

# On the Choice of the Event Trigger in Event-based Estimation

Sebastian Trimpe  
 Autonomous Motion Department  
 Max Planck Institute for Intelligent Systems  
 Tübingen, Germany  
 Email: strimpe@tuebingen.mpg.de

Marco C. Campi  
 Department of Information Engineering  
 University of Brescia  
 Brescia, Italy  
 Email: marco.campi@unibs.it

**Abstract**—In event-based state estimation, the event trigger decides whether or not a measurement is used for updating the state estimate. In a remote estimation scenario, this allows for trading off estimation performance for communication, and thus saving resources. In this paper, popular event triggers for estimation, such as send-on-delta (SoD), measurement-based triggering (MBT), variance-based triggering (VBT), and relevant sampling (RS), are compared for the scenario of a scalar linear process with Gaussian noise. First, the analysis of the information pattern underlying the triggering decision reveals a fundamental advantage of triggers employing the real-time measurement in their decision (such as MBT, RS) over those that do not (VBT). Second, numerical simulation studies support this finding and, moreover, provide a quantitative evaluation of the triggers in terms of their average estimation versus communication performance.

## I. INTRODUCTION

The fundamental problem in event-based state estimation is to decide, whether a measurement shall be used to update a state estimate, or not. The corresponding decision rule is called the event trigger. A prototypical example, which is also considered in this paper, is shown in Fig. 1. A sensor node measures the state of a dynamic process and decides (by means of an event trigger) if a measurement is forwarded to a remote estimator. That is, it decides whether to use the communication link and improve the remote estimate, or to save communication. The cost of communication may stem from blocking the network for other nodes, or an overdimensioned network. In other scenarios, taking a measurement may be associated with the use of energy (e.g. battery-powered sensor nodes) or computation. In general, event-based estimation aims at making a compromise between estimation performance and resource usage by taking or transmitting a measurement only when necessary.

Since resource-constraint problems are typical in engineering, potential applications for event-based methods are numerous. Already today, basic event-based protocols (such as send-on-delta) are commonly used in building automation, [1], [2]. Recent application examples of event-based control include servo motor control [3], process plant control [4]–[7], and control of a laboratory crane [8]. Different methods for event-based estimation were applied for distributed control of a multi-body inverted pendulum in [9]–[11].

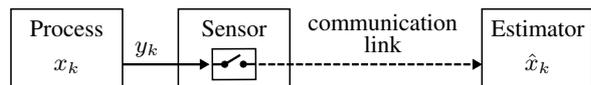


Fig. 1. Remote estimation scenario. The sensor node has periodic access to measurements  $y_k$  of the process  $x_k$ . Using an event trigger, the sensor decides at every step  $k$ , whether to forward  $y_k$  to the remote estimator (which incurs a communication cost) or not (which degrades estimation performance).

An event-based state estimator typically consists of two components: the *estimation* or *filtering algorithm*, which recursively computes state estimates from a model and the asynchronously arriving measurements, and the *event trigger*. While the question of how to design an optimal filtering algorithm for event-based estimation is interesting (see e.g. [12]–[15]), we focus on the design and comparison of event triggers herein. As the filtering algorithm, we simply take a standard Kalman filter (KF), which is suboptimal in this setting, but represents a straightforward and often satisfactory solution in practice (cf. discussion in Sec. II-A).

Several event triggering mechanisms have been suggested for event-based estimation in the last decade. In the following, we mention some popular triggers to be investigated herein. The triggers are exemplified on the scalar linear measurement equation  $y_k = cx_k + w_k$ , where  $x_k$  denotes the state to be estimated,  $y_k$  is the measurement, and  $w_k$  is sensor noise. Note, however, that all triggers are more general and readily extend to vector measurements.

Probably the most basic triggering concept is *send-on-delta* (SoD) [16], or also termed *deadbands* [17]. With SoD, the current measurement  $y_k$  is sent to the estimator whenever the difference to the measurement that was sent last,  $y_{\text{last},k}$ , exceeds a constant threshold  $\delta$ :

$$\text{transmit } y_k \Leftrightarrow |y_k - y_{\text{last},k}| \geq \delta. \quad (1)$$

SoD is a general purpose event trigger because no assumption or knowledge about the underlying process is required.

If a model of the underlying process is available, one may predict the evolution of the measurement and use a prediction  $\hat{y}_k$  instead of  $y_{\text{last},k}$  in the triggering rule. This idea is implemented in *measurement-based triggering* (MBT) [9], [10], or *predicted sampling* [18]:

$$\text{transmit } y_k \Leftrightarrow |y_k - c\hat{x}_{k|k-1}| \geq \delta \quad (2)$$

The work by S. Trimpe was supported by the Max Planck Society.

where  $\hat{x}_{k|k-1}$  is the one-step ahead, model-based prediction of  $x_k$ , and  $\hat{y}_k = c\hat{x}_{k|k-1}$  thus the measurement prediction. In order to realize this triggering law, the sensor in Fig. 1 implements a copy of the state estimator. If no process model for prediction is available, a prediction  $\hat{y}_k$  can also be obtained by approximating the derivatives of the signal  $y_k$  from successive past measurements. This has been called *SoD with linear predictor* [19].

MBT (2) imposes a bound on the measurement prediction error. Placing a bound on the prediction variance as a measure of uncertainty instead, yields the *variance-based triggering* (VBT) law [20]<sup>1</sup>:

$$\text{transmit } y_k \Leftrightarrow \text{Var}[y_k - c\hat{x}_{k|k-1}] = c^2 P_{k|k-1} + r \geq \delta \quad (3)$$

with  $P_{k|k-1}$  the state prediction variance,  $r$  the sensor noise variance, and assuming independence of  $x_k - \hat{x}_{k|k-1}$  and  $w_k$ .

An information-theoretic trigger is employed in *relevant sampling* (RS) [21]. The trigger is obtained by comparing the two state probability density functions (PDFs) that result when  $y_k$  is used in the update ( $f_{\text{upd}}$ ), or when it is not ( $f_{\text{prior}}$ ):

$$\text{transmit } y_k \Leftrightarrow D_{\text{KL}}(f_{\text{upd}}, f_{\text{prior}}) \geq \delta. \quad (4)$$

Here,  $D_{\text{KL}}$  is the Kullback-Leibler (KL) divergence, which can be interpreted as the information loss when the approximation  $f_{\text{prior}}$  is used instead of  $f_{\text{upd}}$ , [22]. In case of Gaussian PDFs  $f_{\text{prior}}$  and  $f_{\text{upd}}$  with respective means  $\hat{x}_{k|k-1}$ ,  $\hat{x}_{\text{upd}}$  and variances  $P_{k|k-1}$ ,  $P_{\text{upd}}$ , (4) can be rewritten as

$$D_{\text{KL}}(f_{\text{post}}, f_{\text{prior}}) = \frac{1}{2} \left( \frac{P_{\text{upd}}}{P_{k|k-1}} - \log \left( \frac{P_{\text{upd}}}{P_{k|k-1}} \right) - 1 + \frac{1}{P_{k|k-1}} (\hat{x}_{k|k-1} - \hat{x}_{\text{upd}})^2 \right). \quad (5)$$

While MBT (2) and VBT (3) depend either on the prediction mean or variance, (5) depends on both.

