On Set-Valued Kalman Filtering and Its Application to Event-Based State Estimation

Dawei Shi, Tongwen Chen, *Fellow, IEEE*, and Ling Shi

***Abstract*—Motivated by challenges in state estimation with event-based measurement updates, the properties of the exact and approximate set-valued Kalman filters with multiple sensor measurements for linear time-invariant systems are investigated in this work. First, we show that the exact and the proposed approximate set-valued filters are independent of the fusion sequence at each time instant. Second, the boundedness of the size of the set of estimation means is proved for the exact set-valued filter. For the approximate set-valued filter, if the closed-loop matrix is contractive, then the set of estimation means has a bounded size asymptotically; otherwise a nonsingular linear transform is constructed such that the size of the set of estimation means for the transformed states is asymptotically bounded. Third, the effect of set-valued measurements on the size of the set of estimation means is analyzed and conditions for performance improvement in terms of smaller size of the set of estimation means are proposed. Finally, the results are applied to event-based estimation, which allow the event-triggering conditions to be designed by considering requirements on performance and communication rates. The efficiency of the proposed results are illustrated by simulation examples and comparison with the approximate event-based MMSE estimator and the Kalman filter with intermittent observations.**

***Index Terms*—Event-based estimation, Kalman filter, Minkowski sum, set-valued estimation.**

I. Introduction

Due to the energy-consumption requirements in wireless communication [1], [2], applications of wireless control and monitoring systems require more energy-efficient data sampling/transmission strategies to reduce the communication burden [3]. The advent of event-based sampling and data-scheduling [4], [5] provides a promising solution to this issue. In return, it has brought on new challenges to the design of control and state estimation techniques for dynamical systems.

The motivation of this work stems from event-based state estimation for discrete-time Gaussian systems in the Bayesian framework. In this scenario, the sensors determine whether to send their current measurements to the estimator by testing whether the so-called "event-triggering conditions" are violated or not [6], [7] (e.g., the "send-on-delta" conditions [7]—a sensor does not transmit its measurement unless the current measurement deviates from the previously transmitted measurement by a specified level). Therefore, when the triggering conditions are satisfied, the estimator knows the measurement information (assuming the reliability of the communication channel); when the sensor decides not to send measurement information, the estimator still knows that the measurement lies within a set characterized by the triggering condition. In many cases, this set is known to the estimator without additional communication, e.g., the set described by the "send-on-delta" conditions. In this context, the challenge mainly arises from the additional information provided by the event-triggering conditions during non-event time instants, which results in a state estimation problem with combined set- and point-valued measurement updates. For the case of periodic point-valued measurements, the Gaussianity of the conditional (*a posteriori*) distributions leads to a simple closed-form Minimum Mean Squared Error (MMSE) estimator, or equivalently, the estimate with the smallest estimation error covariance [8]. Due to the combined set- and point-valued measurements, however, the conditional distributions are no longer Gaussian, and the exact MMSE estimator becomes computationally expensive to calculate, as is similar for the nonlinear Gaussian filters [9]–[11].

To overcome this difficulty, some efforts have been attempted. Utilizing a Gaussian assumption on the distribution of the state conditioned on all past set- and point-valued measurement information, the MMSE estimator was derived in [12], and the tradeoff between communication rate and performance was explicitly analyzed; the extension of the results to more general event-triggering conditions and multiple sensor measurements was considered in [13]. A general description of event-based sampling was proposed in [14], and an event-based estimator with a hybrid update was proposed by approximating the uniform distribution with a sum of a finite number of Gaussian distributions to reduce the computational complexity. The Gaussian assumption was shown to be maintained in [15], where a special event-triggering condition was proposed by introducing randomization in the triggering sets. In [16], a state estimate was obtained by minimizing the maximum possible mean squared error in the presence of both the stochastic uncertainty caused by the noises and the non-stochastic uncertainty introduced by the event-triggering conditions. Alternatively, the maximum likelihood estimation problem for an event-triggering scheme quantifying the magnitude of the innovation

Manuscript received January 21, 2014; revised June 18, 2014 and October 10, 2014; accepted November 4, 2014. Date of publication November 12, 2014; date of current version April 18, 2015. This work was supported by the Natural Sciences and Engineering Research Council of Canada and by an HK RGC GRF Grant 618612. Recommended by Associate Editor J. Daafouz.

D. Shi is with the State Key Laboratory of Intelligent Control and Decision of Complex Systems and the School of Automation, Beijing Institute of Technology, P.R. China. This work was completed when he was a PhD student at the University of Alberta, Canada (e-mail: dshi@ualberta.ca).

T. Chen is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: tchen@ualberta.ca).

L. Shi is with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: eesling@ust.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2014.2370472

of the estimator at each time instant was studied in [17], and the computation of upper and lower bounds for the communication rate was discussed. For further results on event-based estimation, see [18]–[23] and references therein.

The event-based estimation problem can also be considered as a set-valued Kalman filtering problem in the convex Bayesian decision framework [24], which takes the difference and separation between "stochastic uncertainty" and "non-stochastic uncertainty" into account. In [25], it was pointed out that "Ignorance, in its root meaning, means lack of knowledge; uncertainty, on the other hand, typically means lack of precision. By specifying a probability distribution for a random variable, we attempt to characterize uncertainty. If the correct distribution function is unknown, that is a manifestation of ignorance." From this perspective, the statistical information of the noise and the initial states are regarded as "uncertainty" (or "stochastic uncertainty"), while the ambiguous information contained in the event-triggering sets can be considered as "ignorance" (or "non-stochastic uncertainty"), since the estimator's inability of knowing the point-valued measurement information during non-event time instants can be regarded as "lack of knowledge," which is caused by the subjective choice of the event-triggering conditions.

Compared with the existing results in event-based state estimation, the set-valued filtering approach provides an alternative way of exploiting and understanding the additional information contained in the event-triggering conditions. The set-valued Kalman filter was originally introduced by [25], where the standard Kalman filter was extended to the case that a convex set of initial estimate distributions was considered. Although the filters bear good asymptotic properties, they are not applicable to the event-based estimation scenario, since only point-valued measurements are considered. Recently, further relaxation of the assumptions on uniqueness for the *a posteriori* probability distributions was considered in [26] by allowing set-valued measurements and the multiple sensor fusion problem was considered in [27] utilizing the information filter approach. When the set-valued measurements are treated as non-stochastic uncertainty, the choice of different points in the measurement set at each time instant only leads to different values of the estimation mean (which we refer to as "the set of estimation means" hereafter), while the estimation error covariance remains unaffected. These results allow set-valued event-based estimators to be designed; however, several problems remain unexplored with respect to these new set-valued filters, which are of fundamental importance for the study of event-based estimation:

1) For multiple point-valued measurements, the performance is quantified only in terms of the estimation error covariance, and it is known that the fusion sequence used to update the sensor measurement information at the same time instant does not affect the resultant centralized Kalman filter. For set-valued Kalman filters, the overall performance is measured according to two terms, the estimation error covariance and the size of the set of estimation means (e.g., the size of an ellipsoidal set can be quantified by the trace of a positive semidefinite matrix defining the shape of the set in this work). Apparently, the fusion sequence still does not affect the error covariance, but its effect on the size of estimation means is not known.
2) In [25], it was shown that the set of estimation means converges towards a singleton as time goes to infinity for point-valued measurements. However, the asymptotic behavior of the size of the set of estimation means is not clear when set-valued measurements are considered. In addition, since the summation of ellipsoids may not be ellipsoids at all [28], the analytical expression of the exact set-valued estimator cannot be maintained, and consequently the set of estimation means can only be calculated approximately [26]; in this regard, the asymptotic property of the size of the approximate set of estimation means is of importance as well.
3) For standard Kalman filters, it is known that increasing the number of sensors can reduce the estimation error covariance; this result is still valid for the set-valued case. The effect of adding more sensors on the size of the set of estimation means is, however, still unknown.

In this work, we seek to explore the above problems for linear time-invariant systems with an emphasis on event-based estimation. The estimation problem is considered in the multiple sensor scenario, where each sensor is allowed to provide its own set-valued measurement parameterized by ellipsoids. The difference between the results presented in this work and those in [12], [13] is that the additional information introduced by the event-triggering conditions is treated as non-stochastic uncertainty in the current work, which leads to a set-valued estimator, while this information was exploited as stochastic uncertainty in [12], [13], based on which approximate MMSE estimates were developed. The main contributions are summarized as follows.

1) The exact set of estimation means is shown to be invariant with respect to the fusion sequences. Since the exact set-valued filter is normally not implementable, a two-step approximate set-valued estimator is proposed and is shown to be unaffected by the fusion sequences. The approximate estimator proposed here is different from that in [27], which was given in the information filtering form.
2) The boundedness of the size of the set of estimation means for the exact set-valued filter is proved. For the approximate estimator, we show that if the closed-loop matrix is contractive at steady state, then the boundedness of the size of the set of estimation means is guaranteed; otherwise, there exists an invertible linear transformation such that the size of the set of estimation means of the approximate estimator after the transformation is bounded.
3) An upper bound on the steady-state performance in terms of the size of the set of estimation means is proposed, based on which the conditions can be characterized to test whether a smaller upper bound on the size of the set of estimation means at steady state can be achieved by including an additional sensor. For scalar systems, a sufficient condition is provided for guaranteed performance improvement. Based on the developed results, an optimal

event-triggering condition design problem is further formulated and solved.

The application of set-valued filtering to event-based state estimation provides an encouraging approach to exploiting the information contained in the event-triggering conditions. Compared with the approximate event-based MMSE estimates, the centre of the set-valued estimator always serves as a point-valued estimate with the best robustness performance, in the sense that it has the smallest worst-case distance to the Kalman filter with periodic observations; due to the results presented in this work, this worst-case distance is numerically computable. By comparison, the precision of the Gaussian assumptions of the non-Gaussian distributions utilized to develop the approximate event-based MMSE estimates are normally not possible to be verified [13]. In addition, due to the lack of closed-form expressions, the approximate event-based MMSE estimators are computationally expensive to implement for sensors with more than one channel (namely, $m > 1$) [13], although for the case of $m = 1$, the empirical estimation performance of the approximate event-based MMSE estimator is observed to be similar to that of the centre of the set-valued estimator. Compared with the Kalman filter with intermittent observations, the set-valued estimator has a smaller estimation covariance (at the cost of getting a set of estimation means) due to the utilization of the same filtering gain for all estimates in the set. The centre of the set-valued estimator also has improved empirical estimation performance in terms of the average state estimation error compared with the Kalman filter with intermittent observations, as will be shown by the simulation examples.

The remainder of the paper is organized as follows: Section II presents the system description and problem setup. Section III discusses the effect of the fusion sequence and the proposed approximate set-valued estimator. The asymptotic properties of the set-valued estimators are provided in Section IV. Section V presents analysis on performance improvement. The application of the results to set-valued event-based estimation is given in Section VI, followed by the concluding remarks and future work in Section VII.

The following notation and symbols will be used. $\mathbb{R}$ denotes the set of real numbers. $\mathbb{N}$ denotes the set of nonnegative integers. $\mathbb{N}^+$ denotes the set of positive integers. Let $m, n \in \mathbb{N}^+$; $\mathbb{R}^{m\times n}$ denotes the set of m by n real-valued matrices. For brevity, denote $\mathbb{R}^m := \mathbb{R}^{m\times 1}$. For $v \in \mathbb{R}^m$, let $\|v\|$ denote its Euclidean norm. For $Z \in \mathbb{R}^{m\times n}$, $Z^\top$ denotes the transpose of Z, and $\|Z\|_2$ denotes the spectral norm of Z. The symbol I denotes the identity matrix with a context-dependent size. For $X, Y \in \mathbb{R}^{n\times n}$, $X > (\geq) Y$ means $X - Y$ is positive definite (positive semidefinite). For two convex sets $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$, let $\mathcal{X} \oplus \mathcal{Y}$ denote their Minkowski sum, namely, $\mathcal{X} \oplus \mathcal{Y} := \{x + y | x \in \mathcal{X}, y \in \mathcal{Y}\}$. Also, $\bigoplus_{i=1}^n \mathcal{X}_i := \mathcal{X}_1 \oplus \mathcal{X}_2 \oplus \cdots \oplus \mathcal{X}_n$. For $T \in \mathbb{R}^{m\times n}$ and $\mathcal{X} \subseteq \mathbb{R}^n$, define $T\mathcal{X}$ as

$$T\mathcal{X} := \{Tx \in \mathbb{R}^m | x \in \mathcal{X}\}.$$

Given $Y > 0$, an ellipsoidal set (or an ellipsoid) $\mathcal{Y} = \mathcal{E}(c, Y)$ in $\mathbb{R}^m$ is defined as

$$\mathcal{Y} := \mathcal{E}(c, Y) = \left\{y \in \mathbb{R}^m | (y - c)^\top (Y)^{-1}(y - c) \leq 1, Y > 0\right\}$$

if Y is singular and $Y \geq 0$, $\mathcal{Y}$ is parameterized as[1]

$$\mathcal{Y} = \left\{y \in \mathbb{R}^m | \langle l, y\rangle \leq \langle l, c\rangle + \langle l, Yl\rangle^{\frac{1}{2}}, \forall\, l \in \mathbb{R}^m\right\}.$$

In this work, we define the size of an ellipsoidal set $\mathcal{Y}$ as $\mathrm{Tr}Y$, and we say set $\mathcal{Y}$ has a bounded size if $\mathrm{Tr}Y$ is bounded.[2]

Let $m, n, p, q \in \mathbb{N}$ satisfying $m \leq n$ and $p \leq q$; $\mathbb{N}_{m:n}$ denotes the set of integers $\{m, \ldots, n\}$; letting $\{s_i \in \mathbb{N} | i \in \mathbb{N}_{1:r}, r \in \mathbb{N}^+\}$ be an indexed set of integers, $y^{s_{m:n}}$ denotes the set $\{y^{s_m}, \ldots, y^{s_n}\}$, and $y_{p:q}^{s_{m:n}}$ denotes the set $\{y_p^{s_{m:n}}, \ldots, y_q^{s_{m:n}}\}$; similarly, $y^{s_{m:n}} \in \mathcal{Y}^{s_{m:n}}$ denotes the relationship $y^{s_m} \in \mathcal{Y}^{s_m}, \ldots, y^{s_n} \in \mathcal{Y}^{s_n}$, and $y_{p:q}^{s_{m:n}} \in \mathcal{Y}_{p:q}^{s_{m:n}}$ denotes the relationship $y_p^{s_{m:n}} \in \mathcal{Y}_p^{s_{m:n}}, \ldots, y_q^{s_{m:n}} \in \mathcal{Y}_q^{s_{m:n}}$. For a vector-valued random variable x, we use $\mathrm{E}(x)$ and $\mathrm{Cov}(x)$ to denote its mean and covariance, respectively.

