2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

Consistency-preserving Event-triggered Estimation in Sensor Networks

Adam Molin, Henrik Sandberg and Karl Henrik Johansson

Abstract— **This paper is concerned with linear state estimation in sensor networks with an event-triggered exchange of information. It is assumed that each sensor node transmits its local estimate to a fusion center whenever an appropriately chosen error norm exceeds a threshold. The fusion rule is a modified version of the Covariance Intersection Algorithm. We investigate how to incorporate the event information of not having transmitted at the fusion center such that the filter remains unbiased and consistent with regard to its error covariance. An upper bound on the error covariance matrix is derived by exploiting the structure of the posterior probability distribution. This enables us to replace the event information by the virtual transmission of consistent local estimates. Based on the consistency-preserving property of the proposed scheme, we show stability of the event-triggered state estimator in terms of a bounded mean square error.**

I. INTRODUCTION

We consider the problem of state estimation of a linear system in which a fusion center gathers information from multiple sensors. Each sensor node is assumed to be capable of preprocessing measurement information and has the autonomy of deciding when it is worth to send the preprocessed data to the fusion center. In this paper, we take two complementary measures in order to cope with the abundance of information emerging in such large-scale sensor networks.

First, data transmission is carried out by an event-triggered scheme to reduce the overall traffic over the communication network. Event-triggered mechanisms turn out to be optimal when there needs to be achieved a trade-off between resource consumption and real-time requirements, such as the minimization of the mean squared error with communication constraints [1]–[3]. Furthermore, it is well known that event-triggered sampling for control and estimation commonly outperforms time-triggered sampling considerably [4]. Second, instead of implementing a Kalman filter at the fusion center, we use a modified version of the *Covariance Intersection Algorithm (CI)*. CI has been introduced in [5] and is a linear fusion technique that is capable of yielding consistent estimates without relying on the cross-correlations between data sources. Since bookkeeping of correlations between all transmitted sensor data may exceed the numerical capabilities of the fusion center when the number of sensors is large, CI turns out to be a favorable alternative for state estimation in sensor networks despite of its conservatism.

The authors are with the ACCESS Linnaeus Centre, Electrical Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden, {adammol, hsan, kallej}@kth.se.

This work was supported in part by the Swedish Research Council under Grant 2013-5523 and Grant 2014-6282, in part by the ACCESS Linnaeus Centre, and in part by the Knut and Alice Wallenberg Foundation.

The main question addressed in this paper is how to incorporate event information at the fusion center while preserving the structure of the filtering mechanism. According to the triggering rule implemented at each sensor node, the fusion center may include the information that sensors have not provided data into the calculation of their state estimate. Such an idea by itself is not novel and has been investigated by several authors [6]–[13]. In the following, we review the major results in the literature on event-triggered state estimation taking event information into account.

The work in [6] constitutes one of the earlier works on event-triggered state estimation. It proposes a modified Kalman filtering algorithm assuming measurements are sampled by the send-on-delta method. If no measurement is received, then the filter uses the previously sampled measurement for estimation with an increased measurement noise. The new measurement noise is modelled by an additive uniform distribution depending on the threshold of the send-on-delta sampler. In [7] and [8], the authors approximate the prior distribution by a Gaussian distribution which enables an explicit calculation of conditional means and error covariances incorporating the event information. Though being widely used in practice, the Gaussianity assumption prohibits to state general properties of the event-triggered state estimator, such as consistency, which also complicates the analysis of its asymptotic behavior. More sophisticated approximations of the posterior distribution incorporating the event information have been conducted in [9] by using a Gaussian sum filter and in [14] by implementing a particle filter. For calculating the exact posterior probability distribution, the work in [12] has shown that the incorporation of event information is closely related to a virtual measurement channel. The stochastic model of the virtual measurement is given by a modified version of the original measurement equation disturbed by additive white noise whose distribution is uniform and has a support defined by the triggering rule. Another method divides the type of uncertainty arising in the filtering procedure into stochastic noise and membership uncertainty [10], [11]. Event information can then be classified as the latter. The authors pursue a worst-case analysis for the incorporation of event information into the state estimator. Though the approach is capable of addressing a variety of filtering problems beyond event-triggered state estimation [11], it disregards potential structure arising from the statistical properties of the event-triggered scheme due to the nature of membership uncertainty. Hence, the estimates obtained from the worst-case analysis might be overly conservative in certain cases.

The main contribution of this paper is to incorporate event information in the filtering procedure while preserving consistency of the predicted statistics. An estimate is said to be consistent if the predicted error covariance matrix is an upper bound on the true error covariance matrix, see [5], [15]. This basic property ensures that estimates do not become too optimistic over time which might eventually lead to a divergent behavior of the filter [15]. The novel feature of our work is to restrict our attention to an important class of triggering rules and exploit several statistical properties to obtain consistent estimates that take the event information into account. The triggering rule is assumed to be a threshold rule on the weighted 2-norm of the estimation discrepancy between the current estimate at the sensor and the predicted estimate at the fusion center. Furthermore, each sensor node is only concerned with computing the state estimate restricted to its observable part of the space space. We transform the event information into a consistent pair by calculating an upper bound on the error covariance matrix. By only exploiting symmetry and unimodality properties of the posterior probability distribution, we are able to reduce the bound on the variance arising from the uncertainty of the event-trigger by a factor of 3 in the case of scalar local estimates compared to a set membership analysis [11]. For higher-order systems we obtain a similar result with a decreasing factor for the covariance improvement. Another contribution is the modification of the CI for fusing estimates from subspaces of the state. Based on collective observability of the sensor system, it is shown that the modified CI is well-defined. Furthermore the consistency-preserving property remains valid in the modified version, which implies that we are able to guarantee that the fused estimate is consistent based on our derived covariance bounds. Finally, this attribute of the event-triggered state estimator enables us to show stability in the sense of a bounded mean squared error assuming collective observability of the dynamical system with regard to the sensor network.

This paper is organized as follows. Section II introduces the system model and states the assumptions on the pre-processing at the sensor node, the triggering rule and the data fusion technique. A consistency preserving scheme that incorporates event information is derived in Section III, whereas Section IV is concerned with the stability analysis of the event-triggered state estimator. The efficiency of our approach is illustrated by a numerical example in Section V.

II. SYSTEM MODEL

The structure of the multi-sensor system for event-triggered state estimation is illustrated in Fig. 1. In the following, we describe the functional blocks and the transmission scheme in more detail.

