Power Scheduling of Kalman Filtering in Wireless Sensor Networks with Data Packet Drops

Gang Wang, Student Member, IEEE, Jie Chen, Senior Member, IEEE, Jian Sun, Member, IEEE, and Yongjian Cai

Abstract—For a wireless sensor network (WSN) with a large number of low-cost, battery-driven, multiple transmission power leveled sensor nodes of limited transmission bandwidth, then conservation of transmission resources (power and bandwidth) is of paramount importance. Towards this end, this paper considers the problem of power scheduling of Kalman filtering for general linear stochastic systems subject to data packet drops (over a packet-dropping wireless network). The transmission of the acquired measurement from the sensor to the remote estimator is realized by sequentially transmitting every single component of the measurement to the remote estimator in one time period. The sensor node decides separately whether to use a high or low transmission power to communicate every component to the estimator across a packet-dropping wireless network based on the rule that promotes the power scheduling with the least impact on the estimator mean squared error. Under the customary assumption that the predicted density is (approximately) Gaussian, leveraging the statistical distribution of sensor data, the mechanism of power scheduling, the wireless network effect and the received data, the minimum mean squared error estimator is derived. By investigating the statistical convergence properties of the estimation error covariance, we establish, for general linear systems, both the sufficient condition and the necessary condition guaranteeing the stability of the estimator.

Index Terms—Power scheduling, Kalman filtering, data packet drops, wireless sensor networks, linear stochastic systems, stability.

I. INTRODUCTION

With the groundbreaking advances of microsensor technology and wireless communication technology, wireless sensor networks (WSNs) have been found in a plethora of applications. The proposed and/or already deployed applications include, but not limited to, battlefield surveillance, intelligent transportation systems, health care, environment monitoring and control, disaster prevention and recovery, and more efficient electric power grids [1]–[7]. However, there are still some severe limitations in current WSNs that prevent them from better serving the people, such as, limited power at each battery-driven sensor, limited communication ability, limited computation ability and limited wireless bandwidth [7]. These limitations will ineluctably bring some challenging problems to the study of estimation and control over WSNs. Therefore, it is of great significance to investigate how to conserve transmission power and bandwidth while achieving a similar estimation performance.

Towards this end, recurring attention has been paid to the research of remote estimation under communication resources (energy constraint and bandwidth) requirement in the last decade and a multitude of publications can be widely found in the literature; see, for example, [2]–[24] and references therein. Among them, by the desire of conserving transmission energy and bandwidth, various methods regarding measurement quantization, censoring, and dimensionality-reduction were specialized in [2]–[6], [8]–[12]. Another creative method in terms of measurement scheduling has been extensively studied in [13]–[15], [17]–[22] etc.

To be more specific, owing to the power-limited nature of wireless sensors and the fact that replacing the exhausted batteries are costly operations and may even be impossible, only a limited number of measurement transmissions can thereby be made by the wireless devices in most WSNs applications. In [13], optimal measurement scheduling policies were devised for a particular class of scalar Gauss-Markov systems to minimize the terminal estimation error variance over a given time horizon $T$, in which only $p < T$ measurements can be taken and transmitted to the remote estimator side. In practical, most commercially available sensor nodes nowadays have multiple transmission power levels [11] and it is assumed that high transmission power leads to reliable data flow while low transmission power may cause unreliable data flow [14] and therefore data packet drops may occur. The results in [13] were then recently extended to a special class of high-order Gauss-Markov systems in [15], where both the sensor energy constraint and data packet drops were taken into account and furthermore, two scenarios in terms of sensor nodes with limited or sufficient computation capacity are considered. Under some appropriate conditions, the optimal schedulers derived indicate that the $p$ measurement transmissions should be distributed along the last $p$ time steps over the time horizon $T$, that is, from $T - 1 - p$ to $T - 1$. It is worth noticing that the optimal measurement schedulers above are deterministic, which are so-called “offline schedulers,” and therefore, this kind of offline schedulers have the apparent advantage of offline determination of optimal scheduling schemes. Nevertheless, also noticed that the estimation error covariance...
matrix increases drastically for unstable systems in the first $T - p$ time steps owing to no measurements transmitted to the estimator side to update the covariance prediction, which is a disadvantage of these offline schedulers. More discussions and generalizations on offline schedulers can also be found in [14] and [22].

On the other hand, to avoid the disadvantage mentioned above, schedulers taking the current measurement value into consideration were devised in [17]–[21] and considering the modified Kalman filter therein is very much involved with a stochastic variable, these schedulers are called “online schedulers.” The send-on-delta strategy was adopted in [17] to reduce sensor data traffic by transmitting sensor data only if their values change exceeds a prescribed threshold. However, the threshold has no analytic relationship with the estimation performance and no stability and performance analysis were given with respect to the proposed modified Kalman filter. Innovation-based measurement schedulers were primarily constructed in [18], [19] by quantifying the “importance” of every measurement using the normalized measurement innovations. The main idea is that only “important” enough measurements will be transmitted to the estimator side to update the state prediction and covariance prediction, and when the transmission does not occur, the additionally known information based on given threshold of the scheduler will be utilized. Moreover, some stability analysis of Kalman filtering with the aforementioned two stochastic schedulers was presented in [18] and however, only necessary conditions guaranteeing the convergence of expected estimation error covariance were established for systems with full-row-ranked observation matrix therein.

Inspired by those observations, this paper builds on and considerably broads the scope of [15] and [18], where the power scheduler is dependent on the time-horizon $T$ and the covariance increases drastically during the first $T - p - 1$ time steps. In comparison, the main contributions of this work are twofold and summarized as follows.

1) We consider power scheduling problem of remote state estimation of general high-order linear stochastic systems. Data packet drop, a typical and natural phenomenon in wireless networks, is also considered and modeled as one Bernoulli i.i.d. process. See Fig. 1 for an illustration, where the power scheduler is embedded in the sensor node. We devise a component-wise innovation-based power scheduler and the corresponding minimum mean squared error estimator (MMSE).

2) We investigate the statistical convergence properties of the estimation error covariance matrix by constructing one auxiliary function and we establish both the sufficient and necessary condition for convergence of the averaged estimation error covariance. Theorem 1 originally establishes the sufficient condition for mean square stability of estimation error covariance matrix and Theorem 2 extends the results for systems with full-row-ranked observation matrix in the literature to general linear systems. Therefore, this work is an important generalization of and a necessary complementary to the literature of state estimation of WSNs in the sense of both estimation framework and theoretical stability analysis; see, for example, [4], [8], [18], [25]–[27].

Coincidentally, from a mathematical point of view, the stability analysis can be cast into the well-received category of Kalman filtering with incomplete (dropped or delayed) observations primarily studied in [26]–[30] and lately in [32], [33]. Explicit comparisons made between the present work and those pioneering works definitely show the implications and necessity of this work. Part of the material in this paper was presented in [34].

The remainder of this paper is organized as follows. We briefly introduce the measurement model as well as some standard assumptions in Section II and devise the minimum mean squared error estimator with power scheduler in Section III. In Section IV, we provide both sufficient condition and necessary condition that guarantee the convergence of averaged estimation error covariance matrix. Finally, conclusions and current research threads are outlined in Section V.

Notations: Straight boldface denote the multivariate quantities such as vectors (lowercase) and matrices (uppercase). Let $\mathbf{x}$ denote the transpose of a matrix, use $\mathbf{y} \sim \mathcal{N}(\mathbf{\mu}, \Sigma)$ for a normally distributed vector $\mathbf{y}$ with mean $\mathbf{\mu}$ and covariance $\Sigma$. For random vectors $\mathbf{x}$ and $\mathbf{y}$, $E[\mathbf{x}]$ denotes the expectation value of $\mathbf{x}$, and $\mathbf{x}^T \mathbf{y}$ denotes the conditional random vector when $\mathbf{y}$ is given. Furthermore, we use $(\cdot)'$ to denote the transpose of a matrix, use $P > 0$ ($\geq 0$) to represent the positive definite (positive semi-definite) matrix $P$, and use $\text{diag}\{l_1, l_2, \ldots, l_m\}$ to denote the diagonal matrix with the main diagonal elements $l_1, l_2, \ldots, l_m$. $I_n$ denotes the $n \times n$ identity matrix and $0$ denotes the zero matrix of appropriate dimensions. The notation $\otimes$ denotes the Kronecker product of two matrices. The mean square stability of the filter, i.e., $\sup_{k \in \mathbb{N}} \mathbb{E}[P_k] < \infty$, implies there always exists a positive definite matrix $\overline{P}$ such that $P_k \leq \overline{P}$ for all $k \in \mathbb{N}$ [25], where the mathematical expectation is taken with respect to both the random power scheduling process and random packet drop process in this paper. For two positive definite matrices $P$ and $Q$, the matrix inequality $P \geq Q$ means matrix $P - Q$ is positive semidefinite. Similar notations will be made for $P \leq Q, P > Q$ and $P < Q$.

II. PROBLEM FORMULATION AND PRELIMINARIES

Consider the following linear discrete-time stochastic system:

\begin{align*}
x_{k+1} &= A x_k + \omega_k \\
y_k &= C x_k + \nu_k
\end{align*}

Fig. 1: Network architecture.
where $x_k \in \mathbb{R}^n$ is the state vector and $y_k \in \mathbb{R}^m$ is the measurement vector, $\omega_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are Gaussian random vectors with zero-means and covariance matrices $Q \succeq 0$ and $R > 0$, respectively. The initial state $x_0$ is also assumed to be a Gaussian random vector with mean $x_0$ and covariance matrix $P_0 > 0$. It is further posited that the random vectors $\omega_k, v_k, x_0$ are mutually independent.