While several event triggers have been suggested in literature and some comparisons for their use in control are available (e.g. [1], [4], [23]), a systematic comparison in the context of event-based estimation is largely missing. The recent work [18] is an exception. Therein, the authors compare the triggers (1) and (2) in terms of their estimation and communication performance for a scalar, continuous-time, stable process subject to impulse and step disturbances. This paper provides a comparison of the triggers (1)–(4), as well as two novel ones to be derived herein. We consider a linear, discrete-time process (stable and unstable) with stochastic, Gaussian noise; that is, the classic stochastic estimation framework of the KF, [24]. In contrast to [18], we consider as the state estimator a time-varying KF instead of a fixed-gain observer. For the purpose of this paper, we restrict attention to a scalar process (as also in [18]), which may be seen as a first step in the endeavor of understanding the effectiveness of common triggering mechanisms for estimation.

This paper makes two main contributions toward this goal. First, we discuss, at a fundamental level, the information that the decision making agent (i.e. the smart sensor in Fig. 1) has, when making the transmit decision. For this, we consider two information structures: the case, where the value of  $y_k$  is

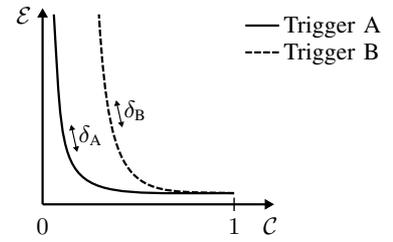


Fig. 2. Estimation-vs-communication graphs used for comparing the effectiveness of two event triggers. Each curve represents the average estimation error  $\mathcal{E}$  (e.g. squared error averaged over a time horizon) over the average communication rate  $\mathcal{C}$  (normalized such that 0 means no, and 1 means full communication) for a specific trigger. Each graph is parametrized by the corresponding triggering threshold  $\delta$ . In this example, Trigger A is more effective than Trigger B.

available at the time of making the triggering decision, and the case, where it is not. The former corresponds to rules like MBT or RS, which depend on the real-time measurement  $y_k$ , while the latter captures VBT, which does not. The outcome of this analysis is 1) fundamental understanding of the value that the measurement  $y_k$  has for the triggering decision; and 2) a unified framework for deriving different event triggers. In particular, we show that the triggers (2), (3), and two new ones, are obtained as (approximate) solutions to optimization problems respecting one of the two information structures and representing a trade-off between some measure of estimation performance and communication cost.

The second main contribution of this work is a quantitative comparison of all mentioned triggers based on simulation studies. We propose to use estimation-vs-communication graphs as shown in Fig. 2 as the method of comparison. These graphs depict the average estimation error achieved by some trigger over the average communication rate. Moving along one of the curves is parametrized by the threshold parameter  $\delta$  of the respective trigger. A graph lying below another one indicates a trigger that is (globally) more effective in making the transmit decision: for the same communication rate, a smaller estimation error is achieved on average.

Taken together, this paper makes a fundamental contribution toward understanding what event trigger should be used for a given event-based estimation problem. Obviously, this is an important question if one wants to deploy an event-based approach in practice and get most out of it in terms of estimation and communication efficiency. While other scenarios are conceivable, two architectural assumptions in this analysis are the use of a standard KF as the estimation algorithm, and the transmission of raw measurements.

*Outline:* In Sec. II, we introduce the considered process, the estimation algorithm, and formalize the triggering decision as an optimization problem. Section III continues with analyzing two information structures (using the actual measurement in the decision, or not) for solving the triggering problem. Using the results of this section, the optimal decision problem from Sec. II is solved in Sec. IV, which yields two novel event triggers. Section V compares the effectiveness of all triggers in terms of their estimation-vs-communication graphs from simulations. The findings of this paper are discussed and summarized in Sec. VI.

<sup>1</sup>The triggering law in [20] is given in a slightly different form, which can readily be transformed into (3) by redefining the triggering threshold.

*Notation:* We use  $f(x)$ ,  $f(x|y)$ , and  $f(x|\mathcal{Y})$  to denote, respectively, the probability density functions (PDFs) of the random variable (RV)  $x$ , of  $x$  conditioned on the RV  $y$ , and of  $x$  conditioned on the set of RVs  $\mathcal{Y}$ .  $\mathcal{N}(x; \mu, \Sigma)$  denotes the PDF of a Gaussian random variable (GRV) with mean  $\mu$  and variance  $\Sigma$ . For  $x$  being distributed according to  $f(x) = \mathcal{N}(x; \mu, \Sigma)$ , we also write  $x \sim \mathcal{N}(\mu, \Sigma)$ .  $\mathbb{E}[\cdot]$  denotes the expected value, and  $\text{Var}[\cdot]$  the variance.

## II. PROBLEM FORMULATION

Consider a scalar, discrete-time, linear system with mutually independent Gaussian noise

$$\begin{aligned} x_k &= ax_{k-1} + v_{k-1} & x_0 &\sim \mathcal{N}(\bar{x}_0, X_0), v_{k-1} \sim \mathcal{N}(0, q) & (6) \\ y_k &= cx_k + w_k & w_k &\sim \mathcal{N}(0, r), & (7) \end{aligned}$$

where  $a, c, q, r, \bar{x}_0, X_0 \in \mathbb{R}$  with  $c \neq 0$ ,  $q \geq 0$ ,  $r > 0$ ,  $X_0 \geq 0$ . Without loss of generality, we take  $c = 1$ .

The state  $x_k$  shall be estimated from measurements  $y_k$  in a scenario, where using a measurement incurs a cost, such as in the scenario in Fig. 1. The event trigger is responsible for making this decision. Let  $\gamma_k \in \{0, 1\}$  denote the binary function indicating whether  $y_k$  is transmitted ( $\gamma_k = 1$ ) or not ( $\gamma_k = 0$ ). Furthermore, let  $\mathcal{Y}_k$  denote the set of triggering decisions and measurements transmitted until time  $k$ ; that is,

$$\tilde{\mathcal{Y}}_k = \{\gamma_\ell \mid 1 \leq \ell \leq k\} \cup \{y_\ell \mid 1 \leq \ell \leq k, \gamma_\ell = 1\}. \quad (8)$$

### A. State Estimator

An optimal Bayesian state estimator computes the PDF of the state  $x_k$  conditioned on the data  $\tilde{\mathcal{Y}}_k$ ,  $f(x_k|\tilde{\mathcal{Y}}_k)$ . In particular, the Bayesian estimator also includes the information that is contained in the event of *not* receiving a measurement. For example, if trigger (2) is used, then  $\gamma_k = 0$  implies that  $y_k \in [c\hat{x}_{k|k-1} - \delta, c\hat{x}_{k|k-1} + \delta]$ . Using this information in a Bayesian update, however, yields a non-Gaussian posterior and thus typically results in intractable algorithms. Approximate algorithms have been suggested in [12]–[15], for example.

