb{H} = \{1, 1.5\}$ with $\Xi = \begin{bmatrix} 0.6 & 0.4 \\ 0.45 & 0.55 \end{bmatrix}$.

Fig. 4 shows the practical run within 50 steps. Fig. 4(a) depicts the real-time system state x and remote state estimation $\hat{x}$, where the 1st (3th) and 2nd (4th) components of the system state are position (velocity) in the x-axis and

y-axis, respectively. Fig. 4(b) plots the packet loss duration time τ, channel gain, the optimal decision a^*, and trace of the estimation error covariance matrix. It can be observed that the optimal scheduling policy performs well in the practical system model.

VI. Conclusion

In this article, we investigated the optimal sensor scheduling for the remote state estimation over an FSMC. Using the formulation of MDP, we showed the existence of an optimal DS policy and derived its double threshold structure. An NSC was obtained to guarantee the stability of the remote estimator.

It is interesting to extend our main results to: 1) other kinds of systems under network-induced factors, such as nonlinear or time-varying systems with time-delay or quantization effects, where how to guarantee the boundness of the estimation error is still the main consideration [38], [45]; 2) more general multisensor multichannel case, in which more efficient algorithms for large-scale systems are appealing [17], [18]; and 3) secure sensor scheduling [5], [16], [39], [46], where a more comprehensive security consideration is necessary for remote state estimation.

Acknowledgment

The authors would like to thank Dr. Wang Jiazheng in Hong Kong University of Science and Technology for his invaluable comments and suggestions of this work.

References

[1] D. Ding, Q.-L. Han, X. Ge, and J. Wang, "Secure state estimation and control of cyber–physical systems: A survey," *IEEE Trans. Syst. Man, Cybern., Syst.*, vol. 51, no. 1, pp. 176–190, Jan. 2021.

[2] K.-D. Kim and P. R. Kumar, "Cyber–physical systems: A perspective at the centennial," *Proc. IEEE*, vol. 100, pp. 1287–1308, May 2012.

[3] C. Qu, W. Chen, J. B. Song, and H. Li, "Distributed data traffic scheduling with awareness of dynamics state in cyber physical systems with application in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 2895–2905, Nov. 2015.

[4] M. M. Sandhu, S. Khalifa, R. Jurdak, and M. Portmann, "Task scheduling for energy-harvesting-based IoT: A survey and critical analysis," *IEEE Internet Things J.*, vol. 8, no. 18, pp. 13825–13848, Sep. 2021.

[5] L. Wang, X. Cao, H. Zhang, C. Sun, and W. X. Zheng, "Transmission scheduling for privacy-optimal encryption against eavesdropping attacks on remote state estimation," *Automatica*, vol. 137, Mar. 2022, Art. no. 110145.

[6] R. Gan, Y. Xiao, J. Shao, and J. Qin, "An analysis on optimal attack schedule based on channel hopping scheme in cyber–physical systems," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 994–1003, Feb. 2021.

[7] Y. Li, S. Zhu, C. Chen, and X. Guan, "Energy-efficient optimal sensor scheduling for state estimation over multihop sensor networks," *IEEE Trans. Cybern.*, early access, Jul. 14, 2021, doi: 10.1109/TCYB.2021.3088321.

[8] H. Zhang and W. X. Zheng, "Robust transmission power management for remote state estimation with wireless energy harvesting," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2682–2690, Aug. 2018.

[9] J. Zhang, J. Sun, and H. Lin, "Optimal DoS attack schedules on remote state estimation under multi-sensor round-robin protocol," *Automatica*, vol. 127, May 2021, Art. no. 109517.

[10] R. Gan, J. Shao, Y. Xiao, H. Zhang, and W. X. Zheng, "Optimizing attack schedules based on energy dispatch over two-hop relay networks," *IEEE Trans. Autom. Control*, vol. 65, no. 9, pp. 3832–3846, Sep. 2020.

[11] J. Qin, M. Li, J. Wang, L. Shi, Y. Kang, and W. X. Zheng, "Optimal denial-of-service attack energy management against state estimation over an SINR-based network," *Automatica*, vol. 119, Sep. 2020, Art. no. 109090.