A. Multiple sensor system

We consider a set of M sensors, indexed by j, $1 \leq j \leq M$, which take measurements y_k^j from a common linear process evolving as

$$x_{k+1} = Ax_k + w_k \tag{1}$$

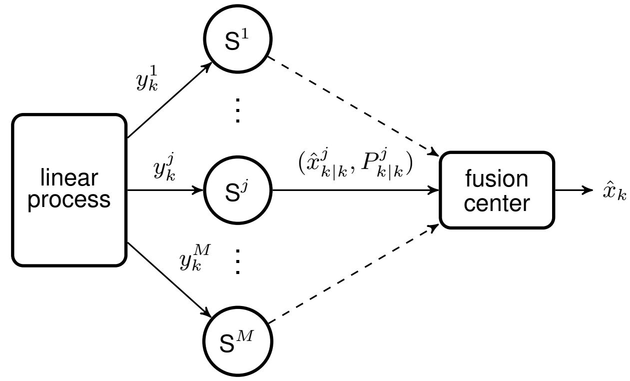

Fig. 1. Multi-sensor system for networked state estimation. Sensors S^j, $1 \leq j \leq M$, sporadically transmit data $(\hat{x}_{k|k}^j, P_{k|k}^j)$ to the fusion center which computes the fused estimate $\hat{x}_k$.

with

$$\begin{aligned} &x_k \in \mathbb{R}^n,\ A \in \mathbb{R}^{n\times n} \\ &x_0 \sim \mathcal{N}(0, R_0), \quad R_0 \in \mathbb{R}^{n\times n} \\ &w_k \sim \mathcal{N}(0, R_w), \quad R_w \in \mathbb{R}^{n\times n} \end{aligned}$$

where R_0 and R_w are assumed to be positive definite. The measurements y_k^j are given by

$$y_k^j = C^j x_k + v_k^j$$

with

$$\begin{aligned} &y_k^j \in \mathbb{R}^{m^j},\ C^j \in \mathbb{R}^{m^j\times n} \\ &v_k^j \sim \mathcal{N}(0, R_v^j), \quad R_v^j \in \mathbb{R}^{m^j\times m^j} \end{aligned} \tag{2}$$

where R_v^j is assumed to be positive definite. We assume that all primitive random variables, given by the initial state x_0, the process noise w_k and the measurement noise v_k^j, are mutually independent.

We use the following terminology from [16] for observability in sensor networks.

Definition 1 (Collective observability): The system given by (1) and (2) is said to be *collectively observable*, when (A, C) is observable with

$$C = \begin{bmatrix} C^1 \\ \vdots \\ C^M \end{bmatrix}.$$

B. Local filtering

In this section, we define the local filtering structure in each sensor node. It is not presumed that the state can be fully recovered at one sensor node by its local measurements, i.e., (A, C^j) is not observable in general. Similar as in [12], we aim at a representation in which a sensor node focuses on the estimation of its observable subspace of the state x_k. For time-invariant linear systems, it is well known (see e.g. [17]) that there exists a non-singular state transformation T^j that separates the state space into an observable and unobservable subspace, such that

$$(T^j)^{-1}AT^j = \begin{bmatrix} A_1^j & 0 \\ A_{21}^j & A_2^j \end{bmatrix}, \quad C^jT^j = \begin{bmatrix} C_1^j & 0 \end{bmatrix}$$
$$T^j = \begin{bmatrix} T_1^j & T_2^j \end{bmatrix}, \quad (T^j)^{-1} = \begin{bmatrix} D_1^j \\ D_2^j \end{bmatrix}$$

with (A_1^j, C_1^j) being observable. Then, the local filter at sensor i estimates the state $x_k^j = D_1^j x_k \in \mathbb{R}^{n_o^j}$ of the subsystem evolving by

$$\begin{aligned} x_{k+1}^j &= A_1^j x_k^j + D_1^j w_k \\ y_k^j &= C_1^j x_{k+1}^j + v_k^j. \end{aligned} \tag{3}$$

The minimum mean square error (MMSE) estimator $\hat{x}_{k|k}^j = \mathsf{E}[x_k^j | Y_k^j]$ of x_k^j is given by the Kalman filter

$$\hat{x}_{k|k}^j = \hat{x}_{k|k-1}^j + K_k^j (y_k^j - C_1^j \hat{x}_{k|k-1}^j) \tag{4a}$$
$$P_{k|k}^j = (I_{n_o^j} - K_k^j C_1^j) P_{k|k-1}^j \tag{4b}$$
$$\hat{x}_{k+1|k}^j = A_1^j \hat{x}_{k|k}^j \tag{4c}$$
$$P_{k+1|k}^j = A_1^j P_{k|k}^j (A_1^j)^\top + D_1^j R_w (D_1^j)^\top \tag{4d}$$

where $K_k^j = P_{k|k-1}^j (C_k^j)^\top (C_1^j P_{k|k-1}^j (C_1^j)^\top + R_v^j)^{-1}$ and $\hat{x}_{0|-1}^j = 0$, $P_{0|-1}^j = D_1^j R_0 (D_1^j)^\top$. I_n denotes the identity matrix in $\mathbb{R}^n$.

C. Event-triggering rule

In the following, we describe the triggering rule of sensors j, $1 \le j \le M$. Let $\delta_k^j \in \{0,1\}$ denote the triggering variable being defined as

$$\delta_k^j = \begin{cases} 1 & \text{transmit } (\hat{x}_{k|k}^j, P_{k|k}^j) \text{ to fusion center} \\ 0 & \text{no transmission.} \end{cases}$$

It should be noted that we disregard packet collisions and transmission delays in this paper. Let τ_k^j and σ_k^j denote the times of the most recent transmission of sensor j before time k and up to and including time k, respectively, i.e.,

$$\begin{aligned} \tau_k^j &= \max\{\ell \mid \delta_\ell^j = 1,\ 0 \le \ell < k\} \\ \sigma_k^j &= \max\{\ell \mid \delta_\ell^j = 1,\ 0 \le \ell \le k\}. \end{aligned}$$