Herein, we opt for ignoring the extra information from  $\gamma_k = 0$  in favor of a tractable algorithm. To this end, we treat the measurements  $y_k$  contained in  $\mathcal{Y}_k$  as if they were obtained from an a-priori fixed sensor schedule (i.e. the sequence  $\gamma_k$  fixed ahead of time). We denote this set of measurements by  $\mathcal{Y}_k$  and shall compute the artificial distribution  $f(x_k|\mathcal{Y}_k)$  instead of  $f(x_k|\tilde{\mathcal{Y}}_k)$ . While obviously being an approximation to the true situation, this yields a straightforward estimation algorithm, since  $f(x_k|\mathcal{Y}_k)$  is Gaussian and can recursively be computed by the standard discrete-time Kalman filter (KF) [24]. The KF equations read

$$\hat{x}_{k|k-1} = a\hat{x}_{k-1|k-1} \quad (9)$$

$$P_{k|k-1} = a^2 P_{k-1|k-1} + q \quad (10)$$

$$L_k = \frac{P_{k|k-1}}{P_{k|k-1} + r} = \frac{a^2 P_{k-1|k-1} + q}{a^2 P_{k-1|k-1} + q + r} \quad (11)$$

$$\hat{x}_{k|k} = \begin{cases} \hat{x}_{k|k-1} & \text{if } \gamma_k = 0 \\ \hat{x}_{k|k-1} + L_k(y_k - \hat{x}_{k|k-1}) & \text{if } \gamma_k = 1 \end{cases} \quad (12)$$

$$P_{k|k} = \begin{cases} P_{k|k-1} & \text{if } \gamma_k = 0 \\ (1 - L_k)P_{k|k-1} & \text{if } \gamma_k = 1 \end{cases} \quad (13)$$

with initialization  $\hat{x}_{0|0} = \bar{x}_0$ ,  $P_{0|0} = X_0$ . As per the above discussion, we have

$$\begin{aligned} f(x_k|\mathcal{Y}_{k-1}) &= \mathcal{N}(x_k; \hat{x}_{k|k-1}, P_{k|k-1}) \\ f(x_k|\mathcal{Y}_k) &= \mathcal{N}(x_k; \hat{x}_{k|k}, P_{k|k}). \end{aligned}$$

In order to distinguish the two paths that the KF can take in (12) and (13), we use  $\hat{x}_k^I$  and  $P_k^I$  to refer to the posterior update in case of no communication of  $y_k$ , and  $\hat{x}_k^II$  and  $P_k^II$  for the case of communication. For simplicity, we later refer to the estimator that updates according to  $\hat{x}_k^I$ ,  $P_k^I$  as *Estimator I*, and the one that updates with and  $\hat{x}_k^II$ ,  $P_k^II$  as *Estimator II*.

### B. Decision Problem

In this subsection, we formulate the triggering decision as a one-step optimal decision problem. Solving this optimization problem shall yield novel triggering rules in Sec. IV and put previously suggested triggers into context, as is discussed later.

Let  $C_k$  be the cost that is incurred when measurement  $y_k$  is transmitted. This cost is problem specific and may be associated to the use of communication bandwidth, computation resources, or energy, for example. In the following, a constant cost  $C_k = C$  is assumed, albeit the presented results readily extend to the case of varying costs.

Suppose that the estimation algorithm (9)–(13) has executed  $k-1$  steps on the estimator node (cf. Fig. 1); that is, the state estimate has been computed with respect to the data  $\mathcal{Y}_{k-1}$ . The sensor node, which also implements (9)–(13), now needs to decide whether or not to transmit the next measurement  $y_k$ . If  $y_k$  is transmitted, the cost  $C$  is incurred. If not, no communication cost  $C$  is incurred, but a price is paid in terms of a deteriorated estimation performance. Namely, the estimation error then is

$$e_k^I := x_k - \hat{x}_k^I, \quad \text{with } \hat{x}_k^I = \mathbb{E}[x_k|\mathcal{Y}_{k-1}] \quad (14)$$

instead of

$$e_k^{II} := x_k - \hat{x}_k^{II}, \quad \text{with } \hat{x}_k^{II} = \mathbb{E}[x_k|y_k, \mathcal{Y}_{k-1}]. \quad (15)$$

Suppose we care to keep the squared estimation error small. We thus define the estimation cost  $E_k$  that is incurred when  $y_k$  is not transmitted as

$$E_k := (e_k^I)^2 - (e_k^{II})^2. \quad (16)$$

That is, the cost is positive if  $(e_k^I)^2$  is larger than  $(e_k^{II})^2$ , which is to be expected on average.

Ideally, we would like to use the error difference (16) directly for deciding whether it is worth transmitting  $y_k$ , or not. However,  $E_k$  cannot be computed by the sensor since the true state  $x_k$  is of course unknown, and so are the actual errors  $e_k^I$  and  $e_k^{II}$ . Instead, we seek a decision law that represents an optimal trade-off *on average* when *conditioning on the available data*. For this, we consider two different information structures:

- (i) the measurement  $y_k$  is known for making the transmit decision at time  $k$ ;
- (ii) the measurement  $y_k$  is not used for the decision.

Intuitively, one may already suspect that (i) is beneficial (“more information can’t hurt”), but we shall make this benefit precise

in the following sections. As shall also be seen, case (ii) can be beneficial if one wants to avoid continuously monitoring the sensor.

Corresponding to the two cases, we define

$$\bar{E}_k^{(i)} := \mathbb{E}[(e_k^I)^2 - (e_k^II)^2 | y_k, \mathcal{Y}_{k-1}] \quad (17)$$

$$\bar{E}_k^{(ii)} := \mathbb{E}[(e_k^I)^2 - (e_k^II)^2 | \mathcal{Y}_{k-1}] \quad (18)$$

and pose the decision problem as the one-step optimization

$$\min_{\gamma_k \in \{0,1\}} \gamma_k C + (1 - \gamma_k) \bar{E}_k \quad (19)$$

where  $\bar{E}_k$  is either  $\bar{E}_k^{(i)}$  or  $\bar{E}_k^{(ii)}$ . The formal solution can be written as

$$\gamma_k = 1 \Leftrightarrow \bar{E}_k \geq C. \quad (20)$$

By evaluating  $\bar{E}_k$  for case (i) and (ii), the corresponding triggering rules will be obtained in Sec. IV.

### III. AVAILABLE INFORMATION FOR MAKING THE TRANSMIT DECISION

Before computing  $\bar{E}_k$  in (20) for the two cases (17) and (18), we first discuss more generally the information that the agent has available for the transmit decision in both cases (i) and (ii). This discussion sheds light on the fundamental benefit of using  $y_k$  in the triggering decision, and also applies to other choices of error metric than (16).

Effectively, the triggering agent is to decide at time  $k$  between  $e_k^I$  and  $e_k^II$  as the estimation error. The information available for making this decision can be characterized by the joint PDF of  $e_k^I$  and  $e_k^II$  conditioned on the available data; i.e.

$$f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1}) \quad \text{and} \quad (21)$$

$$f(e_k^I, e_k^II | \mathcal{Y}_{k-1}) \quad (22)$$

for case (i) and (ii), respectively. These two joint PDFs are computed in the following subsections and interpreted with regards to the triggering decision.

