Stochastic Event-based Sensor Schedules for Remote State Estimation in Cognitive Radio Sensor Networks

Lingying Huang, Jiazheng Wang, Enoch Kung, Yilin Mo, Junfeng Wu and Ling Shi

Abstract—We consider the problem of communication allocation for remote state estimation in a cognitive radio sensor network (CRSN). A sensor collects measurements of a physical plant, and transmits the data to a remote estimator as a secondary user (SU) in the shared network. The existence of the primal users (PUs) brings exogenous uncertainties into the transmission scheduling process, and how to design an event-based scheduling scheme considering these uncertainties has not been addressed in the literature. In this work, we start from the formulation of a discrete-time remote estimation process in the CRSN, and then analyze the hidden information contained in the absence of data transmission. In order to achieve a better tradeoff between estimation performance and communication consumption, we propose both open-loop and closed-loop schedules using the hidden information under a Bayesian setting. The open-loop schedule does not rely on any feedback signal but only works for stable plants. For unstable plants, a closed-loop schedule is designed based on feedback signals. The parameter design problems in both schedules are efficiently solved by convex programming. Numerical simulations are included to illustrate the theoretical results.

Index Terms—Stochastic event-based schedule; Cognitive radio sensor network; Minimum mean squared error; Branch-and-bound algorithm.

I. INTRODUCTION

Recently, cognitive radio (CR) which dynamically assigns the radio resources is applied in 5G Internet of things (IoT) applications [1]. CR, first proposed by Mitola et al. [2] in 1999, is a promising technology to cope with the spectrum scarcity problem. A CR sensor network (CRSN) is a network of dispersed wireless sensor nodes embedded with cognitive radio capability which enables them to dynamically access unused licensed spectrum bands for data transmission while performing conventional wireless sensor nodes’ tasks [3]. An example of that is shown in Fig. 1. If the primary users (PUs), as the licensed user (mobile phone), vacate the spectrum, secondary users (SUs), e.g., the sink, equipped with CR devices can then access the spectrum to transmit packets [4]. Minimizing the communication rate of the SU while satisfying the estimation performance is worth studying in this shared network.

Proper sensor scheduling, which is introduced to cope with limited transmissions, could improve estimation quality. An event-based mechanism achieves better performance than an off-line mechanism [5]–[7] as the online information can be leveraged to improve the estimation quality, has attracted increasing attention in recent years. Astrom and Bernhardsson [8] first showed that an event-based approach outperforms a periodic approach (Riemann sampling) in a first-order stochastic system. The event-triggered mechanisms proposed by Xia et al. [9] and Trimpe et al. [10] require that the sensor has a computational capability to run a local Kalman filter and obtain a local state estimate. In realistic scenarios, however, the sensors may be primitive and have limited computational capability. Based on that condition, Wu et al. [11] derived a minimum mean squared error (MMSE) estimate on the remote estimator under a deterministic event-triggered scheduler. Since finding the exact MMSE estimate is intractable due to the computational complexity, an approximated estimator based on a Gaussian assumption is further derived. To preserve the Gaussian property, stochastic event-triggered sensor schedulers are proposed by Han et al. [12].

Different from traditional studies, in which the radio access network is statically assigned, the existence of PUs introduces an exogenous uncertainty to the SU base scheduling scheme. There is a limited amount of works on optimizing the scheduling scheme of CRSNs. Deng et al. [13] studied how to activate separately non-disjoint sensor groups to extend the network lifetime. Mabrouk et al. [14] introduced opportunistic time slot assignment scheduling scheme to minimize the schedule length and maximize the throughput. All the above setups consider continuous-time measurements of the SU. Minimizing the transmission collision from a probabilistic point of view is important since it is very energy-consuming or even impossible to check the spectrum availability continuously. Moreover, the above studies neglect the information importance.

In this paper, we consider a discrete-time remote estimation process in a CRSN. Unlike previous studies, the SU can check the spectrum availability before each transmission. Moreover, we use the event-triggered mechanism to capture the information importance. To the best of our knowledge, an event-based mechanism for remote state estimation has not been studied in this new but widely-used network structure.
The big challenge is that the exogenous uncertainty in the shared network in addition to the stochastic property of triggering law makes the uncertainties coupled. Kung et al. [13] showed that the Gaussian property cannot be preserved due to coupled uncertainty induced by packet drops. Xu et al. [16] utilizes a Gaussian mixture model to obtain a closed-form MMSE estimator for the packet-dropping scenarios. However, the computational complexity grows exponentially. To cope with this challenge, we utilize the hidden information contained in the absence of transmission data in the CRSN to decouple those uncertainties. To be more specific, since the remote estimator can distinguish the source of the received packet, not triggering is inferred when no packet is received.

Another problem is that, the exogenous uncertainty makes the packet-dropping scenarios. Xu et al. [16] showed that the Gaussian property cannot be preserved due to coupled uncertainty induced by packet drops. Xu et al. [16] utilizes a Gaussian mixture model to obtain a closed-form MMSE estimator for the packet-dropping scenarios. However, the computational complexity grows exponentially. To cope with this challenge, we utilize the hidden information contained in the absence of transmission data in the CRSN to decouple those uncertainties. To be more specific, since the remote estimator can distinguish the source of the received packet, not triggering is inferred when no packet is received.

Another problem is that, the exogenous uncertainty makes the packet-dropping scenarios. Xu et al. [16] showed that the Gaussian property cannot be preserved due to coupled uncertainty induced by packet drops. Xu et al. [16] utilizes a Gaussian mixture model to obtain a closed-form MMSE estimator for the packet-dropping scenarios. However, the computational complexity grows exponentially. To cope with this challenge, we utilize the hidden information contained in the absence of transmission data in the CRSN to decouple those uncertainties. To be more specific, since the remote estimator can distinguish the source of the received packet, not triggering is inferred when no packet is received.
its coverage could send information to the remote estimator \cite{17}. Without loss of generality, we use the aggregated $x_k$ and $y_k$, thus we only need to study one SU case. Let $\eta_k \in \{0, 1\}$ represent the channel’s availability. If $\eta_k = 0$, the channel is occupied by the PU, and vice versa. We assume $\eta_k$ evolves as an i.i.d. Bernoulli random process with $\mathbb{E}[\eta_k] = \lambda \in (0, 1]$, which is widely used in [18–22].

The remote estimator can identify the packet source. If the remote estimator receives no packet, it means that the channel condition is idle for the sensor to transmit packet but the scheduler decides not to send, i.e., $\eta_k = 1$ and $\epsilon_k = 0$. If the remote estimator receives a packet from other sensors, the channel condition is unfavorable at this time step, i.e., $\eta_k = 0$. In this case, there is no information of $\epsilon_k$, and we can set it as $\epsilon_k = 0$. Otherwise, the remote estimator receives the packet from this sensor, i.e., $\eta_k = 1$ and $\epsilon_k = 1$. The following information is available to the estimator at time $k$

$$\mathcal{I}_k \triangleq \{\eta_k\}_0^1 \cup \{\epsilon_k\}_0^1 \cup \{\eta_k \epsilon_k y_k\}_0^1,$$

with $\mathcal{I}_{-1} = \emptyset$. Further define the following notations which will be used in subsequent analysis:

$$\hat{x}_k = \mathbb{E}[x_k | \mathcal{I}_{k-1}], \quad \tilde{y}_k = \mathbb{E}[y_k | \mathcal{I}_{k-1}],$$

$$e_k = x_k - \hat{x}_k, \quad P_k = \mathbb{E}[e_k e_k^\top], \quad \hat{x}_k = \mathbb{E}[x_k | \mathcal{I}_k], \quad e_k = x_k - \hat{x}_k, \quad P_k = \mathbb{E}[e_k e_k^\top].$$

The estimates $\hat{x}_k$ and $\hat{x}_k$ are the \textit{a priori} and the \textit{a posteriori} MMSE estimate, respectively. Meanwhile, $P_k$ and $P_k$ are the \textit{a priori} and the \textit{a posteriori} estimation error covariance, respectively.

We adopt the stochastic event-triggered scheduling schemes in \cite{12} as below. At each time step, the sensor generates an i.i.d. random variable $\zeta_k$ which is uniformly distributed over $[0, 1]$, denoted as $\zeta_k \sim U(0, 1)$. The transmission decision by the sensor, i.e., $\epsilon_k$, follows two event-triggered criteria.