We assume a high transmission energy leads to reliable data flow while a low transmission energy may result in data packet drops during wireless network communications. This assumption is reasonable and motivated by the two facts: Most economically available sensors in the market have multiple transmission energy levels to choose from [11] and higher transmission energy leads to a higher signal-to-noise ratio (SNR) at the remote estimator, which can be simply interpreted as a higher packet arrival rate [35]. Therefore, once communication failure occurs, the whole data packet will drop. For simplicity, the present paper considers that the sensor node has only two transmission power levels [14, 15] and, though, results derived in this paper can be easily generalized to multiple transmission power level case. Specifically, when a high transmission power $\Delta$ is employed, the data packet can be successfully delivered to the estimator side; when a low transmission power $\delta$ is employed, the data packet is supposed to arrive at the estimator side only with a probability $\beta \in (0, 1)$. Similar power scheduling has also been considered in [15] with a different estimation framework.

Before delving into the mechanism of power scheduling, the following two standard assumptions are presented.

**Assumption 1:** $(A, Q^{1/2})$ is controllable and $(C, A)$ is observable.

**Assumption 2:** The covariance $R$ is diagonal, i.e., $R = \text{diag}\{R_1, R_2, \ldots, R_m\}$.

**Remark 1:** In fact, if the measurement noise vectors $v_k$ are white, then covariance matrix $R$ is diagonal. If $v_k$ are not white and $R$ is thus a general positive definite matrix, the idea is primarily to whiten the observations. To this end, we define the square root matrix of a positive definite matrix $R$ as $\tilde{R} := R^{1/2}(R^{1/2})'$. Instead of using $y_k = Cx_k + v_k$, we consider a transformed measurement

$$\hat{y}_k := \tilde{R}^{-1/2}y_k = \tilde{R}^{-1/2}C\hat{x}_k + \tilde{R}^{-1/2}v_k,$$

$$\hat{v}_k := \tilde{R}^{-1/2}v_k$$

where $E[\hat{v}_k\hat{v}_k'] = I_n$. Therefore without loss of generality, we can assume the measurement noise covariance $R$ to be diagonal.

**III. POWER SCHEDULING AND SEQUENTIAL KALMAN FILTERING**

In this paper, a round-robin, slotted-time measurement transmission policy is envisioned such that, only a scalar is allowed to be communicated to the estimator at every transmission and one sampling interval (i.e., one time instant from each $k$ to $k + 1$) can be explicitly partitioned into $m$ (the dimension of measurement vector $y_k$) time slots, and at the $i$th time slot $T_i$, the scheduler located at the sensor node decides whether to use the high or low transmission energy to transmit $y_{ik}^i$, the $i$th component of $y_k$. This sensor scheduling protocol was also used in [6].

It is well acknowledged that the measurement innovation indicates new information of the current measurement that is not contained in all historical measurements and intuitively speaking, a large innovation represents the current measurement is quite different than the predicted measurement and therefore contains much useful information to update the estimate. Thus, we define the measurement of large innovation as “important” measurement and otherwise, less “important” measurement. In this sense, we devise an innovation-based power scheduling policy, which compares the normalized measurement innovation with a given threshold to quantify the “importance” of every measurement and then uses a high (or low) transmission power to communicate the “important” (or less “important”) measurement.

Specifically, at time instant $k$, let the binary random variables $\gamma_{ik}^i$ ($0$ or $1$), $i = 1, 2, \ldots, m$, represent whether the transmission power $\Delta$ or $\delta$ is utilized for transmission of $y_{ik}^i$. Let another sequence of random variables $\beta_{ik}^i = 1$ or $0$ for $i = 1, 2, \ldots, m$, indicate whether the data packet $y_{ik}^i$ arrives at estimator side successfully or not. Throughout this paper, we postulate that the values of $\gamma_{ik}^i$ and $\beta_{ik}^i$ for all $i = 1, 2, \ldots, m$ at every $k$ can be observed; since we can employ TCP-like protocols where the packet acknowledgements are guaranteed at every time instant to notify estimator whether the data packet is received [27, 32]. For future reference, define $\mathcal{I}_k = \{\gamma_{1i}^1y_{1i}^1, (1 - \gamma_{1i}^1)\beta_{1i}^1y_{1i}^1, \gamma_{1i}^2y_{1i}^2, (1 - \gamma_{1i}^2)\beta_{1i}^2y_{1i}^2, \ldots, \gamma_{1i}^m, (1 - \gamma_{1i}^m)\beta_{1i}^m, \gamma_{2i}^1y_{2i}^1, (1 - \gamma_{2i}^1)\beta_{2i}^1, \ldots, \gamma_{2i}^m, (1 - \gamma_{2i}^m)\beta_{2i}^m, \ldots, \gamma_{mi}^1y_{mi}^1, (1 - \gamma_{mi}^1)\beta_{mi}^1, \ldots, \gamma_{mi}^m, (1 - \gamma_{mi}^m)\beta_{mi}^m\}$ and $\mathcal{G}_k = \{y_{ik}^i, \gamma_{ik}^i, \beta_{ik}^i\}, i = 1, 2, \ldots, m$.

For any time instant $k$, denote by $\hat{x}_{ik}^i$ the mean squared error estimate of $x_k$ at estimator based upon all received information at the end of $i$th time slot and likewise, by $P_{ik}^i$ the estimation error covariance, i.e.,

$$\hat{x}_{ik}^i = E[x_k|\mathcal{I}_k],$$

$$P_{ik}^i = E[(x_k - \hat{x}_{ik}^i)(x_k - \hat{x}_{ik}^i)'|\mathcal{I}_k].$$

Let $y_k^i = [y_{1k}^i, y_{2k}^i, \ldots, y_{mk}^i]$ and $\eta_i$, $i = 1, 2, \ldots, m$, be given fixed thresholds. Then the developed power scheduler and the corresponding MMSE estimator are together tabulated as Algorithm 1.

The mechanism of the proposed innovation-based scheduler will be further elaborated in this part. The power scheduler located at the sensor side will sequentially decide whether to adopt a high or low energy to transmit every single component $y_{ik}^1, y_{ik}^2, \ldots, y_{ik}^m$ to remote estimator. For instance, the scheduler is now in a position to make decision on assigning energy for transmission of $y_{ik}^1$ at time slot $i$ of time instant $k$. Suppose the estimator has already reliably sent back current state estimate $\hat{x}_{ik}^i$ and error covariance $P_{ik}^{-1}$ to local sensor side, where, by “reliably” we mean the remote estimator can always adopt a high energy to broadcast information to the sensor so no packet drop happens for the transmission from estimator back to sensor side. This also makes sense because usually wireless sensors consume much less energy for receiving one packet than sending one packet [36]. Thus, after receiving $\hat{x}_{ik}^{-1}$ and
Algorithm 1 (Local Power Scheduler and Remote MMSE Estimator)

**Initialization:**
\[ \hat{x}_{0|0} = x_0, \quad P_{0|0} = P_0. \]

**Time prediction:** given \( \hat{x}_{k-1|k-1} \) and \( P_{k-1|k-1} \), do
\[ \hat{x}_{k|k-1} = A \hat{x}_{k-1|k-1}, \quad P_{k|k-1} = AP_{k-1|k-1}A^\top + Q. \]

**Power scheduling, measurement transmission and measurement update:**

Define \( y_k^i = [y_k^i, y_k^{i+1}, \ldots, y_k^m] \in C^m, \quad C' = \{ C_1', C_2', \ldots, C_m' \}, \)
\[ \hat{x}_{0|k} = \hat{x}_{0|k-1} \quad \text{and} \quad P_{0|k} = P_{k|k-1}. \]
For \( i = 1, 2, \ldots, m \), set

**Estimator side**
\[ \sigma_k^i = \sqrt{C_iP_{i|k}^{-1}C_i' + R_i}, \]
Transmit \( C_i\hat{x}_{k|k-1}, \sigma_k^i \) back to sensor.

**Sensor side**
\[ z_k^i = y_k^i - C\hat{x}_{k|k-1}, \quad \epsilon_k^i = z_k^i / \sigma_k^i, \]

**Power scheduling:**

Define the scheduling variable as \( \gamma_k^i = \begin{cases} 1, & \text{if} \ |\epsilon_k^i| > \eta_i; \\ 0, & \text{otherwise}. \end{cases} \)

**Measurement transmission:**

If \( \gamma_k^i = 1 \), send \( y_k^i \) to the estimator with high transmission power \( \Delta \); otherwise, transmit \( y_k^i \) with low transmission power \( \delta \).

If \( \gamma_k^i = 0 \), then use variable \( \beta_k^i \) to represent whether \( y_k^i \) is successfully transmitted to the remote estimator. Namely, \( \beta_k^i = 1 \) indicates \( y_k^i \) arrives successfully at the remote estimator and \( \beta_k^i = 0 \) means \( y_k^i \) drops during the transmission.