II. Problem Setup

The process is linear time-invariant and evolves in discrete time driven by white noise

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

where $x \in \mathbb{R}^n$ is the state, and $w \in \mathbb{R}^n$ is the noise input, which is zero-mean Gaussian with covariance $Q \geq 0$. We assume (A, Q) is stabilizable.[3] The initial value x_0 of the state is also zero-mean Gaussian, with covariance P_0. The state information is measured using M different sensors, the measurement equations of which are

$$y_k^i = C_i x_k + v_k^i \tag{2}$$

where $y^i \in \mathbb{R}^m$ denotes the output of the ith sensor, $v^i \in \mathbb{R}^m$ is zero-mean Gaussian with covariance R_i for $i \in \mathbb{N}_{1:M}$, and v^i and v^j are uncorrelated if $i \neq j$. In addition, x_0, w, and v^i are uncorrelated with each other. We assume (A, C) is detectable, where $C := [C_1^\top, \ldots, C_M^\top]^\top$; define $R := \mathrm{diag}\{R_1, R_2, \ldots, R_M\}$.

We consider the scenario where the values of the measurement outputs y_k^i are not exactly known, but are only partially known in the sense that only the exact description of sets $\mathcal{Y}_k^i$ is known such that $y_k^i \in \mathcal{Y}_k^i$ for all $i \in \mathbb{N}_{1:M}$. To some extent, this reflects the estimator's inability of telling a point measurement from an uncountable set of measurements, due to the lack of knowledge, e.g., the situation the remote estimator is facing during the non-event instances in an event-based estimation scenario [12]. As a result, the uniqueness of the posteriori probability distributions cannot be maintained, which gives rise to the set-valued Kalman filters [25]. Due to the set-valued measurements from the M sensors at each time instant, one feasible way to update the state estimate is to

[1] Note that the way of parameterizing an ellipsoidal set does not affect the results developed in this work.

[2] Notice that based on this definition, the boundedness of the size of $\mathcal{Y}$ is independent of its centre $c \in \mathbb{R}^m$, since c only describes the relative position of $\mathcal{Y}$. Normally the size of an ellipsoid is given by the maximal eigenvalue of Y. In terms of boundedness, however, these two definitions are equivalent.

[3] Note that this is equivalent to the stabilizability of the pair $(A, \sqrt{Q})$, which can be proved based on the PBH criteria.

fuse the measurement information from the sensors sequentially piece by piece according to some sequence, which (can be chosen either arbitrarily or by design) is mathematically given as

$$s = [s_1, s_2, \ldots, s_M]$$

where $s_i \in \mathbb{N}_{1:M}$ and $s_i \neq s_j$ unless $i = j$, for $i, j \in \mathbb{N}_{1:M}$. We refer to this sequence as "fusion sequence" in this work. Note that in a fusion sequence, each sensor appears once and only once, and the sequence is used to update the information from different sensors measured at the same time instant and does not affect the sensor measurement information in this work.

In standard Kalman filtering, the optimal state prediction $\tilde{x}_k^{s_0}$ that minimizes the estimation error covariance at time instant k is known to satisfy

$$\tilde{x}_k^{s_0} = \mathrm{E}\left(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \ldots, y_{k-1}^{s_{1:M}}\right) \tag{3}$$

where the superscript s_0 is used to indicate that no sensor information measured at time k has been updated; similarly, for $i \in \mathbb{N}_{1:M}$, the optimal state estimate $\tilde{x}_k^{s_i}$ after updating the measurement information from sensors $s_1, s_2, \ldots, s_i$ at time k satisfies

$$\tilde{x}_k^{s_i} = \mathrm{E}\left(x_k | y_0^{s_{1:M}}, \ldots, y_{k-1}^{s_{1:M}}, y_k^{s_{1:i}}\right). \tag{4}$$

The corresponding estimation error covariance satisfies

$$\begin{aligned} P_k^{s_0} &:= \mathrm{Cov}\left(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \ldots, y_{k-1}^{s_{1:M}}\right) \\ &= A P_{k-1}^{s_M} A^\top + Q \end{aligned} \tag{5}$$

$$\begin{aligned} P_k^{s_i} &:= \mathrm{Cov}\left(x_k | y_0^{s_{1:M}}, y_1^{s_{1:M}}, \ldots, y_{k-1}^{s_{1:M}}, y_k^{s_{1:i}}\right) \\ &= P_k^{s_{i-1}} - P_k^{s_{i-1}} C_{s_i}^\top \left(C_{s_i} P_k^{s_{i-1}} C_{s_i}^\top + R_{s_i}\right)^{-1} C_{s_i} P_k^{s_{i-1}} \end{aligned} \tag{6}$$

for $i \in \mathbb{N}_{1:M}$. In set-valued filtering [26], [27], the set-valued measurements are treated as non-stochastic uncertainty; as a result, the choice of different points in the measurement set only leads to different values of the estimation mean, while the estimation error covariance remains unaffected. Specifically, the set of estimation means is defined as

$$\mathcal{X}_k^{s_0} := \left\{\mathrm{E}\left(x_k | y_{0:k-1}^{s_{1:M}}\right) | y_{0:k-1}^{s_{1:M}} \in \mathcal{Y}_{0:k-1}^{s_{1:M}}\right\} \tag{7}$$

$$\mathcal{X}_k^{s_i} := \left\{\mathrm{E}\left(x_k | y_{0:k-1}^{s_{1:M}}, y_k^{s_{1:i}}\right) | y_{0:k-1}^{s_{1:M}} \in \mathcal{Y}_{0:k-1}^{s_{1:M}}, y_k^{s_{1:i}} \in \mathcal{Y}_k^{s_{1:i}}\right\} \tag{8}$$

for $i \in \mathbb{N}_{1:M}$, where $\mathcal{X}_k^{s_0}$ denotes the set of estimation means when no sensor information is fused at time k (namely, the prediction of the state), and for $i \in \mathbb{N}_{1:k}$, $\mathcal{X}_k^{s_i}$ denotes the set of estimation means after fusing the information of sensor $s_1, s_2, \ldots, s_i$ at time instant k. We assume $\mathcal{X}_0^{s_0} = \{0\}$, following the zero-mean Gaussian assumption of x_0. The definition of estimation error covariance still follows that of the standard Kalman filters, which has been given in (5) and (6). In light of the results in [25]–[27], the exact set-valued Kalman filter with multiple sensor measurements is recursively given as

$$\mathcal{X}_k^{s_0} = A \mathcal{X}_{k-1}^{s_M} \tag{9}$$

$$P_k^{s_0} = A P_{k-1}^{s_M} A^\top + Q \tag{10}$$

and for $i \in \mathbb{N}_{0:M-1}$

$$\mathcal{X}_k^{s_{i+1}} = \left(I - K_k^{s_{i+1}} C_{s_{i+1}}\right) \mathcal{X}_k^{s_i} \oplus K_k^{s_{i+1}} \mathcal{Y}_k^{s_{i+1}} \tag{11}$$

where

$$\begin{aligned} K_k^{s_{i+1}} &= P_k^{s_i} C_{s_{i+1}}^\top \left(C_{s_{i+1}} P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}}\right)^{-1} \\ P_k^{s_{i+1}} &= P_k^{s_i} - P_k^{s_i} C_{s_{i+1}}^\top \left(C_{s_{i+1}} P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}}\right)^{-1} C_{s_{i+1}} P_k^{s_i}. \end{aligned} \tag{12}$$

From (12), the filter gain and covariance matrix of the set-valued Kalman filter do not depend on the measurement sets but evolve similarly as those of the standard Kalman filter, for which a point-valued measurement is updated at each time instant. For the scenario of event-based estimation, this implies that the corresponding set-valued event-based estimator responds to the events only by updating the set of estimation means, but the filter gain and covariance matrix evolve as if the point-valued measurements have been received at each time instant.

In Kalman filtering, the confidence on the estimate is fully characterized by the estimation error covariance; while in set-valued Kalman filtering, a set of probability density functions with the same covariance is considered, and the unknown information caused by stochastic uncertainty and non-stochastic uncertainty is treated separately: the confidence on stochastic uncertainty is still quantified as covariance [(10) and (12)]; the confidence on non-stochastic uncertainty is quantified as the size of the set of estimation means [(9) and (11)]. As will be shown in this work, this separation can help provide new insights for event-based estimation problems.

In this work, we assume $\mathcal{Y}_k^i$ are ellipsoidal sets parameterized as

$$\mathcal{Y}_k^i := \mathcal{E}\left(c_k^i, Y_k^i\right). \tag{13}$$

Notice that the parameters of $\mathcal{Y}_k^i$ can be known by the estimator without communication from the sensor. For instance, in an event-based estimation scenario with the aforementioned "send-on-delta" triggering conditions in Section I, c_k^i is the previously transmitted measurement, while Y_k^i may be designed off-line as a constant matrix, like in Examples 2 and 3 in this paper and thus can be known to the estimator beforehand.[4] In the literature, there are alternative ways of describing set-valued measurements, e.g., in terms of parallelotopes and zonotopes [29], [30]. The properties of the resultant estimates are, however, difficult to characterize, due to the lack of intuitive mathematical description of the notion "sizes of the sets."

[4]Note that the estimator may also know Y_k^i when it is time-varying. For instance, if the threshold in the "send-on-delta" condition would vary at each event instant, it could be communicated to the estimator together with the corresponding measurement. In this case, Y_k^i is time-varying and known by the estimator.

Although the Minkowski sum of ellipsoids (which may not be an ellipsoid) is difficult to calculate exactly [28], the ellipsoidal sets are very helpful in analyzing the dynamic behavior of the estimates, since the size and shape of an ellipsoid are uniquely determined by a positive semidefinite matrix. At the same time, outer ellipsoidal approximations are conveniently employed to calculate the set that contains the set of means of the estimates at each time instant, which is calculated according to the following result.

Lemma 1 ([28, Lemma 2.2.1]): Let $p > 0$. We have

$$\begin{aligned}\mathcal{E}(c_1, X_1) \oplus & \mathcal{E}(c_2, X_2) \\ & \subseteq \mathcal{E}\left(c_1 + c_2, (1+p^{-1})X_1 + (1+p)X_2\right). \end{aligned} \tag{14}$$

Normally, p is calculated in some optimal sense. In this work, we take $p = (\mathrm{Tr}X_1)^{1/2}/(\mathrm{Tr}X_2)^{1/2}$, which minimizes the trace of $(1+p^{-1})X_1 + (1+p)X_2$. In this way, we are able to evaluate the outer ellipsoidal approximate estimates

$$\hat{\mathcal{X}}_k^{s_0} := \mathcal{E}\left(\hat{x}_k^{s_0}, X_k\right) \supseteq \mathcal{X}_k^{s_0} \tag{15}$$

of $\mathcal{X}_k^{s_0}$ according to Lemma 1 and (9)–(11), which will be formally introduced in the next section to take account of the effect from sensor fusion sequence.

Based on the introduced notations, we are now in the position to present the problems to be considered in this work:

1) Analyze the effect of the sensor fusion sequence s on the exact and approximate sets of estimation means $\mathcal{X}_k^{s_0}$ and $\hat{\mathcal{X}}_k^{s_0}$, respectively;
2) Analyze the asymptotic behavior of the sizes of $\mathcal{X}_k^{s_0}$ and $\hat{\mathcal{X}}_k^{s_0}$ subject to multiple sensor set-valued measurements.
3) Analyze the effect of including additional sensors on $\hat{\mathcal{X}}_k^{s_0}$.

In addition, after obtaining the solutions to these problems, we will apply them to the analysis and design in event-based state estimation.

III. Sensor Fusion

In this section, we analyze the effect of the fusion sequence on the size of the set of estimation means, based on which a "sequence-independent" separate fusion principle of fusing multiple sensor measurements is proposed. This property is of fundamental importance for the analysis of asymptotic behavior and performance improvement in the multiple-sensor scenario, without which the whole set of fusion sequences (the cardinality of which equals $M!$) would have to be considered to analyze the worst-case behavior.

To aid the analysis, we first present the following lemma on the properties of Minkowski sum.

Lemma 2: Let $\mathcal{X}, \mathcal{Y} \subseteq \mathbb{R}^n$, and let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Then $T(\mathcal{X} \oplus \mathcal{Y}) = (T\mathcal{X}) \oplus (T\mathcal{Y})$.

