Stability Conditions for Remote State Estimation of Multiple Systems over Multiple Markov Fading Channels

Wanchun Liu, *Member, IEEE*, Daniel E. Quevedo, *Fellow, IEEE*, Karl Henrik Johansson, *Fellow, IEEE*, Branka Vucetic, *Fellow, IEEE*, Yonghui Li, *Fellow, IEEE*

***Abstract*— We investigate the stability conditions for remote state estimation of multiple linear time-invariant (LTI) systems over multiple wireless time-varying communication channels. We answer the following open problem: what is the fundamental requirement on the multi-sensor-multi-channel system to guarantee the existence of a sensor scheduling policy that can stabilize the remote estimation system? We propose a novel policy construction and analytical framework and derive the necessary-and-sufficient stability condition in terms of the LTI system parameters and the channel statistics.**

***Index Terms*—Estimation, Kalman filtering, linear systems, stability, mean-square error, Markov fading channel**

I. Introduction

A. Motivation

Industry 4.0, also known as the Fourth Industrial Revolution, is the automation of traditional manufacturing and industrial processes through customized and flexible mass production [1]. Replacing communication cables with wireless devices in conventional factories will be a game-changer: In particular, for automatic control, Industry 4.0 will make use of large-scale, interconnected deployment of massive spatially distributed industrial devices such as sensors, actuators, and controllers. Given their low-cost and scalable deployment, wireless remote state estimation from ubiquitous sensors will play a key role in many industrial control applications, such as smart manufacturing, industrial automation, e-commerce warehouses, and smart grids [2].

However, unlike wired communications, wireless communications are often unreliable and have a limited spectrum for transmission [3]. Consequently, when wireless sensors are deployed in a remote estimation system, scheduling policies need to be designed to allow sensors to update the measurement data over a limited number of frequency channels. The design of such transmission schedules is especially challenging since, due to variability of the environment, wireless channels are time-varying. The transition process of associated channel fading states are commonly modeled as a Markov processes [4], wherein different channel states lead to different packet drop probabilities. In particular, industrial environments are frequently cluttered with moving machines, robots, and vehicles, and the Markov fading channel model is suitable to characterize the time-correlations of the channel states. Experiments in different industrial environments, including Holmen's paper mill in Sweden, have verified the accuracy of the Markov channel modeling [5]. Due to transmission scheduling and packet dropouts, a multi-sensor remote estimator cannot correctly receive all the sensor measurements. This degrades the estimation performance and can even lead to instability, i.e., the expected estimation error covariances may become unbounded.

W. Liu, B. Vucetic, and Y. Li are with School of Electrical and Information Engineering, The University of Sydney, Australia. Emails: {wanchun.liu, branka.vucetic, yonghui.li}@sydney.edu.au. D. E. Quevedo is with the School of Electrical Engineering and Robotics, Queensland University of Technology (QUT), Brisbane, Australia. Email: dquevedo@ieee.org. K. H. Johansson is with School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. Email: kallej@kth.se.

In this note, **we tackle the fundamental problem: what are necessary and sufficient conditions on system parameters that ensure stochastic stability of a multi-sensor remote estimator under multiple Markov fading channels?**

B. Related Works

Existing work related to multi-system remote estimation and control can be divided into two categories: perfect (wired) and imperfect (wireless) communication channels.

Perfect communication channels. Early research in stability analysis of multi-control-loop transmission scheduling over single and multiple independent communication channels involved periodic and aperiodic scheduling policies [6]–[10], assuming perfect communication channels. These works have only determined the sufficient conditions to guarantee the existence of a scheduling policy that can stabilize the networked systems.

Imperfect communication channels. In practice, wireless channels are not error-free, leading to transmission errors and packet dropouts. In industrial control environments comprising moving machines and mobile robots, the channel quality is time-varying [11]. Unlike single systems for which remote estimation and control that have been well investigated (e.g., [12], [13] for independent and identically distributed packet dropout scenarios and [14] for Markovian packet dropout), multi-system remote estimation and control over wireless channels have not drawn as much research attention, until recent efforts motivated by standardization and growing deployment of wireless technology.

Considering a wireless control architecture with multiple control loops over shared wireless fading channels, optimal dynamic transmission scheduling policies were investigated in [15], [16] with different objective functions. The scheduling decision in each step depends on both the wireless channel and control loop states. In [17] and [18], sensor transmission scheduling over single and multiple packet drop channels of remote estimators were investigated, respectively, and some sufficient mean-square stability conditions in terms of the system parameters and optimal transmission scheduling policies were analyzed. If the stability conditions hold, classic Markov decision process (MDP) methods were adopted for finding the optimal scheduling policies (see e.g., [19]–[21]). The follow-up work [22] considered a time-correlated Markov fading channel scenario, and derived a sufficient condition to guarantee the existence of a deterministic and stationary scheduling policy that can stabilize the remote estimator. A deep reinforcement learning method was proposed as well to find the optimal scheduling policy. The approach was further applied for solving a transmission scheduling problem of a fully distributed networked control system in [23].

We also note that the conventional Markov jump linear systems theory [24] can be used to elucidate stability conditions of networked systems using a given transmission policy. However, MJLS theory does not provide much insights about the existence of a dynamic scheduling policy that can stabilize the networked system over shared communication channels.

C. Contributions

In this note, we consider a multi-sensor remote estimator with multiple frequency channels, where individual sensors measure different physical processes. We allow the channels to be time-varying and correlated in both frequency and time, as is common in practical applications [3]. We derive a necessary and sufficient mean-square stability condition in terms of the physical process parameters and the fading channel statistics. Our result establishes that there exists at least one sensor scheduling policy over the frequency channels that can stabilize the remote estimator if and only if the condition holds. The stability condition will provide practical design guidelines for stabilizing multi-sensor remote estimation systems over shared wireless medium in Industry 4.0.

It is worth emphasizing that the stability condition depends on the essential parameters of the physical processes of interest and the communication channels, rather than on the specific scheduling policy employed. In fact, to find merely sufficient conditions for stability, one can construct a specific policy and analyze its properties (e.g., [22], [25]). Such a sufficient condition, however, is commonly not tight and thus cannot be proved to be necessary. The analysis of necessary conditions is challenging as there exists a combinational number of scheduling policies.

To the best of our knowledge, necessary and sufficient stability conditions have not been established before for remote estimation or control over multiple systems and wireless communication channels. In the present work, we consider a general correlated channel, wherein the channel state in each frequency channel is a random variable and correlated with the other channel states. To tackle the challenge, we develop a novel policy construction method and a stochastic estimation-cycle based analytical approach. We develop an asymptotic theory for the spectral radius of a product of non-negative matrices to prove our key result.

Notations: Sets are denoted by calligraphic capital letters, e.g., $\mathcal{A}$. $\mathcal{A}\backslash\mathcal{B}$ denotes set subtraction. Matrices and vectors are denoted by capital and lowercase upright bold letters, e.g., $\mathbf{A}$ and $\mathbf{a}$, respectively. $|\mathcal{A}|$ denotes the cardinality of the set $\mathcal{A}$. $\mathsf{E}[A]$ is the expectation of the random variable A. $(\cdot)^\top$ is the matrix transpose operator. $\|\mathbf{v}\|_1$ is the sum of the vector $\mathbf{v}$'s elements. $|\mathbf{v}| \triangleq \sqrt{\mathbf{v}^\top\mathbf{v}}$ is the Euclidean norm of a vector $\mathbf{v}$. $\mathrm{Tr}(\cdot)$ is the trace operator. $\mathrm{diag}\{\mathbf{v}\}$ denotes the diagonal matrix with the diagonal elements taken from $\mathbf{v}$. $\mathbb{N}$ and $\mathbb{N}_0$ denote the sets of positive and non-negative integers, respectively. $\mathbb{R}^m$ denotes the m-dimensional Euclidean space. $[u]_{B\times B}$ denotes the $B\times B$ matrix with identical elements u. $[\mathbf{A}]_{j,k}$ denotes the element at the jth row and kth column of a matrix $\mathbf{A}$. $\{v\}_{\mathbb{N}_0}$ denotes the semi-infinite sequence $\{v_0, v_1, \dots\}$. $\rho(\mathbf{A})$ is the spectral radius of $\mathbf{A}$, i.e., the largest absolute value of its eigenvalues. $\vartheta(\mathbf{A})$ is the sum of all the elements of $\mathbf{A}$.