If there has not been a transmission yet at sensor j, we define $\tau_k^j = \sigma_k^j = -1$. Both definitions facilitate the subsequent description of the event-triggered system. We will commonly refer to τ_k^j when considering the event rule, whereas σ_k^j is used to describe the event information at the fusion center. In case no transmission occurs at time k, the fusion center predicts the estimate $\hat{x}_{k|k}^j$ by

$$\tilde{x}_{k|k}^j = \mathsf{E}[\hat{x}_{k|k}^j | \hat{x}_{\tau_k^j|\tau_k^j}^j] = (A_1^j)^{k-\tau_k^j} \hat{x}_{\tau_k^j|\tau_k^j}^j. \tag{5}$$

with $\hat{x}_{-1|-1}^j = 0$. It must be noted that the prediction given in the above equation does not include implicit information of not having transmitted in the interval $\{\tau_k^j+1, \ldots, k\}$. The type of event-triggering rule that we focus on in this paper is given by

$$\delta_k^j = \begin{cases} 1 & \|\hat{x}_{k|k}^j - \tilde{x}_{k|k}^j\|_{\Gamma_k^j}^2 \le r_k^j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\Gamma_k^j > 0$, $r_k^j > 0$ and $\|z\|_\Gamma^2 = z^\top \Gamma z$.

Remark 1: The rule given in (6) has been shown to be optimal for certain event-triggered problems for first-order systems penalizing transmissions [1], [2]. For higher-order dynamics, optimal solutions turn out to be a threshold function of $\hat{x}_{k|k}^j - \tilde{x}_{k|k}^j$ whose threshold manifolds have no closed-form solution in general. It is however shown in [18] that the form in (6) yields an efficient and numerically tractable approximation of the optimal solution.

Besides the arguments in the above remark on the form of the event-trigger, we will exploit the given structure extensively in the remainder of this paper. Here, we indicate one of consequences for the predictor at the fusion center, which will be shown in Section III-C. A basic features of the triggering rule is that the predictor defined in (5) is the optimal MMSE estimate of $\hat{x}_{k|k}^j$ at time k incorporating the event information $\delta_{\tau_k^j+1}^j = \cdots = \delta_k^j = 0$. This implies that the conditional mean of $\hat{x}_{k|k}^j$ at the fusion center does not depend on the event information. Unlike the conditional mean, the evolution of the error covariance will highly depend on the event information. While the covariance matrix of the predictor evolves by the prediction update given by (4d), the knowledge of $\delta_{\tau_k^j+1}^j = \cdots = \delta_k^j = 0$ can be exploited to decrease the error covariance. The way how to incorporate event information into the calculation of the error covariance matrix will be the subject of Section III.

D. Data fusion

The task of the fusion center is to estimate the state x_k based on the sporadic data received from the sensor nodes. The fusion rule that we use is a slightly modified version of CI, see [5]. The modification is needed as the estimates to be fused are restricted to their observable state space. Therefore, an appropriate embedding into the original state space must be performed. We define the data $(\breve{x}_{k|k}^j, \breve{P}_{k|k}^j)$ at the fusion center of sensor j as

$$(\breve{x}_{k|k}^j, \breve{P}_{k|k}^j) = \begin{cases} (\hat{x}_{k|k}^j, P_{k|k}^j) & \delta_k^j = 1 \\ (\tilde{x}_{k|k}^j, \tilde{P}_{k|k}^j) & \delta_k^j = 0. \end{cases} \tag{7}$$

While the fusion center uses the local estimates given by (4a)–(4d) in case of $\delta_k^j = 1$, it uses the predicted state $\tilde{x}_{k|k}^j$ given by (5) and the covariance matrix $\tilde{P}_{k|k}^j$ when $\delta_k^j = 0$. The determination of $\tilde{P}_{k|k}^j$ that takes into account the event information will be tackled in the next section.

Remark 2: Effectively, only the conditional mean must be transmitted as the error covariance matrix $P_{k|k}^j$ evolves deterministically by (4b) and (4d).

Based on the data $\{(\breve{x}^j_{k|k}, \breve{P}^j_{k|k})\}_{j\in\{1,\dots,M\}}$, the modified CI yields the fused estimate $\hat{x}_k$ and is defined as

$$\hat{P}_k^{-1} = \sum_{j=1}^{M} \omega^j T_1^j (\breve{P}^j_{k|k})^{-1} (T_1^j)^\top \tag{8}$$

$$\hat{P}_k^{-1}\hat{x}_k = \sum_{j=1}^{M} \omega^j T_1^j (\breve{P}^j_{k|k})^{-1} \breve{x}^j_{k|k} \tag{9}$$

with weights $\omega^j > 0$, $\omega^1 + \cdots + \omega^M = 1$ and assuming that the information matrix $(\breve{P}^j_{k|k})^{-1}$ exists, i.e., $\breve{P}^j_{k|k}$ is invertible for $j \in \{1,\dots,M\}$. The fact that $\hat{P}_k^{-1}$ is non-singular will be shown in Lemma 1 in the next section. Equation (8) can be viewed as a convex combination of information matrices. The transformation matrix T_1^j ensures that the estimate and covariance matrix related to sensor j are embedded appropriately in the original state space.

Besides its numerical advantages compared to the standard Kalman filter, the major advantage of the CI is that it yields consistent estimates without relying on the knowledge of cross-correlations between estimates. In many applications, in which processed information from various sources within a network is fused, the cross-correlations may not be known [19]. In this paper, the fact that the fusion rule does not have to take into account cross-correlations enables us to incorporate event information into the estimation procedure in a straight-forward manner in order to preserve consistency.

III. CONSISTENCY OF ESTIMATES

One of the main prerequisites of a filtering algorithm is that the estimated statistics arising from the estimation procedure reflect the behavior of the true system appropriately. In this paper, the focus is on the first- and second-order moments of the system, i.e., we are primarily interested in the conditional mean $\hat{x}_k$ and its corresponding error covariance matrix. In many cases, the error covariance matrix cannot be determined exactly due to the lack of statistical parameters or due to computational restrictions. The latter is the reason to approximate the error covariance matrix in this paper. When finding a suitable approximation, one needs to account for the well-known fact that the estimation error may diverge if the computed error covariance matrix becomes too optimistic [15]. It is worth to mention that robust formulations aiming at the computation of almost sure confidence intervals, e.g. in [20], [21] fail in the considered setting as the noise variables are not bounded. These facts motivate us to give a formal definition of consistency [5], [15].