**Estimator side**
Measurement update:

For \( i = 1 \) to \( m \), let \( K_k^i = P_{i|k}^{-1}C_i' \left( C_iP_{i|k}^{-1}C_i' + R_i \right)^{-1} \),
\[ \hat{x}_{k|k}^{i-1} = \hat{x}_{k|k}^{i-1} + s \left( \gamma_k^i, \beta_k^i \right) K_k^i \left( y_k^i - C_i\hat{x}_{k|k}^i \right), \]
\[ P_{k|k}^i = P_{k|k}^{i-1} - t \left( \gamma_k^i, \beta_k^i \right) K_k^i C_iP_{k|k}^{i-1} \]
where
\[ s \left( \gamma_k^i, \beta_k^i \right) = \gamma_k^i + \left( 1 - \gamma_k^i \right) \beta_k^i, \]
\[ t \left( \gamma_k^i, \beta_k^i \right) = \gamma_k^i + \left( 1 - \gamma_k^i \right) \left( 1 - \beta_k^i \right). \]

End, do \( \hat{x}_{k|k} = \hat{x}_{k|k}^m \) and \( P_{k|k} = P_{k|k}^m \).

**Remark 2:** It should be noticed that the effect of channel medium between the sensor and estimator has not been considered here. As pointed out in [37], most of the existing work in distributed estimation assumed perfect channels between sensors and the fusion center. However, since only one scalar \( y_k^i \) is transmitted at every time slot, let us consider transmitting \( y_k^i \) to estimator over a channel with channel gain \( f_k^i \). We envision that the channel undergoes slow fading such that the phase of complex channel can be estimated and therefore compensated for at the receiver side, so that \( f_k^i \) defines the real-valued envelop of the complex channel gain [31]. Also, suppose that the channel gain remains invariant over the time slot to send \( y_k^i \). Then the estimator receives a scaled version of \( y_k^i \) corrupted with the channel noise which is independent of the measurement noise. For simplicity, let the channel noise \( n_k^i \) be zero-mean Gaussian white noise with variance \( \sigma_n^2 \), and, let \( \{ n_k^i \} \) be mutually independent for \( i \neq j \).

In light of the round-robin, time-slotted transmission policy and (2), one further arrives at
\[ \begin{align*}
\gamma_k^i &= 1, \quad \text{if} \quad |\epsilon_k^i| > \eta_i; \\
&= 0, \quad \text{otherwise}. \end{align*} \]

**Proposition 1:** It is postulated that the conditional distribution of \( x_k \) given \( \mathcal{I}_{k-1} \) is approximately Gaussian, i.e., the probability density function (pdf) \( f \left( x_k | \mathcal{I}_{k-1} \right) = \mathcal{N} \left( x_k; \hat{x}_{k|k-1}^i, P_{k|k-1}^i \right) \) where \( \hat{x}_{k|k-1}^i \) is a minimum mean squared error estimator.

**Remark 3:** Recall further that the pdf \( f \left( x_k | \mathcal{I}_{k-1} \right) \) is in general non-Gaussian and therefore, (computationally expensive) numerical integrations and (memory intensive) propagation of the posterior pdf are required for the computation of the exact MMSE estimate [4]. However, based upon customary simplifications adopted in nonlinear filtering [38] and Kalman filtering with quantized measurements/innovations [4], [10], [39], the assumption on an approximately Gaussian distribution of the predicted density is made. This assumption can be widely found in the literature; see, for example, [5], [12], [18] and references therein for further discussion.

**Proof of Proposition 1:** Provided that we already have an MMSE estimator \( \hat{x}_{k|k-1}^i \), that is, \( \hat{x}_{k|k-1}^i = \mathbb{E} \left[ x_k | \mathcal{I}_{k-1} \right] \) and \( P_{k|k}^i = \mathbb{E} \left[ (x_k - \hat{x}_{k|k-1}^i)(x_k - \hat{x}_{k|k-1}^i) | \mathcal{I}_{k-1} \right] \). We prove the proposition by conditioning on whether the measurement is received by the estimator. Specifically, when the new measurement \( y_k^i \) is present at the estimator side, that is, the case \( \gamma_k^i = 1 \) or the case \( \gamma_k^i = 0 \) and \( \beta_k^i = 1 \), one can easily verify that
\[ \begin{align*}
\hat{x}_{k|k}^i &= \mathbb{E} \left[ x_k | \mathcal{I}_{k-1}, y_k^i \right] \\
&= \hat{x}_{k|k}^i + K_k^i \left( y_k^i - C_i\hat{x}_{k|k}^i \right), \\
\text{and likewise,}
\end{align*} \]
\[ \begin{align*}
P_{k|k}^i &= \mathbb{E} \left[ (x_k - \hat{x}_{k|k}^i)(x_k - \hat{x}_{k|k}^i) | \mathcal{I}_{k-1}, y_k^i \right] \\
&= P_{k|k}^i - P_{k|k}^i C_i' \left( C_iP_{k|k}^iC_i' + R_i \right)^{-1} C_iP_{k|k}^i. \\
\end{align*} \]
When the estimator does not receive the new measurement \( y_i^* \), that is, \( \gamma_i^* = 0 \) and \( \beta_i^* = 0 \), then it follows that
\[
\hat{x}_{i|k} = E\left[ x_k | \mathcal{I}^{i-1}_k, \gamma_i^* = 0, \beta_i^* = 0 \right]
= E\left[ x_k | \mathcal{I}^{i-1}_k, |\epsilon|^i_k \leq \eta_i \right]
= \int_{-\eta_i}^{\eta_i} \left( \hat{x}_{i|k} + K_i^i \sigma_k^i \epsilon \right) f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k, |\epsilon|^i_k \leq \eta_i \right) d\epsilon. \tag{6}
\]
Here, \( f_{\epsilon_k}(x) \) is the pdf of the random variable \( x \); similarly, \( f_{\epsilon_k}(x|y) \) is the pdf of a random variable \( x \) conditional on variable \( y \). Given \( \mathcal{I}^{i-1}_k \), then \( \epsilon_i^k \) follows Gaussian distribution with zero-mean and unit covariance. Thus, the conditional pdf above follows directly from conditional probability theory:
\[
f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k, |\epsilon|^i_k \leq \eta_i \right) = \begin{cases} \frac{f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k \right)}{\Delta P_i}, & \text{if } |\epsilon|^i_k \leq \eta_i, \\ 0, & \text{otherwise} \end{cases} \tag{7}
\]
where \( \Delta P_i \equiv \Pr \left( |\epsilon|^i_k \leq \eta_i | \mathcal{I}^{i-1}_k \right) = 1 - 2Q(\eta_i) \). Therefore, (6) becomes
\[
\hat{x}_{i|k} = \int_{-\eta_i}^{\eta_i} \left( \hat{x}_{i|k} + K_i^i \sigma_k^i \epsilon \right) f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k \right) \Delta P_i d\epsilon + \int_{-\eta_i}^{\eta_i} \epsilon f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k \right) d\epsilon
= \hat{x}_{i|k}^i + \int_{-\eta_i}^{\eta_i} \epsilon f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k \right) d\epsilon \tag{8}
\]
where the first integration equals to 1 and the second becomes 0 because \( f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k \right) \) is even over the origin-centered symmetrical interval. In the sequel, we compute the covariance \( P_{i|k} \) for \( \gamma_i^* = 0 \) and \( \beta_i^* = 0 \) case as follows:
\[
P_{i|k} = \begin{cases} \gamma_i^* \left( \epsilon_i^k, \beta_i^* \right) \left( P_{i|k}^{i-1} - K_i^i C_i P_{i|k}^{i-1} \right) + \\ \left( 1 - \gamma_i^* \right) \left( \beta_i^* \right) \left( P_{i|k}^{i-1} - \sqrt{\frac{2}{\pi}} \eta_i \exp \left( -\eta_i^2 / 2 \right) \right) \sqrt{1 - 2Q(\eta_i)} \times \\ K_i^i C_i P_{i|k}^{i-1} \end{cases} \tag{9}
\]
where (b) follows directly from \( \hat{x}_{i|k} = \hat{x}_{i|k}^i \) when the new measurement component \( y_i^* \) is not received by the estimator, which has been proved in (8), and (d) is because \( \sigma_k^i \epsilon_i^k = z_i^k = y_k^* - C_i \hat{x}_{i|k} = C_i \left( x_k - \hat{x}_{i|k} \right) + v_k^i \) in Algorithm 1, and (e) is because \( v_k^i \) is zero-mean Gaussian noise with covariance \( R_i \), or, \( E \left[ K_i^i v_k^i (K_i^i v_k^i)^\prime \right] = K_i^i R_i (K_i^i)^\prime \). Meanwhile, we have
\[
E \left[ e^2 | \mathcal{I}^{i-1}_k, |\epsilon|^i_k \leq \eta_i \right] = \int_{-\eta_i}^{\eta_i} e^2 f_{\epsilon_k} \left( \epsilon | \mathcal{I}^{i-1}_k, |\epsilon|^i_k \leq \eta_i \right) d\epsilon
= \frac{1}{1 - 2Q(\eta_i)} \int_{-\eta_i}^{\eta_i} \frac{\epsilon^2}{\sqrt{2\pi}} \exp \left( -\epsilon^2 / 2 \right) d\epsilon
= \frac{1}{1 - 2Q(\eta_i)} \frac{\eta_i}{\sqrt{\pi}} \exp \left( -\eta_i^2 / 2 \right). \tag{10}
\]
Therefore, from (9)–(10) and \( \sigma_k^i \) are mutually independent Bernoulli i.i.d. processes. Define \( N_k = \text{diag} \left\{ t(\gamma_k^i), t(\beta_k^i), \ldots, t(\gamma_m^i) \right\} \). For \( i = 1, 2, \ldots, m, \) let
\[
\mu_i = E[\gamma_i^i] = 2Q(\eta_i) \tag{12}
\]
\[
\nu_i = \frac{\eta_i \exp(-\eta_i^2 / 2)}{\sqrt{\pi / 2} \left( 1 - 2Q(\eta_i) \right)} \tag{13}
\]
which completes the proof.