1) Open-loop scheduler: The sensor makes the decision based on the current raw measurement $y_k$, i.e.,

$$\epsilon_k = \begin{cases} 1, & \text{if } \zeta_k > \exp(-\frac{1}{2}y_k^\top Y y_k), \quad Y > 0, \\ 0, & \text{otherwise}, \end{cases} \quad (2)$$

2) Closed-loop scheduler: The sensor receives a feedback $\tilde{y}_k$ from the remote estimator; then the decision is based on the measurement innovation $z_k = y_k - \tilde{y}_k$ as

$$\epsilon_k = \begin{cases} 1, & \text{if } \zeta_k > \exp(-\frac{1}{2}z_k^\top Z z_k), \quad Z > 0, \\ 0, & \text{otherwise}, \end{cases} \quad (3)$$

The open-loop scheduler is easier to implement since it does not require any feedback. However, open-loop schedulers cannot reduce the communication rate for unstable systems since $\epsilon_k = 1$ almost surely occurs for any given $Y$ after a long time \cite{12}. Thus we need closed-loop schedulers to reduce the communication rate for unstable systems.

**Remark 1.** We choose these scheduling schemes because they preserve the Gaussian property which will be exploited in later work to obtain the linear recursion of update. This refrains from nonlinear complicated and approximate estimation, compared with the other existing event-triggered mechanism, e.g., \cite{11} and \cite{23}.

### B. Problem of Interest

Define the average communication rate as

$$\gamma \triangleq \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[\eta_k \epsilon_k] = \lambda \limsup_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} \mathbb{E}[\epsilon_k]. \quad (4)$$

Since the sequence $\{\eta_k\}_{0}^{\infty}$ has no relationship with the measurement, the iteration of the error covariance is stochastic and cannot be determined offline. Therefore, we are interested in its statistical properties. Define the mean error covariance of the system at time $k$ as $\mathbb{E}[P_k]$. We are interested in the following problem:

**Problem 1.**

$$\min \gamma \quad \text{s.t.} \mathbb{E}[P_k] \leq M,$$

where $M > 0$ is a given matrix-valued bound.

We study two extreme cases to demonstrate that the event-triggered parameter influences $\gamma$ and $\mathbb{E}[P_k]$. Note that if $Y = 0$, $\epsilon_k = 0$ almost surely occurs. Therefore, if $\gamma = 0$, the mean error covariance of the remote estimator diverges for an unstable system, i.e., $\mathbb{E}[P_k] \to \infty$. On the other hand, for sufficiently large $Y$ such that $\exp(-\frac{1}{2}y_k^\top Y y_k) = 0$ almost surely occurs, we have $\gamma = \lambda$. The error covariance converges if $\lambda > 1 - \rho(A)^2$ \cite{24}. The same analysis applies to $Z$. To avoid trivial problems, we assume this condition is satisfied in the following analysis.

It is obvious that the parameter $Y$ or $Z$ introduces an additional degree of freedom to balance the tradeoff between the communication rate and the mean error covariance. However, it is difficult to solve Problem \[1] directly since both the objective and constraint are implicit functions of $Y$ or $Z$. One core problem lies in whether we are able to obtain the explicit expression of the communication rate and the mean error covariance in terms of $Y$ or $Z$. If not, we expect to find some bounds of the mean error covariance. In this paper, we will focus on the derivation of the communication rate and the mean error covariance in terms of $Y$ or $Z$ and then design the parameters to achieve a desired tradeoff. Besides, we will also explore an explicit MMSE estimator since it is also critical for the system implementation.

### III. OPEN-LOOP SCENARIO

In the open-loop scenario, the main difficulty is that, due to the randomness of $\{\eta_k\}_{0}^{\infty}$, only the mean error covariance is bounded.

Since the open-loop scheduler \cite{12} only reduces the communication rate for stable systems, we study the stable system in the open-loop case. We assume in the sequel that the system has already entered into the steady state, which implies that

$$P_0^- = \text{Cov}(x_k) = \Sigma, \text{Cov}(y_k) = \Pi,$$

where $\Sigma = A\Sigma A^\top + Q$, $\Pi = C\Sigma C^\top + R$.

Given $Y$, the average communication rate \cite{12} is

$$\gamma = \lambda \left(1 - (\det(I + P Y))^\frac{1}{2}\right). \quad (5)$$
Define functions $h$, $g_{\theta,W} : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as follows:

$$h(X) \triangleq AXA^T + Q,$$

$$g_{\theta,W}(X) \triangleq AXA^T + Q - \theta AXC^T(CXC^T + W)^{-1}CXA^T,$$

where $X > 0$, $W > 0$ and $\theta \in (0,1]$. The function $h$ can be interpreted as the recursive function of the estimation error covariance matrix when the channel is not available while the function $g$ is the modified algebraic Riccati equation for the Kalman filter with intermittent observations [24]. If $\theta = 1$, $g_{1,W}$ will be written as $g_{W}$ for brevity. The propositions of function $g_{\theta,W}(X)$ are shown in Appendix A.

**Theorem 1.** The MMSE estimate under an open-loop scheduler is computed as follows. Start from the initial condition $\hat{x}_0 = 0$ and $P_0^T = \Sigma$.

**Measurement Update:**

$$K_k = \eta_k P_k^{-1}C^T(CP_k^{-1}C^T + R + (1 - \varepsilon_k)Y^{-1})^{-1},$$

$$\hat{x}_k = \hat{x}_{k-1} + \varepsilon_k K_k y_k - K_k \hat{y}_k = (I - K_k C) x_{k-1} + \varepsilon_k K_k y_k,$$

$$P_k = P_{k-1} - K_k C P_{k-1} C^T,$$

(6)

**Time Update:**

$$\hat{x}_{k+1}^T = A \hat{x}_k, P_{k+1}^T = h(P_k).$$

(7)

**Proof.** See Appendix B.

The proof of this theorem uses the Gaussian property of the distribution proved in [12]. By exploiting the Gaussian property, the recursion of the update is linear, which reduces computational complexities. Thus, it is practical to use in the CRSN.

By exploiting the concavity, monotonicity and limit property, the asymptotic upper and lower bounds on $E[P_k]$ are shown as Lemma 1

**Lemma 1.** The mean error covariance $E[P_k]$ satisfies

$$g_{\hat{P}_k}^T(\Sigma) \leq E[P_k] \leq g_{\hat{P}_k,R+Y^{-1}}^T(\Sigma),$$

where $R_t^{-1} = R^{-1} + (\lambda - \gamma)(R + Y^{-1})^{-1}$.

The asymptotic upper and lower bounds on $E[P_k]$ are

$$\bar{X}_{ol} \leq \lim \inf_{k \to \infty} E[P_k] \leq \lim \sup_{k \to \infty} E[P_k] \leq \bar{X}_{ol},$$

where $\bar{X}_{ol} > 0$ is the unique solution to $\bar{X}_{ol} = g_{R_1}(\bar{X}_{ol})$, and $\bar{X}_{ol} > 0$ is the unique solution to $\bar{X}_{ol} = g_{R_1,R+Y^{-1}}(\bar{X}_{ol})$.

For all schedules satisfy (3) we obtain that $X_{ol} > 0$, where $X_{ol} > 0$ is the unique solution to

$$X_0 = g_{R_{\lambda}/\lambda}(X_0).$$

(9)

**Proof.** See Appendix C.

**Remark 2.** By applying the information filtering and exploiting the convexity of $X^{-1}$, we obtain a different lower bound on $E[P_k]$, i.e., $\bar{X}_{ol}$. We plot it with respect to (w.r.t.) $\gamma$ in Fig. 4. It is different from the lower bound derived in [24], which is denoted as $\bar{X}_{ol}$, where $\bar{X}_{ol} = (1 - \lambda)AX_{ol}A^T + Q$. The matrix $X_\lambda$ is the lower bound of $E[P_k]$ for all schedules. When $\lambda = 1$, the lower bound derived in our paper is larger than $X_{ol}$, i.e., $X_{ol} > Q = X_{ol}$. For scalar systems, we can choose max{$X_{ol}, X_{ol}$} to be the lower bound.

From the above analysis, we relax Problem 1 to bound the asymptotic upper bound on the mean error covariance, i.e., $\bar{X}_{ol}$.