Definition 2 (Consistency): Let $\hat{x}_k$ be an unbiased estimate of the random variable x_k and let $\hat{P}_k$ be an estimate of the corresponding error covariance matrix. Then, the pair $(\hat{x}_k, \hat{P}_k)$ is said to be *consistent* when

$$\mathsf{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^\top] \leq \hat{P}_k.$$

A. Consistency-preserving property of CI

Lemma 1 (Non-singularity of $\hat{P}_k^{-1}$): Let the system (1) and (2) be collective observable and let $\breve{P}^j_{k|k}$ be non-singular for $1 \leq j \leq M$. Then, the covariance intersection algorithm given by (8)–(9) yields a non-singular solution $\hat{P}_k^{-1}$.

Proof: Define the observability matrix for sensor j, $1 \leq j \leq M$, as

$$\mathcal{O}_j = \begin{bmatrix} C^j \\ C^j A \\ \vdots \\ C^j A^{n-1} \end{bmatrix}.$$

The nullspace of $\mathcal{O}_j$ denoted as $\mathcal{R}_{\bar{\mathcal{O}}_j}$ corresponds to the unobservable part of the original state space $\mathbb{R}^n$. Let $\mathcal{R}_{\mathcal{O}_|}$ be the observable subspace of sensor j. From Section II-B, we know that this space is spanned by the columns of T_1^j. Collective observability implies that the nullspace of the aggregated observability matrix $\mathcal{O} = [\mathcal{O}_1^\top, \dots, \mathcal{O}_M^\top]^\top$ only contains the trivial solution. Hence, we have

$$\bigcap_{j=1}^{M} \mathcal{R}_{\bar{\mathcal{O}}_j} = 0.$$

This means on the other hand that the union of the complementary spaces $\mathcal{R}_{\mathcal{O}_j}$ covers the complete state space. Defining

$$\bar{T} = [T_1^1, \dots, T_1^M],$$
$$\bar{\Omega}_k = \mathrm{diag}[\omega^1 (\breve{P}^1_{k|k})^{-1}, \dots, \omega^M \breve{P}^M_{k|k}],$$

we conclude that $P_k^{-1} = \bar{T}\bar{\Omega}_k\bar{T}^\top$ is non-singular because $\bar{\Omega}$ is positive definite and $\bar{T}$ has rank n. ∎

Theorem 1 (Consistency-preserving property of CI): Let the system (1), (2) be collectively observable. If the pair $(\breve{x}^j_{k|k}, \breve{P}^j_{k|k})$ is a consistent estimate of x_k^j with $\breve{P}^j_{k|k}$ being non-singular for $1 \leq j \leq M$, then the CI defined by (8)–(9) yields an unbiased and consistent estimate $(\hat{x}_k, \hat{P}_k)$ of state x_k.

Proof: Using Lemma 1, it is ensured that the inverse of $\hat{P}_k$ is well defined by collective observability of the system and non-singularity of $\breve{P}^j_{k|k}$. As $\breve{x}^j_{k|k}$ is an unbiased estimate of x_k^j, the linear combination (9) also yields an unbiased estimate $\hat{x}_k$. To conclude the proof, we need to show that $\hat{P}_k$ is an upper bound on the error covariance matrix. This follows along the same lines as in [5] for the standard CI and is not repeated here. ∎

B. Incorporation of event information

What remains to be determined is the estimate $\tilde{P}^j_{k|k}$ of the error covariance matrix when sensor j does not transmit information at time k. In the spirit of Theorem 1, the goal is to choose $\tilde{P}^j_{k|k}$ in such way that the pair $(\tilde{x}^j_{k|k}, \tilde{P}^j_{k|k})$ is consistent. At the same time, we aim for a simple computation of $\tilde{P}^j_{k|k}$ that incorporates the event information without introducing too much conservatism. Define the errors for sensor j at time k

$$e_k^j = x_k^j - \tilde{x}^j_{k|k},$$
$$\hat{e}_k^j = x_k^j - \hat{x}^j_{k|k},$$
$$\tilde{e}_k^j = \hat{x}^j_{k|k} - \tilde{x}^j_{k|k}$$

where x_k^j evolves by (3) and $\hat{x}_{k|k}^j$, $\tilde{x}_{k|k}^j$ are defined by (4a) and (5), respectively. Moreover, we define the innovation at time k of sensor j as

$$\tilde{y}_k^j = y_k^j - C_1^j \hat{x}_{k|k-1}^j.$$

Clearly, the event rule (6) can be described as a function of $\tilde{e}_k^j$. Furthermore, the variable $\tilde{e}_k^j$ can be expressed by the innovation sequence $\{\tilde{y}_\ell^j\}_{\ell \in \{\tau_k+1,\ldots,k\}}$. The subsequent lemma is central for the computation of $\tilde{P}_{k|k}^j$. It shows that the error covariance matrix $\mathsf{E}[(e_k^j)(e_k^j)^\top | \hat{x}_k^j, \sigma_k^j]$ can be decomposed into two parts. One arises from the estimation error at the local sensor and the other is due to the uncertainty of the event information.

Lemma 2: Let $\sigma_k^j < k$. Then, the following equality holds for sensor j, $1 \leq j \leq M$,

$$\mathsf{E}[(e_k^j)(e_k^j)^\top | \hat{x}_k^j, \sigma_k^j] = P_{k|k}^j + \mathsf{E}[(\tilde{e}_k^j)(\tilde{e}_k^j)^\top | \hat{x}_k^j, \sigma_k^j] \quad (10)$$

Proof: Let us define $I_k^j = \{\hat{x}_k^j, \sigma_k^j\}$ Then, we can reformulate the covariance matrix of the estimation error e_k^j as

$$\begin{aligned}\mathsf{E}[(e_k^j)(e_k^j)^\top | I_k^j] &= \mathsf{E}[(\hat{e}_k^j + \tilde{e}_k^j)(\hat{e}_k^j + \tilde{e}_k^j)^\top | I_k^j] \\ &= \mathsf{E}[(\hat{e}_k^j)(\hat{e}_k^j)^\top | I_k^j] + \mathsf{E}[(\hat{e}_k^j)(\tilde{e}_k^j)^\top | I_k^j] \\ &\quad + \mathsf{E}[(\tilde{e}_k^j)(\hat{e}_k^j)^\top | I_k^j] + \mathsf{E}[(\tilde{e}_k^j)(\tilde{e}_k^j)^\top | I_k^j]\end{aligned}$$