II. System Model

We consider a remote estimator with N sensors measuring N independent physical process, as illustrated in Fig. 1. The index set of the sensors is denoted as $\mathcal{N} \triangleq \{1, 2, \dots, N\}$. The sensors are connected with a local gateway, which collects the measurements and forwards them to a remote estimator. Connections between sensors and the gateway are reliable and not scheduled, while the gateway to remote estimator communications are wireless and scheduled due to bandwidth limitations. We note that the typical connection density in the Industrial IoT scenario is 10^6/km^2. Bandwidth sharing among a large number of wireless devices is an issue in practice [26].

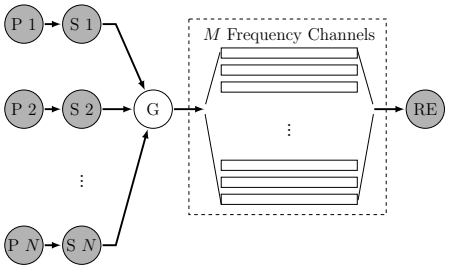

Fig. 1. The multi-sensor-multi-channel remote estimator with a single gateway. Processes, sensors, gateway and remote estimator are denoted as P1, P2, ..., PN, S1, S2, ..., SN, G and RE, respectively.

The discrete-time linear time-invariant (LTI) model of the measurement of each process n is given as [13], [27]

$$\begin{aligned} \mathbf{x}_n(t+1) &= \mathbf{A}_n\mathbf{x}_n(t) + \mathbf{w}_n(t), \\ \mathbf{y}_n(t) &= \mathbf{C}_n\mathbf{x}_n(t) + \mathbf{z}_n(t), \end{aligned} \tag{1}$$

where $\mathbf{x}_n(t) \in \mathbb{R}^{l_n}$ is the process state vector, $\mathbf{A}_n \in \mathbb{R}^{l_n\times l_n}$ the state transition matrix, $\mathbf{y}_n(t) \in \mathbb{R}^{r_n}$ the measurement vector of the sensor attached to the process, $\mathbf{C}_n \in \mathbb{R}^{r_n\times l_n}$ the measurement matrix, $\mathbf{w}_n(t) \in \mathbb{R}^{l_n}$ and $\mathbf{z}_n(t) \in \mathbb{R}^{r_n}$ are the process and measurement noise vectors, respectively. We assume $\mathbf{w}_n(t)$ and $\mathbf{z}_n(t)$ are independent and identically distributed (i.i.d.) zero-mean Gaussian processes with corresponding covariance matrices $\mathbf{W}_n$ and $\mathbf{Z}_n$, respectively. The discrete time slot has the duration of τ_0, which is also the sampling period of the original continuous processes. We note that a shorter sampling period provides more accurate tracking of the original process, but requires larger throughput for data transmission. Without loss of generality, we assume that $\rho^2(\mathbf{A}_1) \geq \rho^2(\mathbf{A}_2) \geq \cdots \geq \rho^2(\mathbf{A}_N)$.

A. Local Estimation

Each sensor adopts a local Kalman filter (KF) to estimate its process before sending to the gateway [14], [27]. We have

$$\begin{aligned} \mathbf{x}_n^s(t|t-1) &= \mathbf{A}_n\mathbf{x}_n^s(t-1) \\ \mathbf{P}_n^s(t|t-1) &= \mathbf{A}_n\mathbf{P}_n^s(t-1)\mathbf{A}_n^\top + \mathbf{W}_n \\ \mathbf{K}_n(t) &= \mathbf{P}_n^s(t|t-1)\mathbf{C}_n^\top(\mathbf{C}_n\mathbf{P}_n^s(t|t-1)\mathbf{C}_n^\top + \mathbf{Z}_n)^{-1} \\ \mathbf{x}_n^s(t) &= \mathbf{x}_n^s(t|t-1) + \mathbf{K}_n(t)(\mathbf{y}_n(t) - \mathbf{C}_n\mathbf{x}_n^s(t|t-1)) \\ \mathbf{P}_n^s(t) &= (\mathbf{I}_n - \mathbf{K}_n(t)\mathbf{C}_n)\mathbf{P}_n^s(t|t-1) \end{aligned}$$

where $\mathbf{I}_n$ is the $l_n \times l_n$ identity matrix, $\mathbf{x}_n^s(t|t-1)$ is the prior state estimate, $\mathbf{x}_n^s(t)$ is the posterior state estimate at time t, $\mathbf{K}_n(t)$ is the Kalman gain. The matrices $\mathbf{P}_n^s(t|t-1)$ and $\mathbf{P}_n^s(t)$ represent the prior and posterior error covariance at the sensor at time t, respectively. The first two equations above present the prediction steps while the last three equations correspond to the updating steps. In particular, $\mathbf{x}_n^s(t)$ is the sensor n's estimate of $\mathbf{x}_n(t)$ at time t, i.e., the pre-filtered measurement of $\mathbf{y}_t$, with the estimation error covariance $\mathbf{P}_n^s(t)$ defined as:

$$\mathbf{P}_n^s(t) \triangleq \mathsf{E}\left[(\mathbf{x}_n^s(t) - \mathbf{x}_n(t))(\mathbf{x}_n^s(t) - \mathbf{x}_n(t))^\top\right].$$

We focus on the *remote* estimation stability and assume that the *local* KFs are stable and operate in steady state [14], [27], i.e., $\mathbf{P}_n^s(t) = \bar{\mathbf{P}}_n, \forall t \in \mathbb{N}, n \in \mathcal{N}$.

B. Markov Channel

We assume that there exist only $M < N$ frequency channels that can be used for transmission of sensor data. The channels are correlated in both time and frequency domains as detailed below.

The M-channel (vector) state $\mathbf{h}(t)$ is modeled as an aperiodic Markov chain with $\bar{M}$ irreducible channel states, $\mathcal{S} \triangleq \{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{\bar{M}}\}$, where $\mathbf{h}_i \triangleq [h_{i,1}, h_{i,2}, \ldots, h_{i,M}]^\top \in \{0,1\}^M$. Here, $h_{i,j} = 0$ or 1 means successful (on) or failed (off) transmission in the jth frequency channel at the ith channel state. Let $\mathbf{M} \in \mathbb{R}^{\bar{M}} \times \mathbb{R}^{\bar{M}}$ denote the state transition probability matrix, where

$$[\mathbf{M}]_{i,j} \triangleq \mathsf{Prob}\left[\mathbf{h}(t) = \mathbf{h}_j | \mathbf{h}(t-1) = \mathbf{h}_i\right] = p_{i,j}.$$

Let $\tilde{\mathcal{S}}_i \in \mathcal{S}$ denote the set of channel state with an 'on' state in the ith frequency channel, i.e.,

$$\tilde{\mathcal{S}}_i \triangleq \{\mathbf{h}_j : h_{j,i} = 1, j \in \bar{\mathcal{M}}\}, i \in \mathcal{M},$$

where $\mathcal{M} \triangleq \{1, 2, \ldots, M\}$ and $\bar{\mathcal{M}} \triangleq \{1, 2, \ldots, \bar{M}\}$.

We make the following assumption on the availability of the channel state.