**Problem 2.**

$$\min_{\gamma} \gamma$$

$$s.t. Y \geq 0, \bar{X}_{ol} \leq M.$$

We observe that for the scalar case ($Y \in \mathbb{R}$), the above problem can be easily solved by convex programming. However, for the general vector cases, Problem 2 is not convex since (5) is a log concave function of matrix $Y$. We need the following lemma to study the general vector case.

**Lemma 2.** [12] Lemma 2 Given $\gamma$ in (5), $\Pi, Y \in \mathbb{S}_+^n$, the following inequality holds:

$$f_1(\text{tr}(\Pi Y)) \leq \gamma \leq f_2(\text{tr}(\Pi Y)),$$

(11)

where $f_1(x) = \lambda(1 - (1 + x)^{-\frac{1}{2}})$, $f_2(x) = \lambda(1 - \exp(-\frac{1}{2}x))$. The equality is only satisfied when $\text{tr}(\Pi Y) = 0$.

Using Lemma 2, the objective of Problem 1 is bounded by two increasing functions. Thus, it can be relaxed into $\min \text{tr}(\Pi Y)$.

**Problem 3.**

$$\min_{\gamma} \text{tr}(\Pi Y)$$

$$s.t. Y \geq 0, \bar{X}_{ol} \leq M, \bar{X}_{ol} = g_{\lambda,R+Y^{-1}}(\bar{X}_{ol}).$$

We transform Problem 3 into an SDP problem using Theorem 2

**Theorem 2.** Problem 3 is equivalent to

$$\min_{S,Y} \text{tr}(\Pi Y)$$

$$s.t. \Psi(S,Y) \geq 0, \left[ \begin{array}{cc} S & I \\ I & M \end{array} \right] \geq 0, Y \geq 0,$$

(12)

where $\Psi(S,Y)$ is defined as (10).

**Proof.** The proof mainly follows two steps. First we prove equivalent LMIs to replace the implicit function $g_{\lambda,R+Y^{-1}}$. Then the function $g_{\lambda,R+Y^{-1}}(X) \leq X$ is transformed into an equivalent SDP constraint. The details are shown in Appendix D.

**Remark 3.** In [2], it also derives an SDP constraint from $g_{\lambda,W}(X) \leq X$, but it neglected the influence of $Q$ and $W$ through relaxation. However, in our case, since $W$ corresponds to the decision variable $Y$, we cannot eliminate the influence.

Define $\gamma^*$ as the communication rate with optimal $Y^*$ in Problem 3. Let the optimal solution to Problem 2 be $Y^{opt}$ and the minimum objective be $\gamma^{opt}$. Define the gap $\kappa$ as $\kappa \triangleq \gamma^* - \gamma^{opt}$. By (11), one has

$$0 < \kappa < \lambda \left(1 + \text{tr}(\Pi Y^*)\right)^{-\frac{1}{2}} - (\det(I + \Pi Y^*))^{-\frac{1}{2}}.$$  

(13)

Fig. 3 shows the relationship between the problems.
The asymptotic upper bound on $\mathbb{E}[P_k^-]$ is

$$\limsup_{k \to \infty} \mathbb{E}[P_k^-] \leq \bar{X}_{cl},$$

where $\bar{X}_{cl} > 0$ is the unique solution to $\bar{X}_{cl} = g_{\lambda,R+Z^{-1}}(\bar{X}_{cl})$.

From Lemma 3 and Lemma 4, denote the upper bound of the average communication rate as

$$\bar{\gamma} = \lambda \left( 1 - \left( \text{det} \left( I + (C\bar{X}_{cl}C^T + R)Z \right) \right)^{-\frac{1}{2}} \right).$$

Substituting the upper bound of the communication rate, we further obtain the asymptotic lower bound on $\mathbb{E}[P_k^-]$ as follows.

**Lemma 5.** The mean error covariance $\mathbb{E}[P_k^-]$ is bounded by

$$g^k_{R_1}(\Sigma) \leq \mathbb{E}[P_k^-],$$

where $R_1^{-1} = \bar{\gamma}R^{-1} + (\lambda - \bar{\gamma})(R + Z)^{-1}$.

The asymptotic lower bound on $\mathbb{E}[P_k^-]$ is

$$\underline{X}_{cl} \leq \liminf_{k \to \infty} \mathbb{E}[P_k^-],$$

where $\underline{X}_{cl} > 0$ is the unique solution to $\underline{X}_{cl} = g_{\lambda,R+Z^{-1}}(\underline{X}_{cl})$.

Meanwhile, one has $\underline{X}_{cl} \geq X_0$, where $X_0$ from (9) is the lower bound of $\mathbb{E}[P_k^-]$ for all schedulers satisfied (5).

**Proof.** The proof of Lemma 3 and Lemma 4 is similar to that of Lemma 2 as shown in Appendix C. Hence, we omit this part. □

**Remark 4.** We observed that the covariance of $z_k$ is smaller than the covariance of $y_k$. Thus, to achieve the same communication rate, the matrix $Z$ under the closed-loop scheduler is larger than $Y$ for the open-loop scheduler. As a result, the closed-loop scheduler achieves better estimation quality compared with the open-loop case.

In the closed-loop scenario, since there is no closed-form of $\lambda$ given $Z$, we relax the objective in Problem 1 by the upper bound $\bar{\gamma}$ and the asymptotic upper bound on the mean error covariance $\bar{X}_{cl}$.

**Problem 4.**

$$\min_{\bar{\gamma}} \bar{\gamma}$$

s.t. $\bar{X}_{cl} \leq M, Z \geq 0$.

We also take Lemma 2 to relax Problem 4 to Problem 5 which is equivalent to Problem 13 in Theorem 4 as further proved. Fig. 4 shows the relationship between each problem directly.
used to reduce the dimension of the optimization variables. In real implementation of the algorithms, half vectorization representation is used to simplify the notation, we use vectorization throughout the paper.

**Problem 5.**
\[
\min_{T,Z} \text{tr}(TZ)
\]
\[
s.t. \tilde{X}_{cl} \leq M, CX_{cl}C + R \leq T, \tilde{X}_{cl} = g_{\lambda,R} + (X_{cl}), Z \geq 0.
\]

**Theorem 4.** Problem 5 is equivalent to
\[
\min_{Z,S,X} f(X, Z) = (CXC^T + R)Z
\]
\[
s.t. \begin{bmatrix} X & I \\ S & M \end{bmatrix} \succeq 0, \begin{bmatrix} S & I \\ I & M \end{bmatrix} \succeq 0, \Psi(S, Z) \geq 0, Z \geq 0.
\]
(18)

**Proof.** See Appendix C.

To simplify the notation, let \( S \) represent the constraints of the decision variables as specified in (18), i.e.,
\[
S = \{(X, Z, S) : \begin{bmatrix} X & I \\ S & M \end{bmatrix} \succeq 0, \begin{bmatrix} S & I \\ I & M \end{bmatrix} \succeq 0, \Psi(S, Z) \geq 0, Z \geq 0\}.
\]
(19)

**B. Jointly Constrained Biconvex Programming**

Problem (18) is a bilinear problem w.r.t. the PSD cone \( X \) and \( Z \). Linear operations, such as matrix vectorization, are applied to convert it into a general form of the jointly bilinear program [25]. Define
\[
\tilde{x} \triangleq \text{vec}(X), \tilde{y} \triangleq \text{vec}(Z)
\]
It is obvious that \( \tilde{x} \) and \( \tilde{y} \) are bijections of \( X \) and \( Z \). Thus, the optimization parameters can be changed to \( (\tilde{x}, \tilde{y}, S) \) and the objective function is a bilinear function w.r.t. \( \tilde{x} \) and \( \tilde{y} \) as follows
\[
f(S, P) = \phi(\tilde{x}, \tilde{y}) = \tilde{x}^T G \tilde{y} + g(\tilde{y}),
\]
where
\[
G = C^T \otimes C^T, g(\tilde{y}) = \text{tr}(RZ).
\]
(20)
(21)

**Lemma 6 (Boundary Solution).** [25][Theorem 1] The jointly bilinear Problem (18) has boundary solutions if the feasible region is compact.