Define the index set $\mathcal{L}_k^j = \{\tau_k^j + 1, \ldots, k\}$ and the measurement sequence $Y_{\mathcal{L}_k^j}^j = \{y_\ell^j\}_{\ell \in \mathcal{L}_k^j}$. Then, we obtain for the first term

$$\begin{aligned}\mathsf{E}[(\hat{e}_k^j)(\hat{e}_k^j)^\top | I_k^j] &= \mathsf{E}[\mathsf{E}[(\hat{e}_k^j)(\hat{e}_k^j)^\top | I_k^j, Y_{\mathcal{L}_k^j}^j] | I_k^j] \\ &= \mathsf{E}[\mathsf{E}[(\hat{e}_k^j)(\hat{e}_k^j)^\top | \hat{x}_k^j, Y_{\mathcal{L}_k^j}^j] | I_k^j] \\ &= \mathsf{E}[P_{k|k}^j | I_k^j] \\ &= P_{k|k}^j.\end{aligned}$$

The first equality is because of the tower property of the conditional expectation. The second equality is due to the fact that σ_k^j is a measurable random variable with respect to $\{\hat{x}_k^j, Y_{\mathcal{L}_k^j}^j\}$. The third equality arises from the basic properties of the Kalman filter: the error covariance matrix is invariant with respect to the measurements used by the filter. For the cross correlation, we obtain

$$\begin{aligned}\mathsf{E}[(\hat{e}_k^j)(\tilde{e}_k^j)^\top | I_k^j] &= \mathsf{E}[\mathsf{E}[\hat{e}_k^j (\tilde{e}_k^j)^\top | I_k^j, Y_{\mathcal{L}_k^j}^j] | I_k^j] \\ &= \mathsf{E}[\mathsf{E}[\hat{e}_k^j | \hat{x}_k^j, Y_{\mathcal{L}_k^j}^j](\tilde{e}_k^j)^\top | I_k^j] \\ &= 0.\end{aligned}$$

In the first equality, we use the tower property of the conditional expectation. The second equality is that σ_k^j is a measurable random variable with respect to $\{\hat{x}_k^j, Y_{\mathcal{L}_k^j}^j\}$. The cross terms vanish eventually due to the unbiasedness of estimate $\hat{x}_{k|k}^j$, i.e., $\mathsf{E}[\hat{e}_k^j | \hat{x}_k^j, Y_{\mathcal{L}_k^j}^j] = 0$. This concludes the proof. ■

In the light of Lemma 2, our remaining concern is the computation of the error covariance matrix $\mathsf{E}[(\tilde{e}_k^j)(\tilde{e}_k^j)^\top | \hat{x}_k^j, \sigma_k^j]$. As the covariance depends on the event information, there will not exist a closed-form solution in general and its computation will involve the knowledge of the conditional probability distribution, which makes the calculation infeasible already for a moderate number of state variables. This motivates us to find an upper bound on $\mathsf{E}[(\tilde{e}_k^j)(\tilde{e}_k^j)^\top | \hat{x}_k^j, \sigma_k^j]$ whose computation scales with the state dimension, n_o^j, at each sensor j, $1 \leq j \leq M$. By the definition of the event rule (6), the event information $\delta_k^j = 0$ corresponds to

$$\|\tilde{e}_k^j\|_{\Gamma_k^j} \leq r_k^j.$$

Given the last transmission of sensor j at time τ^j, the discrepancy between the Kalman estimate at the sensor and the predicted value given by $\tilde{e}_\ell^j$ evolves by the recursion

$$\tilde{e}_{\ell+1}^j = A_1^j \tilde{e}_\ell^j + K_{\ell+1}^j \tilde{y}_{\ell+1}^j$$

for $\ell > \tau^j$ with $\tilde{e}_{\tau^j}^j = 0$. As the innovation process can be regarded as white noise [22], $\tilde{y}_{\ell+1}^j$ is independent of $\tilde{e}_\ell^j$. Clearly, this also holds when conditioned on $\delta_{\tau^j}^j = \cdots = \delta_\ell^j = 0$.

C. Unimodality and symmetry

In the following, we characterize the conditional probability distribution of $\tilde{e}_k^j$ given the event information $\delta_{\tau_k^j+1}^j = \cdots = \delta_k^j = 0$ summarized by σ_k^j.

Let us introduce two definitions from [23] for the characterization.

Definition 3 (Centrally symmetric sets): A set $A \subset \mathbb{R}^n$ is said to be *centrally symmetric* if $x \in A$ implies $-x \in A$. We call a distribution in $\mathbb{R}^n$ *centrally symmetric* if its density function $f(x)$ satisfies $f(x) = f(-x)$.

Definition 4 (Central convex unimodality): A probability distribution in $\mathbb{R}^n$ is said to be *central convex unimodal* if it belongs to the closed convex hull of the set of all uniform distributions on centrally symmetric sets in $\mathbb{R}^n$.

The above definition implies that central convex unimodal distributions are centrally symmetric.

According to the statistical properties of the innovations process, the prior probability distribution of $\tilde{e}_\ell^j$ at time $\ell = \tau^j + 1$ is zero-mean Gaussian with covariance matrix $K_\ell^j (C_1^j P_{\ell|\ell-1}^j (C_1^j)^\top + R_v^j)(K_\ell^j)^\top$. Clearly, Gaussian distributions are central convex unimodal since the sup-level sets of their density function are convex.

By applying the Bayes' rule recursively, the following operations are applied sequentially until arriving at the conditional distribution of $\tilde{e}_k^j$.

- Conditioning on $\delta_\ell^j = 0$, i.e., set the density function to 0 for all $\|\tilde{e}_k^j\|_{\Gamma_k^j} \geq r_k^j$
- Skew the density function by A_1^j
- Add the weighted innovation, i.e., convolute with a zero-mean Gaussian

The last step results from the fact that the density function of two independent random variables can be computed by the convolution of the individual densities. We have omitted here that the first two steps involves normalization, as normalization does not change the shape of the density function and therefore naturally preserves symmetry and unimodality. It is known that the operations defined in the above list preserve central convex unimodality [23]. Therefore, we can conclude that the conditional probability distribution of $\tilde{e}_k^j$ given the event information σ_k^j is central convex unimodal.