Given the new filter formulation in Algorithm 1, the processes \( \{\gamma_k^0\}, \{\beta_k^0\}, \ldots, \{\gamma_k^m\} \) form a sequence of independent and identically distributed (i.i.d.) processes under the Gaussian approximation [18] and also, assume the processes \( \{\beta_k^0\}, \{\beta_k^1\}, \ldots, \{\beta_k^m\} \) are mutually independent Bernoulli i.i.d. processes. Define \( N_k = \text{diag} \left\{ t(\gamma_k^0), t(\gamma_k^1), \ldots, t(\gamma_k^m) \right\} \). For \( i = 1, 2, \ldots, m, \) let
\[
\mu_i = E[\gamma_i^i] = 2Q(\eta_i) \tag{12}
\]
\[
\nu_i = \frac{\eta_i \exp(-\eta_i^2 / 2)}{\sqrt{\pi / 2} \left( 1 - 2Q(\eta_i) \right)} \tag{13}
\]
and $\mathbb{E}[\beta_k] = \beta$; in addition,

$$\mathbb{E} \left[I(\beta_k)\right] = \mathbb{E} \left[\beta_k + (1 - \beta_k) \times \frac{\eta_i \exp(-\eta_i^2/2)}{\sqrt{\pi/2} (1 - 2Q(\eta_i))}\right] = \beta + (1 - \beta) \nu_i = \xi_i,$$

$$(14)$$

$$\mathbb{E} \left[I(\gamma_k, \beta_k)\right] = \mathbb{E} \left[\gamma_k + (1 - \gamma_k)I(\beta_k)\right] = \mu_i + (1 - \mu_i) \xi_i = \lambda_i,$$

$$(15)$$

where, in fact, we have $\nu_i = 1 - 1/\sqrt{2\pi} \int_{\eta_i}^{\infty} \exp(-t^2/2) dt \in [0, 1]$ and $\xi_i \in [0, 1]$. Moreover, $\nu_i$ is one strictly decreasing function in threshold $\eta_i$; this makes sense since the greater the threshold is, the less information will be transmitted through high energy. Then, one can easily verify $0 \leq \lambda_i \leq 1$, and therefore, $\lambda_i$ can be somehow physically interpreted as the normalized averaged information received by remote estimator resulting from the power scheduling and networked effect on transmitting $y_k^i$ and $(1 - \lambda_i)$ quantifies the corresponding averaged information loss rate. All $\lambda_i$s together will govern the mean square stability of estimation error covariance matrix, which will be investigated in the ensuing section. Therefore, we will refer to $\lambda_i$ hereafter other than the specific parameters $\eta_i, \beta$.

IV. STATISTICAL PROPERTIES AND SUFFICIENT, NECESSARY CONVERGENCE CONDITIONS

In this section, the convergence conditions for the expected estimation error covariance will be provided by discussing properties of a constructed function. Denote $\{\gamma_k^1\} := \{\{\gamma_k^1\}^\infty_0, \{\gamma_k^2\}^\infty_0 \ldots \} \in \{\{\beta_k^1\}^\infty_0, \{\beta_k^2\}^\infty_0 \ldots \}$. Since they are inherently stochastic and cannot be determined offline, therefore, only statistical properties can be derived. Before delving into main results, some preliminaries will be given in the following.

Let $S^n_+ = \{S \in \mathbb{R}^{n \times n} | S \succeq 0\}$. Define the function $h: S^n_+ \to S^n_+$ and the function $g_{\lambda_i}: S^n_+ \to S^n_+$ as follows:

$$h(X) \triangleq AXA' + Q \quad (16)$$

$$g_{\lambda_i}(X) \triangleq X - \lambda_i XC'_{\lambda_i} (C_{\lambda_i}XC'_{\lambda_i} + R_i)^{-1} C_{\lambda_i}X \quad (17)$$

$$g_{\lambda_i} \circ h(X) \triangleq g_{\lambda_i}(h(X)) \quad (18)$$

and here denote the notation $\circ$ by the function composite. Therefore, the covariance update in the sequential Kalman filter formulation in Algorithm 1 becomes

$$P_{k|k-1} = h(P_{k-1|k-1})$$

$$P_{k|k} = g_{\lambda_i}(P_{k|k}) = g_{\lambda_i}(P_{k|k-1})$$

$$P_{k|k} = g_{\lambda_i}(P_{k-1|k-1}), i = 2, 3, \ldots, m - 1$$

Let

$$P_{k|k-1} = h(P_{k-1|k-1}) \quad (19)$$

$$P_{k|k} = \mathcal{M}_m(P_{k|k-1}) \triangleq g_{\lambda_m} \circ \cdots \circ g_{\lambda_2} \circ g_{\lambda_1}(P_{k-1|k-1}). \quad (20)$$

Denote the function $\varphi: S^n \to S^n$ by the transformation from $P_{k-1|k-1}$ to $P_{k|k}$, namely,

$$P_{k|k} = \varphi(P_{k-1|k-1}) \triangleq \mathcal{M}_m h(P_{k-1|k-1}). \quad (21)$$

In order to analyze the convergence of the estimation error covariance matrix, we then define the modified algebraic Riccati equation (MARE) in the following way:

$$\varphi(P_k) = g_{\lambda_m} g_{\lambda_{m-1}} \cdots g_{\lambda_1} h(P_k) \quad (22)$$

where we used the simplified notation $P_k = P_{k|k}, k \geq 0$. Meanwhile, as explained, the covariance matrices $\{P_k\}^\infty_0$ depend nonlinearly on the specific realization of the stochastic processes $\{\gamma_k\}$ and $\{\beta_k\}$, so the sequential Kalman filter is inherently stochastic and cannot be determined offline. Then, only statistical properties with respect to the covariance matrices of the proposed sequential Kalman filter can therefore be established.

Remark 4: It is noted in passing that the modified algebraic Riccati equation defined in (22) is a more generalized form than the original MARE specified for Kalman filtering with only one or two lossy channels in [26] and [30], respectively, where the analysis might be much easier than that of (22). Moreover, since the MARE in (22) is sequentially composited by $m$ original MAREs with different parameters, namely, $\lambda_1, \lambda_2, \ldots, \lambda_m$, then the MARE in (22) is also quite different from the MARE discussed in [32] defined for Kalman filtering for multiple-input multiple-output systems with control signals and sensor measurements transmitting across multiple TCP-like erasure channels. Accurately speaking, the MARE in [32] follows directly from that in [26] by replacing the observation matrix $\gamma_k C$ with $\text{diag}(\gamma_k^1, \gamma_k^2, \ldots, \gamma_k^m) C$. Therefore, for the sake of completeness on stability theory of Kalman filtering with intermittent observations, the investigation on properties of the MARE in (22) in the following sections is also of great implications, which significantly contributes to the derivation of sufficient conditions for stability of sequential Kalman filtering with scheduled measurements in [18].

The following lemma on the properties of the auxiliary function $\psi_{\lambda_i}$ is presented before we will formally study the convergence properties of the MARE in (22).

**Lemma 1 (26):** Let the function $\psi_{\lambda_i}$ be

$$\psi_{\lambda_i}(L_i, X) = (1 - \lambda_i)X + \lambda_i \left(E_i X E_i' + L_i R_i L_i'\right), \quad i = 1, \ldots, m \quad (23)$$

where $E_i = I_n + L_i C_i, R_i > 0, X, Y, Z \in S^n_+$. Then the following facts hold:

1) With given $L_i X = -X C_i (C_i X C_i' + R_i)^{-1}$. $g_{\lambda_i}(X) = \psi_{\lambda_i}(L_i X, X)$

2) $g_{\lambda_i}(X) = \min_L \psi_{\lambda_i}(L_i X, X) \leq \psi(L_i, X), \forall L_i$

3) If $X \preceq Y$, then $g_{\lambda_i}(X) \preceq g_{\lambda_i}(Y)$

4) If $\lambda_i \geq \lambda_j$, then $g_{\lambda_i}(X) \preceq g_{\lambda_i}(Y)$

5) If $\tau \in [0, 1]$, then $g_{\lambda_i}(\tau X + (1 - \tau)Y) \geq \tau g_{\lambda_i}(X) + (1 - \tau)g_{\lambda_i}(Y)$

**Proof:** The proofs for these statements are analogous to those of Lemma 1 in [26] with some appropriate notation adaptations.

Notice, that the relationship between the function $g_{\lambda_i}$ and the function $\psi_{\lambda_i}$ has been built, and now, in order to investigate the convergence properties of the MARE in (22), the relationship between the composite function $g_{\lambda_m} g_{\lambda_{m-1}} \cdots g_{\lambda_1}$
and the introduced auxiliary function $\psi_{\lambda_r}$ will be constructed in the following way.

According to (23), observe that the function $\psi_{\lambda_r}(L_i, X)$ is a function with respect to two matrix variables $L_i, X$. With a slight abuse of notation, denote $\psi_{\lambda_r}(L_j, X)$ by the composite function $\psi_{\lambda_r}(\lambda_{j-1}, \psi_{\lambda_r}(L_j, X))$ with respect to the second variable $X$, where $j = 1, \ldots, m - 1$.