Assumption 1 (Known Previous Channel State [22])**.** *At time $t \in \mathbb{N}$, the current channel state is unknown but the previous channel state $\mathbf{h}(t-1)$ is available.*

C. Transmission Scheduling and Remote Estimation

In each time slot, the gateway collects N packets carrying the sensor estimates $\{\hat{\mathbf{x}}_1^s(t), \ldots, \hat{\mathbf{x}}_N^s(t)\}$. It schedules at most M of the packets and sends them through M frequency channels to the remote estimator, as illustrated in Fig. 1. In practice, one could adopt a multiplexing scheme such as the orthogonal frequency-division multiplexing (OFDM) for transmitting multi-stream data in parallel. Each frequency channel can transmit at most one packet. The unscheduled packets are discarded. The communication protocol for gateway-to-remote-estimator transmission is user datagram protocol (UDP) [12], which is widely adopted for real-time communications. We make the following assumption on transmission redundancy.

Assumption 2 (Disabled Redundant Transmissions)**.** *In each time slot, each packet can take at most one frequency channel for transmission.*

Due to the transmission scheduling and the fading channels, packets carrying the estimated N process states may or may not arrive at the remote estimator. Let $\gamma_n(t) = 1$ denote successful detection of sensor n's packet at time t, $n \in \mathcal{N}$. If $\gamma_n(t) = 0$, the packet is not scheduled or is scheduled but with failed detection. It is also assumed that each packet transmission has a unit delay that is equal to the sampling period of the system. The optimal remote estimator in the sense of minimum mean-square error (MMSE) is obtained as [13]

$$\hat{\mathbf{x}}_n(t) = \begin{cases} \mathbf{A}_n \hat{\mathbf{x}}_n(t-1), & \gamma_n(t-1) = 0, \\ \mathbf{A}_n \hat{\mathbf{x}}_n^s(t-1), & \gamma_n(t-1) = 1, \end{cases}$$

and can be simplified as [14]

$$\hat{\mathbf{x}}_n(t) = \mathbf{A}_n^{\phi_n(t)} \hat{\mathbf{x}}_n^s(t - (\phi_n(t))), \tag{3}$$

where $\phi_n(t) \in \mathbb{N}$ is the time duration between the previous successful transmission and the current time t, and can be regarded as the *age-of-information* (AoI).

From the above it follows that the estimation error covariance of process n is given as

$$\mathbf{P}_n(t) \triangleq \mathsf{E}\left[(\hat{\mathbf{x}}_n(t) - \mathbf{x}_n(t))(\hat{\mathbf{x}}_n(t) - \mathbf{x}_n(t))^\top\right] \tag{4}$$

$$= \zeta^{\phi_n(t)}(\bar{\mathbf{P}}_n) \tag{5}$$

where (5) is obtained by substituting (3) and (1) into (4) and

$$\zeta_n(\mathbf{X}) \triangleq \mathbf{A}_n \mathbf{X} \mathbf{A}_n^\top + \mathbf{W}_n$$

$$\zeta_n^1(\cdot) \triangleq \zeta_n(\cdot), \quad \zeta_n^{m+1}(\cdot) \triangleq \zeta_n(\zeta_n^m(\cdot)), \quad m \geq 1.$$

Thus, the quality of the remote estimation error of process n in time slot t can be quantified via $\mathrm{Tr}\left(\mathbf{P}_n(t)\right)$. For ease of exposition, we introduce the following function

$$c_n(i) \triangleq \mathrm{Tr}\left(\zeta_n^i(\bar{\mathbf{P}}_n)\right), \forall i \in \mathbb{N}$$

and note that, using (5), we can write

$$\mathrm{Tr}\left(\mathbf{P}_n(t)\right) \triangleq c_n(\phi_n(t)).$$

Therefore, the estimation quality of process n at time t is a function of its AoI state $\phi_n(t)$.

Propositions 1 and 2 of [14] allow us to state the following property of $c_n(\cdot)$:

Lemma 1. *For any $\epsilon > 0$, there exists $N' > 0$, $\kappa > 0$ and $\eta > 0$ such that*

$$c_n(i) < \kappa\left(\rho^2(\mathbf{A}_n) + \epsilon\right)^i, \forall i > N',$$

and

$$c_n(i) \geq \eta(\rho(\mathbf{A}_n))^{2i}, \forall i > N'.$$

In this work, we solely focus on deterministic stationary scheduling policies. Let $\nu_n(t) \in \{0, 1, \ldots, M\}$, $n \in \mathcal{N}$, denote the selected frequency channel for process n at time t. The sensor is not scheduled for transmission if $\nu_n(t) = 0$. Since available system states include the current AoI states and the previous channel states, scheduling polices $\pi(\cdot)$ can be written as

$$\boldsymbol{\nu}(t) = \pi(\boldsymbol{\phi}(t), \mathbf{h}(t-1)), \tag{6}$$

where $\boldsymbol{\phi}(t) \triangleq [\phi_1(t), \ldots, \phi_N(t)] \in \mathbb{N}^N$ and $\boldsymbol{\nu}(t) = [\nu_1(t), \ldots, \nu_N(t)]$.

Note that $\mathbf{P}_n(t)$ is a countable stochastic process taking values from a countably infinite set

$$\{\zeta_n^1(\bar{\mathbf{P}}_n), \zeta_n^2(\bar{\mathbf{P}}_n), \ldots\}.$$

If $\rho(\mathbf{A}_n) \geq 1$, then this process will grow during periods of consecutive packet dropouts. Since, due to fading, periods of consecutive packet dropouts have unbounded support, at best one can hope for some type of stochastic stability. Our focus is on mean-square stability.

Definition 1 (Average Mean-Square Stability)**.** *The N-sensor remote estimator over M frequency channels described above is average mean-square stabilizable, if there exists a deterministic stationary policy* (6) *such that the sum average estimation mean-square error (MSE) J is bounded, where*

$$J \triangleq \sum_{n=1}^{N} J_n \tag{7}$$

and

$$J_n = \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \mathsf{E}\left[\mathrm{Tr}\left(\mathbf{P}_n(t)\right)\right], n \in \mathcal{N}.$$

III. Key Results

In this section, we present and prove the necessary and sufficient condition for stabilizing the multi-sensor-multi-channel remote estimator in terms of the multi-process parameters and the multi-channel statistics. As will become apparent establishing such a necessary and sufficient stability condition is highly non-trivial as we consider transmission scheduling for multiple sensor packets over multiple fading channels that are correlated in both time and frequency domains.

A. Stability Condition

The physical process and channel parameters jointly determine the stability of the overall remote estimator. Our result is stated in terms of the $\bar{M} \times \bar{M}$ probability matrix $\mathbf{E}(\mathbf{v})$ obtained from the channel state transition probability matrix $\mathbf{M}$:

$$\begin{aligned}[\mathbf{E}(\mathbf{v})]_{i,j} &\triangleq \mathsf{Prob}\left[\mathbf{h}(t)=\mathbf{h}_j, \mathbf{h}_j \notin \tilde{\mathcal{S}}_{v_i} | \mathbf{h}(t-1)=\mathbf{h}_i\right] \\ &= p_{i,j}\mathbb{1}(\mathbf{h}_j \notin \tilde{\mathcal{S}}_{v_i}),\end{aligned}$$

where $\mathbf{v}$ is a $\bar{M}$-length vector and $v_i \in \mathcal{M}$ is the ith element of $\mathbf{v}$ denoting the index of selected frequency channel with the observation of the vector channel state $\mathbf{h}_i$. So $[\mathbf{E}(\mathbf{v})]_{i,j}$ denotes the probability that the current channel state is $\mathbf{h}_j$ and the packet transmission fails given the previously observed channel state $\mathbf{h}_i$ and the selected frequency channel for transmission v_i.

Theorem 1. *Consider $\rho_{\max} \triangleq \rho(\mathbf{A}_1)$, and*

$$\lambda_\infty \triangleq \liminf_{L\to\infty} \lambda_L = \min_{L\in\mathbb{N}} \lambda_L, \tag{8}$$

$$\lambda_L \triangleq \min_{\mathbf{v}_l \in \mathcal{M}^{\bar{M}}} \rho\left(\mathbf{E}(\mathbf{v}_1)\mathbf{E}(\mathbf{v}_2)\cdots\mathbf{E}(\mathbf{v}_L)\right)^{\frac{1}{L}} \tag{9}$$

where $\mathbf{v}_l \triangleq [v_{l,1}, v_{l,2}, \ldots, v_{l,\bar{M}}]^\top \in \mathcal{M}^{\bar{M}}$ is a vector of frequency channel selection at the $\bar{M}$ different channel conditions.