Proof:

$$\begin{aligned}(T\mathcal{X}) \oplus (T\mathcal{Y}) &= \{Tx | x \in \mathcal{X}\} \oplus \{Ty | y \in \mathcal{Y}\} \\ &= \{a + b | a \in \{Tx | x \in \mathcal{X}\}, b \in \{Ty | y \in \mathcal{Y}\}\} \\ &= \{T(x+y) | x \in \mathcal{X}, y \in \mathcal{Y}\} \\ &= T(\mathcal{X} \oplus \mathcal{Y}).\end{aligned}$$

■

Now we show that the fusion sequence does not affect the exact set of means of the estimates. Before that, we first present some insights into the structure of the filter gains and the closed-loop system matrix. For a given fusion sequence s, the closed-loop matrix $\bar{A}_k^{s_0}$ satisfies

$$\bar{A}_k^{s_0} := A \prod_{i=1}^{M} (I - K_k^{s_i} C_{s_i}) \tag{16}$$

and the filter gain $\bar{K}_k^{s_j}$ for the jth sensor satisfies

$$\bar{K}_k^{s_j} := A \left[\prod_{i=j+1}^{M} (I - K_k^{s_i} C_{s_i}) \right] K_k^{s_j}. \tag{17}$$

For these two matrices, we have the following equivalent representations.

Proposition 1: $\bar{A}_k^{s_0} P_k^{s_0} = A P_k^{s_M}$, $\bar{K}_k^{s_j} = A P_k^{s_M} C_{s_j}^\top R_{s_j}^{-1}$.

Proof: First, applying the matrix inversion lemma to (12), we have

$$P_k^{s_{i+1}} = \left(I + P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} C_{s_{i+1}}\right)^{-1} P_k^{s_i}. \tag{18}$$

Similarly, for $K_k^{s_{i+1}}$, we have

$$\begin{aligned}K_k^{s_{i+1}} &= \left(I + P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} C_{s_{i+1}}\right)^{-1} P_k^{s_i} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1} \\ &= P_k^{s_{i+1}} C_{s_{i+1}}^\top R_{s_{i+1}}^{-1}.\end{aligned} \tag{19}$$

Also, from (12) and the fact that $K_k^{s_{i+1}} = P_k^{s_i} C_{s_{i+1}}^\top (C_{s_{i+1}} \times P_k^{s_i} C_{s_{i+1}}^\top + R_{s_{i+1}})^{-1}$, we have

$$P_k^{s_{i+1}} = \left(I - K_k^{s_{i+1}} C_{s_{i+1}}\right) P_k^{s_i}. \tag{20}$$

From (17), we have

$$\begin{aligned}\bar{K}_k^{s_j} &= A \left[\prod_{i=j+1}^{M} (I - K_k^{s_i} C_{s_i}) \right] K_k^{s_j} \\ &= A \left[\prod_{i=j+1}^{M} (I - K_k^{s_i} C_{s_i}) \right] P_k^{s_j} C_{s_j}^\top R_{s_j}^{-1} \\ &= A P_k^{s_M} C_{s_j}^\top R_{s_j}^{-1}\end{aligned} \tag{21}$$

where the last equality is obtained by recursively applying (20). The relation for $\bar{A}_k^{s_0}$ can be obtained following a similar argument. ■

Remark 1: Notice that if $P_k^{s_0}$ is nonsingular, we have $\bar{A}_k^{s_0} = A P_k^{s_M} (P_k^{s_0})^{-1}$. The above result implies that the filter gains can be updated either by calculating the Riccati equation in (12) corresponding to (C_{s_i}, R_{s_i}) sequentially or by lifting all sensor information matrices $\{(C_{s_i}, R_{s_i})\}$ as (C, R) and computing the Riccati equation by replacing C_{s_i} and R_{s_i} with C and R in (12). The calculation of $\bar{A}_k^{s_0}$ is straightforward as it is well know that it satisfies

$$\bar{A}_k^{s_0} = A - A P_k^{s_0} C^\top \left(C P_k^{s_0} C^\top + R\right)^{-1}. \tag{22}$$

The update of the set of estimation means, on the other hand, can only be updated by sequentially fusing the sensor information, although the fusion result is sequence independent, as is shown in Theorem 1 below. For brevity, write

$$\tilde{g}_i(X) := X - X(C^i)^\top \left[C^i X (C^i)^\top + R^i\right]^{-1} C^i X. \quad (23)$$

Before stating the theorem, we first present the following result.

Lemma 3: For $P \geq 0$, $\tilde{g}_1(\tilde{g}_2(P)) = \tilde{g}_2(\tilde{g}_1(P))$.

Proof: For $\tilde{g}_2(\tilde{g}_1(P))$, following some matrix manipulations, we have

$$\tilde{g}_2(\tilde{g}_1(P)) = P - P\left[C_1^\top C_2^\top\right] \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} P \quad (24)$$

where

$$\begin{aligned} X_{11} =& \left(C_1 P C_1^\top + R_1\right)^{-1} \\ &+ \left(C_1 P C_1^\top + R_1\right)^{-1} C_1 P C_2^\top \\ &\times \left\{ C_2 \left[P - P C_1^\top \left(C_1 P C_1^\top + R_1\right)^{-1} C_1 P\right] \right. \\ &\quad \left. C_2^\top + R_2 \right\}^{-1} C_2 P C_1^\top \left[C_1 P C_1^\top + R_1\right]^{-1} \\ X_{12} =& X_{21}^\top \\ =& -\left(C_1 P C_1^\top + R_1\right)^{-1} C_1 P C_2^\top \\ &\times \left\{ C_2 \left[P - P C_1^\top \left(C_1 P C_1^\top + R_1\right)^{-1} C_1 P\right] C_2^\top + R_2 \right\}^{-1} \\ X_{22} =& \left\{ C_2 \left[P - P C_1^\top \left(C_1 P C_1^\top + R_1\right)^{-1} C_1 P\right] C_2^\top + R_2 \right\}^{-1}. \end{aligned}$$

From the matrix inversion lemma, we further have

$$\begin{aligned} \tilde{g}_1(\tilde{g}_2(P)) =& P - P\left[C_1^\top \; C_2^\top\right] \\ &\times \begin{bmatrix} C_1 P C_1^\top + R_1 & C_1 P C_2^\top \\ C_2 P C_1^\top & C_2 P C_2^\top + R_2 \end{bmatrix}^{-1} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} P \\ =& P - P\left[C_1^\top \; C_2^\top\right] \\ &\times \left(\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} P \left[C_1^\top \; C_2^\top\right] + \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \right)^{-1} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} P \\ =& \left(I + P\left[C_1^\top \quad C_2^\top\right] \begin{bmatrix} R_1 & \\ & R_2 \end{bmatrix}^{-1} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \right)^{-1} P \\ =& \left[I + P\left(C_1^\top R_1^{-1} C_1 + C_2^\top R_2^{-1} C_2\right)\right]^{-1} P. \end{aligned}$$

By symmetry, we have

$$\tilde{g}_2(\tilde{g}_1(P)) = \left[I + P\left(C_2^\top R_2^{-1} C_2 + C_1^\top R_1^{-1} C_1\right)\right]^{-1} P$$

which completes the proof. ∎

Based on this result, we are ready to present the theorem.

Theorem 1: Let s^1, s^2 denote two different sensor fusion sequences. We have

1) If $P_{k-1}^{s_0^1} = P_{k-1}^{s_0^2}$, then $P_k^{s_0^1} = P_k^{s_0^2}$.
2) If $\mathcal{X}_{k-1}^{s_0^1} = \mathcal{X}_{k-1}^{s_0^2}$, then $\mathcal{X}_k^{s_0^1} = \mathcal{X}_k^{s_0^2}$.

Proof: From Lemma 3, the estimation error covariance is not affected if we switch the position of any two neighbouring elements s_i^j and s_{i+1}^j in a fusion sequence s^j. The first part of the theorem follows from the fact that starting from s^2, s^1 can be obtained by performing a finite number of position switches of the neighbouring elements.

To prove the second part, first notice that according to Lemma 2, we have

$$\mathcal{X}_k^{s_0^r} = \bar{A}_{k-1}^{s_0^r} \mathcal{X}_{k-1}^{s_0^r} \oplus \bigoplus_{j=1}^{M} \bar{K}_{k-1}^{s_j^r} \mathcal{Y}_{k-1}^{s_j^r} \quad (25)$$

for $r \in \mathbb{N}_{1:2}$. From Proposition 1, $\bar{K}_{k-1}^i$ only depends on C_i and R_i, which are not affected by the relative position of sensor i in the fusion sequence. Since for $i \in \mathbb{N}_{1:M}$, each sensor i appears once and only once in a fusion sequence, different fusion sequences will lead to different permutation of the same set of summands $\{\bar{K}_{k-1}^i \mathcal{Y}_{k-1}^i | i \in \mathbb{N}_{1:M}\}$ in the second term on the right-hand side of (25). Finally, from (22), $\bar{A}_{k-1}^{s_0^r} \mathcal{X}_{k-1}^{s_0^r}$ is unaffected by the fusion sequence either, the conclusion now follows from the commutativity and associativity of Minkowski sums over convex bodies [31], [32]. ∎

Furthermore, note that since (A, C) is detectable

$$\bar{A} = \lim_{k \to \infty} A \left[\prod_{i=1}^{M} \left(I - K_{k-1}^{s_i} C_{s_i}\right) \right]$$

exists and $\bar{A}$ is stable [33], which will be used in the stability analysis in the next section. The above result shows that the estimation performance in terms of either the estimation error covariance or the size of the set of estimation means does not depend on the fusion sequence s for the exact set of means of the estimates. Unfortunately, the exact sets of means of the estimates either in the form (25) or the recursive form (9)–(11) are difficult to obtain analytically when the measurements are given in terms of ellipsoidal sets, since the summation of ellipsoids may not be ellipsoids at all [28], and consequently the analytical expression of the exact set-valued estimator cannot be maintained. Motivated from the above result, however, we propose the following procedure in Algorithm 1 to calculate the outer approximation of the set of estimation means.

Algorithm 1 indicates that for multiple-sensor set-valued filtering, the fusion of covariance and estimation means should be performed separately: the estimation error covariance is updated first (see lines 5–10), where the covariance updates are first calculated by solving the Riccati equation for C and R, and $\bar{A}_k^{s_0}$ and $\bar{K}_k^{s_i}$ are respectively calculated according to (22) and (21); then the update of estimation means is performed (see lines 11–19), where the set of estimation means $\hat{\mathcal{X}}_k^{s_0}$ are calculated by iteratively fusing the summands in (25) in a two-by-two fashion based on Lemma 1 according to an arbitrary fusion sequence s. One may think that it is not necessarily to do so, as is the case for classical Kalman filtering with multiple point-valued measurements. We show that, however, the proposed procedure bears the basic properties of the classical Kalman filter while enjoying the benefits of distributed implementation.

Algorithm 1 Calculation of $\hat{\mathcal{X}}_k^{s_0}$

1: $\hat{\mathcal{X}}_0^{s_0} = \mathcal{E}(0,0)$;
2: $P_0^{s_0} = P_0$;
3: $k = 0$;
4: **while** $k \geq 0$ **do**
5: $\quad P_k^{s_M} = P_k^{s_0} - P_k^{s_0} C^\top (C P_k^{s_0} C^\top + R)^{-1} C P_k^{s_0}$;
6: $\quad P_{k+1}^{s_0} = A P_k^{s_M} A^\top + Q$;
7: $\quad \bar{A}_k^{s_0} = A - A P_k^{s_0} C^\top (C P_k^{s_0} C^\top + R)^{-1}$;
8: $\quad$ **for** $i = 1 : M$ **do**
9: $\quad\quad \bar{K}_k^{s_i} = A P_k^{s_M} C_{s_i} R_{s_i}^{-1}$;
10: $\quad$ **end for**
11: $\quad \bar{\mathcal{X}}_k^{s_0} := \mathcal{E}(\bar{x}_k^{s_0}, \bar{X}_k^{s_0}) = \bar{A}_k^{s_0} \hat{\mathcal{X}}_k^{s_0}$
12: $\quad\quad = \mathcal{E}(\bar{A}_k^{s_0} \hat{x}_k^{s_0}, \bar{A}_k^{s_0} X_k (\bar{A}_k^{s_0})^\top)$;
13: $\quad$ **for** $i = 1 : M$ **do**
14: $\quad\quad p_k^{s_i} = \sqrt{\mathrm{Tr} \bar{X}_k^{s_{i-1}} / \mathrm{Tr} \bar{K}_k^{s_i} Y_k^{s_i} (\bar{K}_k^{s_i})^\top}$;
15: $\quad\quad \bar{x}_k^{s_i} = \bar{x}_k^{s_{i-1}} + \bar{K}_k^{s_i} c_k^{s_i}$;
16: $\quad\quad \bar{X}_k^{s_i} = (1 + 1/p_k^{s_i}) \bar{X}_k^{s_{i-1}} + (1 + p_k^{s_i}) \bar{K}_k^{s_i} Y_k^{s_i} (\bar{K}_k^{s_i})^\top$;
17: $\quad\quad \bar{\mathcal{X}}_k^{s_i} := \mathcal{E}(\bar{x}_k^{s_i}, \bar{X}_k^{s_i})$;
18: $\quad$ **end for**
19: $\quad \hat{\mathcal{X}}_{k+1}^{s_0} := \bar{\mathcal{X}}_k^{s_M}$;
20: $\quad k = k + 1$;
21: **end while**
22: **end**

To see this, we look into the structure of the outer approximation of the set of the means of the estimates with the help of the following lemmas [28].

Lemma 4: Let $\mathcal{E}(a, Q) \subseteq \mathbb{R}^n$. Then $x \in \mathcal{E}(a, Q)$ is equivalent to $Ax + b \in \mathcal{E}(Aa + b, AQA^\top)$.

