An Iterative Algorithm for Optimal Event-Triggered Estimation ⋆

Adam Molin * **Sandra Hirche** *

* *Institute of Automatic Control Engineering, Technische Universität München, D-80290 München, Germany (e-mail: adam.molin@tum.de, hirche@tum.de).*

Abstract: This paper investigates the optimal design of event-triggered estimation for first-order linear stochastic systems. The problem is posed as a two-player team problem with a partially nested information pattern. The two players are given by an estimator and an event-trigger. The event-trigger has full state information and decides whether the estimator shall obtain the current state information by transmitting it through a resource constrained channel. The objective is to find an optimal trade-off between the mean squared estimation error and the expected transmission rate. The proposed iterative algorithm alternates between optimizing one player while fixing the other player. It is shown that the solution of the algorithm converges to a linear predictor and a symmetric threshold policy, if the densities of the initial state and the noise variables are even and radially decreasing functions. The effectiveness of the approach is illustrated via numerical simulations. In case of a multimodal distribution of the noise variables a significant performance improvement can be achieved compared to a separate design that assumes a linear prediction and a symmetric threshold policy.

1. INTRODUCTION

In contrast to periodic estimation, where measurements are sampled within equidistant time-intervals, an event-triggered estimator receives measurement updates in an asynchronous fashion. Event-triggered sampling is also referred to as adaptive sampling in Rabi et al. [2012], Lebesgue sampling in Åström and Bernhardsson [2002] and dead-band control in Otanez et al. [2002], Hirche et al. [2005]. The event-trigger is a preprocessing unit situated at the sensor which decides upon its available information, whether to update the estimator with current information. Event-triggered sampling schemes for estimation are very promising in the context of networked control systems, where estimator and plant are spatially distributed and communication is a sparse resource. Examples for such networked control systems are given by sensor networks, multi-robot systems and distributed power generation networks. The work in Åström and Bernhardsson [2002] and Rabi et al. [2012] showed that event-triggered sampling outperforms periodic sampling with respect to the state estimation error of a first-order linear system in the presence of two different communication constraints. In Rabi et al. [2012], the communication constraint is induced by limiting the number of transmissions during a finite interval, whereas the work in Åström and Bernhardsson [2002] limits the average transmission rate. Differing to these approaches, we extend the standard minimum mean square estimator problem by an additional communication penalty to reflect the communication constraint in the optimization problem. A similar problem is also studied in Xu and Hespanha [2004] and Lipsa and Martins [2011].

⋆ This work was supported in part by the German Research Foundation (DFG) within the Priority Program SPP 1305 "Control Theory of Digitally Networked Dynamical Systems".

Opposed to the aforementioned work which either fixes the estimator, such as Åström and Bernhardsson [2002], Rabi et al. [2012], Xu and Hespanha [2004] or computes the estimator from the choice of the event-trigger, such as Lipsa and Martins [2011], we aim at the joint optimal design of the estimator and the event-trigger. Built on previous work that identifies structural properties of the optimal estimator Molin and Hirche [2010a,b], we formulate a two-player team problem with a nested information pattern, where the players are given by the event-trigger and the estimator. The joint design is motivated by the fact that the choice of the event-trigger may significantly influence the form of the optimal estimator.

The contribution of this paper is two-fold. First, it develops an iterative method for the joint design of event-trigger and estimator for first-order stochastic systems with arbitrary distributions. The algorithm iteratively alternates between optimizing one player while fixing the other player. Similar iterative procedures are shown to be very promising methods for calculating optimal policies for team problems with non-classical information patterns, as studied by Karlsson et al. [2011] for the Witsenhausen counterexample or by Hajek et al. [2008] for the joint optimization of paging and registration policies. It turns out that the proposed iterative method can yield a remarkable decrease of the overall cost compared to a design where the estimator is designed separately of the event-trigger. In such separate design, the optimal estimator takes the form of a linear predictor that assumes that transmission instants are statistically independent of the state, whereas the optimal event-trigger is an even threshold function of the estimation error. In the following, even and symmetric refer to the same meaning.
Second, it is shown that the solution of the algorithm converges to the separate design, when the densities of the

initial state and the noise variables are symmetric and unimodal. This result coincides with results obtained in Lipsa and Martins [2011], which uses majorization theory and rearrangement inequalities to show that there always exists a symmetric threshold policy that outperforms an arbitrary event-triggering law. In fact, we show that symmetric threshold policies are optimal by analyzing the asymptotic behavior of the proposed iterative procedure. Therefore, our approach can be viewed as an alternative line of proof to show that symmetric policies are optimal under the aforementioned assumptions. On the other hand, it turns out that symmetry of the densities is not sufficient to show that the separate design is optimal. In fact, numerical simulations indicate significant improvements of our approach compared to an independent design, when noise densities are symmetric but multimodal.