The remote estimator described in Section II under Assumptions 1 and 2 is stable if and only if

$$\rho_{\max}^2 \lambda_\infty < 1. \tag{10}$$

Theorem 1 shows that the stability depends on the spectral radius of the most unstable process and a complex function of the channel state transition probability matrix. Provided the condition is satisfied, one can always find a scheduling policy that stabilizes the remote estimator; if (10) does not hold, then no stabilizing scheduling policy exists. Theorem 1 does not provide direct insights on the structure of a suitable scheduling policy. However, we will construct a policy with stability guarantees in the proof of the sufficiency part.

We note that λ_∞ can be treated as an inverse quality indicator of the parallel correlated channels: smaller values of λ_∞ indicate a better channel quality. The infinity in λ_∞ takes in to account the infinitely many different orders of matrix products in (9). The last equality in (8) can be easily obtained by using the property that $\rho(\mathbf{D}^L)^{\frac{1}{L}} = \rho(\mathbf{D})$ holds for any square matrix $\mathbf{D}$ and positive integer L.

Remark 1 (Computations)**.** *Although λ_∞ in* (8) *may not be achieved with a finite L, one can approximate λ_∞ by finding the minimum value in $\{\lambda_1, \ldots, \lambda_L\}$ in* (9)*. Since $\lambda_\infty \leq \min\{\lambda_1, \ldots, \lambda_L\}, \forall L \in \mathbb{N}$, we have the sufficient stability condition*

$$\rho_{\max}^2 \min\{\lambda_1, \ldots, \lambda_L\} < 1,$$

which approximates the necessary and sufficient condition (10) *when L is large. Numerical results of the sequence $\{\lambda_1, \ldots, \lambda_L\}$ are included in [28].*

To the best of our knowledge, Theorem 1 is the first necessary and sufficient stability condition for remote estimation over a multi-sensor-multi-channel network in the literature. Before proving our result, we will first establish a relationship to existing results by focusing on a special case. Under the idealized assumption that the Markov channels at different frequencies are independent, a sufficient condition for stability was proved in [22]. It corresponds to Corollary 1 below.

Corollary 1. *Consider the special case that Markov channels at different frequencies are independent, and assume that the $\{0,1\}$ channel state transition probability matrix of frequency channel m is*

$$\mathbf{M}^{(m)} \triangleq \begin{bmatrix} \alpha_{00}^{(m)} & \alpha_{01}^{(m)} \\ \alpha_{10}^{(m)} & \alpha_{11}^{(m)} \end{bmatrix}, m \in \mathcal{M}. \tag{11}$$

The remote estimator described in Section II under Assumptions 1 and 2 is stable if [22]

$$\rho_{\max}^2 \alpha_{00}^{(m^\star)} < 1, \tag{12}$$

where $\alpha_{00}^{(m^\star)} \triangleq \min_m \alpha_{00}^{(m)}$.

The sufficient stability condition (12) *is more restrictive than Theorem 1.*

Proof. See [28] □

Example 1. We numerically compare the sufficient stability condition in [22] and the necessary and sufficient condition in Theorem 1 for a three-plant-two-channel scenario with $\rho_{\max}^2 = 2$. The plant system matrices are

$$\mathbf{A}_1 = \begin{bmatrix} 1.1 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 1.2 & 0.2 \\ 0.2 & 0.9 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 1.3 & 0.2 \\ 0.2 & 1.05 \end{bmatrix},$$

and the measurement matrices and the covariance matrices are equal to the identity matrix, i.e., $\mathbf{C}_n = \mathbf{W}_n = \mathbf{Z}_n = \mathbf{I}$. We consider different channel state transition matrices (11) for the second frequency channel with parameters: (a) $\alpha_{00}^{(2)} = 0.9, \alpha_{11}^{(2)} = 0.9$, (b) $\alpha_{00}^{(2)} = 0.9, \alpha_{11}^{(2)} = 0.1$, (c) $\alpha_{00}^{(2)} = 0.6, \alpha_{11}^{(2)} = 0.9$, and (d) $\alpha_{00}^{(2)} = 0.6, \alpha_{11}^{(2)} = 0.1$. A larger $\alpha_{00}^{(2)}$ and $\alpha_{11}^{(2)}$ lead to a longer channel state memory in the 'off' and the 'on' state, respectively. Then the stability regions in terms of the parameters of first frequency channel state transition matrices, $\alpha_{00}^{(1)}$ and $\alpha_{11}^{(1)}$, are shown in Fig. 2.

It is clear that Theorem 1 has a larger stability region than [22] in case (c), which corresponds to that the second frequency channel has a shorter memory in the bad ('off') state and a longer memory in the good ('on') state. Hence when the quality of the second frequency channel is pretty good, the stability requirement on the first channel based on [22] is more restrictive than Theorem 1. □

We will prove the necessary and sufficiency parts of Theorem 1 in the sequel.

B. Proof of Necessity

1) Policy Construction: To prove the necessity, we consider the scenario that only the estimate of the process with the largest spectral radius is scheduled for transmission in each time slot in

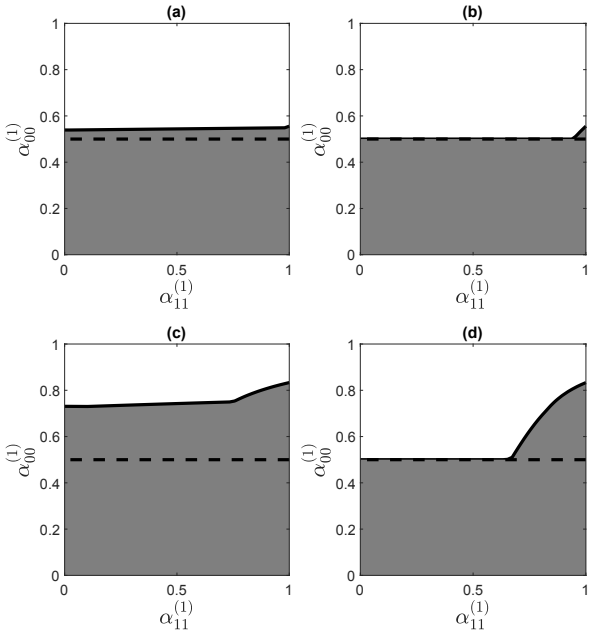

Fig. 2. Comparison between stability conditions of Theorem 1 (solid line) and [22] (dashed line). The gray area denotes the parameter sets that satisfy the stability conditions.

a selected frequency channel, while the other sensors' estimates are perfectly known by the remote estimator and do not need any transmission. In other words, only one packet is scheduled in each time slot. We recall that process 1 has the largest spectral radius, and we drop out the process index n in the following analysis.

The channel scheduling policy (6) is reduced to

$$\nu(t) = \pi(\phi(t), \mathbf{h}(t-1)) \in \mathcal{M}, \tag{13}$$

where $\nu(t)$ and $\phi(t)$ denote $\nu_1(t)$ and $\phi_1(t)$, respectively, for ease of notation.

From (13), once the AoI ϕ is given, the channel selection rule given the previous channel state information can be written as

$$\mathbf{v}(\phi) = [v_1(\phi), \ldots, v_{\bar{M}}(\phi)] \in \mathcal{M}^{\bar{M}},$$

where $v_i(\phi) = \pi(\phi, \mathbf{h}_i)$.