Lemma 5:

$$\bigoplus_{i=1}^{l} \mathcal{E}(c_i, X_i) \subseteq \mathcal{E}(c_0, X_0) \tag{26}$$

with $c_0 = \sum_{i=1}^{l} c_i$

$$X_0 = \left(\sum_{i=1}^{l} q_i \right) \sum_{i=1}^{l} q_i^{-1} X_i \tag{27}$$

for all $q_i > 0$, $i \in \mathbb{N}_{1:l}$.

Following these lemmas and (25), we have the following updating equations of $\hat{\mathcal{X}}_k^{s_0}$ from $\hat{\mathcal{X}}_{k-1}^{s_0} = \mathcal{E}(\hat{x}_{k-1}^{s_0}, X_{k-1})$ and $\mathcal{Y}_k^{s_i} = \mathcal{E}(c_k^{s_i}, Y_k^{s_i})$:

$$\hat{\mathcal{X}}_k^{s_0} = \mathcal{E}\left(\hat{x}_k^{s_0}, X_k\right) \tag{28}$$

$$\hat{x}_k^{s_0} = \bar{A}_{k-1}^{s_0} \hat{x}_{k-1}^{s_0} + \sum_{j=1}^{M} \bar{K}_{k-1}^{s_j} c_{k-1}^{s_j} \tag{29}$$

$$\begin{aligned} X_k = &\left(\sqrt{\mathrm{Tr} \bar{A}_{k-1}^{s_0} X_{k-1} (\bar{A}_{k-1}^{s_0})^\top} + \sum_{j=1}^{M} \sqrt{\mathrm{Tr} \bar{K}_{k-1}^{s_j} Y_{k-1} (\bar{K}_{k-1}^{s_j})^\top} \right) \\ &\times \left[\left(\sqrt{\mathrm{Tr} \bar{A}_{k-1}^{s_0} X_{k-1} \left(\bar{A}_{k-1}^{s_0}\right)^\top} \right)^{-1} \bar{A}_{k-1}^{s_0} X_{k-1} \left(\bar{A}_{k-1}^{s_0}\right)^\top \right. \\ &+ \sum_{j=1}^{M} \left(\sqrt{\mathrm{Tr} \bar{K}_{k-1}^{s_j} Y_{k-1} \left(\bar{K}_{k-1}^{s_j}\right)^\top} \right)^{-1} \\ &\left. \times \bar{K}_{k-1}^{s_j} Y_{k-1} \left(\bar{K}_{k-1}^{s_j}\right)^\top \right]. \end{aligned} \tag{30}$$

Using a similar argument as that in Section IV-A of [27], (30) can be evaluated in an iterative way as in lines 11–19 of Algorithm 1 according to an arbitrary sequence. This not only helps to reduce the computational complexity at the fusion centre through distributed computation (the acknowledgement or computation of $\bar{A}_k^{s_0}$ and $\bar{K}_k^{s_i}$ at sensor i would be necessary), but also guarantees the invariance of outer ellipsoidal approximation of the set of estimation means with respect to the fusion sequence. In [27], to possess this property, a different filter form, namely, the information filter [34], was considered. On the other hand, the filter form utilized here inherits the original form of Kalman filter with multiple point-valued measurements, due to the separate covariance and estimate updating procedure in Algorithm 1. Notice that in fact, at steady state, only the update of the estimation means (lines 11–19) is necessary, since the solution to the Riccati equation converges to its unique stabilizing solution, and therefore the algorithm can be implemented in a completely distributive way without considering covariance update at the steady state.

IV. Asymptotic Properties of the Set of Means of the Estimates

In this section, the objective is to discuss the asymptotical boundedness properties of both the exact and outer-approximate sets of estimation means for the multiple sensor case. We first focus on the single sensor case and analyze the asymptotic properties of the set-valued mean evolutions, and then extend the results to multiple sensor case. When there is only one sensor, the equations are given by

$$\mathcal{X}_k^0 = A \mathcal{X}_{k-1}^1 \tag{31}$$

$$\mathcal{X}_k^1 = (I - K_k C) \mathcal{X}_k^0 \oplus K_k \mathcal{Y}_k. \tag{32}$$

In the prediction form, we have

$$\mathcal{X}_{k+1}^0 = \bar{A}_k \mathcal{X}_k^0 \oplus \bar{K}_k \mathcal{Y}_k \tag{33}$$

where $\bar{A}_k = A(I - K_k C)$, $\bar{K}_k = A K_k$. Correspondingly, let $\hat{\mathcal{X}}_k^0 = \mathcal{E}(\hat{x}_k^0, X_k)$ and $\mathcal{Y}_k = \mathcal{E}(c_k, Y_k)$, and the approximate estimate is given by

$$\hat{\mathcal{X}}_{k+1}^0 = \mathcal{E}\left(\bar{A}_k \hat{x}_k^0 + \bar{K}_k c_k, X_{k+1}\right) \tag{34}$$

$$\begin{aligned} X_{k+1} = &\left(1 + \frac{\sqrt{\mathrm{Tr} \bar{K}_k Y_k \bar{K}_k^\top}}{\sqrt{\mathrm{Tr} \bar{A}_k X_k \bar{A}_k^\top}} \right) \bar{A}_k X_k \bar{A}_k^\top \\ &+ \left(1 + \frac{\sqrt{\mathrm{Tr} \bar{A}_k X_k \bar{A}_k^\top}}{\sqrt{\mathrm{Tr} \bar{K}_k Y_k \bar{K}_k^\top}} \right) \bar{K}_k Y_k \bar{K}_k^\top. \end{aligned} \tag{35}$$

The objective of this section is to show the boundedness of the sizes of the sequence of sets $\{\mathcal{X}_k^0\}$ and the possible boundedness of the sizes of the sequences of sets $\{\hat{\mathcal{X}}_k^0\}$. Before continuing, we give the following lemma.

Lemma 6: Let $Q \geq 0$, and $0 \leq P < I$. Then $\mathrm{Tr}\, QP \leq \mathrm{Tr}\, Q$.

Proof: Since $0 \leq P < I$, from [35, Theorem 4.1.5], there exists a real orthogonal matrix U such that $U^\top = U^{-1}$ and

$$U^\top PU = U^{-1}PU = \mathrm{diag}\{p_1, p_2, \ldots, p_n\}$$

satisfying $0 \leq p_i < 1$, p_i being the eigenvalues of P. Let $p^* = \max_{i \in \mathbb{N}_{1:n}} p_i < 1$. We have

$$\mathrm{Tr}\, QP = \mathrm{Tr}\, U^{-1}QUU^{-1}PU = \mathrm{Tr}\, U^{-1}QU\mathrm{diag}\{p_i\}.$$

Since $U^{-1}QU = U^\top QU \geq 0$, the diagonal elements of $U^{-1}QU$ are nonnegative. Thus, we have

$$\mathrm{Tr}\, U^{-1}QU\mathrm{diag}\{p_i\} \leq \mathrm{Tr}\, U^{-1}QUp^*I = p^*\mathrm{Tr}\, Q \leq \mathrm{Tr}\, Q.$$

Notice that since $p^* < 1$, the equality holds if and only if $Q = 0$. ■

Now we are ready to present the first result on the asymptotic properties of the sizes of the sets of the means.

Theorem 2: Assume the pair (A, C) is detectable and (A, Q) is stabilizable. Let $\bar{A} := \lim_{k\to\infty} \bar{A}_k$ and $\bar{K} := \lim_{k\to\infty} \bar{K}_k$.

1) The sizes of the sequence of sets $\{\mathcal{X}_k^0\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.
2) If $\|\bar{A}\|_2 < 1$, the sizes of the sequence of ellipsoids $\{\hat{\mathcal{X}}_k^0\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.
3) If $\|\bar{A}\|_2 \geq 1$, there exists an invertible linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ such that the sizes of the sequence of the set of estimation means $\{\hat{\tilde{\mathcal{X}}}_k^0\}$ for the transformed state $\tilde{x}_k := Tx_k$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k\}$ with bounded sizes.

Proof: Since (A, C) is detectable and (A, Q) is stabilizable, the solution to the Riccati equation converges to the unique stabilizing solution. Thus, $\bar{A} = \lim_{k\to\infty} \bar{A}_k$ and $\bar{K} = \lim_{k\to\infty} \bar{K}_k$ are well defined, and satisfy $\bar{A} = A - \bar{K}C$ and $\bar{K} = A\bar{P}C^\top(C\bar{P}C^\top + R)^{-1}$, $\bar{P}$ being the stabilizing solution to the Riccati equation

$$P = APA^\top - APC^\top(CPC^\top + R)^{-1}CPA^\top + Q.$$

We will prove parts (2) and (3) before proving the result in part (1).

First we show that if $\|\bar{A}\|_2 < 1$, the evolution of the size of the outer approximation of the Minkowski sum in (34), (35) is asymptotically bounded. Since the evolution of (35) does not affect the evolution of the covariance and (A, C) is detectable, it suffices to consider the steady-state Kalman filter gain, which is equivalent to the consideration of $\bar{A}$. At steady state, taking traces on both sides of (35), we have

$$\begin{aligned}\mathrm{Tr}\, X_{k+1} &= \left(1 + \frac{\sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top}}{\sqrt{\mathrm{Tr}\, \bar{A}X_k\bar{A}^\top}}\right)\mathrm{Tr}\, \bar{A}X_k\bar{A}^\top \\ &\quad + \left(1 + \frac{\sqrt{\mathrm{Tr}\bar{A}X_k\bar{A}^\top}}{\sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top}}\right)\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top \\ &= \left(\sqrt{\mathrm{Tr}\, \bar{A}X_k\bar{A}^\top} + \sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top}\right)^2. \qquad (36)\end{aligned}$$

Since $X_{k+1} \geq 0$, we have $\mathrm{Tr}\, X_{k+1} \geq 0$. Thus

$$\begin{aligned}\sqrt{\mathrm{Tr}\, X_{k+1}} &= \sqrt{\mathrm{Tr}\, \bar{A}X_k\bar{A}^\top} + \sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top} \\ &= \sqrt{\mathrm{Tr}\, X_k\bar{A}^\top\bar{A}} + \sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top} \\ &\leq \sqrt{a^*}\sqrt{\mathrm{Tr}\, X_k} + \sqrt{\mathrm{Tr}\, \bar{K}Y_k\bar{K}^\top} \qquad (37)\end{aligned}$$

for some $a^* \in (0, 1)$, which follows from $\|\bar{A}\|_2 < 1$ and Lemma 6. This implies the boundedness of $\{\sqrt{\mathrm{Tr}\, X_k}\}$, given the boundedness of $\{Y_k\}$.

Now we consider the case $\|\bar{A}\|_2 \geq 1$. Since $\bar{A}$ is stable, there exists $P_s > 0$ such that P_s is the solution to the Lyapunov equation $\bar{A}^\top P\bar{A} - P + I = 0$, which implies $P_s \geq I > 0$. Now we introduce a linear transformation $T = P_s^{1/2} : \mathbb{R}^n \to \mathbb{R}^n$ and let $\tilde{x}_k = Tx_k$. Apparently $\hat{x}_k$ evolves according to

$$\tilde{x}_{k+1} = \tilde{A}\tilde{x}_k + \tilde{B}w_k, \quad y_k = \tilde{C}\tilde{x}_k + v_k$$

where $\tilde{A} = TAT^{-1}$, $\tilde{B} = T$, $\tilde{C} = CT^{-1}$. Furthermore, it is easy to verify that $\bar{\tilde{K}} := \tilde{A}\tilde{P}\tilde{C}^\top(\tilde{C}\tilde{P}\tilde{C}^\top + R)^{-1} = T\bar{K}$ and $\tilde{P} = T\bar{P}T^\top$, $\tilde{P}$ being the stabilizing solution to the Riccati equation

$$P = \tilde{A}P\tilde{A}^\top - \tilde{A}P\tilde{C}^\top(\tilde{C}P\tilde{C}^\top + R)^{-1}\tilde{C}P\tilde{A}^\top + TQT^\top.$$

Define $\bar{\tilde{A}} := \tilde{A} - \bar{\tilde{K}}\tilde{C}$, we have $\bar{\tilde{A}} = T(A - \bar{K}C)T^{-1} = P_s^{1/2}\bar{A}P_s^{-1/2}$. Thus

$$\begin{aligned}\bar{\tilde{A}}^\top\bar{\tilde{A}} &= P_s^{-1/2}\bar{A}^\top P_s\bar{A}P_s^{-1/2} \\ &= P_s^{-1/2}(P_s - I)P_s^{-1/2} = I - P_s^{-1} < I\end{aligned}$$

which implies $\|\bar{\tilde{A}}\|_2 < 1$. Therefore the conclusion of part (3) follows from the same argument used for proof of part (2).