The remainder of this paper is organized into four sections. In Section 2, we introduce the stochastic system model and describe the problem setting. Section 3 contains the main results of this paper and studies the joint design of event-trigger and estimator. In Section 4, numerical simulations are conducted to validate the proposed method.

Notation. In this paper, $(\cdot)$ denotes a free parameter. The expectation operator is denoted by $\mathsf{E}_f[\cdot]$ and the conditional expectation is denoted by $\mathsf{E}_f[\cdot|\cdot]$, where the underlying probability measure P_f is parameterized by the policy f. The variable X^k denotes the sequence of variables $[x_0, \dots, x_k]$. The indicator function is denoted by $\mathbb{1}_A(x)$ taking a value of 1 if $x \in A$ and 0 otherwise. The complement of a set A is denoted by A^c. The convolution of two real-valued function f and g is denoted by $f * g$.

2. PROBLEM FORMULATION

We consider the following stochastic scalar discrete-time process $\mathcal{P}$ driven by noise w_k

$$x_{k+1} = ax_k + w_k, \tag{1}$$

where $a \in \mathbb{R}\backslash\{0\}$. The system noise process w_k takes values in $\mathbb{R}$ and is an i.i.d. (independent identically distributed) random process described by the probability density function ϕ_w, which is zero-mean and has finite variance. The initial state, x_0 is statistically independent of w_k and is described by density function ϕ_{x_0}, which has a mean $\bar{x}_0$ and a finite variance. System parameters and statistics are known to the event-trigger and estimator.

The system model is illustrated in Fig. 1. The process $\mathcal{P}$ outputs the state x_k. The event-trigger $\mathcal{E}$ decides upon its available information whether or not to transmit the current state to the remote state estimator $\mathcal{S}$. We define the output of the event-trigger as

$$\delta_k = \begin{cases} 1 & \text{update } x_k \text{ sent} \\ 0 & \text{otherwise} \end{cases}$$

The channel $\mathcal{N}$ can be viewed as a δ_k-controlled erasure channel whose outputs are described by

$$z_k = \begin{cases} x_k & \delta_k = 1 \\ \varnothing & \delta_k = 0 \end{cases} \tag{2}$$

where $\varnothing$ is the erasure symbol. As it will be useful for subsequent analysis, we define the last update time τ_k as

$$\tau_k = \max\{\kappa | \delta_\kappa = 1,\ \kappa < k\} \tag{3}$$

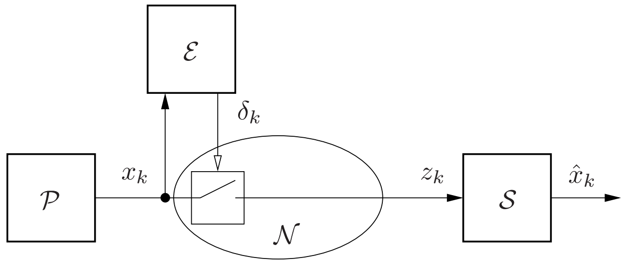

Fig. 1. System model of the networked control system with plant $\mathcal{P}$, event-trigger $\mathcal{E}$, state estimator $\mathcal{S}$, and communication channel $\mathcal{N}$.

with $\tau_k = -1$, if no transmissions have occurred prior to k. The variable τ_k can be described by the following δ_k-controlled difference equation

$$\tau_{k+1} = \begin{cases} k & \delta_k = 1 \\ \tau_k & \delta_k = 0 \end{cases} \quad \tau_0 = -1. \tag{4}$$