Given the channel selection vector $\mathbf{v}(\phi)$, we define the successful transmission probability matrix $\tilde{\mathbf{E}}(\mathbf{v}(\phi)) \in \mathbb{R}^{\bar{M}} \times \mathbb{R}^{\bar{M}}$, where

$$\begin{aligned}[\tilde{\mathbf{E}}(\mathbf{v}(\phi))]_{i,j} &\triangleq \mathsf{Prob}\left[\mathbf{h}(t) = \mathbf{h}_j, \mathbf{h}_j \in \tilde{\mathcal{S}}_{v_i(\phi)} | \mathbf{h}(t-1) = \mathbf{h}_i\right] \\ &= p_{i,j} \mathbb{1}(\mathbf{h}_j \in \tilde{\mathcal{S}}_{v_i(\phi)}).\end{aligned}$$

In other words, $[\tilde{\mathbf{E}}(\mathbf{v}(\phi))]_{i,j}$ is the probability that the current channel state is $\mathbf{h}_j$ and the transmission is successful in the selected $v_i(\phi)$th frequency channel given that the previous channel state is $\mathbf{h}_i$. Accordingly, we define the failed transmission probability matrix

$$\mathbf{E}(\mathbf{v}(\phi)) \triangleq \mathbf{M} - \tilde{\mathbf{E}}(\mathbf{v}(\phi)) \in \mathbb{R}^{\bar{M}} \times \mathbb{R}^{\bar{M}}.$$

2) Analysis of the Average Cost: Similar to [14], we consider an estimation cycle based analysis method. Each estimation cycle starts after a successful transmission and ends at the next one, and thus the AoI state ϕ at the beginning of each estimation cycle is equal to 1. T_k is the sum of transmissions in the kth estimation cycle. C_k is the sum MSE in the kth estimation cycle and is a function of T_k as

$$C_k = g(T_k) \triangleq \sum_{j=1}^{T_k} c(j). \tag{14}$$

The channel state before the kth cycle is denoted as $\mathbf{b}_k \in \mathcal{S}$, and a successful transmission occurs at $\mathbf{b}_k$. Similar to Lemma 1 in [14], we have the following Markovian property of the pre-cycle channel states.

Lemma 2. *$\{\mathbf{b}\}_{\mathbb{N}_0}$ is a time-homogeneous ergodic Markov chain with $\bar{M}_1 \leq \bar{M}$ irreducible states of $\mathcal{S}$. The state transition matrix of $\{\mathbf{b}\}_{\mathbb{N}_0}$ is $\mathbf{G}'$, which is the $\bar{M}_1$-by-$\bar{M}_1$ matrix taken from the top-left corner of*

$$\mathbf{G} = \sum_{j=1}^{\infty} \tilde{\mathbf{\Xi}}(j),$$

where

$$\tilde{\mathbf{\Xi}}(j) = \mathbf{\Xi}(j-1)\tilde{\mathbf{E}}(\mathbf{v}(j)), j = 1, 2, \ldots. \tag{15}$$

and

$$\mathbf{\Xi}(j) = \begin{cases} \mathbf{I}, & j = 0 \\ \prod_{i=1}^{j} \mathbf{E}(\mathbf{v}(i)), & j > 0. \end{cases}$$

The stationary distribution of $\{\mathbf{b}\}_{\mathbb{N}_0}$ is $\boldsymbol{\beta} \triangleq [\beta_1, \ldots, \beta_{\bar{M}_1}]^\top$, which is the unique null-space vector of $(\mathbf{I} - \mathbf{G}')^\top$ and $\beta_i > 0, \forall i \in \bar{\mathcal{M}}_1$, where $\bar{\mathcal{M}}_1 \triangleq \{1, 2, \ldots, \bar{M}_1\}$.

Remark 2. *Only the first $\bar{M}_1$ channel states can be a pre-cycle state, and thus the last $(\bar{M} - \bar{M}_1)$ columns of $\mathbf{G}$ are all zeros.*

From (7) it follows that the average estimation MSE can be rewritten as

$$J = \limsup_{K\to\infty} \frac{\frac{1}{K}(C_1 + C_2 + \cdots + C_K)}{\frac{1}{K}(T_1 + T_2 + \cdots + T_K)} = \frac{\mathsf{E}[C]}{\mathsf{E}[T]},$$

where the last equality holds because the distributions of T_k and C_k depend on $\mathbf{b}_k$, which is ergodic, and hence the unconditional distributions of T_k and C_k are also ergodic; time averages are equal to the ensemble averages and we drop the time indexes. Then, we have

$$\mathsf{E}[T] = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} T_k = \sum_{m=1}^{\bar{M}} \beta_m \mathsf{E}[T | \mathbf{b} = \mathbf{h}_m], \tag{16}$$

and

$$\mathsf{E}[C] = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} C_k = \sum_{m=1}^{\bar{M}} \beta_m \mathsf{E}[C | \mathbf{b} = \mathbf{h}_m], \tag{17}$$

where β_m is defined in Lemma 2 when $m \in \bar{\mathcal{M}}_1$, and $\beta_m = 0$ when $m \in \bar{\mathcal{M}}_0 \triangleq \bar{\mathcal{M}} \backslash \bar{\mathcal{M}}_1 = \{\bar{M}_1 + 1, \ldots, \bar{M}\}$.

From the definition of estimation cycle and the property of channel state transition, the conditional probability of the length of an estimation cycle is obtained as

$$\mathsf{Prob}[T = i | \mathbf{b} = \mathbf{h}_m] = \vartheta\left(\mathbf{L}_m \tilde{\mathbf{\Xi}}(i)\right) \triangleq \sum_{k=1}^{\bar{M}} \left[\tilde{\mathbf{\Xi}}(i)\right]_{m,k}, \forall m \in \bar{\mathcal{M}} \tag{18}$$

where $\mathbf{L}_m$ is an all-zero matrix except for the mth diagonal element, which equals to 1.

Taking (18) into (16) and into (17), then after some algebraic manipulations, one can obtain

$$\mathsf{E}[T] = \sum_{m=1}^{\bar{M}} \beta_m \left(\sum_{i=1}^{\infty} i\vartheta\left(\mathbf{L}_m \tilde{\mathbf{\Xi}}(i)\right)\right),$$
$$\mathsf{E}[C] = \sum_{m=1}^{\bar{M}} \beta_m \left(\sum_{i=1}^{\infty} g(i)\vartheta\left(\mathbf{L}_m \tilde{\mathbf{\Xi}}(i)\right)\right). \tag{19}$$

From the definition of $g(i)$ in (14) and the property of $c(i)$ in Lemma 1, $g(i)$ grows exponentially fast with i. It is easy to verify the property below.

Lemma 3. *$J < \infty$ if and only if $\mathsf{E}[C] < \infty$.*

From Lemma 3, it follows that it suffices to only investigate conditions such that $\mathsf{E}[C] < \infty$ in the following.

3) Proof of Necessity: We define a set of channel selection vectors $\mathcal{F} \subset \mathcal{M}^{\bar{M}}$, where for any $\mathbf{v} \in \mathcal{F}$ we have $\tilde{\mathbf{E}}(\mathbf{v}) = \mathbf{0}$ leading to zero chance of successful transmission in any of the frequency channels. Thus, for any $\mathbf{v} \in \tilde{\mathcal{F}} \triangleq \mathcal{M}^{\bar{M}} \backslash \mathcal{F}$, we have $\tilde{\mathbf{E}}(\mathbf{v}) \neq \mathbf{0}$. It is clear that if $\tilde{\mathcal{F}} = \emptyset$, the packet dropout occurs all the time at all frequency channels. From the definition (9), we can prove that $\lambda_\infty = 1$ as $\mathbf{E}(\mathbf{v})$ is a stochastic matrix for any $\mathbf{v} \in \mathcal{M}^{\bar{M}}$. Thus, the necessary condition of stability is straightforward as $\rho^2(\mathbf{A}) < 1$, which can be written as $\rho^2(\mathbf{A})\lambda_\infty < 1$. In what follows, we will focus on the scenario with $\tilde{\mathcal{F}} \neq \emptyset$.