In the sequel, let

$$\mathcal{T}_s(L_1, L_2, \ldots, L_s, X) = \psi_{\lambda_s}\psi_{\lambda_{s-1}} \ldots \psi_{\lambda_1}(L_1, L_2, \ldots, L_s, X), \quad s = 1, 2, \ldots, m.$$ \hspace{1cm} (24)

Then,

$$\mathcal{T}_0 = \psi_{\lambda_1}(L_1, X_{\tau-1}) = (1 - \lambda_1)X_{\tau-1} + \lambda_1 \left( E_s \mathcal{T}_{s-1} \mathcal{E}_s' + L_s R_s L_s' \right)$$

$$= \sum_{j=1}^{s} \prod_{i=j+1}^{s-1} (1 - \lambda_i) \lambda_j (E_j \mathcal{T}_{j-1} \mathcal{E}_j' + L_j R_j L_j')$$

$$+ \sum_{i=1}^{s-1} (1 - \lambda_i) X + \lambda_i \left( E_s \mathcal{T}_{s-1} \mathcal{E}_s' + L_s R_s L_s' \right)$$

$$= \sum_{j=0}^{s} \prod_{i=j+1}^{s-1} (1 - \lambda_i) \lambda_j (E_j \mathcal{T}_{j-1} \mathcal{E}_j' + L_j R_j L_j')$$

$$+ \lambda_s \left( E_s \mathcal{T}_{s-1} \mathcal{E}_s' + L_s R_s L_s' \right), \quad s = 1, 2, \ldots, m.$$ \hspace{1cm} (24)

where, to make the expression more concrete, we defined $\lambda_0 = 1$ and $E_0 = I_n, R_0 = 0, \mathcal{T}_{-1} = X, \mathcal{T}_0(X) = X$.

For the sake of brevity, denote

$$\eta_{j,s} = \prod_{i=j+1}^{s} (1 - \lambda_i) \lambda_j, \quad j = 0, 1, \ldots, s - 1,$$

$$\eta_{s,s} = \lambda_s.$$ \hspace{1cm} (25)

More importantly, it is easy to exploit the fact that the sum of $s + 1$ coefficients $\eta_{j,s}, j = 0, 1, \ldots, s$ is identically 1, i.e.,

$$\sum_{j=0}^{s} \eta_{j,s} = \sum_{j=0}^{s} \left[ \prod_{i=j+1}^{s} (1 - \lambda_i) \lambda_j \right] + \eta_{s,s}$$

$$= (1 - \lambda_s) + \lambda_s = 1.$$ \hspace{1cm} (26)

Alternatively, (24) can be given by

$$\mathcal{T}_{-1} = X, \mathcal{T}_0 = X,$$

$$\mathcal{T}_s = \sum_{j=0}^{s} \eta_{j,s}^2 \left( E_j \mathcal{T}_{j-1} \mathcal{E}_j' + L_j R_j L_j' \right), \quad s = 1, 2, \ldots, m.$$ \hspace{1cm} (26)

Therefore,

$$\mathcal{T}_m = \sum_{j=0}^{m-1} \eta_{j,m}^2 \left( E_j \mathcal{T}_{j-1} \mathcal{E}_j' + L_j R_j L_j' \right), \quad j = 0, \ldots, m.$$ \hspace{1cm} (27)

where $\eta_{m,0}^2 = 1, E_0 = I_n, R_0 = 0$ and $X \geq 0$, and $\mathcal{T}_j$ is defined in (26) with $\eta_{j,m}$ given by (25).

Remark 5: Note that the auxiliary function $\mathcal{T}_m$ defined by (27) is of similar form to that in [32, 33]. To be more specific, the latter auxiliary function is referred to as follows:

$$\phi(K_1, K_2, \ldots, K_{2^m}, P) = \sum_{i=0}^{2^m} \Phi_i (F_i Y F_i' + V_i)$$ \hspace{1cm} (28)

with $F_i = A + K_i C_i, V_i = K_i R_i K_i' + Q$ and constants $\tau_j$. One can easily observe that there are $2^m$ terms in (28), and the computation burden will become catastrophic when the dimension $m$ of the measurement vector tends to be very large. Meanwhile, only $\frac{(m+1)m}{2}$ terms will be needed for the sequential Kalman filter in this paper, which may significantly reduce the computation burden and is therefore of great importance.

We are now in a position to establish some properties of the function $\mathcal{T}_m(L_1, L_2, \ldots, L_m, X)$ in form of lemmas in the following.

Lemma 2: Consider the function $\mathcal{T}_m(L_1, L_2, \ldots, L_m, X)$ as stated by (27) with $E_j = I_n + L_j C_j$. Assume $X, Y, Z \in \mathbb{S}_n^+$. Then, the following facts hold:

1) With given $L_j X = -T_j X \{ C_j T_j X + R_j \}^{-1}, j = 1, 2, \ldots, m, \mathcal{M}_m(X) = \mathcal{T}_m(L_1 X, L_2 X, \ldots, L_m X), \quad \text{where} \quad \mathcal{T}_1 X = \mathcal{T}_m(L_1 X, L_2 X, \ldots, L_m X)$

2) $\mathcal{M}_m(X) = \mathcal{T}_m(L_1, L_2, \ldots, L_m, X) \leq \mathcal{T}_m(L_1, L_2, \ldots, L_m, X), \forall L_1, L_2, \ldots, L_m \in \mathbb{R}^{n \times 1}$

3) If $X \preceq Y$, then $\mathcal{M}_m(X) \preceq \mathcal{M}_m(Y)$

4) If $\tau \in [0, 1]$, then $\mathcal{M}_m(\tau X + (1 - \tau) Y) \succeq \tau \mathcal{M}_m(X) + (1 - \tau) \mathcal{M}_m(Y)$

5) $\mathcal{M}_m(X) \preceq \prod_{j=1}^{m} (1 - \lambda_j) X$

6) For a random variable $X, \quad \prod_{j=1}^{m} (1 - \lambda_j) \mathbb{E}[X] \succeq \mathbb{E}[\mathcal{M}_m(X)] \preceq \mathcal{M}_m(\mathbb{E}[X])$.

Proof:

1) Fact 1) together with Fact 2) is equivalent to showing the minimizer and the minimum value of matrix-valued function $\mathcal{T}_s, \forall s = 1, 2, \ldots, m$, with respect to multiple vector-valued variables $L_1, L_2, \ldots, L_s \in \mathbb{R}^{n \times 1}$. For convenience of notation, denote $L_s = (L_1, L_2, \ldots, L_s)$.

We first make extensive use of differential of general matrix-valued function $F$ with respect to a matrix argument $X$: see, for instance, [40].

Definition 1: Let $F$ be a differentiable $m \times n$ real matrix function of a $p \times q$ matrix of real variables $X$. The Jacobian matrix of $F$ at $X$ is given by the $mn \times pq$ matrix

$$\mathcal{D}X F(X) = \frac{\partial \vec{vec} F(X)}{\partial (\vec{vec} X)}.$$ 

Then by vectorizing the differential $d \mathcal{T}_s$, it gives that:

$$d \mathcal{T}_s = J_{1,s} d \mathcal{V}_1 + J_{2,s} d \mathcal{V}_2 + \ldots + J_{s,s} d \mathcal{V}_s,$$

where $J_{j,s}$ is defined as $J_{j,s} = J_{j,s}(L_1, L_2, \ldots, L_s) = dL_j \mathcal{T}_s, i = 1, 2, \ldots, s$. To make the results more concrete, let us define:

$$G_{j,j} \triangleq I_n \otimes I_n = I_n^2$$

$$G_{j,t} \triangleq \eta_{j,t}^2 I_n^2 + \sum_{i=j+1}^{t} \eta_{i,t}^2 (E_i \otimes E_i) G_{j,i-1},$$

$$t = j + 1, j + 2, \ldots, s.$$
Therefore, after complicated and tedious matrix computations, the Jacobian matrices can be obtained as follows:

\[
 J_{j,s} = \left( \eta_{j,s} I_{n^2} + \sum_{k=j+1}^{n} \eta_{k,s} (E_k \otimes E_k) G_{j,k} \right) \\
 \times \left( (E_j T_{j-1} C_{j-1} + L_j R_j) \otimes I_n \right) \\
 + I_n \otimes \left( (E_j T_{j-1} C_{j-1} + L_j R_j) \right)
\]

where intentionally, \( \eta_{j,s} \) was not replaced by 1 for the compactness of the structure of \( J_{j,s} \).

By solving \( J_{j,s} = 0, j = 1, 2, \ldots, s - 1 \), it follows straightforwardly that

\[
 E_j T_{j-1} C_{j-1} + L_j R_j = 0 \\
 \iff L_{s,j} = -T_{j-1} C_{j-1} (C_j T_{j-1} C_{j-1} + R_j)^{-1} \triangleq L_{s,j}^X
\]

where \( T_{j-1} = T_{j-1} \left( L_1^X, L_2^X, \ldots, L_{j-1}^X \right) = T_{j-1}^X \). Then similarly, by solving \( J_{j,s} = 0 \), it gives that

\[
 E_j T_{s-1} C_{s-1} + L_s R_s = 0 \\
 \iff L_{s,s} = -T_{s-1} C_{s-1} (C_s T_{s-1} C_{s-1} + R_s)^{-1} \triangleq L_{s,s}^X
\]

where \( T_{s-1} = T_{s-1} \left( L_1^X, L_2^X, \ldots, L_{s-1}^X \right) = T_{s-1}^X \). It should be clearly noticed that \( L_{s,t}^X = L_{s,t}^X, \forall t \geq s \), and then plugging \( L_1^X, L_2^X, \ldots, L_m^X \) into (27) verifies that \( M_m(X) = T_m(L_1^X, L_2^X, \ldots, L_m^X) \).