Admissible event-triggers are given by mappings of their past history to

$$\delta_k = f_k(X^k), \quad k = 0, \dots, N-1.$$

The state estimator $\mathcal{S}$ outputs the state estimate $\hat{x}_k$ and is given by mappings g_k defined by

$$\hat{x}_k = g_k(Z^k), \quad k = 0, \dots, N-1.$$

The design objective is to jointly design the event-trigger $f = [f_0, \dots, f_{N-1}]$ and the estimator $g = [g_0, \dots, g_{N-1}]$ that minimize cost J.

$$J = \mathsf{E}_{f,g}\left[\sum_{k=0}^{N-1} |x_k - \hat{x}_k|^2 + \lambda\delta_k\right]. \tag{5}$$

The per-stage cost of J is composed of the squared estimation error $|x_k - \hat{x}_k|_2^2$ and a communication penalty $\lambda\delta_k$. The weight λ determines the amount of penalizing transmissions over the channel $\mathcal{N}$.

3. JOINT DESIGN OF EVENT-TRIGGER AND ESTIMATOR

3.1 Preliminaries

We begin with a characterization of the optimal estimator.

Lemma 1. For any given event-trigger f, the optimal state estimator g^* is given by the least squares estimator

$$\hat{x}_k = g_k^*(Z^k) = \mathsf{E}_f[x_k|Z^k], \quad k = 0, \dots, N-1.$$

Proof. Fix an arbitrary event-trigger f. The communication penalty term $\mathsf{E}_f\left[\sum_{k=0}^{N-1}\lambda\delta_k\right]$ is then constant and can be omitted from the optimization. In the remaining estimation problem the mean squared error $\mathsf{E}_f\left[\sum_{k=0}^{N-1}|x_k - \hat{x}_k|^2\right]$ is to be minimized. The optimal solution for this problem is given by the least squares estimator $\mathsf{E}_f[x_k|Z^k]$, Bertsekas [2007]. This completes the proof.

In the following, we define the linear predictor $\hat{x}_k^{\mathrm{LP}}$ by the following recursion

$$\hat{x}_{k+1}^{\mathrm{LP}} = \begin{cases} x_{k+1} & \delta_k = 1 \\ a\hat{x}_k^{\mathrm{LP}} & \delta_k = 0 \end{cases} \tag{6}$$

with $\hat{x}_0^{\mathrm{LP}} = \bar{x}_0$.

Remark 1. The linear predictor can be regarded as the optimal estimator, when having no information about the choice of the event-trigger f and assuming that transmission instances are statistically independent of the state evolution. This also implies that the linear predictor is optimal in the case, when transmission instances are selected in advance.

Similar to Lipsa and Martins [2011], Molin and Hirche [2009], let us rewrite the optimization problem by defining

$$e_k = x_k - a\hat{x}_{k-1}^{\mathrm{LP}}, \quad k = 1, \dots, N-1$$

and $e_0 = w_{-1}$, where we define $w_{-1} = x_0 - \bar{x}_0$. The variable e_k defines our new state to be estimated and follows the recursion

$$e_{k+1} = h_k(e_k, \delta_k, w_k) = (1-\delta_k)ae_k + w_k. \tag{7}$$

Further, we define $\hat{e}_k$ to be the least squares estimate $\mathsf{E}[e_k|\tilde{Z}^k]$, where $\tilde{z}_k$ is defined accordingly as

$$\tilde{z}_k = \begin{cases} e_k & \delta_k = 1 \\ \varnothing & \delta_k = 0 \end{cases}$$

The next lemma gives us further insights into the structure of $\hat{e}_k$.

Lemma 2. Let the event-trigger f be fixed. Then, the least squares estimate of e_k is given by

$$\hat{e}_k = \begin{cases} e_k & \delta_k = 1 \\ \alpha_k(\tau_k) & \delta_k = 0 \end{cases}$$

where τ_k is defined by (3) and $\alpha_k(\tau_k)$ is defined by

$$\alpha_k(\tau_k) = \mathsf{E}_f\left[\sum_{l=\tau_k}^{k-1} a^{k-l-1}w_l|\delta_{\tau_k+1} = 0, \dots, \delta_k = 0\right]. \tag{8}$$

Proof. Clearly, we have $\hat{e}_k = e_k$ for $\delta_k = 1$, as $e_k \in \tilde{Z}^k$. For $\delta_k = 0$, τ_k is a sufficient statistics for $\hat{e}_k$. The mapping α_k is determined by applying recursively (7) with $e_{\tau_k+1} = w_{\tau_k}$. This completes the proof. □

The function α in Lemma 2 can be interpreted as a bias term to improve the state estimate by incorporating additional information $\delta_{\tau_k+1} = \cdots = \delta_k = 0$ at time k.