We categorize all possible scheduling policies into two types:

Definition 2 (Type-I and II Policies)**.** *For a type-I policy, there exists $\bar{\phi} > 0$ such that*

$$\mathbf{v}(\phi) \in \mathcal{F}, \forall \phi > \bar{\phi},$$

For a type-II policy, if $\mathbf{v}(\phi) \in \tilde{\mathcal{F}}$, one can always find $\phi' > \phi$ such that $\mathbf{v}(\phi') \in \tilde{\mathcal{F}}$.

A type-I policy has a strictly zero chance of successful transmission when the AoI is larger than a threshold, while a type-II policy still has a non-zero success probability when the AoI is arbitrarily large. Thus, to stabilize the system, a type-I policy should guarantee successful transmission within the first $\bar{\phi}$ attempts, while it is not necessary for a type-II policy.

We first prove the necessary condition for type I and then type II policies. Before proceeding further, we define the following

$$\lambda_{m,\infty} \triangleq \liminf_{L\to\infty} \lambda_{m,L}, m \in \bar{\mathcal{M}}, \tag{20}$$

and

$$\lambda_{m,L} \triangleq \min_{\mathbf{v}_l \in \mathcal{M}^{\bar{M}}} \rho\left(\mathbf{L}_m \mathbf{E}(\mathbf{v}_1)\mathbf{E}(\mathbf{v}_2)\cdots\mathbf{E}(\mathbf{v}_L)\right)^{\frac{1}{L}}.$$

Type-I Policy. Since the scheduling policy consistently chooses channel selection vectors from $\mathcal{F}$ in the high AoI scenario that leads to zero chance of success, it is clear that a necessary condition to stabilize the system is that the transmission process has a zero probability to fail for consecutive $\bar{\phi}$ times at the beginning of an estimation cycle. This can be written as

$$\max_{m\in\bar{\mathcal{M}}} \vartheta(\mathbf{L}_m \mathbf{E}(\mathbf{v}(1))\mathbf{E}(\mathbf{v}(2))\cdots\mathbf{E}(\mathbf{v}(\bar{\phi}))) = \mathbf{0},$$

and hence $\max_{m\in\bar{\mathcal{M}}} \lambda_{m,\infty} = 0$. In the following, we focus on the type-II policy.

Type-II Policy. We categorize the channel states $\mathbf{h}$ in two cases: (i) the pre-cycle states $\mathbf{h} = \mathbf{h}_m$ with $\beta_m > 0$, i.e., $m \in \bar{\mathcal{M}}_1$, and (ii) the non-pre-cycle states $\mathbf{h} = \mathbf{h}_m$ with $\beta_m = 0$, i.e., $m \in \bar{\mathcal{M}}_0$. In other words, an estimation cycle can and cannot start after a case (i) and a case (ii) channel state, respectively. Due to the ergodicity of the channel states, both the cases of channel states occur with non-zero probabilities. Then, we analyze the necessary conditions to (i) make the average sum MSE of an estimation cycle in (19) bounded that starts after a pre-cycle state

$$\sum_{j=1}^{\infty} g(j)\vartheta\left(\mathbf{L}_m \tilde{\mathbf{\Xi}}(j)\right) < \infty \tag{21}$$

and to (ii) make the average sum MSE of an estimation cycle bounded that contains a non-pre-cycle state.

(i) Assume that the channel state $\mathbf{h}_m$ is a pre-cycle state. Using Perron–Frobenius Theorem [29], we have

$$\vartheta\left(\mathbf{L}_m \mathbf{\Xi}(i)\right) \geq \rho\left(\mathbf{L}_m \mathbf{\Xi}(i)\right), \tag{22}$$

thus, there exists an element in $\mathbf{L}_m\mathbf{\Xi}(i)$, e.g., the m'th element of the mth row, such that

$$\left[\mathbf{L}_m \mathbf{\Xi}(i)\right]_{m,m'} \geq \frac{1}{\bar{M}} \rho\left(\mathbf{L}_m \mathbf{\Xi}(i)\right). \tag{23}$$

From (20), there is a constant $\kappa_0 > 0$ such that $\rho\left(\mathbf{L}_m \mathbf{\Xi}(i)\right) \geq \kappa_0 (\lambda_{m,\infty})^i$. Let $\mathbf{F}(i) \in \mathbb{R}^{\bar{M}\times\bar{M}}$ denote a degraded matrix of $\mathbf{L}_m\mathbf{\Xi}(i)$, where $[\mathbf{F}(i)]_{m,m'} = [\mathbf{L}_m\mathbf{\Xi}(i)]_{m,m'}$ and the other elements of $\mathbf{F}(i)$ are zeros. From (23), it is clear that

$$\vartheta\left(\mathbf{L}_m \mathbf{\Xi}(i)\right) \geq \vartheta\left(\mathbf{F}(i)\right) \geq \frac{\kappa_0}{\bar{M}} \left(\lambda_{m,\infty}\right)^i. \tag{24}$$

Based on the property of type-II policy, we define an infinite sequence of AoI variables $\{\tilde{\phi}_1, \tilde{\phi}_2, \dots\}$, where $\tilde{\phi}_i < \tilde{\phi}_j$ if $i < j$, and $\mathbf{v}(\phi) \in \tilde{\mathcal{F}}$ if and only if $\phi \in \{\tilde{\phi}\}_{\mathbb{N}}$. Also, we define the operator $\lceil x \rceil_{\mathcal{A}}$ as the smallest value in the set $\mathcal{A}$ that is no smaller than x. Building on $\{\tilde{\phi}_1, \tilde{\phi}_2, \dots\}$ and a constant $L \in \mathbb{N}$, we construct a sequence of AoI as $\{\varphi_1, \varphi_2, \dots\}$, where

$$\varphi_i = \begin{cases} \lceil L \rceil_{\{\tilde{\phi}\}_{\mathbb{N}}}, & i = 1 \\ \lceil \varphi_{i-1} + L \rceil_{\{\tilde{\phi}\}_{\mathbb{N}}}, & i > 1, \end{cases} \tag{25}$$

and thus $\mathbf{v}(\varphi_i) \in \tilde{\mathcal{F}},\ \varphi_{i+1} - \varphi_i \geq L, \forall i \in \mathbb{N}$.

Then, we introduce the following technical lemma.

Lemma 4. *There is a constant $L > 0$ such that the transmission is successful with at least a non-zero probability ξ within the next L steps for any channel selection vector at the Lth step $\mathbf{v}_L \in \tilde{\mathcal{F}}$, no matter what the current channel state is and the first $(L-1)$-step channel selection vectors are, i.e.,*

$$\xi \triangleq \min_{\mathbf{v}_1,\mathbf{v}_2,\dots,\mathbf{v}_{L-1}\in\mathcal{M}^{\bar{M}},\mathbf{v}_L\in\tilde{\mathcal{F}}} \min_{m\in\bar{\mathcal{M}}} \max_{l=1,\dots,L} \vartheta(\mathbf{L}_m \mathbf{E}(\mathbf{v}_1)\mathbf{E}(\mathbf{v}_2)\cdots\mathbf{E}(\mathbf{v}_{l-1})\tilde{\mathbf{E}}(\mathbf{v}_l)) > 0,$$

where $\mathbf{E}(\mathbf{v}_0) \triangleq \mathbf{I}$.