Note that \( S \) (19) is closed, convex set. The feasible region being compact is a necessary condition to implement the jointly constrained biconvex programming. Hence, we study the bounds of each element of \( \tilde{x} \) and \( \tilde{y} \). Equivalently, we study the compact region of each element of \( X \) and \( Z \). The main difficulty remains that from the initial set \( S \), there is no obvious upper bound of the elements of matrix \( Z \) (\( \tilde{y} \)) in the constraints. The upper bound of \( Z \) is from the objective which aims to minimize \( \text{tr}((CXC^T + R)Z) \). Therefore, we derive a necessary condition w.r.t. \( Z \) for the optimal solution as an upper bound requirement. The details are shown in Lemma 7.

**Lemma 7.** The optimal solution \((X^*, Z^*)\) belongs to the set
\[
\Omega = \{(X, Z) : j_{ij} \leq X_{ij} \leq J_{ij}, d_{ij} \leq Z_{ij} \leq z^*\},
\]
where
\[
\begin{align*}
  j_{ij} &= \begin{cases} (X_0)_{ii}, & i = j, \\
                 - (M_{ii}M_{jj})^2, & \text{else,}
\end{cases} \\
  d_{ij} &= \begin{cases} 0, & i = j, \\
                  -z^*, & \text{else,}
\end{cases}
\end{align*}
\]
and \( z^* \) satisfies the following optimization problem
\[
\begin{align*}
  z^* &= \min_{x,y} \text{tr}(CMC^T + R)
  \\
  \text{s.t.} & \begin{bmatrix} S & I \\ I & M \end{bmatrix} \geq 0, \Psi(S, Z) \geq 0, Z \geq 0.
\end{align*}
\]
(22)

**Proof.** See Appendix D.

Adding a linear constraint \( \tilde{z} = G\tilde{y} \) to those defining \( S \), the bounds
\[
\begin{align*}
  n_k &= \min \{(G\tilde{y})_k : d_{ij} \leq Z_{ij} \leq z^*\}, \\
  N_k &= \max \{(G\tilde{y})_k : d_{ij} \leq Z_{ij} \leq z^*\}
\end{align*}
\]
replace the bounds on \( \tilde{y} \) in defining \( \Omega \).

**C. Branch-and-bound (B&B) Method**

B&B method intends to construct an increasing convex underestimator with an associated decreasing upper bound for the jointly constrained bilinear problem. Before the presentation of the convex underestimator, we first introduce a general lemma for any \( \Omega \). The convex envelope of \( \tilde{x}^T \tilde{z} \) over \( \Omega \) is the pointwise supremum of all convex functions which underestimate \( \tilde{x}^T \tilde{z} \) over \( \Omega \), denoted by \( \text{Vex}_{\Omega}(\tilde{x}^T \tilde{z}) \).

**Lemma 8 (Convex Envelop).** [25] Corollary If \( x, y \in \mathbb{R}^n \) and \( (x, y) \in \Omega \), where \( \Omega = \{(x, y) : t \leq x \leq T, d \leq y \leq D\} \), \( \Omega_i = \{(x_i, y_i) : t_i \leq x_i \leq T_i, d_i \leq y_i \leq D_i\} \),
\[
\text{Vex}(x^T y) = \sum_{i=1}^n \text{Vex}(x_iy_i),
\]
\[
\text{Vex}(x_iy_i) = \max \{d_i x_i + t_i y_i - d_i t_i, D_i x_i + T_i y_i - T_i D_i\}.
\]
Moreover, \( x^T y \geq \text{Vex}_{\Omega}(x^T y) \) for all \( (x, y) \in \Omega \), and the “=“ holds iff \((x, y) \in \partial\Omega \).

Let \( \psi^1(\tilde{x}, \tilde{y}) = \text{Vex}_{\Omega} (\tilde{x}^T \tilde{z}) + g(\tilde{y}) \) and \( \Omega^1 = \Omega \) from Lemma 7. Note that \( \psi^1 \) is the convex underestimator of \( \phi \) and, furthermore, agrees with \( \phi \) on \( \partial\Omega \). Solving the convex problem \( v^1 = \min \psi^1(\tilde{x}, \tilde{y}), s.t.(X, Z, S) \in S \cap \Omega^1 \) (denoted as \( \mathcal{P}^1 \)) yields the optimal value \( v^1(\psi^1) = \psi^1(\tilde{x}^1, \tilde{y}^1) \). If \( v^1 \) is \( \gamma^1 \), where \( \gamma^1 = (\tilde{x}^1, \tilde{y}^1) \) is a solution to (18). Otherwise, one has
\[
\Delta_i^1 = \tilde{x}_i^1 \tilde{z}_i^1 - \text{Vex}(\tilde{x}_i^1 \tilde{z}_i^1) > 0, \text{ for some } i.
\]
(24)

We choose the index I which produces the largest difference \( \Delta_i^1 \), and split the I-th rectangle into four subrectangles \( \Omega^{2t} \) according to the rule illustrated in Fig 5.
The result of this splitting serves to set up four new subproblems, \( \psi^{2t}(\bar{x}, \bar{y}), \) s.t. \((X, Z, S) \in S \cap \Omega_t^{2t} \) (denoted as \( \mathcal{P}_t^{2t} \)) at stage 2, where \( \psi^{2t}(\bar{x}, \bar{y}) = \text{Vex}_{\Omega_t^{2t}}(\bar{x}^T z) + g(\bar{y}) \). More general, at stage \( k \), the convex problem denoted as Problem \( \mathcal{P}_t^{kt} \) follows

\[
\begin{align*}
\mathcal{P}_t^{kt} : \min_{X,Z,S} \psi^{kt}(\bar{x}, \bar{y}), \text{s.t. } (X, Z, S) &\in S \cap \Omega_t^{kt},
\end{align*}
\]

where \( \psi^{kt}(\bar{x}, \bar{y}) = \text{Vex}_{\Omega_t^{kt}}(\bar{x}^T z) + g(\bar{y}) \). The optimal solution to \( \mathcal{P}_t^{kt} \) is \( \psi^{kt} = \psi^{kt}(\tilde{x}^k, \tilde{y}^k) \) and \( Y^k = \phi(\tilde{x}^k, \tilde{y}^k) \). The upper bound at stage \( k \) is \( Y^k \triangleq \min \{ \phi : \text{all the visited optimal solution at stage } k \} \), and may be expressed recursively as

\[
Y^k = \min \{ Y^{k-1}, Y^1, Y^{k-2}, Y^3, Y^k \}. \tag{26}
\]

If \( \psi^{kt} > Y^k \), the subspace will be eliminated from the further consideration. Therefore, we only record the boundary \( \Omega_t^{kt} \) and the optimal solution \( (\tilde{x}^k, \tilde{y}^k) \) as an open node \((k, t)\) if \( \psi^{kt} \leq Y^k \). The lower bound at stage \( k \) is

\[
\nu^k \triangleq \min \{ \nu^{lj} : \text{node } (l, j) \text{ is open at stage } k \}. \tag{27}
\]

Moving from stage \( k \) to stage \( k + 1 \) involves the selection of an open node (lines 1 – 3), and the creation of four new nodes from that node (lines 4 – 9) as shown in Algorithm 1.

Since the procedure converges to a globally optimal solution [25], once we have any \( \nu^k = Y^k \), the optimal solution is obtained as \( (\tilde{x}^k, \tilde{y}^k) \). Due to the computational consideration, the algorithm can be terminated at a prespecified \( \epsilon \) degree of accuracy whenever \( \nu^k \geq Y^k - \epsilon \). The algorithm is summarized in Algorithm 2.

\[ \text{Algorithm 2} \]

\[ \epsilon \text{— accuracy B&B Algorithm} \]

**Input:** \( A, C, Q, R, W, U, \Omega \);

**Output:** \( X^*, Z^*, v^* \)

1: Initialize: \( \Omega_1 \leftarrow \Omega, k \leftarrow 1; \)
2: Solve the convex problem \( \mathcal{P}^1 \) and obtain the optimal solution \( (\tilde{x}^1, \tilde{y}^1) \) with the lower bound \( v_1 \) and the upper bound \( \Upsilon^1 \);
3: while \( v^k < \Upsilon^k - \epsilon \) do
4: Move from stage \( k \) to stage \( k + 1 \) by Algorithm 1
5: for \( t = 1 : 4 \) do
6: Solve each of the four problem \( \mathcal{P}^{kt} \) in turn and obtain the point \( (\tilde{x}^{kt}, \tilde{y}^{kt}) \) with the value \( v^{kt} \) and \( \Upsilon^{kt} \);
7: For \( v^{kt} \leq \Upsilon^k \), record it as an open node \((k, t)\);
8: Update \( Y^k \) by (26) and \( v^k \) by (27);
9: end while
10: Let \( v^* = v^k, \text{vec}(X^*) = \tilde{x}^{kt} \) and \( \text{vec}(Z^*) = \tilde{y}^{kt} \).