[12] T.-Y. Zhang and D. Ye, "False data injection attacks with complete stealthiness in cyber–physical systems: A self-generated approach," *Automatica*, vol. 120, Oct. 2020, Art. no. 109117.

[13] L. Peng, X. Cao, and C. Sun, "Optimal transmit power allocation for an energy-harvesting sensor in wireless cyber–physical systems," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 779–788, Feb. 2021.

[14] S. Wu, X. Ren, Q.-S. Jia, K. H. Johansson, and L. Shi, "Learning optimal scheduling policy for remote state estimation under uncertain channel condition," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 2, pp. 579–591, Jun. 2020.

[15] H. Yuan, Y. Xia, M. Lin, H. Yang, and R. Gao, "Dynamic pricing-based resilient strategy design for cloud control system under jamming attack," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 1, pp. 111–122, Jan. 2020.

[16] L. Huang, K. Ding, A. S. Leong, D. E. Quevedo, and L. Shi, "Encryption scheduling for remote state estimation under an operation constraint," *Automatica*, vol. 127, May 2021, Art. no. 109537.

[17] G. Pang, W. Liu, Y. Li, and B. Vucetic, "DRL-based resource allocation in remote state estimation," 2022, *arXiv:2205.12267*.

[18] T. Farjam, H. Wymeersch, and T. Charalambous, "Distributed channel access for control over unknown memoryless communication channels," *IEEE Trans. Autom. Control*, early access, Nov. 22, 2021, doi: 10.1109/TAC.2021.3129737.

[19] H. Hmedi, J. Carroll, and A. Arapostathis, "Optimal sensor scheduling under intermittent observations subject to network dynamics," *IEEE Trans. Autom. Control*, early access, Feb. 14, 2022, doi: 10.1109/TAC.2022.3151578.

[20] C. O. Savage and B. F. La Scala, "Optimal scheduling of scalar Gauss–Markov systems with a terminal cost function," *IEEE Trans. Autom. Control*, vol. 54, no. 5, pp. 1100–1105, May 2009.

[21] M. M. Vasconcelos and U. Mitra, "Data-driven sensor scheduling for remote estimation in wireless networks," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 2, pp. 725–737, Jun. 2021.

[22] H. Liu, "SINR-based multi-channel power schedule under DoS attacks: A Stackelberg game approach with incomplete information," *Automatica*, vol. 100, pp. 274–280, Feb. 2019.

[23] S. Knorn, S. Dey, A. Ahlén, and D. E. Quevedo, "Optimal energy allocation in multisensor estimation over wireless channels using energy harvesting and sharing," *IEEE Trans. Autom. Control*, vol. 64, no. 10, pp. 4337–4344, Oct. 2019.

[24] L. An and G.-H. Yang, "Optimal transmission power scheduling of networked control systems via fuzzy adaptive dynamic programming," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 6, pp. 1629–1639, Jun. 2021.

[25] X. Ren, J. Wu, S. Dey, and L. Shi, "Attack allocation on remote state estimation in multi-systems: Structural results and asymptotic solution," *Automatica*, vol. 87, pp. 184–194, Jan. 2018.

[26] Y. Qi, P. Cheng, and J. Chen, "Optimal sensor data scheduling for remote estimation over a time-varying channel," *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4611–4617, Sep. 2017.

[27] J. Wei and D. Ye, "On two sensors scheduling for remote state estimation with a shared memory channel in a cyber–physical system environment," *IEEE Trans. Cybern.*, early access, Sep. 29, 2021, doi: 10.1109/TCYB.2021.3112677.

[28] W. Liu, D. E. Quevedo, Y. Li, K. H. Johansson, and B. Vucetic, "Remote state estimation with smart sensors over Markov fading channels," *IEEE Trans. Autom. Control*, vol. 67, no. 6, pp. 2743–2757, Jun. 2022.

[29] J. Chakravorty and A. Mahajan, "Remote estimation over a packet-drop channel with Markovian state," *IEEE Trans. Autom. Control*, vol. 65, no. 5, pp. 2016–2031, May 2020.