As a consequence of symmetry, we have that $\mathsf{E}[\tilde{e}_k^j|\sigma_k^j] = 0$. This implies that the linear predictor $\tilde{x}_{k|k}^j$ in (5) is identical with the MMSE estimate of $\hat{x}_{k|k}^j$ at time k given the event information σ_k^j, i.e.,

$$\tilde{x}_{k|k}^j = \mathsf{E}[\hat{x}_{k|k}^j|\hat{x}_{\sigma_k^j|\sigma_k^j}^j, \sigma_k^j].$$

D. Bounds on the error covariance matrix

Based on the symmetry and unimodality properties derived in the last section, we aim at finding upper bounds on the corresponding covariance matrices $\mathsf{E}[(\tilde{e}_k^j)(\tilde{e}_k^j)^\top|\sigma_k^j]$.

1) Scalar subsystems ($n_o^j = 1$): For illustrative purposes, we will first consider first-order subsystems at sensor j, i.e., the dimension of the observable subspace at sensor j denoted as n_o^j is 1. The following lemma gives a bound on the variance of the error $\tilde{e}_k^j$.

Lemma 3: Let $n_o^j = 1$ and $\sigma_k^j < k$. Then, the variance of $\tilde{e}_k^j$ given the event information σ_k^j is upper bounded by

$$\mathsf{E}[(\tilde{e}_k^j)^2|\sigma_k^j] \leq \frac{1}{3}(r_k^j)^2(\Gamma_k^j)^{-1} \tag{11}$$

Proof: Because of the assumption $\sigma_k^j < k$, we have $\|\tilde{e}_k^j\|_{\Gamma_k^j} \leq r_k^j$. Therefore, the support of the density function of $\tilde{e}_k^j$ given σ_k^j will be given by

$$\mathcal{D}_1 = \Big[-\frac{r_k^j}{\sqrt{\Gamma_k^j}}, \frac{r_k^j}{\sqrt{\Gamma_k^j}} \Big].$$

Moreover, we can assert from the previous discussion that the density function of $\tilde{e}_k^j$ given σ_k^j is even and unimodal, having its peak at 0. Among these density functions having their support on $\mathcal{D}_1$, it is well known that the uniform distribution on $\mathcal{D}_1$ maximizes the variance [24]. This gives us the bound in (11). ■

Remark 3: In a worst case analysis as performed in [11] and [10], the error covariance matrix is bounded by $(r_k^j)^2(\Gamma_k^j)^{-1}$. This implies that taking into account symmetry and unimodality improves the bound at least by a factor of 3 for scalar subsystems. This bound could further be improved by incorporating σ_k^j, the system matrix A_1^j, the Kalman gain and the statistical properties of the innovations process.

Remark 4: The inequality (11) is tight in the sense that we can choose the system parameters in such way that the error covariance matrix approaches the bound arbitrarily close.

2) Higher-order subsystems ($n_o^j > 1$): In the general case, we obtain the following bound on the error covariance matrix.

Lemma 4: Let $\sigma_k^j < k$. Then, the covariance matrix of $\tilde{e}_k^j$ given the event information σ_k^j is upper bounded by

$$\mathsf{E}[(\tilde{e}_k')(\tilde{e}_k')^\top|\sigma_k^j] \leq \frac{n_o^j}{2+n_o^j}(r_k^j)^2(\Gamma_k^j)^{-1}. \tag{12}$$

Proof: Let us fix an arbitrary j and define the bijective state transformation

$$\tilde{e}_k' = \frac{1}{r_k^j}\Gamma_k^{\frac{1}{2}}\tilde{e}_k^j$$

where $\Gamma_k^j = (\Gamma_k^{\frac{1}{2}})^\top\Gamma_k^{\frac{1}{2}}$. Then, the event information $\delta_k^j = 0$ translates into

$$\|\tilde{e}_k'\| \leq 1.$$

As the above transformation is linear and bijective, it will preserve central convex unimodality. Denote the conditional error covariance matrix as

$$\Sigma_{\tilde{e}_k'} = \mathsf{E}[(\tilde{e}_k')(\tilde{e}_k')^\top|\sigma_k^j].$$

Let us take the eigenvalue decomposition

$$\Sigma_{\tilde{e}_k'} = V\Delta V^\top$$

where $V^\top V = I_{n_o^j}$ with $V = [v_1, \ldots, v_{n_o^j}]$ being composed of the eigenvectors v_i of $\Sigma_{\tilde{e}_k'}$ with corresponding eigenvalue $\lambda_i \geq 0$ and $\Delta = \mathrm{diag}[\lambda_i]$.

Define the hyper-rectangle

$$\mathcal{R}_1 = \{\zeta \in \mathbb{R}^{n_o^j} | \|\zeta\|_\infty \leq 1\}$$

and consider its V-transformed version

$$\mathcal{R}_V = \{\zeta \in \mathbb{R}^{n_o^j} | \|V^\top\zeta\|_\infty \leq 1\}.$$

Consider the uniform distribution on $\mathcal{R}_V$. Compare the marginals of this distribution with the conditional probability distribution of $\tilde{e}_k'$ projected onto v_i. We make two observations. On the one hand, the marginal of the uniform distribution is again a uniform distribution on the interval $[-1, 1]$. On the other hand, we know that the marginal distribution of a central convex unimodal distribution is again central convex unimodal [23]. Therefore, the marginal of the conditional probability distribution of $\tilde{e}_k'$ is central convex unimodal with support on $[-1, 1]$. Similar as in the scalar case, we conclude that the variance of the uniform distribution on $[-1, 1]$ is not smaller than any symmetric unimodal distribution with support on $[-1, 1]$. As this holds true for all $1 \leq i \leq n_o^j$, we conclude that the covariance matrix of the uniform distribution on $\mathcal{R}_V$ is a bound on the error covariance matrix $\Sigma_{\tilde{e}_k'}$.
As we do not know V, we need to take an overapproximation of the covariance matrices resulting from any choice of V. This can be done by considering the uniform distribution on the n_o^j-ball with radius $\sqrt{n_o^j}$. Its covariance matrix is identical with the right-hand side of (12) when transformed back into the original coordinates system referring to $\tilde{e}_k^j$. This concludes the proof. ■

Remark 5: Similar as in the scalar case, we can improve the bound on the error covariance matrix by a factor of $\frac{n_o^j}{2+n_o^j}$ with respect to the a worst case analysis as performed in [11]. Hence, the benefit of exploiting unimodality and symmetry vanishes when n_o^j grows in our case. It is believed that incorporating more statistical details of the linear system will lead to a significant improvement.