#### A. Case (i): Exploiting real-time measurement $y_k$

For computing (21), we rewrite the estimation errors as

$$e_k^I = x_k - \hat{x}_{k|k-1} = z_k - w_k \quad (23)$$

$$\begin{aligned} e_k^II &= x_k - \hat{x}_{k|k-1} - L_k(y_k - \hat{x}_{k|k-1}) \\ &= (1 - L_k)z_k - w_k \end{aligned} \quad (24)$$

where we used  $x_k = y_k - w_k$  and introduced  $z_k := y_k - \hat{x}_{k|k-1}$  for the KF innovation. Notice that the agent can compute  $z_k$  from the data  $y_k, \mathcal{Y}_{k-1}$ ; that is, the uncertainty in  $e_k^I$  and  $e_k^II$  when conditioned on  $y_k, \mathcal{Y}_{k-1}$  solely stems from  $w_k$ . In the appendix, we show that the conditional PDF of  $w_k$  is Gaussian:

$$f(w_k | y_k, \mathcal{Y}_{k-1}) = \mathcal{N}\left(w_k; \frac{r}{P_{k|k-1} + r} z_k, r - \frac{r^2}{P_{k|k-1} + r}\right). \quad (25)$$

From this, it follows that  $f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1})$  is Gaussian likewise. Using (23), (24), (25), we compute its mean as

$$\begin{aligned} \mathbb{E}\left[\begin{bmatrix} e_k^I \\ e_k^II \end{bmatrix} \middle| y_k, \mathcal{Y}_{k-1}\right] &= \begin{bmatrix} z_k \\ (1 - L_k)z_k \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathbb{E}[w_k | y_k, \mathcal{Y}_{k-1}] \\ &= \begin{bmatrix} \frac{P_{k|k-1} + r - r}{P_{k|k-1} + r} z_k \\ \frac{P_{k|k-1} + r - P_{k|k-1} - r}{P_{k|k-1} + r} z_k \end{bmatrix} = \begin{bmatrix} L_k z_k \\ 0 \end{bmatrix}. \end{aligned} \quad (26)$$

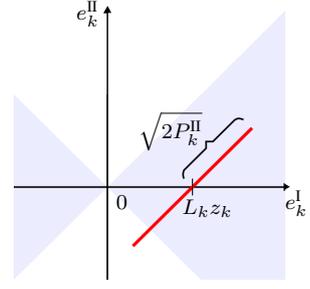


Fig. 3. Visualization of the conditional PDF  $f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1})$  (30). The PDF is one-dimensional; it is shown in red for one standard deviation  $(P_k^II)^{1/2}$ . The PDF depends on the measurement  $y_k$  through  $z_k$ . The light blue area indicates points where Estimator II is superior to Estimator I, i.e.  $(e_k^II)^2 < (e_k^I)^2$ .

Similarly, we obtain the variance terms. First,

$$\begin{aligned} \text{Var}[e_k^II | y_k, \mathcal{Y}_{k-1}] &= \mathbb{E}[(e_k^II - \mathbb{E}[e_k^II | y_k, \mathcal{Y}_{k-1}])^2 | y_k, \mathcal{Y}_{k-1}] \\ &= \mathbb{E}[\left((1 - L_k)z_k - w_k\right)^2 | y_k, \mathcal{Y}_{k-1}] \\ &= (1 - L_k)^2 z_k^2 - 2(1 - L_k)z_k \mathbb{E}[w_k | y_k, \mathcal{Y}_{k-1}] \\ &\quad + \mathbb{E}[w_k^2 | y_k, \mathcal{Y}_{k-1}] \\ &= \frac{P_{k|k-1} r}{P_{k|k-1} + r} = (1 - L_k) P_{k|k-1} \end{aligned} \quad (27)$$

where we used

$$\mathbb{E}[w_k | y_k, \mathcal{Y}_{k-1}] = \frac{r}{P_{k|k-1} + r} z_k = (1 - L_k) z_k \quad (28)$$

$$\begin{aligned} \mathbb{E}[w_k^2 | y_k, \mathcal{Y}_{k-1}] &= (\mathbb{E}[w_k | y_k, \mathcal{Y}_{k-1}])^2 + \text{Var}[w_k | y_k, \mathcal{Y}_{k-1}] \\ &= (1 - L_k)^2 z_k^2 + r - \frac{r^2}{P_{k|k-1} + r}. \end{aligned} \quad (29)$$

Using the result (27), we obtain next

$$\begin{aligned} \text{Var}[e_k^I | y_k, \mathcal{Y}_{k-1}] &= \mathbb{E}[(e_k^I - \mathbb{E}[e_k^I | y_k, \mathcal{Y}_{k-1}])^2 | y_k, \mathcal{Y}_{k-1}] \\ &= \mathbb{E}[\left((1 - L_k)z_k - w_k\right)^2 | y_k, \mathcal{Y}_{k-1}] = (1 - L_k) P_{k|k-1} \end{aligned}$$

and, for the cross-covariance,

$$\begin{aligned} \mathbb{E}[(e_k^I - \mathbb{E}[e_k^I | y_k, \mathcal{Y}_{k-1}])(e_k^II - \mathbb{E}[e_k^II | y_k, \mathcal{Y}_{k-1}]) | y_k, \mathcal{Y}_{k-1}] \\ = \mathbb{E}[\left((1 - L_k)z_k - w_k\right)^2 | y_k, \mathcal{Y}_{k-1}] = (1 - L_k) P_{k|k-1}. \end{aligned}$$

In summary,

$$f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1}) = \mathcal{N}\left(\begin{bmatrix} e_k^I \\ e_k^II \end{bmatrix}; \begin{bmatrix} L_k z_k \\ 0 \end{bmatrix}, \begin{bmatrix} P_k^II & P_k^II \\ P_k^II & P_k^II \end{bmatrix}\right) \quad (30)$$

with  $P_k^II$  as defined in (13). The PDF (30) characterizes the information available for making the transmit decision and thus deciding whether the estimation error on the estimator side becomes  $e_k^I$  or  $e_k^II$ . Notice that the PDF depends on the real-time measurement  $y_k$  through  $z_k$ .

*Interpretation:* From (30), we see that  $f(e_k^II | y_k, \mathcal{Y}_{k-1}) = \mathcal{N}(0, P_k^II)$ ; that is, the conditional error of Estimator II is unbiased with variance  $P_k^II$ . This makes perfect sense: it is simply the standard KF update incorporating the additional measurement  $y_k$ . Given the new measurement  $y_k$ , the error  $e_k^I$  of Estimator I is biased, since the additional information is not taken into account. The bias corresponds to the correction term  $L_k z_k$  that Estimator II would apply.

A visualization of the conditional PDF (30) is shown in Fig. 3. Some properties can be seen from this:

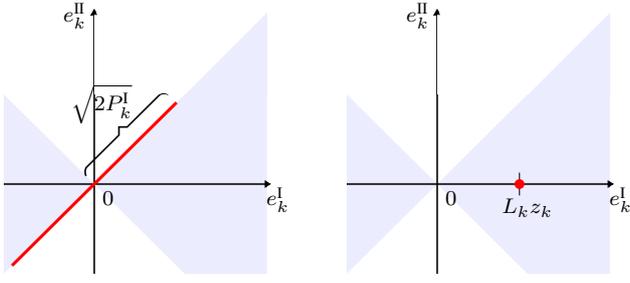


Fig. 4. Conditional PDF  $f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1})$  as in Fig. 3 for the special cases  $L_k = 0$  (LEFT) and  $L_k = 1$  (RIGHT). On the RIGHT, the support of the PDF reduces to a point, corresponding to the Dirac measure.