By the same method in the open-loop case, the optimality gap
\[
\kappa = \gamma^* - \gamma_{\text{opt}}
\]

\[ = \lambda \left( (1 + v^*)^{-\frac{1}{2}} - \left( \det \left( I + (CX^*C^T + R)Z^* \right) \right)^{-\frac{1}{2}} \right). \]

V. SIMULATION

Numerical examples are provided to illustrate the results.

A. Policy Assessment

We consider a scalar stable system with parameters \( A = 0.8, C = 1, Q = 1, R = 1 \) and \( \lambda = 0.8 \). We compare our stochastic event-triggered schedulers with two other offline schedulers as follows.

1) Random offline scheduler: The sensor transmits packets with probability \( \frac{1}{T} \) at each time step in random scheduling;

2) Periodic offline scheduler: The sensor sends the data using the optimal offline periodic scheduling [1] with rate \( \frac{1}{T} \).

We adopt the Monte Carlo method with 150000 independent iterations to calculate the mean estimation error covariance, which is shown in Fig. 6. The stochastic event-triggered policies proposed in our work not only outperform the random offline scheduler, but also reduce the mean error covariance compared to the optimal offline periodic scheduler, especially when the communication rate is not sufficient to allow the persistent data transmissions transmit packets. Moreover, the closed-loop scheduler is better than the open-loop scheduler especially for \( \gamma \in [0.1, 0.4] \) in this case.

B. Performance Bounds

We consider the network availability rate \( \lambda = 0.8 \). Fig. 7 demonstrate the asymptotic bounds of mean error covariance.
Fig. 6. Empirical mean error covariance under three scheduling strategies versus effective communication rate.

Fig. 7. Trace of the asymptotic upper bound $\bar{X}_{ol}$ and the lower bound $X_{ol}$ of the open-loop scheduler versus empirical data.

in Lemma 1 for a stable system with parameters (same parameters as [12] for comparison)

\[
A = \begin{bmatrix} 0.8 & 1 \\ 0 & 0.95 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, R = 1,
\]

using an open-loop scheduler by 60000 simulation runs. We observe that when the communication rate $\gamma$ is closer to $\lambda$, the traces of the bounds for both cases are tighter. Similar results exist for an unstable system under the closed-loop scheduler.

C. Design of Event-triggered Parameter

We assume $\lambda = 0.8$. To compare this result with the schedulers proposed in [12], we use the same system parameters

\[
A = \begin{bmatrix} 0.8 & 1 \\ 0 & 0.95 \end{bmatrix}, C = \begin{bmatrix} 0.5 & 0.3 \\ 0 & 1.4 \end{bmatrix}, Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\]

with the open-loop scheduler. Note that

\[
X_0 = \begin{bmatrix} 2.4353 & 0.3976 \\ 0.3976 & 1.3756 \end{bmatrix}.
\]

Set the system quality constraint as $M = X_0 + \varpi I$, where $\varpi$ is a positive real number. The suboptimal solution in Theorem 2 is obtained under different values of $\varpi$, and it is shown in Fig. 8 by the red dotted line. The same as in [12], the suboptimal solution equals the optimal solution when $\varpi$ is large enough, though the equivalent point of $\varpi$ is larger than that in [12].

Moreover, the suboptimal solution which follows Theorem 4 using a B&B algorithm is shown by purple dotted line in Fig. 8. We observe that to achieve the same estimation quality, the upper bound of the communication rate using the closed-loop scheduler is much smaller than the communication rate using the open-loop scheduler. The suboptimal solution is also equivalent to the optimal solution when $\varpi$ is large enough.

D. Comparison between Different Access Probabilities $\lambda$

In this subsection, we illustrate the scheduling performance by varying $\lambda$. A scalar stable system with parameters $A = 0.8, C = 1, Q = 1$ and $R = 1$ by an open-loop scheduler is considered. We adopt the Monte Carlo method using 50 independent sample paths with 3000 time steps each to calculate the mean error covariance. The results are shown in Fig. 9. Two lower bounds are also plotted:

1) The lower bound (black filled dots) from Lemma 1 with $\gamma = \lambda$, i.e., $X_0$ [9].
2) The lower bound (red hollow dots) in [24], i.e., $X_p$ [10].

The black dots in Fig. 9 are closer to the empirical results for fixed $\lambda$ compared with the red dots when $\lambda = 0.4, 0.6, 0.8, 1$. If $\lambda = 0.2$, the red dot is better than the black one. This coincides with the result that the lower bound derived in our paper is larger than the previous one in [24] especially when $\lambda$ is large. The unstable system under closed-loop scheduler has similar results.

VI. CONCLUSION

In this work, we developed stochastic event-triggered schedulers for remote estimation in which the network access is
uncertain. We started from the formulation of a discrete-time remote estimation process in the CRSN, and then analyze the hidden information contained in the absence of data transmission. In order to achieve a better tradeoff between estimation performance and communication consumption, we proposed both open-loop and closed-loop schedules. Utilize the hidden information, the MMSE estimators for both schedules were derived. The problem to minimize the average communication rate while upholding a level of quality was studied. We proposed a suboptimal expression to design event parameter in the open-loop scenario by solving an SDP problem. Since the closed-form of the communication rate cannot be obtained in the closed-loop scenario, a jointly bilinear problem is used to minimize the upper bound of the communication rate satisfying the quality constraint; the related global optimal boundary solution is obtained by B&B algorithm. Numerical examples were provided to illustrate our results. Future work includes safety issues and multiple sensor scheduler in this system structure. It is also interesting to consider other network channel models, e.g., multi-state Markov chain.

REFERENCES

[1] M. Hasegawa, H. Hirai, K. Nagano, H. Harada, and K. Aihara, “Optimization for centralized and decentralized cognitive radio networks,” Proc. IEEE, vol. 102, no. 4, pp. 574–584, 2014.

[2] J. Mitola, G. Q. Maguire et al., “Cognitive radio: making software radios more personal,” IEEE Personal Communications, vol. 6, no. 4, pp. 13–18, 1999.

[3] O. B. Akan, O. B. Karli, and O. Ergul, “Cognitive radio sensor networks,” IEEE network, vol. 23, no. 4, pp. 34–40, 2009.

[4] L. Kakalou, K. E. Pansinnis, P. Krawiec, and R. Badea, “Cognitive radio network and network service chaining toward 5G: challenges and requirements,” IEEE Communications Magazine, vol. 55, no. 11, pp. 145–151, 2017.

[5] C. Yang and L. Shi, “Deterministic sensor data scheduling under limited communication resource,” IEEE TSP, vol. 59, no. 10, pp. 5050–5056, 2011.

[6] L. Shi, P. Cheng, and J. Chen, “Sensor data scheduling for optimal state estimation with communication energy constraint,” Automatica, vol. 47, no. 8, pp. 1693–1698, 2011.

[7] Y. Mo, E. Garone, and B. Sinopoli, “On infinite-horizon sensor scheduling,” Systems & Control Letters, vol. 67, pp. 65–70, 2014.

[8] K. J. Astrom and B. M. Bernhardsson, “Comparison of Riemann and Lebesgue sampling for first order stochastic systems,” in Proc. IEEE CDC, vol. 2, 2002, pp. 2011–2016.

[9] M. Xia, V. Gupta, and P. J. Antsaklis, “Networked state estimation over a shared communication medium,” IEEE TAC, vol. 62, no. 4, pp. 1729–1741, 2017.

[10] S. Trimp, “Stability analysis of distributed event-based state estimation,” in Proc. IEEE CDC, 2014, pp. 2013–2019.

[11] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, “Event-based sensor data scheduling: Trade-off between communication rate and estimation quality,” IEEE TAC, vol. 58, no. 4, pp. 1041–1046, 2013.

[12] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, “Stochastic event-triggered sensor schedule for remote state estimation,” IEEE TAC, vol. 60, no. 10, pp. 2661–2675, 2015.