By taking Lemma 2–4 into account, we arrive at the following result on consistency.

Theorem 2: Let $\sigma_k^j < k$, $1 \le j \le M$. If we set

$$\tilde{P}_{k|k}^j = P_{k|k}^j + \frac{n_o^j}{2+n_o^j}(r_k^j)^2(\Gamma_k^j)^{-1}, \tag{13}$$

then, the pair $(\tilde{x}_{k|k}^j, \tilde{P}_{k|k}^j)$ is a consistent estimate of x_k^j.

Remark 6: Because of the above theorem and the consistency-preserving property of the modified CI given in Theorem 1, we can conclude that the estimate $(\hat{x}_k, \hat{P}_k)$ at the fusion center is consistent when using (13) in (7).

IV. STABILITY ANALYSIS

In this section, we investigate the asymptotic behavior of the event-triggered state estimator. Our notion for stability of the filter at the fusion center is given by means of the bounded mean square error. By using the concept of consistency and the bound on the error covariance matrix derived in Theorem 2, we can state the following assertion on stability.

Theorem 3: Let the system (1) and (2) be collective observable and let the time-varying event-triggering parameters be uniformly bounded by $\Gamma_k^j \ge \Gamma^j > 0$ and $r_k^j \le r^j$. Suppose that $\tilde{P}_{k|k}^j$ is defined by (13) for $1 \le j \le M$. Then, the covariance intersection algorithm given by (8)–(9) yields a stable estimate $\hat{x}_k$ in the sense of a uniformly bounded mean square error in the limit, i.e.,

$$\limsup_{k\to\infty} \mathsf{E}[\|x - \hat{x}_k\|^2] < \infty.$$

Proof: Because R_0, R_w and R_v^j are positive definite, $P_{k|k}^j$ can be assumed to be positive definite for any time k. This together with collective observability and the definition of $\tilde{P}_{k|k}^j$ implies that $\hat{P}_k^{-1}$ is non-singular due to Lemma 1 for any time k. As we have the identity

$$\mathsf{E}[\|x - \hat{x}_k\|^2] = \mathrm{tr}[\mathsf{E}[(x-\hat{x}_k)(x-\hat{x}_k)^\top]],$$

it suffices to show that the error covariance matrix is bounded in the limit when k approaches infinity. Furthermore, we know that the estimate $(\hat{x}_k, \hat{P}_k)$ is consistent because of the consistency-preserving property of the modified CI in (8)–(9) due to Theorem 1 and the consistency guarantee due to Theorem 2. Due to the definition of consistency, it remains to show that we have a bound on $\hat{P}_k$ in the limit $k \to \infty$. Define P_∞^j to be the limit error covariance matrix of the local Kalman filter at sensor j given by (4a)–(4d). The limit exists by construction due to observability of (A_1^j, C_1^j) for each $1 \le j \le M$. By taking the uniform bounds of the triggering rule into account, we arrive at the following bound on the mean square error in the limit

$$\limsup_{k\to\infty} \mathsf{E}[\|x - \hat{x}_k\|^2] \le \mathrm{tr}\left[\left(\sum_{j=1}^M \omega^j T_1^j \left(P_\infty^j + \frac{n_o^j}{2+n_o^j}(r^j)^2(\Gamma^j)^{-1}\right)^{-1}(T_1^j)^\top\right)^{-1}\right] \tag{14}$$

This concludes the proof. ■

Remark 7: In addition to the stability guarantee, we are also able to quantify a bound on the error covariance matrix given by (14).

V. NUMERICAL EXAMPLE

In the following, we provide a numerical example to demonstrate the benefits of our approach. Suppose the dynamical system

$$x_{k+1} = \begin{bmatrix} 1.25 & 0.25 \\ 0 & 1.05 \end{bmatrix} x_k + w_k$$

with $R_0 = I_2$ and $R_w = \frac{1}{4}I_2$. The system is observed by $M = 10$ sensors with measurement matrices

$$C^1 = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad C^2 = \begin{bmatrix} 4/\sqrt{41} & 5/\sqrt{41} \end{bmatrix},$$
$$C^j = \begin{bmatrix} 0.1j & 1-0.1j \end{bmatrix}, \quad 3 \le j \le 10.$$

and with a measurement noise variance $R_v^j = 2$ for $1 \le j \le 10$. The systems (A, C^1) and (A, C^2) are not observable with $n_o^1 = n_o^2 = 1$. The triggering threshold is given by (6) with $\Gamma_k^j = I_{n_o^j}$ and the threshold r_k^j is chosen to be identical for any $k \ge 0$ and any $j \in \{1, \dots, 10\}$. The weights of the CI algorithm are assumed to be identical for every sensor in all cases.

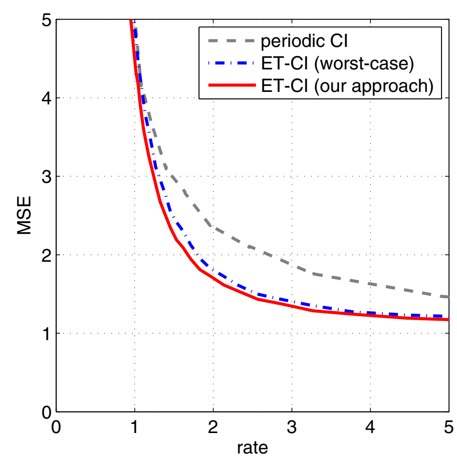

Fig. 2. Trade-off curve of the overall transmission rate and the mean square error (MSE).

Fig. 2 depicts the trade-off between the mean square error (MSE) and the average total rate for three different schemes. The values on the trade-off curve are determined by computing the empirical means of the MSE and the total rate for a fixed transmission scheme over a horizon of 50 with 10,000 trials. By either varying the thresholds for the event-triggered strategy or the sampling period in the periodic case, different transmission rates are achieved. The red solid line refers to our approach using the event-triggered CI (ET-CI) with the covariance bound (13). In the case of the dashed blue line, we assume a worst case bound omitting the correction factor $n_o^j/(2+n_o^j)$ in (13). The gray dashed line refers to periodically sampled sensors using the CI. In this case the sensors are sampled at the same sampling period $T_S \in \{1, 2, \dots, 10\}$ resulting in a total rate of M/T_S. The phase shift between transmissions of sensor j and $j+1$ is 1 for $1 \leq j < 10$. When comparing our event-triggered strategy with the periodic scheme, we observe a performance improvement ranging from 15% to 30% within the rate range of $[1,\ 4]$ taking its maximum at a total rate of 2. Compared with the worst-case event-triggered strategy, we obtain a performance gain between 3% and 8% in the rate range of $[1,\ 4]$ taking its maximum at a total rate of 1.2.