Proof. See [28]. □

From Lemma 4, we can find a constant L to construct the AoI sequence $\{\varphi\}_{\mathbb{N}}$ in (25). Then, using Lemma 1 and (24), we have

$$\begin{aligned}
&\sum_{j=\varphi_i-L}^{\varphi_i} g(j)\vartheta\left(\mathbf{L}_m\tilde{\mathbf{\Xi}}(j)\right) \geq g(\varphi_i)\vartheta\left(\mathbf{L}_m\tilde{\mathbf{\Xi}}(\varphi_i)\right)\\
&\geq g(\varphi_i-L)\vartheta\Big(\mathbf{L}_m\mathbf{\Xi}(\varphi_i-L)\mathbf{E}(\mathbf{v}(\varphi_i-L+1))\\
&\mathbf{E}(\mathbf{v}(\varphi_i-L+2))\cdots\mathbf{E}(\mathbf{v}(\varphi_i-1))\tilde{\mathbf{E}}(\mathbf{v}(\varphi_i))\Big)\\
&\geq g(\varphi_i-L)\vartheta\left(\mathbf{L}_m\mathbf{\Xi}(\varphi_i-L)\right)\xi\\
&\geq \frac{\kappa_0\xi}{\bar{M}}(\rho(\mathbf{A})^2)^{\varphi_i-L}\left(\lambda_{m,\infty}\right)^{\varphi_i-L}.
\end{aligned}$$

Now the average sum MSE per cycle in (21) is

$$\begin{aligned}
&\sum_{j=1}^{\infty} g(j)\vartheta\left(\mathbf{L}_m\tilde{\mathbf{\Xi}}(j)\right) \geq \sum_{i=1}^{\infty}\sum_{j=\varphi_i-L}^{\varphi_i} g(j)\vartheta\left(\mathbf{L}_m\tilde{\mathbf{\Xi}}(j)\right)\\
&\geq \sum_{i=1}^{\infty}\frac{\kappa_0\xi}{\bar{M}}(\rho(\mathbf{A})^2)^{\varphi_i-L}\left(\lambda_{m,\infty}\right)^{\varphi_i-L}.
\end{aligned} \tag{26}$$

To make the last sum in (26) bounded, we must have $\lim_{i\to\infty}(\rho(\mathbf{A})^2)^{\varphi_i}\left(\lambda_{m,\infty}\right)^{\varphi_i} = 0$, i.e., $(\rho(\mathbf{A})^2)\lambda_{m,\infty} < 1$. Thus, by considering all the pre-cycle channel states, $\max_{m\in\bar{\mathcal{M}}_1}\rho(\mathbf{A})^2\lambda_{m,\infty} < 1$ holds if $\mathsf{E}[C]$ in (19) is bounded.

(ii) Assume now that the channel state $\mathbf{h}_{m'}$ is a non-pre-cycle state. Since the channel state transition process is an ergodic Markov chain, given a state $\mathbf{h}_m$ with $m \in \bar{\mathcal{M}}_1$, it will take finite steps to arrive at $\mathbf{h}_{m'}, m' \in \bar{\mathcal{M}}_0$ with a positive probability no matter what the channel scheduling policy is. Mathematically, there exists $L' \geq 1$ such that

$$\tilde{\beta}_{m,m'} \triangleq [\mathbf{L}_m\mathbf{\Xi}(L')]_{m,m'} > 0.$$

Then, it is straightforward to have

$$\vartheta(\mathbf{L}_m\mathbf{\Xi}(i)) \geq \tilde{\beta}_{m,m'}\vartheta(\mathbf{L}_{m'}\mathbf{E}(\mathbf{v}_{L'+1})\mathbf{E}(\mathbf{v}_{L'+2})\cdots\mathbf{E}(\mathbf{v}_i)), \forall i \geq L'. \tag{27}$$

We see that the right-hand side of (27) has a similar format to the left-hand side of (22) in case (i). Then by taking (27) into (21), following similar steps as in case (i) and considering all the non-pre-cycle channel states, we obtain that $\max_{m'\in\bar{\mathcal{M}}_0}\rho(\mathbf{A})^2\lambda_{m',\infty} < 1$ holds if $\mathsf{E}[C]$ in (19) is bounded.

From cases (i) and (ii), a necessary condition to bound $\mathsf{E}[C]$ can be uniformly written as

$$\max_{m\in\bar{\mathcal{M}}}\rho(\mathbf{A})^2\lambda_{m,\infty} < 1.$$

From the technical lemma below, it is easy to prove that $\max_{m\in\bar{\mathcal{M}}}\lambda_{m,\infty} = \lambda_\infty$. Then, the necessary condition of type-I and II policies can be jointly written as $\rho(\mathbf{A})^2\lambda_\infty < 1$, which completes the proof of the necessity of Theorem 1.

Lemma 5. *Given a sequence of N-by-N matrices $\{\mathbf{M}_1, \mathbf{M}_2, \dots\}$, the following equation holds*

$$\max_{i\in\mathcal{N}}\liminf_{L\to\infty}\rho\left(\mathbf{L}_i\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_L\right)^{\frac{1}{L}} = \liminf_{L\to\infty}\rho\left(\mathbf{M}_1\mathbf{M}_2\cdots\mathbf{M}_L\right)^{\frac{1}{L}},$$

where $\mathbf{L}_i \in \mathbb{R}^{N\times N}$ has the ith diagonal element equal to 1 and all other elements equal to zero.

Proof. See [28]. □

C. Proof of Sufficiency

1) Policy Construction: We consider a *persistent serial scheduling* policy that persistently schedules the transmission of one sensor at a time until it is successful and then schedules the next sensor and so on. Although it seems that such a policy cannot take advantage of the parallel channels, we will show that the policy stabilizes the remote estimator if condition (10) holds. The policy can be written as

$$\pi(\boldsymbol{\phi}(t), \mathbf{h}(t-1)) = \begin{cases} \pi_1(\phi_1(t), \mathbf{h}(t-1)), & \text{if } \phi_n(t)=1, n=N\\ \pi_{n+1}(\phi_{n+1}(t), \mathbf{h}(t-1)), & \text{if } \phi_n(t)=1, n<N\\ \pi(\boldsymbol{\phi}(t-1), \mathbf{h}(t-2)), & \text{else}\end{cases}$$

where the policy

$$\pi_n(\phi_n(t), \mathbf{h}(t-1)) = [\underbrace{0,\dots,0}_{n-1}, \nu_n(t), \underbrace{0,\dots,0}_{N-n}], n \in \mathcal{N} \tag{28}$$

denotes scheduling the nth sensor on the $\nu_n(t)$th frequency channel. Without loss of generality, the initial scheduling policy is given by $\pi(\boldsymbol{\phi}(1), \mathbf{h}(0)) = \pi_1(\phi_1(1), \mathbf{h}(0))$. With a slight abuse of notation, we drop out the zeros in (28) so that the frequency channel selection rule for sensor n is rewritten as

$$\pi_n(\phi_n(t), \mathbf{h}(t-1)) = \nu_n(t) \in \mathcal{M}, n \in \mathcal{N}.$$

Once the AoI of sensor n, ϕ_n, is given, the channel selection rule with different previous channel states is denoted as

$$\mathbf{v}_n(\phi_n) \triangleq [\pi_n(\phi_n, \mathbf{h}_1), \pi_n(\phi_n, \mathbf{h}_2), \dots, \pi_n(\phi_n, \mathbf{h}_{\bar{M}})] \in \mathcal{M}^{\bar{M}}.$$

Moreover, we assume that the frequency channel selection rule $\mathbf{v}_n(\phi_n), \forall n \in \mathcal{N}$, is a *periodic policy* in terms of ϕ_n with L potential channel selection vectors taken from the set $\mathcal{V} \triangleq \{\tilde{\mathbf{v}}_n(1), \tilde{\mathbf{v}}_n(2), \dots, \tilde{\mathbf{v}}_n(L)\}$, where $\tilde{\mathbf{v}}_n(L) \in \mathcal{M}^{\bar{M}}$. From the definition of λ_∞ in (8), for an arbitrarily small $\epsilon > 0$, we can find a constant L and a L-length channel-selection-vector set $\mathcal{V}$ satisfying the condition that

$$\rho\left(\mathbf{E}(\tilde{\mathbf{v}}_n(1))\mathbf{E}(\tilde{\mathbf{v}}_n(2))\cdots\mathbf{E}(\tilde{\mathbf{v}}_n(L))\right)^{\frac{1}{L}} \leq \lambda_\infty + \epsilon.$$

The periodic channel selection policy of sensor n is defined as

$$\mathbf{v}_n(\phi_n) = \tilde{\mathbf{v}}_n(\phi_n \bmod L),$$

where $(a \bmod b)$ denotes the remainder of the Euclidean division of a by b if the remainder is non-zero, otherwise $(a \bmod b) = b$.