2) We show this fact by mathematical induction. When \( m = 1 \), one can easily verify that \( L_1^X \) minimizes \( T_1(L_1, X) \). Suppose now that it holds for \( m \) that is, the point \( (L_1^X, L_2^X, \ldots, L_m^X) \) minimizes \( T_k(L_1, L_2, \ldots, L_k, X) \). Then for \( k = k + 1 \),

\[
 T_{k+1} = \left( 1 - \lambda_{k+1} \right) T_k + \lambda_{k+1} E_{k+1} T_{k+1} E_{k+1} + L_{k+1} R_{k+1} L_{k+1} \]

and

\[
 D_{T_k} T_{k+1} = (1 - \lambda_{k+1}) (I_n \otimes I_n) + \lambda_{k+1} (E_{k+1} \otimes E_{k+1}) > 0
\]

so one necessary condition for some point \( (L_1^X, L_2^X, \ldots, L_k^X, L_{k+1}) \) minimizing \( T_{k+1} \) is that the point should also minimize \( T_k \), or, \( (L_1^X, L_2^X, \ldots, L_k^X) \) minimizes \( T_k \). Therefore, \( (L_1^X, L_2^X, \ldots, L_k^X) = (L_1^X, L_2^X, \ldots, L_k^X) \), or, \( T_k = T_k^X \) when minimizing \( T_{k+1} \). Given that \( L_{k+1} \) is independent of \( T_k \) and, \( T_k^X > 0, R_{k+1} > 0 \), and meanwhile, \( T_{k+1} \) is quadratic and convex in the variable \( L_{k+1} \), and therefore, the minimizer for \( T_{k+1} \) can be found by letting

\[
 D_{L_{k+1}} T_{k+1} = \lambda_{k+1} \left[ \left( E_{k+1} T_{k+1} C_{k+1} + L_{k+1} R_{k+1} \right) \otimes I_n \right] \\
 + I_n \otimes \left( E_{k+1} T_{k+1} C_{k+1} + L_{k+1} R_{k+1} \right)
\]

which leads to the unique solution \( L_{k+1}^X = -T_{k+1} C_{k+1} (C_{k+1} + L_{k+1} R_{k+1})^{-1} \). Therefore, the point \( (L_1^X, L_2^X, \ldots, L_k^X, L_{k+1}^X) \) minimizes \( T_{k+1} \). This completes the proof.

3) Observe that the function \( T_m \) is affine in the variable \( X \). Let \( X \leq Y \), and it yields that

\[
 M_m(X) = T_m(L_1^X, L_2^X, \ldots, L_m^X, X) \\
 \leq T_m(L_1^Y, L_2^Y, \ldots, L_m^Y, X) \\
 \leq T_m(L_1^Y, L_2^Y, \ldots, L_m^Y, Y) \\
 \leq \ldots \leq T_m(L_1^Y, L_2^Y, \ldots, L_m^Y, Y)
\]

where (a) is because \( L_m^X \) minimizes the function \( T_m \) with respect to variables \( L_1, L_2, \ldots, L_m \), then for any \( L_m \neq L_m^X \), say \( L_m = L_m^X \), that is, (a) holds true. (b) is due to \( T_m \) is affine in the variable \( X \) and (c) follows straightforwardly from Fact 2 above.

4) Let \( Z = \tau X + (1 - \tau) Y \), where \( \tau \in [0, 1] \). Notice that

\[
 M_1(Z) = T_1(L_1^Z, Z) = \eta_{0,1}^Z + \eta_{1,1}^Z \left[ (I_n + L_1^Z C_1) Z (I_n + L_1^Z C_1) \right] + \tau L_2^Z R_1 (L_1^Z) + (1 - \tau) L_2^Z R_1 (L_2^Z) \]

\[
 = \tau \left[ \eta_{0,1}^Z X + \eta_{1,1}^Z (I_n + L_1^Z C_1) X (I_n + L_1^Z C_1) \right] + L_2^Z R_1 (L_1^Z) + (1 - \tau) \left[ \eta_{0,1}^Z Y + \eta_{1,1}^Z \right] \times (I_n + L_1^Z C_1) Y (I_n + L_1^Z C_1) + L_1^Z R_1 (L_1^Z) \]

\[
 = \tau T_1(L_1^Z, X) + (1 - \tau) T_1(L_1^Z, Y) \]

\[
 \geq \tau M_1(X) + (1 - \tau) M_1(Y) = \tau M_1(X) + (1 - \tau) M_1(Y).
\]

Assume that \( M_s(Z) \geq \tau M_s(X) + (1 - \tau) M_s(Y) \). Then, we have

\[
 M_{s+1}(Z) = T_{s+1}(L_1^Z, L_2^Z, \ldots, L_s^Z) = \left( 1 - \lambda_{s+1} \right) (I_n \otimes I_n) + \lambda_{s+1} (E_{s+1} \otimes E_{s+1}) \]

\[
 + L_{s+1}^Z R_{s+1} (L_{s+1}^Z) \]

\[
 \geq \left( 1 - \lambda_{s+1} \right) \left[ \tau M_s(X) + (1 - \tau) M_s(Y) \right] + \lambda_{s+1} \left[ (I_n + L_{s+1}^Z C_{s+1}) M_s(X) \right] + \left( 1 - \tau \right) M_s(Y) \]

\[
 + (1 - \tau) \left[ (I_n + L_{s+1}^Z C_{s+1}) M_s(Y) \right] + (\tau + 1 - \tau) L_{s+1}^Z R_{s+1} (L_{s+1}^Z) \]

\[
 = \tau T_{s+1}(L_1^Z, L_2^Z, \ldots, L_s^Z, L_{s+1}^Z, X) + (1 - \tau) T_{s+1}(L_1^Z, L_2^Z, \ldots, L_s^Z, L_{s+1}^Z, Y) \]

\[
 \geq \tau M_{s+1}(X) + (1 - \tau) M_{s+1}(Y). \]

Therefore, the fact holds true.

5) Note that

\[ \mathcal{M}_m(X) = \mathcal{T}_m \left( L^X_1, L^X_2, \ldots, L^X_m, X \right) \]
\[ = \prod_{j=1}^m \left( E_{0j} - R_{0j} L_{0j} \right) \]
\[ + \sum_{j=1}^m \eta^2_{j,m} \left[ E^X_j E_{j-1} (E^X_j)' + L^X_j R_{j} (L^X_j)' \right] \]
\[ \geq \prod_{j=1}^m (1 - \lambda_j) X \]

where \( \eta^2_{j,m} = \prod_{j=1}^m (1 - \lambda_j), E_0 = I_n, R_0 = 0 \) and \( E^X_j E_{j-1} (E^X_j)' + L^X_j R_{j} (L^X_j)' \geq 0 \), \( \forall j = 1, 2, \ldots, m \).

6) The first inequality follows straightforwardly from Fact 5) above and linearity of expectation, that is,

\[ \mathbb{E} [\mathcal{M}_m(X)] \geq \prod_{j=1}^m (1 - \lambda_j) \mathbb{E} [X] . \]

The second inequality is due to Fact 4) above which implies the concavity of the function \( \mathcal{M}_m(X) \), and therefore in the light of Jensen’s inequality, it readily gives that

\[ \mathcal{M}_m(\mathbb{E}[X]) \geq \mathbb{E}[\mathcal{M}_m(X)] . \]

To take \( P_{k+1} \) into consideration, the auxiliary function \( \phi_m \) can be given in the following way:

\[ \phi_{-1} = h(X), \phi_0 = h(X), \]
\[ \phi_s = \sum_{j=0}^s \eta^2_{j,s} (E_j \phi_{j-1} (E_j' + L_j R_j L_j)'), s = 1, 2, \ldots, m-1, \]
\[ \phi_m = \sum_{j=0}^m \eta^2_{j,m} (E_j \phi_{j-1} (E_j' + L_j R_j L_j')). \]

Lemma 3: Consider the function \( \phi_m(1_{L_1}, L_2, \ldots, L_m, X) \) as stated by (29) with \( E_j = I_{n_j}, L_j \subset C_j \). Assume \( X, Y, Z \in \mathbb{R}_{+}^{n} \).

Then, the following facts hold:

1) With given \( L^X_j = -\phi_{j-1} C_j' (C_j \phi_{j-1} C_j' + R_j)^{-1}, j = 1, 2, \ldots, m, \) \( \phi(X) = \phi_m(L^X_1, L^X_2, \ldots, L^X_m, X) \), where \( \phi_{j-1} \).

2) Assume \( \phi_m(1_{L_1}, L_2, \ldots, L_m, X) \leq \phi_m(L_1, 1_{L_2}, \ldots, 1_{L_m}, X) \).

3) If \( X \leq Y \), then \( \phi(X) \leq \phi(Y) \).

4) If \( r \in [0, 1] \), then \( \phi(rX + (1-r)Y) \geq r\phi(X) + (1-r)\phi(Y) \).

5) \( \phi(X) \geq \prod_{j=1}^m (1 - \lambda_j) A X A' + Q \).

6) If \( \bar{X} \geq \phi(\bar{X}) \), then \( \bar{X} > 0 \).