Rather than regarding α as a function of k and τ_k, we will interpret α as a vector in $\mathbb{R}^{\frac{1}{2}N(N+1)}$ by reindexing its entries appropriately.

It is straightforward to see that the estimation error $e_k - \hat{e}_k$ and $x_k - \hat{x}_k$ are identical random variables for a fixed event-trigger f, as e_k corresponds to a translatory coordinate transformation of x_k shifted by $-a\hat{x}_{k-1}^{\mathrm{LP}}$ which is known since the sequence $[\delta_0, \dots, \delta_{k-1}$ is measurable with respect to Z^k. Therefore, our initial optimization problem with cost function J can be rewritten as

$$\min_f \mathsf{E}_f\left[\sum_{k=0}^{N-1}(1-\delta_k)|e_k - \alpha_k(\tau_k)|^2 + \lambda\delta_k\right]. \tag{9}$$

It can be observed that the running cost reduces to λ and is therefore independent of the current α_k in the case $\delta_k = 1$. Because of the introduction of the state e_k, the event-trigger f is given by a mapping from E^k to $\{0,1\}$. Since there always exists a bijection from X^k to E^k given the variables $\delta_0, \dots, \delta_{k-1}$, this change of variables does not put any restrictions on the further analysis keeping in mind that any policy expressed in E^k can also be written as a function in X^k.

3.2 Iterative procedure

What prevents a further study of the optimization problem (9) is the fact that the value α_k at τ_k depends on the particular policy f chosen up to time k. Therefore, methods like dynamic programming are not directly applicable to solve (9). In order to overcome this burden, we relax optimization problem (9) by considering the variable α_k as a new decision variable being a function of τ_k. Then, the optimization problem is given by

$$\min_{f,\alpha} J \tag{10}$$

with

$$J(f,\alpha) = \mathsf{E}_f\left[\sum_{k=0}^{N-1}(1-\delta_k)|e_k - \alpha_k(\tau_k)|^2 + \lambda\delta_k\right]. \tag{11}$$

The optimization problem (10) enlarges the set of possible solutions compared to optimization problem (9), because it omits the constraint for α given by (8). By considering optimization problem (10), we are able to specify the structure of the optimal event-trigger, which is given by the following lemma.

Lemma 3. Let α be fixed. Then, for all $k \in \{0, \dots, N-1\}$ the variables e_k and τ_k are a sufficient statistics for the optimal event-trigger f_k .

Proof. The evolution of the pair (e_k, τ_k) can be regarded as a δ_k-controlled Markov process defined by (4) and (7). The running cost of J at time k is a function of the pair (e_k, τ_k), input δ_k and noise w_k. By Bertsekas [2007], this problem can be solved by dynamic programming with (e_k, τ_k) being the state, which is a sufficient statistics of the optimal solution f_k. This completes the proof. □

Lemma 3 implies that the optimal event-trigger is a function of e_k and τ_k. It can be observed that for a fixed event-trigger f, the optimal map α can be calculated by (8). On the other hand, for any fixed map α, the optimal event-trigger f can be calculated by dynamic programming. We therefore define the running cost and the Bellman operator as follows

$$c_k^{\alpha_k}(e_k, \tau_k, \delta_k) = (1-\delta_k)|e_k - \alpha_k(\tau_k)|^2 + \lambda\delta_k$$

$$\mathcal{T}_k^{\alpha_k} J_{k+1}(\cdot) = \min_{\delta_k \in \{0,1\}} c_k^{\alpha_k}(\cdot, \delta_k) + \mathsf{E}\left[J_{k+1}(e_{k+1}, \tau_{k+1})|\cdot, \delta_k\right]$$

The value function J_k being a function of the augmented state (e_k, τ_k) is determined by recursive application of the Bellman equation given by

$$J_k = \mathcal{T}_k^{\alpha_k} J_{k+1}$$

with $J_N \equiv 0$, where the argument in the minimization yields the optimal event-trigger f and we have

$$J(f,\alpha) = \mathsf{E}_f[J_0(e_0, -1)].$$

This observation motivates us to propose the following iterative procedure sketched in Fig. 2, which alternates between optimizing f while fixing policy α and vice versa.