[13] R. Deng, J. Chen, C. Yuen, P. Cheng, and Y. Sun, “Energy-efficient cooperative spectrum sensing by optimal scheduling in sensor-aided cognitive radio networks,” IEEE Transactions on Vehicular Technology, vol. 61, no. 2, pp. 716–725, 2012.

[14] O. Mabrouk, H. Idoudi, I. Amdouni, R. Soua, P. Minet, and L. Saidane, “OTICOR: Opportunistic time slot assignment in cognitive radio sensor networks,” in Proc. IEEE, 2014, pp. 790–797.

[15] E. Kung, J. Wu, D. Shi, and L. Shi, “On the nonexistence of event-based triggers that preserve gaussian state in presence of packet-drop,” in Proc. ACC, 2017, pp. 1233–1237.

[16] L. Xu, Y. Mo, and L. Xie, “Remote state estimation with stochastic event-triggered sensor schedule in the presence of packet drops,” in Proc. ACC, IEEE, 2019, pp. 5780–5785.

[17] C.-H. Chen and C.-L. Wang, “Power allocation for ofdm-based cognitive radio systems under primary user activity,” in 2010 IEEE 71st Vehicular Technology Conference. IEEE, 2010, pp. 1–5.

[18] Y. Saleem and M. H. Rehmani, “Primary radio user activity models for cognitive radio networks: A survey,” Journal of Network and Computer Applications, vol. 43, pp. 1–16, 2014.

[19] A. Ganti, E. Modiano, and J. N. Tsitsiklis, “Transmission scheduling for multi-channel satellite and wireless networks,” in Proc. Annual Allerton Conference on Communication Control and Computing, vol. 40, no. 3, 2002, pp. 1319–1328.

[20] ——, “Optimal transmission scheduling in symmetric communication models with intermittent connectivity,” IEEE IEEE Trans. Inf., vol. 53, no. 3, pp. 998–1008, 2007.

[21] A. Banaei and C. N. Georghiades, “Throughput analysis of a randomized sensing scheme in cell-based ad-hoc cognitive networks,” in Proc. IEEE ICC, 2009, pp. 1–6.

[22] J. Gambini, O. Simeone, Y. Bar-Ness, U. Spagnolini, and T. Yu, “Packet-wise vertical handover for unlicensed multi-standard spectrum access with cognitive radios;” IEEE Trans. Wirel. Commun, vol. 7, no. 12, pp. 5172–5176, 2008.

[23] X. Wang, Y. Sun, and N. Hovakimyan, “Asynchronous task execution in networked control systems using decentralized event-triggering,” Systems & Control Letters, vol. 61, no. 9, pp. 936–944, 2012.

[24] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, “Kalman filtering with intermittent observations;” IEEE TAC, vol. 49, no. 9, pp. 1453–1464, 2004.

[25] F. A. Al-Khayyal and J. E. Falk, “Jointly constrained biconvex programming,” Mathematics of Operations Research, vol. 8, no. 2, pp. 273–286, 1983.

[26] H. V. Henderson and S. Searle, “Vec and vech operators for matrices, with some uses in Jacobians and multivariate statistics,” Mathematics of Operations Research, vol. 17, no. 1, pp. 63–81, 1979.

[27] C. Yang, J. Wu, W. Zhang, and L. Shi, “Schedule communication for decentralized state estimation,” IEEE TSP, vol. 61, no. 10, pp. 2525–2535, 2013.

[28] S. Lehmich, P. Neff, and J. Lankeit, “On the convexity of the function det(C) → f(det(C)) on positive-definite matrices,” Mathematics and Mechanics of Solids, vol. 19, no. 4, pp. 369–375, 2014.
APPENDIX

A. Propositions of function \( g_{\theta,W} \)

We first prove some useful properties of the matrix function \( g_{\theta,W}(X) \).

**Proposition 1.** For all \( X_1, X_2 \in \mathbb{S}^n_+ \), we have the following properties of \( g_{\theta,W} \):

1) Monotonicity: If \( X_1 \geq X_2 \), then \( g_{\theta,W}(X_1) \geq g_{\theta,W}(X_2) \).
2) Existence and uniqueness of a fixed point: There exists a unique positive-definite \( X_{\theta,W} \) such that \( X_* = g_{\theta,W}(X_*) \).
3) Limit property of the iterated function: \( g_{\theta,W}(X) \to X_{\theta,W} \), for any \( X \in \mathbb{S}^n_+ \) as \( k \to \infty \).
4) Concavity: For all \( \alpha \in [0,1] \), \( g_{\theta,W}(\alpha X_1 + (1-\alpha)X_2) \geq \alpha g_{\theta,W}(X_1) + (1-\alpha)g_{\theta,W}(X_2) \). Therefore, by Jensen’s inequality, one has \( \mathbb{E}(g_{\theta,W}(X)) \leq g_{\theta,W}(\mathbb{E}(X)) \).
5) Monotonicity on \( W \): For any \( W_1 \geq W_2 \), \( g_{\theta,W_1}(X) \geq g_{\theta,W_2}(X) \).

**Proof.** 1) - 4) are proved in [24] in detail.
5) For any \( W_1 \geq W_2 \), \( CXC^T + W_1 \geq CXC^T + W_2 \geq 0 \), \( (CXC^T + W_1)^{-1} \leq (CXC^T + W_2)^{-1} \), \( \theta AXC^T(CXC^T + W_1)^{-1}CA^T \leq \theta AXC^T(CXC^T + W_2)^{-1}CA^T \) and \( g_{\theta,W_1}(X) \geq g_{\theta,W_2}(X) \). The second equation holds because of the monotonicity property of the iterated function.

B. **Proof of Theorem 7**

Since the process \( x_0 \) has a prior Gaussian distribution, i.e., \( x_0 \sim N(0,\Sigma) \), one can prove the MMSE estimate in a recursive way. Assume \( x_k \) has a prior Gaussian distribution as \( x_k \sim N(\hat{x}_k,P_k^-) \). We need to prove the estimation update of \( x_k \) based on the new update \( \hat{x}_k, \eta_k \) and \( \hat{\eta}_k \). The above analysis, the above recursive equations are satisfied, where \( K_k = \eta_k K_{\theta k} \). This completes the measurement update proof.

Then we consider the pdf of the time update. It is a Gaussian process \( f(x_{k+1}|z_k) = f(A\hat{x}_k + w_k|z_k) = N(A\hat{x}_k, AP_k A^T + Q) \), which is directly derived given that \( x_k \) and \( w_k \) are mutually independent Gaussian. This is the same as equations in [7]. Thus, the proof is completed.

C. **Proof of Lemma 7**

We prove Lemma 1 by induction. For simplicity, denote \( U_k = g_{\Lambda k,R+Y^{-1}}(S) \).

Clearly, \( \mathbb{E}[P_k^-] = U_0 = \Sigma \). Assume \( \mathbb{E}[P_{k+1}^-] \leq U_k \). Then the statement is equal to proving that \( \mathbb{E}[P_{k+1}^-] \leq U_{k+1} \) from equations (6) and (7), one obtains

\[
\begin{align*}
\mathbb{E}[P_{k+1}^-] &= \mathbb{E}[g_{\Lambda k,R+Y^{-1}}(P_k^-)] 
\leq g_{\Lambda k,R+Y^{-1}}(U_k) = U_{k+1},
\end{align*}
\]

where the first inequality holds from the fifth statement in Proposition 1. The second inequality holds due to the concavity of function \( g_{\theta,W} \) and the last inequality holds recalling that \( g_{\theta,W} \) is a monotonically increasing function.

From the above analysis, \( \mathbb{E}[P_k^-] \leq U_k \) for all \( k \) by induction. Moreover, by Proposition 1, \( U_k \to X_{ol} \), as \( k \to \infty \), which implies that

\[
\lim_{k \to \infty} \sup \mathbb{E}[P_k^-] \leq X_{ol}.
\]

On the other hand, to derive the lower bound, let us define \( S_k \leq P_k^- \), \( S_k^- \leq (P_k^-)^{-1} \).