7) For a random variable \( X, \prod_{j=1}^m (1 - \lambda_j) A \mathbb{E}[X] A' + Q \leq \mathbb{E}[\phi(X)] \leq \mathbb{E}[\mathbb{E}[X]] . \)

Proof: We only prove Fact 6) because the others can be derived directly from Lemma 2.

6) According to Fact 7), above, it gives that \( \bar{X} \geq \phi(\bar{X}) \geq \prod_{j=1}^m (1 - \lambda_j) A \bar{X} A' + Q \). Since \( (A, Q^\frac{1}{2}) \) is controllable, then there must exist an \( \bar{X} > 0 \) subject to the Lyapunov equation \( \bar{X} = \prod_{j=1}^m (1 - \lambda_j) A \bar{X} A' + Q \) if \( \prod_{j=1}^m (1 - \lambda_j) A \) is asymptotically stable. Accordingly, it follows that

\[ \bar{X} - \bar{X} > \prod_{j=1}^m (1 - \lambda_j) A (\bar{X} - \bar{X}) A' \]

implying there exists a \( Q > 0 \) such that

\[ \bar{X} - \bar{X} = \prod_{j=1}^m (1 - \lambda_j) A (\bar{X} - \bar{X}) A' + Q . \]

Therefore, \( \bar{X} - \bar{X} > 0 \), or \( \bar{X} > \bar{X} > 0 \). This completes the proof.

Remark 6: Observe that if we substitute \( X = P_{k+1} \) into Fact 7) in Lemma 3, it follows that \( \prod_{j=1}^m (1 - \lambda_j) A \mathbb{E}[P_k] A' + Q \leq \mathbb{E}[\phi(P_k)] \leq \mathbb{E}[\mathbb{E}[P_k]] \). Since \( \mathbb{E}[P_{k+1}] = \mathbb{E}[P_k] \) and \( \mathbb{E}[P_k] = \mathbb{E}[\mathbb{E}[P_k]] \), then \( \prod_{j=1}^m (1 - \lambda_j) A \mathbb{E}[P_k] A' + Q \leq \mathbb{E}[P_{k+1}] \leq \mathbb{E}[\mathbb{E}[P_k]] \). That is, the expected value of \( P_{k+1} \) can be lower-bounded and upper-bounded by \( \prod_{j=1}^m (1 - \lambda_j) A \mathbb{E}[P_k] A' + Q \) and \( \mathbb{E}[\mathbb{E}[P_k]] \) both as functions of \( \mathbb{E}[P_k] \), respectively.

To facilitate the convergence analysis, let us define the linear part of function \( \phi_m(L_1, L_2, \ldots, L_m, X) \) in terms of variable \( X \) as another auxiliary function, namely

\[ \mathcal{L}_m(Y) = \sum_{j=0}^m \eta^2_{j,m} (E_j \phi_{j-1} (E_j' + L_j R_j L_j')). \]

where \( \phi_{j-1}, \phi_{j-1} = 0, 1, \ldots, m \) are defined in (29). Then, the following lemma can be readily presented.

Lemma 4: Consider the function \( \mathcal{L}_m(Y) \) as stated in (30). If there exists a positive definite matrix \( \mathcal{Y} > \mathcal{L}_m(Y) \).

1) \( \forall W \geq 0, \lim_{k \to \infty} \mathcal{L}_m(W) = 0 \)

2) Given \( U > 0 \), the following sequence \( Y_{k+1} = \mathcal{L}_m(Y_k) + U \)

initialized at \( Y_0 \geq 0 \). Then, the sequence \( Y_k \) is bounded.

Proof:

1) Note that \( \mathcal{L}_m(Y) \) is affine in \( Y \) and \( \mathcal{L}_m(Y) \geq 0, \forall Y \geq 0 \), and \( \mathcal{L}_m(Y) \geq \mathcal{L}_m(Z) \), for \( Y \geq Z \). There exist constants \( 0 \leq r < 1 \) and \( t \geq 0 \) such that \( \mathcal{L}_m(\mathcal{Y}) \leq r \mathcal{Y} < \mathcal{Y} \) and \( \mathcal{W} \leq t \mathcal{Y} \), respectively. Then

\[ 0 \leq \mathcal{L}_m(W) \leq t \mathcal{Y} \leq \mathcal{L}_m(\mathcal{Y}) < 1. \]

Therefore, \( \lim_{k \to \infty} \mathcal{L}_m(W) = 0 \) given that \( 0 \leq r < 1 \).

2) Based on (31) above, for any initialization \( U_0 \geq 0 \) and any \( U > 0 \), there always exist two constants \( t_{Y_0} \geq 0 \) and \( t_U \geq 0 \) such that \( Y_0 \leq t_{Y_0} \mathcal{Y} \) and \( U \leq t_U \mathcal{Y} \), which are independent of \( k \). Therefore, similar arguments in
Lemma 5: Consider the function \( \phi_m(\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, \mathbf{X}) \) defined in (29). Assume there exist \( m \) gain matrices \( \mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m \) and a positive definite matrix \( \mathbf{P} \) such that

\[
\mathbf{P} > 0 \quad \text{and} \quad \mathbf{P} > \phi_m \left( \mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, \mathbf{P} \right).
\]

Then, the sequence \( P_k = \phi^k(P_0) \) is bounded for any given \( P_0 \). That is, there exists a positive definite matrix \( M_{P_0} > 0 \) depending on \( P_0 \) such that

\[
P_k \leq M_{P_0}, \quad \forall k \geq 0.
\]

Proof: Observe that \( \phi_m(\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, \mathbf{Y}) = \mathcal{L}_m(\mathbf{Y}) + Q + N_m \), where \( N_m := \sum_{j=0}^{m} \eta_{j,m}^2 (E_j L_{j-1}^t E_j + L_j R_j L_j^t) \geq 0 \) with \( \eta_0 = 0 \), \( Q \geq 0 \), and \( R_j \geq 0, j = 0, 1, \ldots, m \). Therefore,

\[
\tilde{P} > \phi_m \left( \mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, \tilde{P} \right)
= \mathcal{L}_m(\tilde{P}) + Q + N_m
\geq \mathcal{L}_m(\tilde{P}).
\]

That is, \( \tilde{P} > \mathcal{L}_m(\tilde{P}) \), hence, the function \( \mathcal{L}_m(\mathbf{Y}) \) satisfies the condition of Lemma 4. Considering the definition of \( \tilde{P} \), it yields that

\[
P_{k+1} = \varphi(P_k) \leq \phi_m(\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, P_k)
= \mathcal{L}_m(P_k) + Q + N_m
= \mathcal{L}_m(P_k) + U
\]

where \( U := Q + N_m \geq 0 \). Then based on fact 2) in Lemma 4, it can be concluded that the sequence \( P_k \) is bounded for any \( k \geq 0 \).

Lemma 6: Let \( Y_{s+1} = f(Y_s) \) and \( Z_{s+1} = f(Z_s) \). Suppose that the function \( f(\mathbf{Y}) \) is monotonically increase in \( \mathbf{Y} \). Then:

\[
Y_1 \geq Y_0 \implies Y_{s+1} \geq Y_s, \forall s \geq 0
\]
\[
Y_1 \leq Y_0 \implies Y_{s+1} \leq Y_s, \forall s \geq 0
\]
\[
Y_0 \leq Z_0 \implies Y_k \leq Z_s, \forall s \geq 0.
\]

Proof: The three statements can be similarly proved by mathematical induction. Thus, due to page limitation, we here only prove the first one. Since \( Y_1 \geq Y_0 \), then the first statement is true for \( k = 0 \). Then assume that \( Y_{t+1} \geq Y_t \) holds, so \( Y_{t+2} = f(Y_{t+1}) \geq f(Y_t) = Y_{t+1} \) holds owing to the monotonicity of function \( f(\mathbf{Y}) \).

After building these lemmas above, we are now in a position to establish the sufficient condition for mean square stability of the averaged estimation error covariance matrix.

\[
\text{Theorem 1 (Sufficient condition): Consider the function } \phi_m = \sum_{j=0}^{m} \eta_{j,m}^2 (E_j \phi_{j-1} E_j + L_j R_j L_j^t) \text{ defined in (29). If there exist } m \text{ matrices } L_j, j = 1, 2, \ldots, m \text{ and a positive definite matrix } \mathbf{P} \text{ such that }
\]

\[
\tilde{P} > 0 \text{ and } \tilde{P} > \phi_m (\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_m, \tilde{P}).
\]

Then, the following facts are true:

1) The MARE converges for any initial condition \( P_0 \geq 0 \) and the limit \( \lim_{t \to \infty} P_t = \lim_{k \to \infty} \phi^k_m(P_0) = \tilde{P} \) is independent of the initial condition \( P_0 \).
2) \( \tilde{P} \) is the unique positive definite fixed point of the MARE.