The averages of the predicted MSE at the fusion center are illustrated in Fig. 3. These average bounds are obtained by calculating the empirical mean of $\mathrm{tr}[\hat{P}_k]$ for a fixed transmission scheme over a horizon of 50 with $10,000$ trials. We observe a significant difference between our approach using the covariance bound (13) and the worst-case computation. The improvement on the average MSE estimate ranges from 20% to 55% within the rate range of $[1,\ 4]$, which is growing for a decreasing total rate.

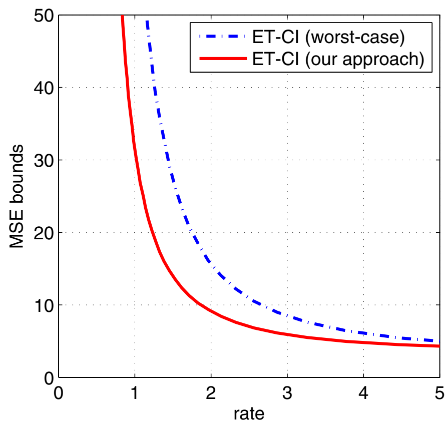

Fig. 3. Bounds on the mean square error (MSE) depending the overall transmission rate.

VI. Summary

This paper demonstrates how event information can be incorporated consistently for event-triggered state estimation within a sensor network. The key feature is to exploit symmetry and unimodality properties of the error distribution in order to give non-trivial bounds on the error covariance matrix that take into account the event information. These bounds ensure that the event-triggered state estimator is stable in the sense of bounded mean square error.

Future work includes the consideration of cross-correlations between sensor estimates to improve performance at the fusion center, the incorporation of more statistical information of the dynamical system in order to obtain tighter bounds on the error covariance matrix and the study of unreliable communication channels for the transmission of sensor data.

References

[1] G. Lipsa and N. Martins, "Remote state estimation with communication costs for first-order LTI systems," *Automatic Control, IEEE Transactions on*, vol. 56, no. 9, pp. 2013–2025, 2011.

[2] A. Molin and S. Hirche, "An iterative algorithm for optimal event-triggered estimation," in *IFAC Conference on Analysis and Design of Hybrid Systems*, pp. 64–69, 2012.

[3] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," *Automatic Control, IEEE Transactions on*, vol. 59, pp. 3266–3281, Dec 2014.

[4] K. J. Åström and B. M. Bernhardsson, "Comparison of Riemann and Lebesgue sampling for first order stochastic systems," in *Proceedings of the 41th IEEE Conference on Decision and Control (CDC'02)*, 2002.

[5] S. J. Julier and J. K. Uhlmann, "A non-divergent estimation algorithm in the presence of unknown correlations," in *Proceedings of the American Control Conference*, pp. 2369–2373, 1997.

[6] Y. S. Suh, V. H. Nguyen, and Y. S. Ro, "Modified Kalman filter for networked monitoring systems employing a send-on-delta method," *Automatica*, vol. 43, no. 2, pp. 332–338, 2007.

[7] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *Automatic Control, IEEE Transactions on*, vol. 58, no. 4, pp. 1041–1046, 2013.

[8] D. Shi, T. Chen, and L. Shi, "An event-triggered approach to state estimation with multiple point-and set-valued measurements," *Automatica*, vol. 50, no. 6, pp. 1641–1648, 2014.

[9] J. Sijs and M. Lazar, "Event based state estimation with time synchronous updates," *Automatic Control, IEEE Transactions on*, vol. 57, no. 10, pp. 2650–2655, 2012.

[10] J. Sijs, B. Noack, and U. D. Hanebeck, "Event-based state estimation with negative information," in *Information Fusion (FUSION), 2013 16th International Conference on*, pp. 2192–2199, IEEE, 2013.

[11] B. Noack, F. Pfaff, and U. D. Hanebeck, "Combined stochastic and set-membership information filtering in multisensor systems," in *Information Fusion (FUSION), 2012 15th International Conference on*, pp. 1218–1224, IEEE, 2012.

[12] G. Battistelli, A. Benavoli, and L. Chisci, "Data-driven communication for state estimation with sensor networks," *Automatica*, vol. 48, no. 5, pp. 926–935, 2012.

[13] C. Ramesh, H. Sandberg, and K. H. Johansson, "Multiple access with attention-based tournaments for monitoring over wireless networks," in *European Control Conference*, pp. 4302–4307, EUCA, 2009.

[14] M. G. Cea and G. C. Goodwin, "Event based sampling in non-linear filtering," *Control Engineering Practice*, vol. 20, no. 10, pp. 963–971, 2012.

[15] A. H. Jazwinski, *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[16] M. Farina, G. Ferrari-Trecate, and R. Scattolini, "Distributed moving horizon estimation for linear constrained systems," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2462–2475, 2010.

[17] P. J. Antsaklis and A. N. Michel, *Linear systems*. Springer Science & Business Media, 2006.

[18] R. Cogill, "Event-based control using quadratic approximate value functions," in *Proceedings of the 48th IEEE Conference on Decision and Control (CDC'09)*, pp. 5883–5888, IEEE, 2009.

[19] L. Chen, P. O. Arambel, and R. K. Mehra, "Estimation under unknown correlation: covariance intersection revisited," *Automatic Control, IEEE Transactions on*, vol. 47, no. 11, pp. 1879–1882, 2002.

[20] G. Reissig, "Computing abstractions of nonlinear systems," *Automatic Control, IEEE Transactions on*, vol. 56, pp. 2583–2598, Nov 2011.

[21] L. El Ghaoui and G. Calafiore, "Robust filtering for discrete-time systems with bounded noise and parametric uncertainty," *Automatic Control, IEEE Transactions on*, vol. 46, no. 7, pp. 1084–1089, 2001.

[22] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

[23] S. W. Dharmadhikari and K. Joag-Dev, *Unimodality, convexity, and applications*, vol. 27. Academic Press, 1988.

[24] H. Gray and P. Odell, "On least favorable density functions," *SIAM Review*, vol. 9, no. 4, pp. 715–720, 1967.