There are three cases of the recursive function of \( P_k \) as follows

\[
P_k = \begin{cases} P_k^-, & \text{if } \eta_k = 0, \\ P_k^--P_k^-(C + CP_k^-(C^T)^{-1}CP_k^-)^{-1}C P_k^-, & \text{if } \eta_k, \varepsilon_k = 1, \\ P_k^--P_k^-(C + Y^{-1} + CP_k^-(C^T)^{-1}CP_k^-)^{-1}C P_k^-, & \text{else.} 
\end{cases}
\]

Inverting both side of (25), we have

\[
S_k = \begin{cases} S_k^-, & \text{if } \eta_k = 0, \\ S_k^- + C^T R^{-1}C, & \text{if } \eta_k = 1, \varepsilon_k = 1, \\ S_k^- + C^T (R + Y^{-1})^{-1}C, & \text{if } \eta_k = 1, \varepsilon_k = 0. 
\end{cases}
\]

Aggregating (29), one has

\[
S_k = S_k^- + \eta_k (1-\varepsilon_k)C^T (R + Y^{-1})^{-1}C + \eta_k \varepsilon_k C^T R^{-1}C.
\]

Taking the expectation of both sides, one has

\[
\mathbb{E}[\eta_k (1-\varepsilon_k)] = \mathbb{P}(\eta_k = 1, \varepsilon_k = 0) = \lambda - \gamma, \\
\mathbb{E}[\eta_k \varepsilon_k] = \mathbb{P}(\eta_k = 1, \varepsilon_k = 1) = \gamma.
\]

Thus, we obtain

\[
\mathbb{E}[S_k] = \mathbb{E}[S_k^-] + C^T R_1^{-1}C.
\]

Meanwhile, the third equation in (7) is the same as

\[
S_{k+1}^- = (AS_k^-A^T + Q)^{-1},
\]

from which \( S_{k+1}^- \) is concave w.r.t. \( S_k \). By Jensen’s inequality, the following inequality holds:

\[
\mathbb{E}[S_{k+1}^-] \leq (A \mathbb{E}[S_k^-]A^T + Q)^{-1} = \Lambda_{R_1}(\mathbb{E}[S_k^-]),
\]

where \( \Lambda_W(X) \triangleq [A(X + C^T W^{-1}C)^{-1}A^T + Q]^{-1} \), \( W > 0 \) for simplicity.

For any \( X_1 \geq X_2 \geq 0 \), the following equation holds:

\[
[\Lambda_W(X_1)]^{-1} = g_W(X_1^{-1}) \leq g_W(X_2^{-1}) = [\Lambda_W(X_2)]^{-1},
\]

\( \Leftrightarrow \Lambda_W(X_1) \geq \Lambda_W(X_2) \).
Hence, the monotonicity of \( \Lambda_W \) is proven.

Furthermore, \( [\Lambda_W(X^{-1})]^{-1} = g_W(X) \) holds by applying the matrix inversion lemma and the result directly follows

\[
[\Lambda_W(X^{-1})]^{-1} = g_W(X).
\]

Based on the monotonicity of \( \Lambda_{R_1} \) and \( S_0 = \Sigma^{-1} \), we obtain

\[
\mathbb{E}[S_k] \leq \Lambda_{R_1}(\mathbb{E}[S_{k-1}]) \leq \cdots \leq \Lambda_{R_1}^k(\Sigma^{-1}).
\]

Since \( f(X) = X^{-1}, X \geq 0 \) is convex w.r.t \( X \), and by Jensen’s inequality, one has

\[
\mathbb{E}[P_k] = \mathbb{E}[(S_k^{-1})] \geq \mathbb{E}[(S_{k-1}^{-1})] \geq [\Lambda_{R_1}^k(\Sigma^{-1})]^{-1} = g_{R_1}^k(\Sigma).
\]

Denote \( D_k \triangleq g_{R_1}^k(\Sigma) \). From the above analysis, one has \( \mathbb{E}[P_{k-1}] \geq D_k \) for all \( k \). By Proposition 1, \( D_k \to X_{ol} \) as \( k \to \infty \), which implies that

\[
\lim_{k \to \infty} \mathbb{E}[P_{k-1}] \geq X_{ol}.
\]

From the fifth statement in Proposition 1, \( X_{ol} \geq X_0 \) always holds as \( R_1 \geq \frac{R}{X} \). The proof is done.

### D. Proof of Theorem 2

The proof of Theorem 2 follows the following two steps.

First, we prove an equivalent set of constraints to replace the implicit constraint \( X_{ol} \leq M \). Second, the set of constraints are transformed to an SDP constraint.

Firstly, the following two statements are equivalent:

1) \( X_{ol} \leq M \),

2) There exists \( 0 < X \leq M \) such that \( g_{\lambda,R+Y^{-1}}(X) \leq X \).

To prove 1) \( \Rightarrow \) 2): It is obvious that the second statement can be obtained from the first, i.e., \( X_{ol} \) is a feasible solution to \( X \).

To prove 2) \( \Rightarrow \) 1): Recall that \( g_{\lambda,W}(X) \) is a monotonically increasing function in \( X \) from Proposition 1.

\[
M \geq X \geq g_{\lambda,R+Y^{-1}}(X) \geq g_{\lambda,R+Y^{-1}}^2(X) \geq \ldots \geq \lim_{k \to \infty} g_{\lambda,R+Y^{-1}}^k(X) = X_{ol}.
\]

Then the first statement is obtained from the second. Thus, these two statements are equivalent.

The constraints of Problem 3 are rewritten as follows:

\[
Y \geq 0, 0 < X \leq M, g_{\lambda,R+Y^{-1}}(X) \leq X. \tag{30}
\]

Secondly, the main difficulty is to transform the last inequality into an equivalent SDP constraint. Since the last inequality cannot be changed to linear form based on \( X \), we transform it to linear form based on the inverse of \( X \), i.e., \( S \). To maintain the parameter utility, the second inequality should also be changed to the linear form based on \( S \). Taking the inverse of both sides of the second inequality in (30), we obtain \( S \geq M^{-1}. \) It is straightforward to see that

\[
S \geq M^{-1} \iff \begin{bmatrix} S & I \\ I & M \end{bmatrix} \geq 0,
\]

by Schur complement since \( S = X^{-1} > 0 \).

The left-hand part of the problem is to transform the third inequality in (30) to an SDP form in \( S \). By rearranging the term, one has

\[
X - (1 - \lambda)AXA^\top - Q - \lambda(AXA^\top - AXC^\top)(CXC^\top + R + Y^{-1})^{-1}AXA^\top
= X - (1 - \lambda)AXA^\top - Q
- \lambda A(S + C^\top(R + Y^{-1})^{-1}C)^{-1}A^\top \geq 0,
\]

where the equality follows the matrix inversion lemma.

Since \( S > 0, Y \geq 0, R > 0 \), the following equation holds

\[
S + C^\top(R + Y^{-1})^{-1}C > 0,
\]

then by applying the Schur complement to (31), the third inequality in (30) is the same as

\[
\begin{bmatrix}
X - (1 - \lambda)AXA^\top - Q & \sqrt{\lambda A} \\
\sqrt{\lambda A}^\top & S + C^\top(R + Y^{-1})^{-1}C
\end{bmatrix} \geq 0.
\tag{33}
\]

We obtain \( X - Q - (1 - \lambda)AXA^\top \geq 0 \) from (33). Meanwhile, as \( S = X^{-1} > 0 \), the following inequality holds:

\[
\begin{bmatrix}
X - Q & \sqrt{\lambda A} \\
\sqrt{\lambda A}^\top & S
\end{bmatrix} \geq 0.
\tag{34}
\]

Given that \( X - Q \geq 0 \) from (34) and \( X > 0 \), it is straightforward to see that

\[
\begin{bmatrix}
X & I \\
I & Q^{-1}
\end{bmatrix} \geq 0.
\tag{35}
\]

Combining (33), (34) and (35), one has

\[
\Theta \triangleq \begin{bmatrix}
S & 0 & 0 & 0 \\
0 & I & 0 & 0 \\
0 & 0 & I & 0 \\
0 & 0 & 0 & I
\end{bmatrix} \otimes \begin{bmatrix}
S & 0 & 0 & 0 \\
0 & I & 0 & 0 \\
0 & 0 & I & 0 \\
0 & 0 & 0 & I
\end{bmatrix} \geq 0 \iff
\begin{bmatrix}
S & \sqrt{\lambda A}S & \sqrt{\lambda I}S \\
\sqrt{\lambda A}^\top S & S & \sqrt{\lambda I}S \\
S & \sqrt{\lambda I}S & \sqrt{\lambda I}S \\
\sqrt{\lambda I}^\top S & \sqrt{\lambda I}^\top S & \sqrt{\lambda I}^\top S
\end{bmatrix} \geq 0,
\tag{37}
\]

where \( \Gamma_{22} \triangleq S + C^\top(R + Y^{-1})^{-1}C \).