Finally we prove the result in part (1). The case of $\|\bar{A}\|_2 < 1$ follows from the result in part (2), since $\{\hat{\mathcal{X}}_k^0\}$ provides an outer approximation of $\{\mathcal{X}_k^0\}$. To prove the case of $\|\bar{A}\|_2 \geq 1$, we analyze the relationship between the exact Minkowski sum for the original state estimate and that of the transformed state estimate. Similar to (33), we have

$$\tilde{\mathcal{X}}_{k+1}^0 = \bar{\tilde{A}}_k\tilde{\mathcal{X}}_k^0 \oplus \bar{\tilde{K}}_k\mathcal{Y}_k \qquad (38)$$

where $\bar{\tilde{A}}_k = \tilde{A} - \bar{\tilde{K}}_kC$, $\bar{\tilde{K}}_k = \tilde{A}\tilde{P}_{k-1}\tilde{C}^\top(\tilde{C}\tilde{P}_{k-1}\tilde{C}^\top + R)^{-1}$ and $\tilde{P}_k$ being the solution to the Riccati equation $\tilde{P}_{k+1} = \tilde{A}\tilde{P}_k\tilde{A}^\top - \tilde{A}\tilde{P}_k\tilde{C}^\top(\tilde{C}\tilde{P}_k\tilde{C}^\top + R)^{-1}\tilde{C}\tilde{P}_k\tilde{A}^\top + TQT^\top$ subject to $\tilde{P}_0 = TP_0T^\top$. At time $t = 0$, $\tilde{\mathcal{X}}_0^0 = \{Tx_0\} = T\mathcal{X}_0^0$. Note that following a similar argument as that in the proof of part (3), $\bar{\tilde{A}}_k = T\bar{A}_kT^{-1}$ and $\bar{\tilde{K}}_k = T\bar{K}_k$. Now assume at time $t = k$, the relationship $\tilde{\mathcal{X}}_k^0 = T\mathcal{X}_k^0$ holds. We have

$$\tilde{\mathcal{X}}_{k+1}^0 = T\bar{A}_kT^{-1}\tilde{\mathcal{X}}_k^0 \oplus T\bar{K}_k\mathcal{Y}_k. \qquad (39)$$

Following the definition Minkowski sum

$$\begin{aligned}\tilde{\mathcal{X}}_{k+1}^0 &:= \left\{T\bar{A}_kT^{-1}\tilde{x} + T\bar{K}_ky|\tilde{x}\in\tilde{\mathcal{X}}_k^0, y\in\mathcal{Y}_k\right\}\\ &= \left\{T\bar{A}_kx + T\bar{K}_ky|x\in\mathcal{X}_k^0, y\in\mathcal{Y}_k\right\}\\ &= \left\{T(a+b)|a\in\left\{\bar{A}_kx|x\in\mathcal{X}_k^0\right\}, b\in\left\{\bar{K}_ky|y\in\mathcal{Y}_k\right\}\right\}\\ &= T\left(\bar{A}_k\mathcal{X}_k^0\oplus\bar{K}_k\mathcal{Y}_k\right) = T\mathcal{X}_{k+1}^0.\end{aligned}$$

Thus, $\tilde{\mathcal{X}}_k^0 = T\mathcal{X}_k^0$ for all k. Since T is nonsingular, the boundedness of $\{\mathcal{X}_k^0\}$ is equivalent to that of $\{\tilde{\mathcal{X}}_k^0\}$. The conclusion follows from part (3) of the theorem and the fact that $\{\hat{\tilde{\mathcal{X}}}_k^0\}$ provides an outer approximation of $\{\tilde{\mathcal{X}}_k^0\}$. ■

Remark 2: From the above proof, a quantitative relationship of the size of the set of estimation means with the sets of measurements, the statistical properties of the noises and the system matrices can be obtained. To see this, assume there exists an upper bound $\bar{Y}\geq Y_k$ for all $k\in\mathbb{N}$. For the case $\|\bar{A}\|_2<1$, from (37), we have

$$\sqrt{\operatorname{Tr}X_{k+1}}\leq\|\bar{A}\|_2\sqrt{\operatorname{Tr}X_k}+\sqrt{\operatorname{Tr}\bar{K}\bar{Y}\bar{K}^\top}. \quad (40)$$

Thus, we have

$$\lim_{k\to\infty}\sqrt{\operatorname{Tr}X_k}\leq\frac{\sqrt{\operatorname{Tr}\bar{K}\bar{Y}\bar{K}^\top}}{(1-\|\bar{A}\|_2)} \quad (41)$$

where the system and noise parameters are reflected in $\bar{K}$ and $\bar{A}$ (Recall that $\bar{K}=A\bar{P}C^\top(C\bar{P}C^\top+R)^{-1}$ and $\bar{A}=A-\bar{K}C$, respectively, $\bar{P}$ being the stabilizing solution to the Riccati equation $P=APA^\top-APC^\top(CPC^\top+R)^{-1}CPA^\top+Q$). This implies that an upper bound on the size of the set of estimation means at steady state can be provided based on the upper bound of the shape matrix Y_k, the system matrices and the covariance matrices of the noises. The same analysis applies to the case of $\|\bar{A}\|_2\geq 1$ by introducing the linear transformation $T=P_s^{1/2}$. In addition, it is technically difficult to prove the boundedness of $\{\hat{\mathcal{X}}_k^0\}$ based on the boundedness of $\{\hat{\tilde{\mathcal{X}}}_k^0\}$, since considering the trace operations in (35), the relationship between $\hat{\mathcal{X}}_k^0$ and $\hat{\tilde{\mathcal{X}}}_k^0$ is very complicated.

Remark 3: As we take $p=(\operatorname{Tr}X_1)^{1/2}/(\operatorname{Tr}X_2)^{1/2}$ in Lemma 1, the obtained outer ellipsoidal approximation of the set of the estimation means is the tightest approximation in the sense of minimizing $\operatorname{Tr}[(1+p^{-1})X_1+(1+p)X_2]$. Even for this tightest approximation, however, it is not clear whether its size is bounded or not as the time goes to infinity. We show in Theorem 2 that when $\bar{A}$ is contractive, the boundedness can be proved. Although it is difficult to obtain the boundedness result when $\bar{A}$ is non-contractive, we show that in this case, we are able to find a constant nonsingular linear transformation T such that the boundedness can be guaranteed if we transform the states by the transformation T. In this way, if $\bar{A}$ is non-contractive, we can apply the set-valued filtering technique to the transformed system to ensure boundedness; as a point-valued estimator is normally needed for control and monitoring purposes, it suffices to apply the inverse transformation to the centre of the set of estimation means to get the estimates of the original states.

Another consequence of the above result is that for first-order systems with constant size of measurement set, we are able to exactly characterize the size of the set of means of the estimate at steady state.

Corollary 1: For $n=m=1$, and $Y_k=Y$. The size of $\{\mathcal{X}_k\}$ converges to $|\bar{K}\sqrt{Y}|/(1-|\bar{A}|)$.

Proof: The proof of this result follows from inequality (37) and the fact that $|\bar{A}|<1$ always holds for $n=1$. ■

The next result generalizes Theorem 2 to the multiple sensor case, utilizing the properties of the outer-approximate estimate set.

Corollary 2: Consider the exact and approximate multiple sensor set-valued estimators in (9)–(12) and (28)–(30), respectively. Assume (A,C) is detectable and (A,Q) is stabilizable. Let $\bar{A}:=\lim_{k\to\infty}\bar{A}_k^{s_0}$.

1) The sizes of the sequence of sets $\{\mathcal{X}_k^{s_0}\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.
2) If $\|\bar{A}\|_2<1$, the sizes of the sequence of ellipsoids $\{\hat{\mathcal{X}}_k^{s_0}\}$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.
3) If $\|\bar{A}\|_2\geq 1$, there exists an invertible linear transformation $T:\mathbb{R}^n\to\mathbb{R}^n$ such that the sizes of the set of mean of the estimates $\{\hat{\tilde{\mathcal{X}}}_k^{s_0}\}$ for the transformed state $\tilde{x}_k:=Tx_k$ are asymptotically bounded for all measurement set sequences $\{\mathcal{Y}_k^{s_i}\}$ with bounded sizes.

Proof: To show this result, we establish the relationship between the single sensor case and the multiple sensor case. By calculating the traces of both sides for (30), it is not difficult to verify that

$$\operatorname{Tr}X_k=\left(\sqrt{\operatorname{Tr}\bar{A}_{k-1}^{s_0}X_{k-1}\left(\bar{A}_{k-1}^{s_0}\right)^\top}+\sum_{i=1}^{M}\sqrt{\operatorname{Tr}\bar{K}_{k-1}^{s_i}Y_{k-1}^{s_i}\left(\bar{K}_{k-1}^{s_i}\right)^\top}\right)^2 \quad (42)$$

and thus

$$\sqrt{\operatorname{Tr}X_k}=\sqrt{\operatorname{Tr}\bar{A}_{k-1}^{s_0}X_{k-1}\left(\bar{A}_{k-1}^{s_0}\right)^\top}+\sum_{i=1}^{M}\sqrt{\operatorname{Tr}\bar{K}_{k-1}^{s_i}Y_{k-1}^{s_i}\left(\bar{K}_{k-1}^{s_i}\right)^\top}. \quad (43)$$

Noticing the relationship with (37) and the boundedness of $\{Y_k^{s_i}\}$, the results are proved with a similar argument as that in the proof of Theorem 2. ■

Remark 4: Note that similar to the single sensor case, a quantitative relationship of the size of the set of estimation means with the sets of measurements, the statistical properties of the noises and the system matrices can also be obtained. Assume for $i\in\mathbb{N}_{1:M}$, there exist upper bounds $\bar{Y}^i\geq Y_k^i$ for all $k\in\mathbb{N}$. Write $\bar{K}^i:=\lim_{k\to\infty}\bar{K}_k^i$, which satisfies $\bar{K}^i=A\bar{P}C_iR_i^{-1}$ according to Proposition 1, where $\bar{P}$ is the stabilizing solution to the Riccati equation $P=APA^\top-APC^\top(CPC^\top+R)^{-1}CPA^\top+Q$ (Recall that $C=[C_1^\top,\ldots,C_M^\top]^\top$ and $R=\operatorname{diag}\{R_1,R_2,\ldots,R_M\}$ for the multiple sensor

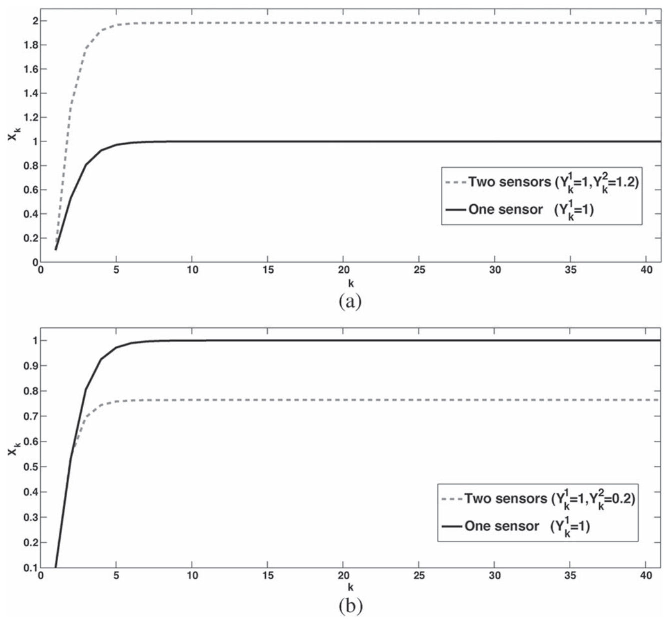

Fig. 1. Comparison of the sizes of the sets of estimation means for different choices of Y_k^2. (a) Performance comparison for $Y_k^1 = 1$, $Y_k^2 = 1.2$. (b) Performance comparison for $Y_k^1 = 1$, $Y_k^2 = 0.2$.

case). Following the analysis in Remark 2 and from (43), we have for the case of $\|\bar{A}\|_2 < 1$

$$\sqrt{\operatorname{Tr} X_{k+1}} \leq \|\bar{A}\|_2 \sqrt{\operatorname{Tr} X_k} + \sum_{i=1}^{M} \sqrt{\operatorname{Tr} \bar{K}^i \bar{Y}^i (\bar{K}^i)^\top}. \quad (44)$$

Thus, we have

$$\lim_{k\to\infty} \sqrt{\operatorname{Tr} X_k} \leq \frac{\sum_{i=1}^{M} \sqrt{\operatorname{Tr} \bar{K}^i \bar{Y}^i (\bar{K}^i)^\top}}{\left(1 - \|\bar{A}\|_2\right)}. \quad (45)$$

The same analysis applies to the case of $\|\bar{A}\|_2 \geq 1$ by introducing the linear transformation T.

V. Performance Improvement

Now we analyze the effect of including more sensors on the estimation performance in the set-valued estimation framework. Adding sensors always reduces the estimation error covariance, following the monotonicity properties of the solutions to the Riccati equations. Adding sensors, however, does not necessarily reduce the size of the set of the means of the estimates, as is shown in the following example.

Example 1: Consider the system in (1) with $n = 1$, $m = 1$, $A = 1.3$, $Q = 1.2$, $C_1 = 1$, $C_2 = 0.6$, $R_1 = 1.9$, $R_2 = 0.7$. Assume $Y_k^1 = 1$, we consider two choices of Y_k^2: 1) $Y_k^2 = 1.2$ and 2) $Y_k^2 = 0.2$. The performance in terms of the size of the set of estimation means obtained by using sensor 1 alone and using sensor 1 and sensor 2 are shown in Fig. 1(a) and (b), respectively. It is shown that when $Y_k^2 = 0.2$, the addition of sensor 2 helps to improve the estimation performance; the choice of $Y_k^2 = 1.2$, however, deteriorates the performance in terms of a larger size of the set of estimation means.