[30] A. S. Leong, A. Ramaswamy, D. E. Quevedo, H. Karl, and L. Shi, "Deep reinforcement learning for wireless sensor scheduling in cyber–physical systems," *Automatica*, vol. 113, Mar. 2020, Art. no. 108759.

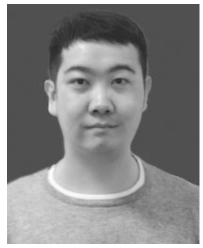
[31] J. Wang, X. Ren, S. Dey, and L. Shi, "Remote state estimation with usage-dependent Markovian packet losses," *Automatica*, vol. 123, Jan. 2021, Art. no. 109342.

[32] X. Ren, J. Wu, K. H. Johansson, G. Shi, and L. Shi, "Infinite horizon optimal transmission power control for remote state estimation over fading channels," *IEEE Trans. Autom. Control*, vol. 63, no. 1, pp. 85–100, Jan. 2018.

[33] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[34] A. R. Mesquita, J. P. Hespanha, and G. N. Nair, "Redundant data transmission in control/estimation over lossy networks," *Automatica*, vol. 48, pp. 1020–1027, Aug. 2012.

[35] Z. Ren, P. Cheng, J. Chen, L. Shi, and H. Zhang, "Dynamic sensor transmission power scheduling for remote state estimation," *Automatica*, vol. 50, no. 4, pp. 1235–1242, Apr. 2014.
[36] J. Wei and D. Ye, "Multi-sensor scheduling for remote state estimation over a temporally correlated channel," *IEEE Trans. Ind. Informat.*, early access, May 3, 2022, doi: 10.1109/TII.2022.3171612.
[37] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.
[38] L. Li, D. Yu, Y. Xia, and H. Yang, "Remote nonlinear state estimation with stochastic event-triggered sensor schedule," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 734–745, Mar. 2019.
[39] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal DoS attack scheduling in wireless networked control system," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 3, pp. 843–852, May 2016.
[40] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-state Markov modeling of fading channels—A survey of principles and applications," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 57–80, Sep. 2008.
[41] R. Poisel, *Modern Communications Jamming: Principles and Techniques*. Boston, MA, USA: Artech House, 2011.
[42] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York, NY, USA: Wiley, 2014.
[43] H. Minc, *Nonnegative Matrices*. New York, NY, USA: Wiley, 1988.
[44] W. Yang, Y. Zhang, C. Yang, Z. Zuo, and X. Wang, "Online power scheduling for distributed filtering over an energy-limited sensor network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, pp. 4216–4226, May 2018.
[45] L. Zou, Z. Wang, J. Hu, and H. Dong, "Ultimately bounded filtering subject to impulsive measurement outliers," *IEEE Trans. Autom. Control*, vol. 67, no. 1, pp. 304–319, Jan. 2022.
[46] Y. Li, L. Shi, P. Cheng, J. Chen, and D. E. Quevedo, "Jamming attacks on remote state estimation in cyber–physical systems: A game-theoretic approach," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2831–2836, Oct. 2015.



Jiang Wei received the B.S. degree in automation from the Inner Mongolia University of Technology, Hohhot, China, in 2017, and the M.S. degree in control engineering from Northeastern University, Shenyang, China, in 2019, where he is currently pursuing the Ph.D. degree in control science and engineering with the College of Information Science and Engineering.

His current research interests include remote state estimation and resource allocation in cyber–physical systems.

Dan Ye (Senior Member, IEEE) received the B.S. and M.S. degrees in mathematics and applied mathematics from Northeast Normal University, Changchun, China, in 2001 and 2004, respectively, and the Ph.D. degree in control theory and engineering from Northeastern University, Shenyang, China, in 2008.

She was a Lecturer with Northeastern University from 2008 to 2010. She is currently a Professor with the College of Information Science and Engineering, Northeastern University. Her research interests include fault-tolerant control, robust control, adaptive control, and security of cyber–physical systems.