- The probability mass of  $f(e_k^I, e_k^II | y_k, \mathcal{Y}_{k-1})$  is concentrated in one dimension. Estimator II is superior (i.e.  $(e_k^II)^2 < (e_k^I)^2$ ) whenever  $(e_k^I, e_k^II)$  lies in the light blue area. Since the larger part of the probability mass is always in this area, Estimator II will yield a superior estimate with high probability, as expected. It may so happen, however, that the concrete representation of the errors  $(e_k^I, e_k^II)$  is outside the light blue area, and thus  $(e_k^II)^2 > (e_k^I)^2$ . This corresponds to a measurement  $y_k$  with a large noise component  $w_k$ , which would impair the prior estimate. In general, the advantage of Estimator II over Estimator I is larger, the larger the correction term  $|L_k z_k|$ .
- From (11), we see that  $L_k \in [0, 1)$ . For increasing measurement noise  $r$  and all other parameters constant,  $L_k$  decreases, which lowers the advantage of Estimator II over Estimator I. This is expected since the measurement carries “less value” for the estimator because of increased noise. In the limit as  $r \rightarrow \infty$ ,  $L_k \rightarrow 0$  and Estimator II is just as good as Estimator I (see Fig. 4 LEFT).
- In contrast, for decreasing  $r$ ,  $L_k$  increases and Estimator II gains in advantage over Estimator I. In the limit as  $r \rightarrow 0$ ,  $L_k$  tends to its maximum 1 (see Fig. 4 RIGHT). In this case, the knowledge of  $(e_k^I, e_k^II)$  is exact (the PDF (30) reduces to a Dirac delta).

### B. Case (ii): Without real-time measurement

In this section, we compute (22) and compare the PDF to the previous case (30). For the estimation errors, we have

$$e_k^I = x_k - \hat{x}_{k|k-1} \quad (31)$$

$$e_k^II = (1 - L_k)(x_k - \hat{x}_{k|k-1}) - L_k w_k. \quad (32)$$

The variables  $(e_k^I, e_k^II)$  are jointly Gaussian. Using conditional independence of  $x_k - \hat{x}_{k|k-1}$  and  $w_k$ , we find the conditional mean to be 0, and compute the variance as

$$\begin{aligned} \text{Var} \begin{bmatrix} e_k^I \\ e_k^II \end{bmatrix} \Big| \mathcal{Y}_{k-1} &= \begin{bmatrix} P_{k|k-1} & (1 - L_k)P_{k|k-1} \\ * & (1 - L_k)^2 P_{k|k-1} + L_k^2 r \end{bmatrix} \\ &= \begin{bmatrix} P_{k|k-1} & (1 - L_k)P_{k|k-1} \\ * & (1 - L_k)P_{k|k-1} \end{bmatrix}. \end{aligned} \quad (33)$$

Therefore,

$$f(e_k^I, e_k^II | \mathcal{Y}_{k-1}) = \mathcal{N} \left( \begin{bmatrix} e_k^I \\ e_k^II \end{bmatrix}; \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} P_k^I & P_k^{II} \\ P_k^{II} & P_k^II \end{bmatrix} \right) \quad (34)$$

where  $P_k^I$  and  $P_k^{II}$  are as defined in (13).

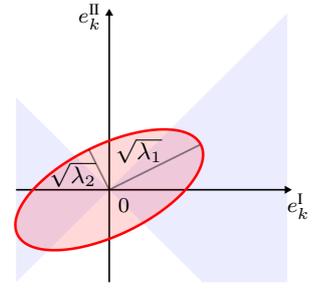


Fig. 5. Visualization of the conditional PDF  $f(e_k^I, e_k^II | \mathcal{Y}_{k-1})$  (34) (1-sigma ellipse in red). The direction of the main axis of the ellipse is  $[(L_k + S_k)/(2 - 2L_k), 1]$  where  $S_k := (5L_k^2 - 8L_k + 4)^{1/2}$ . The eigenvalues of the covariance matrix in (33) are  $\lambda_1 = 0.5(2 - L_k + S_k)P_{k|k-1}$  and  $\lambda_2 = 0.5(2 - L_k - S_k)P_{k|k-1}$ . The PDF does not depend on the measurement  $y_k$ .

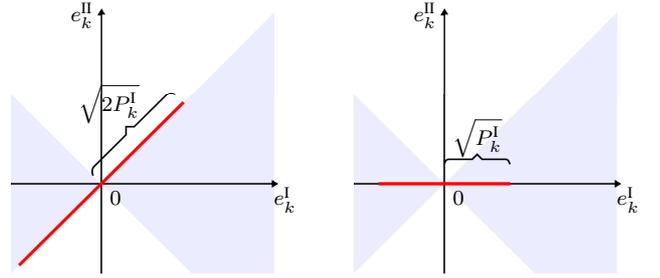


Fig. 6. Conditional PDF  $f(e_k^I, e_k^II | \mathcal{Y}_{k-1})$  as in Fig. 5 for the special cases  $L_k = 0$  (LEFT) and  $L_k = 1$  (RIGHT).

*Interpretation:* In contrast to (30), we see from (34) that both marginal PDFs  $f(e_k^I | \mathcal{Y}_{k-1}) = \mathcal{N}(0, P_k^I)$  and  $f(e_k^II | \mathcal{Y}_{k-1}) = \mathcal{N}(0, P_k^{II})$  are unbiased; they only differ in their variance. Since  $L_k \in [0, 1)$ ,  $P_k^{II} = (1 - L_k)P_k^I < P_k^I$ ; that is, the uncertainty of Estimator II is reduced, as expected.

The PDF (34) is visualized in Fig. 5 for general  $L_k$ , and in Fig. 6 for the extreme cases  $L_k = 0$  and  $L_k = 1$ . We make the following observations (cf. observations in Sec. III-A):

- Except for the limiting cases  $L_k = 0$  or 1, the joint PDF is two dimensional. Most of the probability mass is in the area where  $(e_k^II)^2 < (e_k^I)^2$ , as is expected since using  $y_k$  is generally better than not using it (as far as estimation performance is concerned). When comparing to Fig. 3, we see that conditioning on  $y_k$  effectively reduces the uncertainty from two to one dimension.
- For  $r \rightarrow \infty$  and  $L_k \rightarrow 0$  (cf. Fig. 6 LEFT), the measurement is of no value. This means that Estimator I is as good as Estimator II, and also (34) becomes identical to (30) (cf. Fig. 6 LEFT and Fig. 4 LEFT).
- For  $r \rightarrow 0$  and  $L_k \rightarrow 1$  (Fig. 6 RIGHT), the PDF (34) is again one dimensional. In this case,  $(e_k^II)^2 < (e_k^I)^2$  with probability one.

The key difference between the PDFs (30) and (34) is that the first includes the information carried by the measurement  $y_k$  (through  $z_k$ ): the larger  $z_k$ , the greater the advantage of Estimator II (cf. Fig. 3) and thus of transmitting  $y_k$ . The PDF (34) ignores this information.

#### IV. OPTIMAL EVENT TRIGGERS

Using the results from the previous section, the errors (17) and (18) can readily be computed. This yields the optimal event triggers according to (20).