Motivated by the above example, given an existing sensor 1, it is interesting to characterize conditions on properties of sensor 2 such that improved performance can be guaranteed. Now we make the problem more explicit. Suppose we have a linear system originally measured only by sensor 1, namely, (1) and (2) with $M = 1$. Now we introduce sensor 2 and measure the system state using two sensors. We want to compare the size of the set of estimation means obtained only using sensor 1 with that using sensors 1 and 2 together. First, we need to quantify the performance. To do this, we focus on the steady-state behavior of the size of the set of estimation means by assuming that the closed-loop matrix $\bar{A}$ under consideration satisfies $\|\bar{A}\|_2 < 1$ (Note that if this condition is not satisfied by the original system, we can introduce the linear transformation T in the proof of Theorem 2 such that the transformed closed-loop matrix satisfies this condition). For simplicity, we assume that the shape matrix $Y_k^{s_i}$ of $\mathcal{Y}_k^{s_i}$ satisfies $\lim_{k\to\infty} Y_k^{s_i} = Y^{s_i}$, and further assume (A, Q) is reachable, which guarantees the positive definiteness of P^{s_0} (see the corollary [36, p. 710]). From (43), we have

$$\sqrt{\operatorname{Tr} X_k} \leq \left\|\bar{A}_{k-1}^{s_0}\right\|_2 \sqrt{\operatorname{Tr} X_{k-1}} + \sum_{i=1}^{M} \sqrt{\operatorname{Tr} \bar{K}_{k-1}^{s_i} Y_{k-1}^{s_i} \left(\bar{K}_{k-1}^{s_i}\right)^\top}. \quad (46)$$

Since $X_k \geq 0$ and $Y^{s_i} \geq 0$, the solution to the following equation serves as an upper bound for the size of the set of means of estimates at steady state

$$\bar{x} = \|\bar{A}^{s_0}\|_2 \bar{x} + \sum_{i=1}^{M} \sqrt{\operatorname{Tr} \bar{K}^{s_i} Y^{s_i} (\bar{K}^{s_i})^\top}. \quad (47)$$

From Proposition 1, this is equivalent to

$$\bar{x} = \left\|AP^{s_M} (P^{s_0})^{-1}\right\|_2 \bar{x} + \sum_{i=1}^{M} \sqrt{\operatorname{Tr} AP^{s_M} C_{s_i}^\top R_{s_i}^{-1} Y^{s_i} R_{s_i}^{-1} C_{s_i} P^{s_M} A^\top} \quad (48)$$

where P^{s_0} is the stabilizing solution to the algebraic Riccati equation

$$P = APA^\top - APC^\top (CPC^\top + R)^{-1} CPA^\top + Q \quad (49)$$

and

$$P^{s_M} = P^{s_0} - P^{s_0} C^\top (CP^{s_0} C^\top + R)^{-1} CP^{s_0}. \quad (50)$$

Notice that $P^{s_0} = AP^{s_M} A^\top + Q$. Thus from Proposition 1, we have

$$\begin{aligned} \bar{A}^{s_0} &= AP^{s_M} (P^{s_0})^{-1} \\ &= AP^{s_M} (AP^{s_M} A^\top + Q)^{-1}. \end{aligned} \quad (51)$$

For the case of one sensor ($M = 1$), we denote $P^1 := P^{s_1}$ for brevity, and the steady-state performance is

$$\bar{x}^1 = \frac{\sqrt{\operatorname{Tr} AP^1 C_1^\top R_1^{-1} Y^1 R_1^{-1} C_1 P^1 A^\top}}{1 - \left\|AP^1 (AP^1 A^\top + Q)^{-1}\right\|_2}. \quad (52)$$

When sensor 2 is included (namely, $M = 2$), we denote $P^2 := P^{s_2}$ for brevity, and the steady-state performance becomes

$$\begin{aligned}\bar{x}^{1,2} = \Bigg[&\sqrt{\operatorname{Tr} AP^2C_1^\top R_1^{-1}Y^1R_1^{-1}C_1P^2A^\top} \\ &+ \sqrt{\operatorname{Tr} AP^2C_2^\top R_2^{-1}Y^2R_2^{-1}C_2P^2A^\top}\Bigg] \Bigg/ \quad (53)\\ &\left[1 - \left\|AP^2(AP^2A^\top + Q)^{-1}\right\|_2\right]. \quad (54)\end{aligned}$$

Notice that for sensor i, the C_i and R_i matrices are fixed and cannot be adjusted; the only adjustable parameter[5] is Y^i, which controls the size and shape of the set of measurement. Therefore, at steady state, the parameters A, C_i, R_i, P^i are constant. In this way, it is easier to check whether a choice of Y^2 will lead to improved performance in terms of the size of the set of estimation means. In particular, the condition becomes easier to verify when Y^i's has special structures, e.g., $Y^i = \eta^i I$, which can be used for design purposes. Finally, we consider scalar systems, namely, $n = m = 1$, and have the following result.

Proposition 2: For $n = m = 1$, if $Y^2/Y^1 < [(P^1 - P^2)C_1R_1^{-1}/P^2C_2R_2^{-1}]^2$, then adding sensor 2 improves the steady-state performance in terms of both the estimation error covariance and the size of the set of the means of the estimates.

Proof: When $n = m = 1$, (52) and (54) reduce to

$$\bar{x}^1 = \frac{|A|P^1C_1^\top R_1^{-1}\sqrt{Y^1}}{1 - |A|P^1(P^0)^{-1}} = \frac{|A|P^1C_1^\top R_1^{-1}\sqrt{Y^1}}{1 - |A|P^1(AP^1A + Q)^{-1}} \quad (55)$$

and

$$\bar{x}^{1,2} = \frac{|A|P^2\left(C_1^\top R_1^{-1}\sqrt{Y^1} + C_2^\top R_2^{-1}\sqrt{Y^2}\right)}{1 - |A|P^2(AP^2A + Q)^{-1}} \quad (56)$$

respectively. Since $C_1R_1^{-1}C_1 < C^\top R^{-1}C$, from the monotonicity properties of the solutions to the Riccati (49) and (50) (Lemma 3 of [37]), we have $P^1 > P^2$. Therefore

$$\begin{aligned}&\frac{Q}{AP^1A + Q} < \frac{Q}{AP^2A + Q} \\ &\quad\Rightarrow \frac{|A|P^1}{AP^1A + Q} > \frac{|A|P^2}{AP^2A + Q} \\ &\quad\Rightarrow 1 - \frac{|A|P^1}{AP^1A + Q} < 1 - \frac{|A|P^2}{AP^2A + Q} \\ &\quad\Rightarrow \frac{1}{1 - \frac{|A|P^1}{AP^1A+Q}} > \frac{1}{1 - \frac{|A|P^2}{AP^2A+Q}} \quad (57)\end{aligned}$$

where the fact that $0 < |A|P^1/(AP^1A + Q) < 1$ and $0 < |A|P^2/(AP^2A + Q) < 1$ are utilized in the last line, due to (51) and the stability of the Kalman filter. Since $Y^2/Y^1 < ((P^1 - P^2)C_1R_1^{-1}/P^2C_2R_2^{-1})^2$, we have

$$|A|P^1C_1R_1^{-1}\sqrt{Y^1} > |A|P^2\left(C_1R_1^{-1}\sqrt{Y^1} + C_2R_2^{-1}\sqrt{Y^2}\right).$$

[5]This can be achieved by changing the event-triggering conditions in the microprocessors on the sensor side.

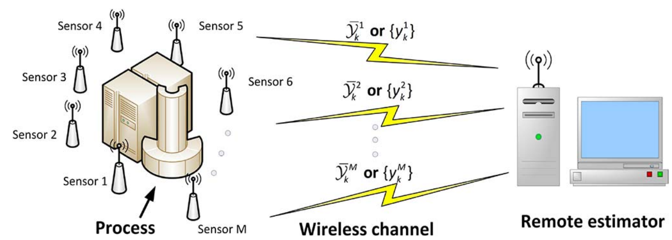

Fig. 2. Multiple-sensor event-based remote estimation architecture.

Combining with (57), we have $\bar{x}^1 > \bar{x}^{1,2}$, which completes the proof. ■

Remark 5: The intuition provided in the above result is that to achieve improved performance, the accuracy of sensor 2 should exceed certain level determined by that of sensor 1, although this does not require the accuracy of sensor 2 should be better compared with that of sensor 1.

Remark 6: As the confidence on stochastic and non-stochastic uncertainties is parameterized separately as covariance and the size of the set of estimation means, evaluation of the overall performance of a set-valued estimator is more complicated compared with its point-valued counterpart. Adding a sensor can always reduce the estimation error covariance, but can decrease, slightly or even severely increase the size of the set of estimation means. One possible approach of evaluating the overall or combined performance is to explore the equivalence relationship between stochastic and non-stochastic uncertainties in some sense, which can be potentially pursued based on ideas of the probabilistic approach or randomized algorithms utilized in control and estimation of uncertain systems [38], [39].

It is straightforward to observe that similar phenomenon exists for the multiple sensor case, and conditions for performance improvement can be obtained in a similar way. On the other hand, when taking all Y^i's as tuning parameters, the above analysis can be utilized to formulate design problems such that pre-specified performance can be achieved, as will be shown in the next section.

VI. Application to Event-Based State Estimation

In this section, we show how the results obtained in this work can be applied in remote event-based state estimation. Furthermore, an optimization problem is formulated and solved to design the event-triggering conditions by considering requirements on estimation performance and communication rates.

A. Analysis and Parameter Design

We consider the system in (1) measured by M sensors described in (2), which communicate with the remote state estimator through a wireless channel (see Fig. 2). We focus on the scenario that the communication channel is reliable with no packet dropouts, which is the case for shared networks using TDMA protocols or point-to-point communication links [40]. Due to the event trigger, the sets $\mathcal{Y}_k^i$ have more detailed parameterizations. At each time instant, the sensors measure the current state and decide whether to send the current measurement or not according to the values of binary decision variables γ_k^i's that are determined by pre-specified triggering conditions. We consider a relatively general description of the triggering conditions:

$$\gamma_k^i = \begin{cases} 0, & \text{if } y_k^i \in \bar{\mathcal{Y}}_k^i \\ 1, & \text{if } y_k^i \notin \bar{\mathcal{Y}}_k^i \end{cases} \tag{58}$$

where

$$\bar{\mathcal{Y}}_k^i = \left\{ y \in \mathbb{R}^m \middle| \left(y - \bar{c}_k^i\right)^\top \left(\bar{Y}_k^i\right)^{-1} \left(y - \bar{c}_k^i\right) \le 1 \right\}. \tag{59}$$

Note that the necessity of transmitting $\bar{Y}_k^i$ and $\bar{c}_k^i$ to the estimator during the non-event instants depends on the specific triggering conditions under consideration, as will be shown in Section VI-B. In this case, when $\gamma_k^i = 1$, the remote estimator receives the point-valued measurement information from sensor i, and thus the set of measurement information is given by a singleton $\mathcal{Y}_k^i = \{y_k^i\}$; when $\gamma_k^i = 0$, the set of measurement information is implicitly given by $\mathcal{Y}_k^i = \bar{\mathcal{Y}}_k^i$.

From the results obtained in Sections III and IV, it is known that:

1) The performance of the exact and approximate set-valued event-based estimators are invariant with respect to the fusion sequence. Notice that the counter part for either the exact or approximate MMSE event-based estimator is very difficult to be systematically verified [13].
2) For the event-based set-valued estimator, the set of estimation means is asymptotically bounded, and the outer approximations of the sets are bounded as well, which can be calculated according to arbitrary fusion sequences at each time instant.

On the other hand, it is still not clear how to design the event-triggering conditions so that the requirements on communication rate and estimation performance can be simultaneously considered, which is one of the main concerns in event-based control and estimation [12], [41]. In the following, we show how the analysis in Section V can be utilized in parameter design problems for guaranteed worst-case estimation performance and optimized communication rate.

For convenience of design and implementation, we consider the parameters $\bar{Y}_k^i$'s to be time invariant, namely, $\bar{Y}_k^i = \bar{Y}^i$. From Remark 4, it can be observed that increasing $\bar{Y}^i$ will lead to the decrease of the estimation performance in terms of the upper bound on the size of the set of estimation means at the steady state. On the other hand, from the literature of event-based estimation [12], it is known that the increase of $\bar{Y}^i$ leads to the reduction of the communication rate.[6] Therefore $\bar{Y}^i$'s can serve as tuning parameters for the tradeoff between estimation performance and communication rate. Note that the estimation performance here considers the size of the set of estimation means only, since the covariance is independent of $\bar{Y}^i$ in the set-valued filtering framework.

[6]Considering the scope of this work, we omit the analysis of the exact relationship between the communication rates and $\bar{Y}^i$, although, in fact, this analysis can be done following the approach in [12] with the difference that no Gaussian assumptions are required under the framework of set-valued filtering in this work.