Since \( \Gamma_{22} \) is not linear in \( Y \), we expand \( (R + Y^{-1})^{-1} \) by using the matrix inversion lemma, where

\[
(R + Y^{-1})^{-1} = R^{-1} - R^{-1}(Y + R^{-1})^{-1}R^{-1}.
\]
Then one has
\[
\Gamma = \begin{bmatrix}
S & \sqrt{\lambda S}A & S + C^\top R^{-1}C & 0 & 0 \\
\sqrt{\lambda A^\top S} & S + C^\top R^{-1}C & 0 & 0 & 0 \\
\sqrt{1 - \lambda A^\top S} & S & 0 & 0 & 0 \\
S & 0 & 0 & Q^{-1} & 0 \\
\end{bmatrix}
\]
\[
= \begin{bmatrix}
0 & 0 \\
C^\top R^{-1} & (Y + R^{-1})^{-1} \\
0 & 0 \\
0 & 0 \\
\end{bmatrix}
\]
\[
\geq 0.
\]
(38)

As \(Y \geq 0, R > 0, (Y + R^{-1})^{-1} \geq 0\) holds. The above inequality (38) can also be viewed as a Schur complement, where
\[
\hat{A} = \begin{bmatrix}
S & \sqrt{\lambda S}A & S + C^\top R^{-1}C & 0 & 0 \\
\sqrt{\lambda A^\top S} & S + C^\top R^{-1}C & 0 & 0 & 0 \\
\sqrt{1 - \lambda A^\top S} & S & 0 & 0 & 0 \\
S & 0 & 0 & Q^{-1} & 0 \\
\end{bmatrix}, \quad \hat{B} = \begin{bmatrix}
0 & 0 \\
C^\top R^{-1} & (Y + R^{-1})^{-1} \\
0 & 0 \\
0 & 0 \\
\end{bmatrix}, \quad \hat{C} = Y + R^{-1} > 0.
\]

Given that \(\hat{C} > 0\), then \(\Gamma = \hat{A} - \hat{B}\hat{C}^{-1}\hat{B}^\top \geq 0\) if and only if \(\Psi(S, Y) = \begin{bmatrix}
\hat{A} & \hat{B} \\
\hat{B}^\top & \hat{C} \\
\end{bmatrix} \geq 0\). The proof is done.

**E. Proof of Theorem 3**

From Theorem 1 one has
\[
f(z_k | \mathcal{I}_{k-1}) = f(y_k - \hat{y}_k | \mathcal{I}_{k-1}) = f(Ce_k + v_k | \mathcal{I}_{k-1})
\]
\[
= \mathcal{N}(0, C^2P_kC^\top + R),
\]
where the second equation holds as \(\hat{y}_k = C\hat{x}_k\) from Theorem 1, the last equation holds as \(E[e_k^2 | \mathcal{I}_{k-1}] = 0\), and \(e_k, v_k\) are mutually independent Gaussian variables.

For the measurement update, performing a similar analysis, we can obtain (15). Note that, substituting \(y_k\) by \(\hat{y}_k\) in (6), \(\hat{x}_k = \hat{x}_k + s_kK_k\hat{y}_k - K_kE[z_k | \mathcal{I}_{k-1}] = \hat{x}_k + s_kK_k\hat{y}_k\), which is consistent with (15). We omit the remainder proof as it is straightforward.

**F. Proof of Lemma 2**

We have
\[
\gamma = E[\Pr(\eta_k | e_k = 1 | \mathcal{I}_{k-1})]
\]
\[
= E[\Pr(\eta_k = 1) \Pr(\zeta_k \leq \exp(-\frac{1}{2}\hat{Z}Z_k) | \mathcal{I}_{k-1})]
\]
\[
= \lambda E \left[ 1 - (\det (I + (CP_kC^\top + R)Z))^{-\frac{1}{2}} \right].
\]

To prove the second inequality, it suffices to prove the concavity of (16). By Jensen’s inequality, it suffices to prove the convexity of function \(f(CXC^\top Z + RZ + I) \leq (\det (I + (CP_kC^\top + R)Z))^{-\frac{1}{2}}\).

The convexity holds for a composition with affine functions; therefore, it is equivalent to prove that
\[
f(X) = |\det(X)|^{-\frac{1}{2}}, \text{ for } X \geq I,
\]
is convex.

Define \(b : (0, \infty) \rightarrow \mathbb{R}\) and \(b(s) \triangleq s^{-\frac{1}{2}}\). Lehmic et al. (28) states the convexity of \(f\) on the set \(X \in S^n_{+} + I\) is equivalent to
\[
nsb''(s) + (n - 1)b'(s) \geq 0, \text{ and } b'(s) \leq 0, \text{ for all } s > 0.
\]
Since \(b'(s) = -\frac{1}{2}s^{-\frac{3}{2}} \leq 0\) and
\[
nsb''(s) + (n - 1)b'(s) = ns\frac{3}{4}s^{-\frac{3}{2}} - \frac{1}{2}s^{-\frac{3}{2}} = \frac{3}{4}s^{-\frac{3}{2}} + (n - 1)\frac{1}{2}s^{-\frac{3}{2}} \geq 0,
\]
the proof is completed.

**G. Proof of Theorem 4**

An equivalent statement of the constraints of Problem 5 is as follows. The following two statements are equivalent:
1) \(X_{cl} \leq M, CX_{cl}C^\top + R \leq T\),
2) There exists \(S^{-1} = X \geq M\) such that
\[
\Psi(S, Z) \geq 0, CXC^\top + R \leq T.
\]

1) \(\Rightarrow\) 2): Let \(X\) be equal to \(X_{cl}\); it is obvious that \(X_{cl}\) is a feasible matrix satisfying the second statement.

2) \(\Rightarrow\) 1): From the similar proof in Appendix D one has
\[
g_{\lambda, R + Z^{-1}}(X) \leq X, \quad X_{cl} \leq X \leq M;
\]
therefore, \(CX_{cl}C^\top + R \leq CXC^\top + R \leq T\).

On the other hand, it is well known that replacing \(S^{-1} = X\) by \(S^{-1} < X\) does not affect the solution to the optimization problem (18) since \(\text{tr}((CXC^\top + R)Z) \geq \text{tr}((CS^{-1}C^\top + R)Z)\). Therefore, \(S^{-1} = X\) is satisfied for at least one optimal solution to the optimization problem, which completes the proof.

**H. Proof of Lemma 7**

By Lemma 4 and Lemma 5, the feasibility condition for the problem is that \(X_0 \leq X_{cl}\). Moreover, from the proof of Theorem 4 one has \(X_0 \leq X_{cl} \leq X \leq M\). It is sufficient to obtain that \(l_{ii} \leq x_{ii} \leq m_{ii}\). Furthermore, since every principal sub-matrix is positive definite for a positive semidefinite matrix, we have \(|x_{ij}| \leq (x_{ii}x_{jj})^{\frac{1}{2}} \leq (m_{ii}m_{jj})^{\frac{1}{2}}\).

As the objective function in equation (18) satisfies
\[
\min_{X, S} \text{tr}(CXC^\top + RZ) \leq \min_{X, S} \text{tr}(Z) \text{tr}(CXC^\top + R) \leq \min_{X, S} \text{tr}(CXC^\top + R),
\]
the following equation holds
\[
z^* \geq \frac{\min_{X, S} \text{tr}(CXC^\top + R)}{\text{tr}(CXC^\top + R)}, \text{ s.t. } (X, Z, S) \in S,
\]
where \(z^* \geq \text{tr}(Z)\) is the upper bound of \(\text{tr}(Z)\) on \(S\). As the molecule \(\min_{Z, X, S} \text{tr}(CXC^\top + R) \leq \min_{X, S} \text{tr}(CXC^\top + R) \geq \text{tr}(CX_0C^\top + R)\), let \(z^*\) be the solution to (22), and we can prove that \(z^*\) satisfies (39) from the above analysis. Therefore, one has \(0 \leq z_{ij} \leq \text{tr}(Z) \leq z^*\). Moreover, \(|z_{ij}| \leq (z_{ii}z_{jj})^{\frac{1}{2}} \leq z^*\). This completes the proof.