2) Analysis of the Average Cost: To analyze the average cost of the average estimation MSE of the nth process J_n, the estimation cycle starts after a successful transmission of process n and ends at the following one. $T_{n,k}$ and $C_{n,k}$ denote the sum of transmissions and the sum MSE of the kth estimation cycle:

$$C_{n,k} = g(T_{n,k}) \triangleq \sum_{j=1}^{T_{n,k}} c_n(j), \tag{29}$$

$$T_{n,k} = \sum_{i=1}^{N} T_{(n,i),k},$$

where $T_{(n,i),k}$ is the time duration scheduled for the sensor i's transmission during the kth estimation cycle of process n.

Similar to the single-sensor case, it can be proved that the time average of $\{C_n\}_{\mathbb{N}_0}$ is equal to its ensemble average, and J_n is bounded if $\mathsf{E}[C_n]$ is. In the following, we drop out the time index

k and will analyze $\mathsf{E}[C_n]$. For the ease of notation but without loss of generality, we analyze $\mathsf{E}[C_N]$ for the Nth process:

$$\mathsf{E}[C_N]=\sum_{t_1=1}^{\infty}\cdots\sum_{t_{N-1}=1}^{\infty}\mathsf{E}\left[C_N|T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{N-1}\right]\times\mathsf{Prob}\left[T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{N-1}\right].$$

Let $\mathbf{b}_N$ denote the channel state before an estimation cycle of process N. From (18), it is easy to obtain the conditional probability

$$\begin{aligned}&\mathsf{Prob}\left[T_{(N,N)}=t_N|T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{n-1}\right]\\&=\vartheta(\mathrm{diag}\{[\varsigma_1,\ldots,\varsigma_{\bar{M}}]\}\tilde{\boldsymbol{\Xi}}(t_N)),\end{aligned}\tag{30}$$

where

$$\varsigma_m\triangleq\mathsf{Prob}\left[\mathbf{b}_N=\mathbf{h}_m|T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{n-1}\right],m\in\bar{\mathcal{M}}.$$

Recall that $\tilde{\boldsymbol{\Xi}}(\cdot)$ was defined in (15).

Thus, from (29) and (30), the conditional expectation is simplified as

$$\begin{aligned}&\mathsf{E}\left[C_N|T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{N-1}\right]\\&=\sum_{t_N=1}^{\infty}g\left(\sum_{i=1}^{N}t_i\right)\\&\times\mathsf{Prob}\left[T_{(N,N)}=t_N|T_{(N,1)}=t_1,\ldots,T_{(N,N-1)}=t_{n-1}\right]\\&=\sum_{t_N=1}^{\infty}g\left(\sum_{i=1}^{N}t_i\right)\vartheta(\mathrm{diag}\{[\varsigma_1,\ldots,\varsigma_{\bar{M}}]\}\tilde{\boldsymbol{\Xi}}(t_N)).\end{aligned}\tag{31}$$

3) Proof of Sufficiency: Using Lemma 1 and the conditional expectation (31), we can obtain an upper bound of $\mathsf{E}[C_N]$ and derive the sufficient condition making the average cost bounded. The detailed proof is included in [28] due to the space limitation.

Remark 3. *The policy above with persistent sensor scheduling and periodic channel selection is a stability-guaranteeing policy. Note that such a policy, which does not utilize the parallel frequency channels, is only constructed for the proof of the sufficiency of Theorem 1, and is not optimal. Once the stability condition is satisfied, we can find the optimal policies in* (6)*, e.g., by designing suitable MDP problems, see for example [22].*

IV. Conclusions

We have derived the necessary and sufficient stability condition of the multi-plant remote estimation system over Markov fading channels, which is more effective than existing sufficient conditions available in the literature. For future work, we will consider stability analysis of a multi-control-loop system over shared wireless channels. In addition, it is also worth extending to practical nonlinear systems, e.g., building on the literature of fuzzy systems (see e.g., [30], [31] and the references therein).

References

[1] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing Letters*, vol. 3, pp. 18 – 23, 2015.

[2] P. Park, S. Coleri Ergen, C. Fischione, C. Lu, and K. H. Johansson, "Wireless network design for control systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 2, pp. 978–1013, Second Quarter 2018.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[4] 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.

[5] T. Olofsson, A. Ahlén, and M. Gidlund, "Modeling of the fading statistics of wireless sensor network channels in industrial environments," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3021–3034, Jun. 2016.

[6] D. Hristu-Varsakelis, "Feedback control systems as users of a shared network: communication sequences that guarantee stability," in *Proc. IEEE CDC*, 2001, pp. 3631–3636.

[7] H. Rehbinder and M. Sanfridson, "Scheduling of a limited communication channel for optimal control," in *Proc. IEEE CDC*, 2000, pp. 1011–1016.

[8] L. Zhang and D. Hristu-Varsakelis, "Communication and control co-design for networked control systems," *Automatica*, vol. 42, no. 6, pp. 953–958, 2006.

[9] G. C. Walsh, Hong Ye, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 3, pp. 438–446, 2002.

[10] T. Iwaki, J. Wu, Y. Wu, H. Sandberg, and K. H. Johansson, "Multi-hop sensor network scheduling for optimal remote estimation," *Automatica*, vol. 127, p. 109498, 2021.

[11] A. Ahlen, J. Åkerberg, M. Eriksson, A. J. Isaksson, T. Iwaki, K. H. Johansson, S. Knorn, T. Lindh, and H. Sandberg, "Toward wireless control in industrial process automation: A case study at a paper mill," *IEEE Control Syst. Mag.*, vol. 39, no. 5, pp. 36–57, 2019.

[12] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, Jan 2007.

[13] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.

[14] 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.

[15] K. Gatsis, M. Pajic, A. Ribeiro, and G. J. Pappas, "Opportunistic control over shared wireless channels," *IEEE Trans. Autom. Control*, vol. 60, no. 12, pp. 3140–3155, Dec. 2015.

[16] M. Eisen, M. M. Rashid, K. Gatsis, D. Cavalcanti, N. Himayat, and A. Ribeiro, "Control aware radio resource allocation in low latency wireless control systems," *IEEE Internet Things J.*, vol. 6, no. 5, pp. 7878–7890, 2019.

[17] A. S. Leong, S. Dey, and D. E. Quevedo, "Sensor scheduling in variance based event triggered estimation with packet drops," *IEEE Trans. Autom. Control*, vol. 62, no. 4, pp. 1880–1895, 2017.

[18] 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, 2018.

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

[20] 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.

[21] 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.

[22] 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, p. 108759, 2020.

[23] W. Liu, K. Huang, D. E. Quevedo, B. Vucetic, and Y. Li, "Deep reinforcement learning for wireless scheduling in distributed networked control," *submitted to Automatica*, 2021. [Online]. Available: https://arxiv.org/pdf/2109.12562.pdf

[24] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov jump linear systems*. Springer Science & Business Media, 2006.

[25] 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, 2018.

[26] P. Valerio, "Industry 4.0 requires massive iot and seamless connectivity," *IoT Times*, 2021. [Online]. Available: https://iot.eetimes.com/industry-4-0-requires-massive-iot-and-seamless-connectivity

[27] L. Shi and L. Xie, "Optimal sensor power scheduling for state estimation of Gauss-Markov systems over a packet-dropping network," *IEEE Trans. Signal Process.*, vol. 60, no. 5, pp. 2701–2705, May 2012.

[28] 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. [Online]. Available: https://arxiv.org/abs/2104.04181

[29] H. Minc, *Nonnegative Matrices*. Wiley, 1988.

[30] X. Liu, Z. Gao, and A. Zhang, "Observer-based fault estimation and tolerant control for stochastic takagi–sugeno fuzzy systems with brownian parameter perturbations," *Automatica*, vol. 102, pp. 137–149, 2019.

[31] M. Naghdi and M. A. Sadrnia, "A novel fuzzy extended state observer," *ISA Transactions*, vol. 102, pp. 1–11, 2020.