##### A. Case (i)

Using the result (30), we obtain

$$\begin{aligned}\bar{E}_k^{(i)} &= \mathbb{E}[(e_k^I)^2 - (e_k^{II})^2 | y_k, \mathcal{Y}_{k-1}] \\ &= (\mathbb{E}[e_k^I | y_k, \mathcal{Y}_{k-1}])^2 + \text{Var}[e_k^I | y_k, \mathcal{Y}_{k-1}] \\ &\quad - (\mathbb{E}[e_k^{II} | y_k, \mathcal{Y}_{k-1}])^2 - \text{Var}[e_k^{II} | y_k, \mathcal{Y}_{k-1}] \\ &= (L_k z_k)^2 + P_k^I - 0 - P_k^{II} = L_k^2 z_k^2.\end{aligned}\quad (35)$$

The transmit decision (20) thus becomes

$$\gamma_k = 1 \Leftrightarrow L_k^2 z_k^2 = L_k^2 |y_k - \hat{x}_{k|k-1}|^2 \geq C. \quad (36)$$

Except for the time-varying weight  $L_k$ , this triggering rule corresponds to the measurement-based triggering rule (2) (recall that  $c = 1$  was assumed w.l.o.g.). The communication cost  $C$  has the role of the triggering threshold  $\delta$ . We thus refer to trigger (36) as *optimal measurement-based trigger* (OMBT).

##### B. Case (ii)

With (34), we compute

$$\begin{aligned}\bar{E}_k^{(ii)} &= \mathbb{E}[(e_k^I)^2 - (e_k^{II})^2 | \mathcal{Y}_{k-1}] = P_k^I - P_k^{II} \\ &= L_k P_{k|k-1} = L_k^2 (P_{k|k-1} + r).\end{aligned}\quad (37)$$

The corresponding triggering rule (20) is

$$\gamma_k = 1 \Leftrightarrow L_k P_{k|k-1} \geq C. \quad (38)$$

From (37), it can be seen that this triggering rule corresponds to variance-based triggering (3) except for the factor  $L_k^2$ . We thus call (38) *optimal variance-based trigger* (OVBT).

The rule (38) does not depend on  $y_k$ . Given the decision  $\gamma_k$  at time  $k$ , the KF variance and gain at time  $k+1$  are determined according to (10), (11), (13); hence,  $\gamma_{k+1}$  can also be computed from (38). By continuing this recursion, all triggering decisions can thus be precomputed off-line from the problem data ( $a$ ,  $q$ ,  $r$ , and  $\bar{X}_0$ ) prior to seeing any data. Therefore, it is possible to precompute the sending sequence and implement it as a time-based schedule on the sensor, without the need to implement a copy of the estimator (9)–(13) and the trigger (38) on the sensor. This may be preferred if the sensor is not equipped with sufficient computational resources. As is discussed in [20], the triggering sequence  $\gamma_k$  is typically periodic with VBT for a time-invariant system with stationary noise.

#### V. COMPARISON OF TRIGGER EFFECTIVENESS BASED ON NUMERICAL EXAMPLES

In this section, the established triggers (1), (2), (3), (4), and the novel triggers (36), (38) are compared in terms of their average estimation-vs-communication trade-off, based on numerical simulations.

##### A. Method of comparison

For comparing the triggers, we generate estimation-vs-communication graphs as in Fig. 2 from simulation examples. For this, the event-based estimation system given by the process (6), (7) (with specific parameters), the estimator (9)–(13) and a specific choice of trigger is simulated over a suitable simulation horizon  $K$ . Each simulation is repeated  $N_{\text{sim}}$  times, and average estimation error  $\mathcal{E}$  and average communication rate  $\mathcal{C}$  are computed as follows:

$$\mathcal{E} = \text{avg}_{i=1, \dots, N_{\text{sim}}}(\mathcal{E}_i) \quad \text{with} \quad \mathcal{E}_i = \text{avg}_{k=1, \dots, K}(e_k^2) \quad (39)$$

$$\mathcal{C} = \text{avg}_{i=1, \dots, N_{\text{sim}}}(\mathcal{C}_i) \quad \text{with} \quad \mathcal{C}_i = \text{avg}_{k=1, \dots, K}(\gamma_k) \quad (40)$$

where avg is the numerical average over the indicated data points, and  $e_k := x_k - \hat{x}_{k|k}$  is the estimation error. The pair  $(\mathcal{E}, \mathcal{C})$  quantifies the average estimation/communication performance achieved by a specific triggering rule and choice of threshold  $\delta$ . Computing  $(\mathcal{E}, \mathcal{C})$  over a suitable range of  $\delta$ 's yields the desired estimation-vs-communication graph.

In this study, we do not pay attention to the specific values of  $\delta$ . Accordingly, the concrete values of  $\delta$  do not appear in the graphs like Fig. 2. This is motivated by the role of  $\delta$  in practice: it is typically used as tuning parameter to obtain a satisfactory trade-off between estimation performance and communication (in simulations or experiments). In this sense, the actual value of  $\delta$  is irrelevant. This also means that two triggers  $t_1(y_k) \geq \delta_1$  and  $t_2(y_k) \geq \delta_2$  related by an invertible transformation  $g$ ,  $t_1(y_k) = g(t_2(y_k))$ , will yield the same estimation-vs-communication graph. This is because

$$t_1(y_k) \geq \delta_1 \Leftrightarrow g(t_1(y_k)) \geq g(\delta_1) \Leftrightarrow t_2(y_k) \geq g(\delta_1) \quad (41)$$

and by choosing  $\delta_2 := g(\delta_1)$ , the rule  $t_2(y_k) \geq \delta_2$  will trigger if, and only if,  $t_1(y_k) \geq \delta_1$  does. For example, if  $L_k$  in (36) is constant, then (36) is equivalent to (2) as far as their estimation-vs-communication graph  $(\mathcal{E}, \mathcal{C})$  is concerned.

While not of interest for the comparison of average values herein, the actual value of  $\delta$  can be relevant as a bound on certain estimation errors. For example, the SoD rule (1) guarantees that the difference of the current measurement  $y_k$  and the last known value  $y_{\text{last},k}$  never exceeds  $\delta$ . For (3), it is shown in [20, Corollary 1] that the error variance  $P_{k|k-1}$  has an upper bound given in terms of  $\delta$ .

##### B. Unstable process

As a first example, we consider an unstable process (6), (7) with parameters

$$\text{Example 1: } a = 1.01, c = q = r = 1, \bar{x}_0 = X_0 = 1.$$

Due to the unstable dynamics, some communication is required in steady-state to keep the estimation error bounded.