Proof: 1) To begin with, we verify the convergence of the MARE sequence initialized at \( Q_0 = 0 \) and therefore \( Q_k = \varphi^k(0) \). Then it directly follows that \( 0 = Q_0 \leq \varphi(0) = Q_1 \), and in the light of Fact 3 in Lemma 3, it gives that

\[
Q_1 = \varphi(Q_0) \leq \varphi(Q_1) = Q_2.
\]

From Lemma 6 and according to Lemma 5, a monotonically nondecreasing sequence of matrices follow straightforwardly from a simple inductive argument and the sequence is also upper-bounded, that is,

\[
0 = Q_0 \leq Q_1 \leq Q_2 \leq \ldots \leq M_{Q_0}.
\]

Here, one can easily verify that the monotonically nondecreasing and upper-bounded sequence converges from the Bolzano-Weierstrass theorem, that is,

\[
\lim_{k \to \infty} Q_k = \tilde{P}
\]

where \( \tilde{P} \geq 0 \) is a fixed point of the following modified Riccati iteration

\[
\hat{P} = \varphi(\hat{P}).
\]

Then, we show that the modified Riccati iteration initialized at \( S_0 \geq \tilde{P} \) also converges to the same point \( \tilde{P} \). By resorting to (30), it gives that

\[
\tilde{P} = \varphi(\tilde{P}) = \mathcal{L}_{\tilde{P}}(\hat{P}) + Q + N_{\tilde{P}} \geq \mathcal{L}_{\tilde{P}}(\tilde{P})
\]

where \( \mathcal{L}_{\tilde{P}}(\mathbf{Y}) = \sum_{j=0}^{m} \eta_{j,m}^2 \left( E_j \tilde{P} \phi_{j-1} \left( E_j^t \right) \right) \). Therefore, the function \( \mathcal{L}_{\tilde{P}} \) satisfies the condition of Lemma 4. Accordingly, we realize that

\[
\lim_{k \to \infty} \left( \mathcal{L}_{\tilde{P}}(\mathbf{Y}) \right) = 0, \quad \forall \mathbf{Y} \geq 0.
\]

Assume that \( S_0 \geq \tilde{P} \) and then,

\[
S_1 = \varphi(S_0) \geq \varphi(\tilde{P}) = \tilde{P}
\]

where is due to the monotonically increase property of the function \( \varphi(\mathbf{X}) \) and (33). By induction, it establishes that

\[
S_k \geq \tilde{P}, \quad \forall k > 0.
\]
Meanwhile, we have
\[ 0 \leq S_{k+1} - \hat{P} = \varphi(S_k) - \varphi(\hat{P}) = \phi_m(L_{S_k}^1, L_{S_k}^2, \ldots, L_{S_k}^m, S_k) - \phi_m(L_{\hat{P}}^1, L_{\hat{P}}^2, \ldots, L_{\hat{P}}^m, \hat{P}) \leq \phi_m(L_{\hat{P}}^1, L_{\hat{P}}^2, \ldots, L_{\hat{P}}^m, S_k) - \phi_m(L_{\hat{P}}^1, L_{\hat{P}}^2, \ldots, L_{\hat{P}}^m) = \sum_{j=0}^{m} \eta_j \epsilon_j^2 \left( \phi_j^S - \phi_j^P \right) \left( E_j^P \right) \]
\[ = \mathcal{L}_m^P(S_k - \hat{P}). \]

Then, since \( \lim_{k \to \infty} \mathcal{L}_m^P(S_k - \hat{P}) = 0 \), it directly follows that \( \lim_{k \to \infty} (S_{k+1} - \hat{P}) = 0 \). That is, we have shown \( S_k \to \hat{P} \) as \( k \to \infty \) when \( S_0 \geq \hat{P} \).

In the following, we are ready to justify that the modified Riccati iteration \( P_k = \varphi^k(P_0) \) converges to \( \hat{P} \) for all initial conditions \( P_0 \geq 0 \). Let \( Q_0 = 0 \) and \( S_0 = \hat{P} + P_0 \). Then consider the three Riccati iterations initialized at \( Q_0, P_0 \) and \( S_0 \), respectively. Clearly, \( Q_k \leq P_k \leq S_k \), and in the light of Lemma 6, it gives that \( 0 \leq Q_k \leq P_k \leq S_k, \forall k \geq 0 \). Given that both the sequence \( Q_k \) and the sequence \( S_k \) converge to \( \hat{P} \), consequently, we have \( \lim_{k \to \infty} P_k = \hat{P} \).

2) Let us further postulate there exists another positive semi-definite matrix \( \hat{P} \geq 0 \) such that \( \hat{P} = \varphi(\hat{P}) \). Let us consider the Riccati iteration initialized at \( \hat{P} \), and therefore, we can derive the following sequence
\[ \hat{P}, \hat{P}, \hat{P}, \ldots. \]

From analysis above, it has been shown that every Riccati iteration converges to the same limit \( \hat{P} \). Therefore, we have \( \hat{P} = \hat{P} \).

In the sequel, we will provide an example of a scalar-state vector-observation system to justify the existence of sufficient condition in Theorem 1.

**Example:** We consider the following system
\[ x_{k+1} = ax_k + \omega_k, \]
\[ y_k = Cx_k + \nu_k, \]
where \( a = 1.2 \), \( C' = [c_1, c_2] = [1, 1] \), noise covariances are \( q = 1 \) and \( R = \text{diag}(1, 2) = \text{diag}(0.1, 1) \). For simplicity, consider \( \lambda_1 = \lambda_2 = 0 \), and let \( l_1, l_2 \) be, for instance, such that \( l_1 = -1, -2.8276 \leq l_2 < 0.8276 \) or \( l_2 = -1, -2.8276 < l_1 < 0.8276 \). Then one can always find \( p > 0 \) such that \( l_1, l_2, p \) satisfy condition (32) in Theorem 1. That is, the expected estimation error covariance matrix will converge.

In the ensuing part, we will present one necessary condition for ensuring mean square stability of expected estimation error covariance matrix which extends the result in [18] to general linear systems with data packet drops.

**Theorem 2** (Necessary condition): Consider system (1) and Algorithm 1. Assume that \( A \) is unstable, that \( (A, Q^{1/2}) \) is controllable and that \( (C, A) \) is observable. If \( E[P_k] \leq M_{P_0}, \forall k \geq 0 \) holds for any initial condition \( P_0 \geq 0 \), then \( \lambda_1, \lambda_2, \ldots, \lambda_m \) defined in (15) should satisfy the following condition
\[ \prod_{i=1}^{m} \left( 1 - \lambda_i \right) \leq \frac{1}{\left( \max_i \left| \sigma_i(A) \right| \right)^2} \tag{34} \]
where \( \sigma_i(A), i = 1, 2, \ldots, n \), are all eigenvalues of square matrix \( A \), and \( M_{P_0} > 0 \) depends on the initial condition \( P_0 \geq 0 \).

**Proof:** The proof follows straightforwardly from Fact 7) in Lemma 3.

**V. Concluding Remarks**

In this paper we devised a measurement innovation component-wise based power scheduler for wireless sensors in terms of optimally deciding whether to use a high or low transmission power to communicate each component of a measurement to the remote estimator side. The high transmission power is used to transmit the well-defined “important” measurements and low transmission power to transmit the less “important” measurements. Meanwhile, the high power transmission power is assumed to lead to reliable data flow while the low transmission power leads to unreliable data flow, that is, data packet drops. Under this new framework, the MMSE estimator was derived. Then convergence analysis of the averaged estimation error covariance was provided and moreover, both the sufficient condition and necessary condition guaranteeing its convergence were established for general linear stochastic systems. Since the assumption of modeling the arrival of measurements as independent Bernoulli i.i.d. processes can be clearly improved upon, and therefore, future work will concentrate on accounting for communication channel modeling in this filtering framework [11], [31].

**Acknowledgment**

The authors would like to express thanks to the anonymous reviewers for their insightful and constructive comments that helped improving the quality of this paper.

**References**

[1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless sensor networks: a survey,” Comput. networks, vol. 38, no. 4, pp. 393–422, Mar. 2002.

[2] A. Ribeiro and G. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks—Part I: Gaussian case,” IEEE Trans. Signal Process., vol. 54, no. 3, pp. 1131–1143, Mar. 2006.

[3] ——, “Bandwidth-constrained distributed estimation for wireless sensor networks—Part II: Unknown probability density function,” IEEE Trans. Signal Process., vol. 54, no. 7, pp. 2784–2796, Jul. 2006.

[4] A. Ribeiro, G. Giannakis, and S. Roumeliotis, “SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations,” IEEE Trans. Signal Process., vol. 54, no. 12, pp. 4782–4795, Dec. 2006.

[5] E. Msechu, S. Roumeliotis, A. Ribeiro, and G. Giannakis, “Decentralized quantized Kalman filtering with scalable communication cost,” IEEE Trans. Signal Process., vol. 56, no. 8, pp. 3727–3741, Aug. 2008.

[6] E. Msechu and G. Giannakis, “Sensor-centric data reduction for estimation with WSNs via censoring and quantization,” IEEE Trans. Signal Process., vol. 60, no. 1, pp. 400–414, Jan. 2012.

[7] Q. Jia, L. Shi, Y. Mo, and B. Sinopoli, “On optimal partial broadcasting of wireless sensor networks for Kalman filtering,” IEEE Trans. Autom. Control, vol. 57, no. 3, pp. 715–721, Mar. 2012.

[8] A. Ribeiro, I. Schizas, S. Roumeliotis, and G. Giannakis, “Kalman filtering in wireless sensor networks,” IEEE Control Syst. Mag., vol. 30, no. 2, pp. 66–86, Apr. 2010.

[9] I. Schizas, G. Giannakis, and Z. Luo, “Distributed estimation using reduced-dimensionality sensor observations,” IEEE Trans. Signal Process., vol. 55, no. 8, pp. 4284–4299, Aug. 2007.

[10] I. Xiao, A. Ribeiro, Z. Luo, and G. Giannakis, “Distributed compression-estimation using wireless sensor networks,” IEEE Signal Process. Mag., vol. 23, no. 4, pp. 27–41, Jul. 2006.