First we introduce the constraints on the estimation performance. Observing that the measurement set $\mathcal{Y}_k^i$ is time varying (which can be $\bar{\mathcal{Y}}_k^i$ or $\{y_k^i\}$ depending on the value of γ_k^i), we consider the worst-case transient performance, namely, $\gamma_k^i = 0$ for a large number of consecutive k's such that the measurement set is always given by $\bar{\mathcal{Y}}_k^i$ during this period. From Corollary 2, the upper bound on the size of the set of estimation means will evolve towards an equilibrium, which thus quantifies the worst-case performance. We still assume $\|\bar{A}\|_2 < 1$; in case that $\|\bar{A}\|_2 \ge 1$, the results developed in this section can be applied by introducing the linear transformation T defined in the proof of Theorem 2 to the system. From (48), the worst-case performance bound is given by

$$\bar{x} = \frac{\sum_{i=1}^M \sqrt{\operatorname{Tr} AP^M C_i^\top R_i^{-1} \bar{Y}^i R_i^{-1} C_i P^M A^\top}}{1 - \left\| AP^M (AP^M A^\top + Q)^{-1} \right\|_2} \tag{60}$$

where P^M, C_i, and R_i are used instead of P^{s_M}, C_{s_i}, and R_{s_i} for notational brevity, since P^{s_M} is independent of the fusion sequence. To guarantee the worst-case performance, we specify an upper bound $\bar{x}^*$ and enforce the constraint $\bar{x} \le \bar{x}^*$. From (60), direct verification of this constraint is not computationally efficient for design purposes. Alternatively, using the Cauchy-Schwarz inequality

$$\begin{aligned} \bar{x} &= \frac{\sum_{i=1}^M \sqrt{\operatorname{Tr} \bar{Y}^i R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \left\| AP^M (AP^M A^\top + Q)^{-1} \right\|_2} \\ &\le \frac{\sum_{i=1}^M \sqrt{\operatorname{Tr} \bar{Y}^i} \sqrt{\operatorname{Tr} R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \left\| AP^M (AP^M A^\top + Q)^{-1} \right\|_2} \end{aligned} \tag{61}$$

thus the performance inequality can be indirectly enforced by requiring

$$\frac{\sum_{i=1}^M \sqrt{\operatorname{Tr} \bar{Y}^i} \sqrt{\operatorname{Tr} R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}}}{1 - \left\| AP^M (AP^M A^\top + Q)^{-1} \right\|_2} \le \bar{x}^* \tag{62}$$

which is a linear constraint of $\sqrt{\operatorname{Tr} \bar{Y}^i}$. On the other hand, we also include requirements on the upper bounds of the communication rates of each sensor by considering $\operatorname{Tr} \bar{Y}_i \ge \underline{\eta}^i \ge 0$, which is equivalent to $\sqrt{\operatorname{Tr} \bar{Y}_i} \ge \sqrt{\underline{\eta}^i}$. The objective of the parameter design is to minimize the communication rate, which is done by maximizing $\sum_{i=1}^M \operatorname{Tr} \bar{Y}_i$. To summarize, the parameter design problem is formulated as the following optimization problem:

$$\begin{aligned} \max_{a_1, a_2, \ldots, a_M} \quad & \sum_{i=1}^M a_i^2 \\ \text{s.t.} \quad & \sum_{i=1}^M b_i a_i \le \bar{x}^*, \quad a_i \ge \sqrt{\underline{\eta}_i},\ i = 1, 2, \ldots, M \end{aligned} \tag{63}$$

where $a_i = \sqrt{\operatorname{Tr} \bar{Y}^i}$ and $b_i = \sqrt{\operatorname{Tr} R_i^{-1} C_i P^M A^\top AP^M C_i^\top R_i^{-1}} / (1 - \|AP^M (AP^M A^\top + Q)^{-1}\|_2) > 0$ are used for notational brevity. Note that we safely ignored the case $b_i = 0$, since from Proposition 1 and the definition of matrix spectral norm, $b_i = 0$ if and only if the steady-state Kalman filter gain $\bar{K}^i$ corresponding to sensor i is zero, which implies that the consideration of sensor i will affect neither the estimation error covariance nor the size of the set of estimation means. To solve this problem, we further consider the following equivalent representation:

$$\begin{aligned} \max_{p_1,p_2,\ldots,p_M} \quad & \sum_{i=1}^{M} \left(p_i + \sqrt{\underline{\eta}^i}\right)^2 \\ \text{s.t.} \quad & \sum_{i=1}^{M} b_i p_i \le q, \quad p_i \ge 0, \ i = 1,2,\ldots,M \end{aligned} \tag{64}$$

where $q = \bar{x}^* - \sum_{i=1}^{M} b_i \sqrt{\underline{\eta}_i}$. Notice that this problem is feasible if and only if $q \ge 0$, which should be taken as the guideline in choosing the specifications of $\underline{\eta}^i$ and $\bar{x}^*$ in problem (63). Since this problem is a maximization problem of a positive semidefinite quadratic function over a polytope, the optimal solution is at one of the vertices, which are composed by the origin $p_i = 0$ and points of the form $p_i = q/b_i$, $p_j = 0$ for $j \neq i$ and i, $j \in \mathbb{N}_{1:M}$ for this case. Let $i^* = \arg\max_{i \in \mathbb{N}_{1:M}} q/b_i + \sqrt{\underline{\eta}^i}$. Since $b_i > 0$ and $\eta_i > 0$, the optimal value function of this problem equals $\left(q/b_{i^*} + \sqrt{\underline{\eta}^{i^*}}\right)^2 + \sum_{j=1, j\neq i^*}^{M} \underline{\eta}^j$ with optimizer $p_{i^*} = q/b_{i^*}$, $p_i = 0$ for $i \neq i^*$. This implies that the set of optimal parameters should be chosen as

$$\operatorname{Tr} \bar{Y}^i = \begin{cases} \underline{\eta}^i, & \text{if } i \neq i^*; \\ \left(\left(\sqrt{\underline{\eta}^i} + \left(\bar{x}^* - \sum_{j=1, j\neq i^*}^{M} b_j \sqrt{\underline{\eta}_j}\right)/b_i\right)^2, & \text{if } i = i^*. \end{cases} \tag{65}$$

Based on the value of $\operatorname{Tr} \bar{Y}^i$, $\bar{Y}^i$ can be chosen to satisfy further requirements, e.g., relative importance of different sensor channels. For the case of $m = 1$, $\bar{Y}^i$ reduces to a positive scalar, then the analysis here provides a complete parameter design procedure.

B. Examples

Example 2: In this example, we apply the set-valued estimation approach to the scenario of event-based state estimation with one sensor and interpret the difference of the obtained results from the existing results applicable to the same scenario [13], [42]. Consider a second-order system with parameter matrices

$$A = \begin{bmatrix} 0.5 & 0.3 \\ -0.1 & 0.8 \end{bmatrix}, \ Q = \begin{bmatrix} 0.202 & 0.053 \\ 0.053 & 0.136 \end{bmatrix}, \ C_1 = [0\ 1]$$

and $R_1 = 0.2$. We consider the "send-on-delta" triggering condition [7] of the following form:

$$\gamma_k^1 = \begin{cases} 0, & \text{if } \left(y_k^1 - y_{\tau_k^1}^1\right)^2 \le \delta \\ 1, & \text{otherwise} \end{cases} \tag{66}$$

where τ_k^1 denotes the last time instant when the measurement of the sensor is transmitted. For this system, $\|\bar{A}\|_2 = 0.51$, and thus the boundedness of the size of the set of estimation means can be guaranteed by Theorem 2 without introducing the linear transformations. Two other approaches are also applied for performance comparison, including the approximate event-based MMSE estimator [13] applicable to this event-triggering condition and the Kalman filter with intermittent observations [42], which corresponds to the MMSE estimator with an event-dependent filter gain obtained by only considering the measurement information received at the event instants and ignoring the information contained in event-triggering conditions during the non-event instants. To consider the performance of the estimators under different average communication rates, the estimators are implemented for δ equal to 0.1 and 1.2, the resultant average communication rates of which equal 0.621 and 0.171, respectively. The estimation error plots are shown in Figs. 3 and 4.

It is observed when the average communication rate is relatively high ($\delta = 0.1$), the size of the set of estimation means of the set-valued estimator is small, and the performance in terms of estimation error of the set-valued estimator can be

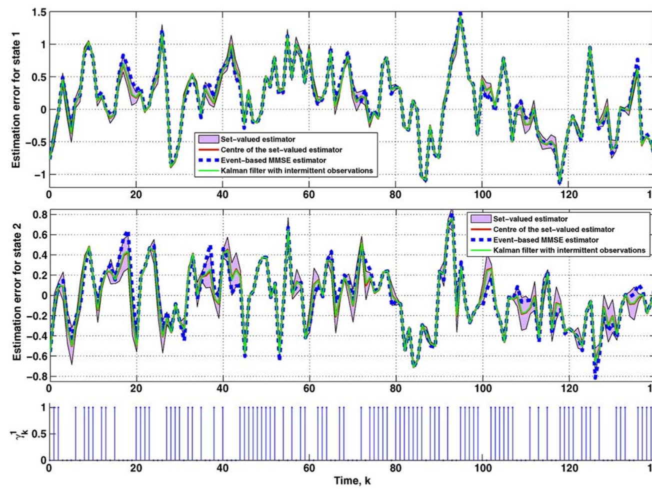

Fig. 3. Performance comparison of the different estimation strategy for $\delta = 0.1$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).

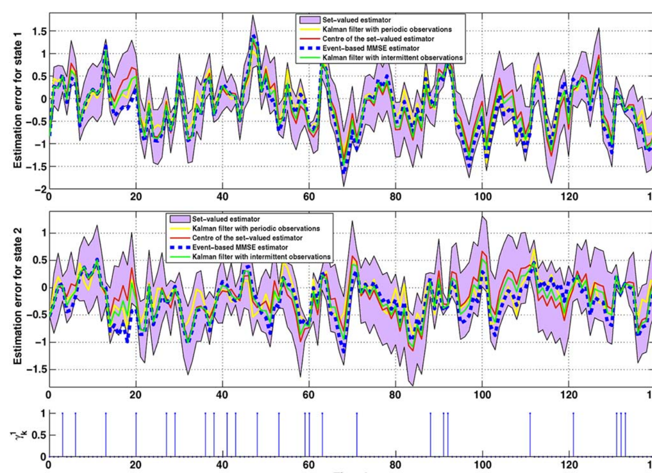

Fig. 4. Performance comparison of the different estimation strategy for $\delta = 1.2$ (the sets of estimation means are calculated by projecting the two-dimensional ellipsoids on one dimension).

characterized by the centre of the set of the estimation means. The average estimation errors[7] of the event-based MMSE estimator, the Kalman filter with intermittent observations and the centre of the set of the estimation means are numerically evaluated as 0.5589, 0.5781, and 0.5581, respectively. Under a lower average communication rate, however, the effect of separate parameterization of stochastic and non-stochastic uncertainty becomes more apparent. The exploration of set-valued information as non-stochastic uncertainty leads to a set of estimates with the same filtering gain that contains the estimates corresponding to all point-valued measurements lying in the event-triggering sets during non-event instants, including the MMSE estimate obtained by using the exact point-valued sensor measurements for all time instants (namely, the Kalman filter with periodic observations, see Fig. 4). In this case, it is not possible to tell which one in the set is associated with the smallest estimation error (without knowing the real state). The alternative answer, however, is that the centre of the set-valued estimator always serves as a point-valued estimate with the best robustness performance, in the sense that it has the smallest worst-case distance to the Kalman filter with periodic observations. This worst-case distance is known to be bounded (Theorem 2) and based on Remark 2, the asymptotic upper bounds are calculated as 0.2918 and 1.0110 for δ being 0.1 and 1.2, respectively. On the other hand, the precisions of the Gaussian approximations of the non-Gaussian distributions that were utilized to derive the event-based MMSE estimators are normally not possible to be verified, which is the basic motivation and theoretical benefit of utilizing the set-valued estimation approach. Finally, for $\delta = 1.2$, the average estimation errors of the event-based MMSE estimator, the Kalman filter with intermittent observations and the centre of the set of the estimation means are calculated as 0.6007, 0.6357, and 0.6018, respectively. This implies that for the event-triggering conditions considered, the centre of the set-valued estimator, which can be viewed as a point-valued estimator, empirically achieves similar improved performance in terms of average estimation error as that of the event-based MMSE estimator, compared with the Kalman filter with intermittent observations.

Example 3: In this example, we apply the developed event-trigger parameter design procedure in Section VI-A to a third-order system,[8] which is obtained by discretizing the benchmark model for a three-blade horizontal-axis turbine with a full converter coupling [43] with sampling time $T_s = 2.5$ s and including a system noise term

$$x_{k+1} = \begin{bmatrix} 0.9 & 0 & -1.5 \\ 66.1 & 0.3 & 2103.6 \\ 0 & 0 & 0.2 \end{bmatrix} x_k + \begin{bmatrix} 0 & -4.1 \\ 4 & -464.7 \\ 0 & 0 \end{bmatrix} u_k + w_k.$$

The input signal is generated according to the data provided in [43]. Four sensors are used to measure the state information

$$\begin{aligned} y_k^1 &= [1 \quad 0 \quad 0]\, x_k + v_k^1 \\ y_k^2 &= [1 \quad 0 \quad 0]\, x_k + v_k^2 \\ y_k^3 &= [0 \quad 0.1 \quad 0]\, x_k + v_k^3 \\ y_k^4 &= [0 \quad 0.1 \quad 0]\, x_k + v_k^4 \end{aligned}$$

with measurement noise covariances $R_1 = 0.03$, $R_2 = 0.05$, $R_3 = 0.17$, and $R_4 = 0.18$, respectively, and the system noise covariance is

$$Q = \begin{bmatrix} 0.2023 & 0.0530 & 0 \\ 0.0530 & 0.1360 & 0 \\ 0 & 0 & 0.1000 \end{bmatrix}.$$

We still consider the "send-on-delta" triggering conditions, namely

$$\gamma_k^i = \begin{cases} 0, & \text{if } y_k^i \in \bar{\mathcal{Y}}_k^i \\ 1, & \text{if } y_k^i \notin \bar{\mathcal{Y}}_k^i \end{cases} \tag{67}$$

where $\bar{\mathcal{Y}}_k^i = \{y \in \mathbb{R}^m | (y - y_{\tau_k^i}^i)^\top (\bar{Y}^i)^{-1} (y - y_{\tau_k^i}^i) \leq 1\}$, τ_k^i denoting the last time instant when the measurement of sensor i is transmitted. In this case, no communication is needed during the non-event instants, since $\bar{Y}^i$'s are constant and $y_{\tau_k^i}^i$'s are always known to the estimator. For this system, $\|\bar{A}\|_2 = 2103.6$. To guarantee the boundedness of the set of estimation means, we calculate the linear transformation

$$T = \begin{bmatrix} 80 & -0.1 & 974 \\ 0 & 1 & -380 \\ 0 & 0 & 1823.3 \end{bmatrix}$$

according to the proof of Theorem 2 and apply the estimation procedure to the transformed system. Furthermore, for problem (63), the b_i coefficients are calculated as $b_1 = 3.6049 \times 10^8$, $b_2 = 3.1543 \times 10^8$, $b_3 = 4.0788 \times 10^7$, and $b_4 = 3.852 \times 10^7$, respectively. The values for η_i's are specified as $\eta_1 = 0.5$, $\eta_2 = 0.4$, $\eta_3 = 30$, $\eta_4 = 28$, and $\bar{x}^* = 9 \times 10^8$. The event-triggering conditions are calculated according to (65) as $\bar{Y}_1 = 0.5$, $\bar{Y}_2 = 0.4$, $\bar{Y}_3 = 35.1317$, and $\bar{Y}_4 = 28$. The set-valued event-based estimator is then implemented and the estimation performance is shown in Fig. 5, which is obtained by applying inverse transformation T^{-1} to the estimates. The plot of sensor transmissions are shown in Fig. 6, where the average communication rates for the four sensors equal 0.577, 0.632, 0.950, and 0.955, respectively. From Fig. 5, it is observed that bounded envelopes for the estimates are always obtained, and the centers of the ellipsoids also serve as efficient point-valued estimates for the state variables. Notice that although the constraint in (62) guarantees worst-case performance, it also implicitly helps to control the transient performance. Another way to quantify the transient performance is to consider probabilistic performance constraints (combined with the average communication rates), which is the topic of our future work.