The estimation-vs-communication graphs obtained from  $N_{\text{sim}} = 2000$  simulations with a horizon  $K = 1000$  for the considered triggers are shown in Fig. 7. For suitably small  $\delta$ , all event-based estimators reduce to the standard KF receiving a measurement at every time step (full communication case,  $\mathcal{C} = 1$ ). Therefore, all triggering schemes yield the same estimation performance at  $\mathcal{C} = 1$ . When increasing the thresholds  $\delta$ , communication  $\mathcal{C}$  reduces and the estimation error  $\mathcal{E}$  grows, as expected. In the limit as

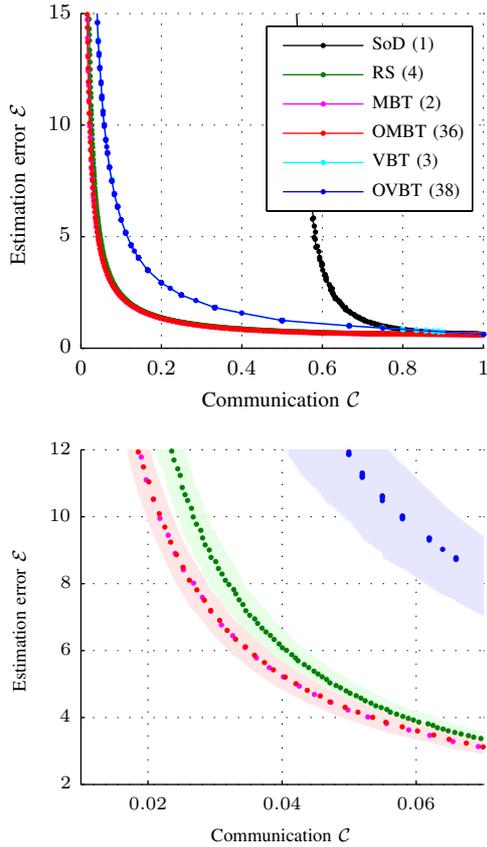


Fig. 7. Average estimation error  $\mathcal{E}$  over average communication  $\mathcal{C}$  for Example 1 and different triggering schemes. The bottom graph shows a detail of the top one, augmented with  $\pm$  one standard deviation of the estimation error for each graph (i.e. the shaded area corresponds to  $\mathcal{E} \pm \text{std}(\mathcal{E}_i)$ ). The graph for MBT (magenta) is overlapped with its optimal counterpart OMBT (red); and, similarly, VBT is hidden by OVBT.

$\mathcal{C} \rightarrow 0$ ,  $\mathcal{E} \rightarrow \infty$  because of the unstable process dynamics in this example. However, the triggering methods differ in the way that  $\mathcal{E}$  increases when reducing communication, which indicates the triggers' different effectiveness in the estimation-vs-communication trade-off.

In Fig. 7, we notice that SoD (1) shows the worst performance of all triggers. Because of the unstable process, the triggering rule is not effective after a while since  $x_k \rightarrow \infty$  implies  $|y_k - y_{k-1}| \rightarrow \infty$ , and thus  $\gamma_k = 1$  for all  $k$  after some step. All other triggers are based on some quantity of the estimation error (mean or variance), which is bounded as long as the communication rate is sufficiently high.

The optimal triggers OMBT (36) and OVBT (38) derived in Sec. IV essentially yield the same performance as the original triggers (2) and (3). This indicates a minor effect of the time-varying scaling  $L_k^2$  in the triggering rules (36) and (38). In general, the variance-based triggers (3) and (38) are worse than the measurement-based triggers (2) and (36). This reflects the analysis in Sec. III: the variance-based triggers do not exploit the information from the measurement  $y_k$ .

In this example, the MBT and OMBT are slightly more effective than RS (see bottom graph in Fig. 7). Depending on the choice of problem parameters (noise variances), we have

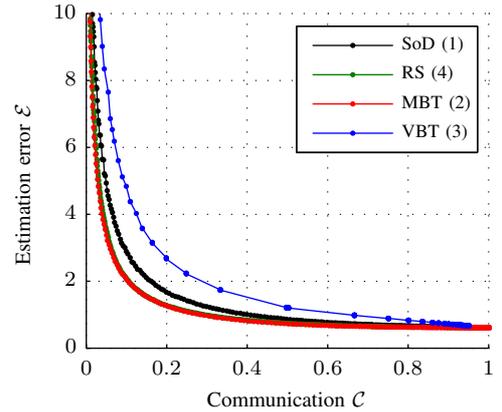


Fig. 8. Estimation-vs-communication for Example 2 and four different triggering schemes. The graphs for the optimal rules (36) and (38) are not shown since they overlap with those for the triggers (2) and (3), respectively.

observed that this difference may be more pronounced, but also cases exist where RS is slightly superior.

### C. Stable process

As a second example, we consider a stable process:

*Example 2:*  $a = 0.98$ ,  $c = q = r = 1$ ,  $\bar{x}_0 = X_0 = 1$ .

The estimation-vs-communication graphs with  $K = 200$ , and  $N_{\text{sim}} = 2000$  are depicted in Fig. 8.

For this example, SoD performs significantly better compared to the unstable Example 1. Still, MBT and RS, which are tailored to the estimation problem, are superior. As in the previous example, OMBT and OVBT are essentially identical to MBT and VBT.

## VI. DISCUSSION AND REMARKS

The evolution of the classic Kalman filter for the linear Gaussian case with periodic communication is predetermined, in the sense that variance and filter gains are independent of any real-time measurement. Therefore, one may be tempted to think that there is no benefit in using the real-time measurement in the triggering decision for event-based estimation. The analysis in Sec. III clearly shows that this is a misconception. Knowing the actual measurement allows for a more informed decision because the measurement carries some information about the actual estimation error. Thus, triggers that exploit the real-time measurement (such as MBT (2) and RS (4)) are more effective in terms of achieving a better estimation-vs-communication trade-off than those that do not (VBT (3)). If, on the other hand, it is desired to offline design a time-based transmit schedule, VBT is useful, as mentioned in Sec. IV and extensively discussed in [20].

While SoD (1) is based on the real-time measurement, its performance in the context of state estimation was found to be inferior compared to the other triggers in the simulations in Sec. V. In particular, SoD is ineffective for an unstable process. This may be expected since SoD is a general purpose trigger, which requires no assumptions on the underlying process. All other triggers are based on some quantity related to the

estimation problem (such as estimation error or variance) and are thus more tailored to the problem at hand.

Using an optimal decision framework, two new event triggers were obtained in Sec. IV. These optimal triggers (36) and (38) were found to resemble MBT and VBT, respectively. In fact, for a constant KF gain, they are equivalent in terms of their average estimation/communication performance. Even with varying gains, no significant difference was found in the scalar simulations in Sec. V. Investigating whether the same holds for the vector case is future work.

The optimal decision problem (19) discussed herein, is a one-step problem, while the actually problem of finding an optimal trade-off between estimation and communication is a dynamic optimization problem. Clearly, a triggering decision at some time step will also have an effect on the estimation performance at future steps due to the underlying dynamics. Quantifying the sub-optimality of the one-step approximation to the dynamic problem is an interesting open question.

From the analysis and simulations herein, it may be concluded that measurement-based triggers in the form of (2) or (36) are effective for event-based state estimation. Relevant sampling (4) essentially achieved the same performance, while being computationally slightly more expensive.