[7]For both $\delta = 0.1$ and $\delta = 1.2$, the average estimation error of an estimator is calculated by performing the simulation for 10 000 steps and calculating the sum of the 2-norm of the estimation error divided by 10 000.

[8]Note that the results in [13] and [42] are not considered for comparison here, as the example is devoted to illustrating the proposed event-triggering condition design procedure.

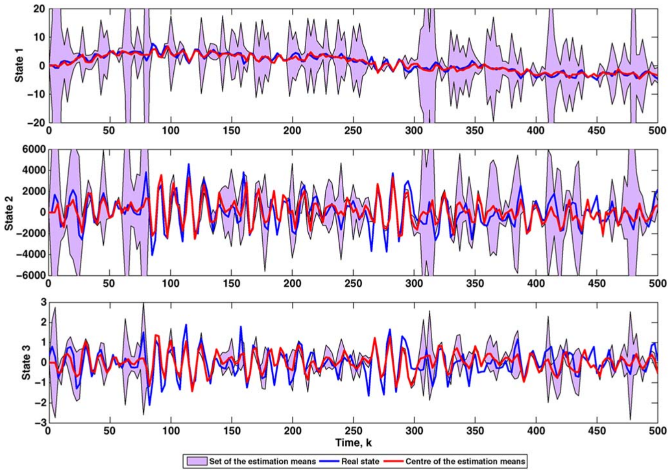

Fig. 5. Performance of the set-valued state estimation strategy (the envelopes are calculated by projecting the three-dimensional ellipsoids on one dimension).

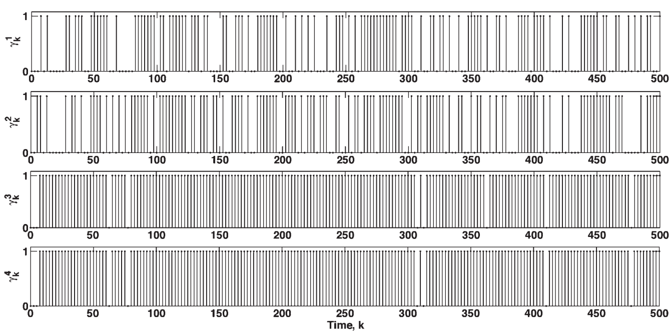

Fig. 6. Plot of sensor transmissions.

VII. Conclusion

In this work, the properties of set-valued Kalman filters with multiple sensor measurements are explored, which help provide further insights on event-based state estimation. Despite the distinct nature of the filter, it is shown that the important features of the classic Kalman filter, namely, the invariance of the estimation performance with respect to fusion sequences, the asymptotic boundedness of the performance measures (under certain assumptions, e.g., detectability and stabilizability), are maintained by both the exact set-valued filter and the proposed approximate set-valued filter. On the other hand, we show that the inclusion of more sensors does not necessarily reduce the size of the set of estimation means, and certain conditions need to be satisfied to guarantee performance improvement, which is utilized to formulate design problems in event-based estimation.

The developed results of the properties of set-valued Kalman filters are also applicable to the scenario of state estimation with quantized measurements. In this case, the measurement space is divided into a number of small quantization regions by the quantizer; we do not know the exact value of the measurement but know the region where it lies. For the case of $m = 1$, the results developed can be applied directly, as the regions corresponding to the same quantized values on $\mathbb{R}$ can be directly parameterized by one-dimensional ellipsoids; for the case of $m > 1$, outer-ellipsoidal approximations are needed to bound the m-dimensional quantization regions,[9] which are normally not in the form of ellipsoidal sets. As soon as these outer-ellipsoidal approximations are obtained, the results developed can be directly applied.

In the present event-based estimation framework, the communication channel is assumed to be reliable; a further step is to investigate the effect of packet dropouts, which is a non-trivial extension of the current work. Future research work also includes the consideration of transient behavior the size of the set of estimation means in performance analysis, sensor scheduling within finite horizon, and the consequent applications in event-based estimation.

Acknowledgment

The authors would like to thank the Associate Editor and the anonymous reviewers for their suggestions which have improved the quality of the work.

References

[1] V. C. Gungor and G. P. Hancke, “Industrial wireless sensor networks: Challenges, design principles, technical approaches,” *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4258–4265, Oct. 2009.
[2] V. C. Gungor, B. Lu, and G. P. Hancke, “Opportunities and challenges of wireless sensor networks in smart grid,” *IEEE Trans. Ind. Electron.*, vol. 57, no. 10, pp. 3557–3564, Oct. 2010.
[3] O. C. Imer and T. Basar, “Optimal estimation with limited measurements,” in *Proc. 44th IEEE Conf. Decision Control, Eur. Control Conf.*, Dec. 2005, pp. 1029–1034.
[4] K. J. Åström and B. M. Bernhardsson, “Comparison of Riemann and Lebesgue sampling for first order stochastic systems,” in *Proc. 41st IEEE Conf. Decision Control*, Dec. 2002, pp. 2011–2016.
[5] J. Wu, Y. Yuan, H. Zhang, and L. Shi, “How can online schedules improve communication and estimation tradeoff?” *IEEE Trans. Signal Process.*, vol. 61, no. 7, pp. 1625–1631, 2013.
[6] K. E. Årzén, “A simple event-based PID controller,” in *Proc. 14th IFAC World Congr.*, Beijing, China, 1999, vol. 18, pp. 423–428.
[7] M. Miskowicz, “Send-on-delta concept: An event-based data reporting strategy,” *Sensors*, vol. 6, no. 1, pp. 49–63, 2006.
[8] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.
[9] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, “A new method for the nonlinear transformation of means and covariances in filters and estimators,” *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.
[10] K. Ito and K. Xiong, “Gaussian filters for nonlinear filtering problems,” *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.
[11] I. Arasaratnam and S. Haykin, “Cubature kalman filters,” *IEEE Trans. Autom. Control*, vol. 54, no. 6, pp. 1254–1269, Jun. 2009.
[12] J. Wu, Q. Jia, K. Johansson, and L. Shi, “Event-based sensor data scheduling: Trade-off between communication rate and estimation quality,” *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 1041–1045, Apr. 2013.
[13] 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.
[14] J. Sijs and M. Lazar, “Event based state estimation with time synchronous updates,” *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2650–2655, Oct. 2012.
[15] D. Han, Y. Mo, J. Wu, B. Sinopoli, and L. Shi, “Stochastic event-triggered sensor scheduling for remote state estimation,” in *Proc. 52nd IEEE Conf. Decision and Control*, 2013, pp. 6079–6084.
[16] J. Sijs, B. Noack, and U. D. Hanebeck, “Event-based state estimation with negative information,” in *Proc. 16th Int. Conf. Information Fusion*, 2013.
[17] D. Shi, T. Chen, and L. Shi, “Event-triggered maximum likelihood state estimation,” *Automatica*, vol. 50, no. 1, pp. 247–254, 2014.
[18] L. Li, M. Lemmon, and X. Wang, “Event-triggered state estimation in vector linear processes,” in *Proc. 2010 American Control Conf.*, Jul. 2010, pp. 2138–2143.

[9]Note that as long as the quantization regions are given, the outer-ellipsoidal approximations are straightforward to be obtained.

[19] M. Lemmon, "Event-triggered feedback in control, estimation, optimization," in *Networked Control Systems*, vol. 406, A. Bemporad, M. Heemels, and M. Johansson, Eds. Springer, 2010, ser. Lecture Notes in Control and Information Sciences.
[20] S. Trimpe and R. D'Andrea, "An experimental demonstration of a distributed and event-based state estimation algorithm," in *Proc. 18th IFAC World Congr.*, Milano, Italy, Aug. 2011, pp. 8811–8818.
[21] J. Weimer, J. Araujo, and K. H. Johansson, "Distributed event-triggered estimation in networked systems," in *Proc. 4th IFAC Conf. Analysis and Design of Hybrid Systems*, Eindhoven, The Netherlands, Jun. 2012, pp. 178–185.
[22] S. Trimpe, "Event-based state estimation with switching static-gain observers," in *Proc. 3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, Santa Barbara, CA, USA, Sep. 2012, pp. 91–96.
[23] M. Rabi, G. Moustakides, and J. Baras, "Adaptive sampling for linear state estimation," *SIAM J. Control and Optimiz.*, vol. 50, no. 2, pp. 672–702, 2012.
[24] W. C. Stirling and D. R. Morrell, "Convex Bayes decision theory," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 1, pp. 173–183, 1991.
[25] D. R. Morrell and W. C. Stirling, "Set-valued filtering and smoothing," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 1, pp. 184–193, 1991.
[26] B. Noack, V. Klumpp, and U. D. Hanebeck, "State estimation with sets of densities considering stochastic and systematic errors," in *Proc. 12th Int. Conf. Information Fusion (Fusion 2009)*, Seattle, WA, USA, Jul. 2009, pp. 1751–1758.
[27] B. Noack, F. Pfaff, and U. D. Hanebeck, "Combined stochastic and set-membership information filtering in multisensor systems," in *Proc. 15th Int. Conf. Information Fusion (Fusion 2012)*, Singapore, Jul. 2012, pp. 1218–1224.
[28] A. Kurzhanski and I. Vályi, *Ellipsoidal Calculus for Estimation and Control*. Cambridge, MA, USA: Birkhauser, 1996.
[29] L. Chisci, A. Garulli, and G. Zappa, "Recursive state bounding by parallelotopes," *Automatica*, vol. 32, no. 7, pp. 1049–1055, 1996.
[30] T. Alamo, J. M. Bravo, and E. F. Camacho, "Guaranteed state estimation by zonotopes," *Automatica*, vol. 41, no. 6, pp. 1035–1043, 2005.
[31] R. Schneider, *Convex Bodies: The Brunn-Minkowski Theory*. Cambridge, U.K.: Cambridge University Press, 1996.
[32] E. Lutwak, "Intersection bodies and dual mixed volumes," *Adv. Math.*, vol. 71, pp. 232–261, 1988.
[33] G. De Nicolao, "On the time-varying Riccati difference equation of optimal filtering," *SIAM J. Control Optimiz.*, vol. 30, no. 6, pp. 1251–1269, 1992.
[34] A. G. O. Mutambara, *Decentralized Estimation adn Control for Multisensor Systems*. Boca Raton, FL, USA: CRC Press, 1998.
[35] R. A. Horn and C. R. Johnson, *Matrix Analysis,* 2nd ed. Cambridge, MA, USA: Cambridge University Press, 2013.
[36] S. Bittanti, P. Colaneri, and G. De Nicolao, "The difference periodic Riccati equation for the periodic prediction problem," *IEEE Trans. Autom. Control*, vol. 33, no. 8, pp. 706–712, Aug. 1988.
[37] D. Shi and T. Chen, "Optimal periodic scheduling of sensor networks: A branch and bound approach," *Syst. Control Lett.*, vol. 62, no. 9, pp. 732–738, 2013.
[38] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, 2010.
[39] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized Algorithms for Analysis and Control of Uncertain Systems*. London, U.K.: Springer-Verlag, 2005.
[40] D. Antunes and W. P. M. H. Heemels, "Rollout event-triggered control: Beyond periodic control performance," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3296–3311, 2014.
[41] X. Meng and T. Chen, "Optimal sampling and performance comparison of periodic and event based impulse control," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 3252–3259, 2012.
[42] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.
[43] P. F. Odgaard, J. Stoustrup, and M. Kinnaert, "Fault-tolerant control of wind turbines: A benchmark model," *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 4, pp. 1168–1182, 2013.

Dawei Shi received the B.Eng. degree in electrical engineering and automation from the Beijing Institute of Technology, Beijing, China, in 2008, and the Ph.D. degree in control systems from the University of Alberta, Edmonton, AB, Canada, in 2014.

Since September 2014, he has been an Associate Professor at the School of Automation, Beijing Institute of Technology. His research interests include event-based control and estimation, robust model predictive control and tuning, and wireless sensor networks.

Dr. Shi is a reviewer for a number of international journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica, and Systems and Control Letters*. In 2009, he received the Best Student Paper Award in IEEE International Conference on Automation and Logistics.

Tongwen Chen (F'06) received the B.Eng. degree in automation and instrumentation from Tsinghua University, Beijing, China, in 1984, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1988 and 1991, respectively.

He is presently a Professor of Electrical and Computer Engineering at the University of Alberta, Edmonton, AB, Canada. His research interests include computer and network based control systems, process safety and alarm systems, and their applications to the process and power industries.

Dr. Chen has served as an Associate Editor for several international journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, *Automatica, and Systems and Control Letters*. He is a Fellow of IFAC and the Canadian Academy of Engineering.

Ling Shi received the B.S. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology, Kowloon, in 2002 and the Ph.D. degree in control and dynamical systems from California Institute of Technology, Pasadena, in 2008.

He is currently an Associate Professor in the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology. His research interests include networked control systems, wireless sensor networks, and distributed control.