#### APPENDIX

Before computing  $f(w_k|y_k, \mathcal{Y}_{k-1})$  in this section, we first determine  $f(y_k, w_k|\mathcal{Y}_{k-1})$ . From (7),

$$\begin{bmatrix} y_k \\ w_k \end{bmatrix} = \begin{bmatrix} x_k + w_k \\ w_k \end{bmatrix} = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix} \begin{bmatrix} x_k \\ w_k \end{bmatrix}. \quad (42)$$

Since  $f(x_k|\mathcal{Y}_{k-1})$  and  $f(w_k|\mathcal{Y}_{k-1}) = f(w_k)$  are Gaussian and independent,  $f(x_k, w_k|\mathcal{Y}_{k-1})$  is jointly Gaussian. From (42), it follows that  $f(y_k, w_k|\mathcal{Y}_{k-1})$  is also jointly Gaussian. Thus, it suffices to compute conditional mean and variance:

$$\mathbb{E} \begin{bmatrix} y_k \\ w_k \end{bmatrix} \Big| \mathcal{Y}_{k-1} = \begin{bmatrix} \hat{x}_{k|k-1} + \mathbb{E}[w_k] \\ \mathbb{E}[w_k] \end{bmatrix} = \begin{bmatrix} \hat{x}_{k|k-1} \\ 0 \end{bmatrix} \quad (43)$$

where we used independence of  $w_k$  and  $\mathcal{Y}_{k-1}$ , and

$$\begin{aligned} \text{Var} \begin{bmatrix} y_k \\ w_k \end{bmatrix} \Big| \mathcal{Y}_{k-1} &= \mathbb{E} \begin{bmatrix} y_k - \mathbb{E}[y_k|\mathcal{Y}_{k-1}] \\ w_k - \mathbb{E}[w_k|\mathcal{Y}_{k-1}] \end{bmatrix} (\dots)^T \Big| \mathcal{Y}_{k-1} \\ &= \mathbb{E} \begin{bmatrix} (x_k - \hat{x}_{k|k-1})^2 + 2(x_k - \hat{x}_{k|k-1})w_k + w_k^2 & * \\ (x_k - \hat{x}_{k|k-1})w_k + w_k^2 & w_k^2 \end{bmatrix} \Big| \mathcal{Y}_{k-1}, \end{aligned} \quad (44)$$

where  $*$  is a placeholder for the symmetric entry. For the individual terms in (44), we have

$$\mathbb{E} [(x_k - \hat{x}_{k|k-1})^2 | \mathcal{Y}_{k-1}] = P_{k|k-1} \quad (45)$$

$$\mathbb{E} [(x_k - \hat{x}_{k|k-1})w_k | \mathcal{Y}_{k-1}] = 0 \quad (46)$$

$$\mathbb{E} [w_k^2 | \mathcal{Y}_{k-1}] = \mathbb{E} [w_k^2] = r. \quad (47)$$

In (46), we used conditional independence of  $x_k - \hat{x}_{k|k-1}$  and  $w_k$ . In summary, we obtain

$$f(y_k, w_k|\mathcal{Y}_{k-1}) = \mathcal{N} \left( \begin{bmatrix} y_k \\ w_k \end{bmatrix}; \begin{bmatrix} \hat{x}_{k|k-1} \\ 0 \end{bmatrix}, \begin{bmatrix} P_{k|k-1} + r & r \\ r & r \end{bmatrix} \right).$$

It follows from this (cf. [24, p. 39]) that  $f(w_k|y_k, \mathcal{Y}_{k-1})$  is a GRV with mean and variance as given in (25).

#### REFERENCES

- [1] V. Vasyutynskyy and K. Kabitzsch, "Towards comparison of deadband sampling types," in *IEEE International Symposium on Industrial Electronics*, June 2007, pp. 2899–2904.
- [2] V. Vasyutynskyy and J. Plönnigs, "Survey within the project 'Agent-based Middleware for E-Services' in the user group 'Building Automation'," TU Dresden, Dec. 2005. [Online]. Available: <http://iis807.inf.tu-dresden.de/~ames>
- [3] J. Sandee, W. Heemels, S. Hulsenboom, and P. van den Bosch, "Analysis and experimental validation of a sensor-based event-driven controller," in *Proc. of the American Control Conference*, Jul. 2007, pp. 2867–2874.
- [4] J. Sánchez, M. Á. Guarnes, and S. Dormido, "On the application of different event-based sampling strategies to the control of a simple industrial process," *Sensors*, vol. 9, no. 9, pp. 6795–6818, 2009.
- [5] D. Lehmann and J. Lunze, "Extension and experimental evaluation of an event-based state-feedback approach," *Control Engineering Practice*, vol. 19, no. 2, pp. 101–112, 2011.
- [6] J. Araújo, M. Mazo Jr, A. Anta, P. Tabuada, and K. H. Johansson, "System architectures, protocols and algorithms for aperiodic wireless control systems," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 175–184, Feb 2014.
- [7] M. Sigurani, C. Stöcker, L. Grüne, and J. Lunze, "Experimental evaluation of two complementary decentralized event-based control methods," *Control Engineering Practice*, vol. 35, pp. 22–34, 2015.
- [8] F. Altaf, J. Araújo, A. Hernandez, H. Sandberg, and K. H. Johansson, "Wireless event-triggered controller for a 3D tower crane lab process," in *Proc. 19th Med. Conf. on Control & Autom.*, 2011, pp. 994–1001.
- [9] S. Trimpe and R. D'Andrea, "An experimental demonstration of a distributed and event-based state estimation algorithm," in *Proc. of the 18th IFAC World Congress*, Aug. 2011, pp. 8811–8818.
- [10] S. Trimpe, "Event-based state estimation with switching static-gain observers," in *Proc. of the 3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, Sep. 2012, pp. 91–96.
- [11] S. Trimpe and R. D'Andrea, "Reduced communication state estimation for control of an unstable networked control system," in *Proc. of the 50th IEEE Conf. on Decision and Control*, Dec. 2011, pp. 2361–2368.
- [12] T. Henningsson, "Event-based control and estimation with stochastic disturbances," Ph.D. dissertation, Lund Univ., Nov. 2008.
- [13] T. Henningsson and K. Åström, "Log-concave observers," in *Proc. of the 17th Int. Symp. on Math. Theory of Netw. and Systems*, Jul. 2006.
- [14] J. Sijs and M. Lazar, "On event based state estimation," in *Hybrid Systems: Computation and Control*, ser. Lecture Notes in Computer Science. Springer, 2009, vol. 5469, pp. 336–350.
- [15] J. Sijs, B. Noack, and U. Hanebeck, "Event-based state estimation with negative information," in *Proc. of the 16th International Conference on Information Fusion*, Jul. 2013, pp. 2192–2199.
- [16] M. Miskowicz, "Send-on-delta concept: an event-based data reporting strategy," *Sensors*, vol. 6, no. 1, pp. 49–63, 2006.
- [17] P. Otanez, J. Moyne, and D. Tilbury, "Using deadbands to reduce communication in networked control systems," in *Proc. of the American Control Conference*, 2002, pp. 3015–3020.
- [18] J. Sijs, L. Kester, and B. Noack, "A study on event triggering criteria for estimation," in *Proc. of the 17th International Conference on Information Fusion*, Jul. 2014, pp. 1–8.
- [19] Y. Suh, "Send-on-delta sensor data transmission with a linear predictor," *Sensors*, vol. 7, no. 4, pp. 537–547, 2007.
- [20] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," *IEEE Transaction on Automatic Control*, vol. 59, no. 12, pp. 3266–3281, Dec. 2014.
- [21] J. Marck and J. Sijs, "Relevant sampling applied to event-based state-estimation," in *4th International Conference on Sensor Technologies and Applications*, Jul. 2010, pp. 618–624.
- [22] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 2006.
- [23] M. Miskowicz, "Event-based sampling strategies in networked control systems," in *10th IEEE Workshop on Factory Communication Systems*, May 2014, pp. 1–10.
- [24] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Mineola, New York: Dover Publications, 2005.