Algorithm 1 describes the iterative procedure. With slight abuse of notation, we declared τ_k as a second subscript instead of an argument of α_k.

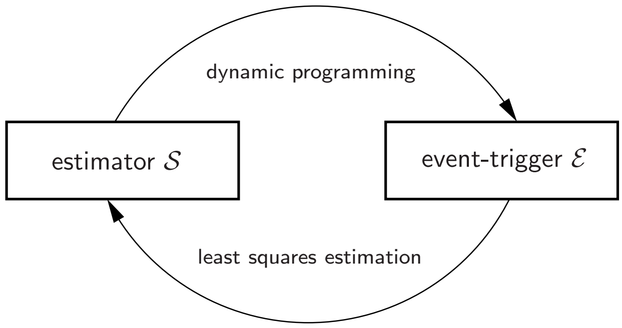

Fig. 2. Iterative scheme to calculate event-trigger $\mathcal{E}$ and estimator $\mathcal{S}$.

Algorithm 1 Iterative procedure to calculate (f, α)

Require: $\alpha^0_{k,\tau_k} \in \mathbb{R}, \quad k = 0, \ldots, N-1, \tau_k = -1, \ldots, k-1$
$i \leftarrow 0$

repeat
  $k = N,\ J_N \equiv 0$
  repeat
    $k \leftarrow k - 1$
    $J_k \leftarrow \mathcal{T}_k^{\alpha_k^i} J_{k+1}$
    $f_k^i(e_k, \tau_k) \in \operatorname{argmin}_{\delta_k \in 0,1} c_k^{\alpha_k^i}(e_k, \tau_k, \delta_k) + \mathsf{E}\left[J_{k+1}(e_{k+1}, \tau_{k+1}) | e_k, \tau_k, \delta_k\right]$
  until $k = 0$
  $\alpha^{i+1}_{k,\tau_k} \leftarrow \mathsf{E}_{f^i}\left[\sum_{l=\tau_k}^{k-1} a^{k-l-1} w_l | \delta_{\tau_k+1} = 0, \ldots, \delta_k = 0\right]$
  $i \leftarrow i + 1$
until convergence

As the cost J decreases or is at least kept constant in each step of the iteration, the sequence $[(f^0, \alpha^0), (f^1, \alpha^1), \ldots]$ produces a non-increasing succession of costs J.

In the following subsection, we are interested in the convergence properties of the proposed iterative algorithm for symmetric unimodal distributions.

3.3 Symmetric unimodal distributions

In the following, we consider the iterative procedure described in previous subsection as a discrete-time dynamical system and consider α as the state. By using Lyapunov stability theory we show that $\alpha \equiv 0$ is a globally asymptotically stable equilibrium point when the initial state e_0 and the noise process $\{w_k\}$ have symmetric unimodal density functions. The next lemma finds a potential equilibrium point only by assuming symmetric distributions.

Lemma 4. Let the initial state e_0 and the noise process $\{w_k\}$ have symmetric distributions. Then $\alpha^* \equiv 0$ is a fixpoint of the Algorithm 1. The policy of the event-trigger f^* that corresponds to α^* is an even mapping of e_k and independent of τ_k for $k = 0, \ldots, N-1$.

Proof. Let us choose the map α^0 to be 0 for all k and all τ in the initialization of Algorithm 1. The cost function J reduces then to

$$J(f, \alpha^0) = \mathsf{E}_f\left[\sum_{k=0}^{N-1} (1-\delta_k)|e_k|^2 + \lambda\delta_k\right]$$

where e_k evolves by recursion (7). Therefore, the resulting optimal f_k^0 is only a function of e_k for all $k = 0, \ldots, N-1$. In the following, we first show that the application of the Bellman operator $\mathcal{T}_k^0$ preserves symmetry of the value function J_{k+1} for any k. Given an even value function J_{k+1}, the conditional expectation $\mathsf{E}\left[J_{k+1}(e_{k+1}, \tau_{k+1})|\cdot, \delta_k\right]$ preserves symmetry for both $\delta_k = 0$ and $\delta_k = 1$. Adding the cost $c_k^0(\cdot, \delta_k)$ also preserves symmetry, because the sum of two even functions is again even. Taking the pointwise minimum of two even functions yields an even function. Therefore, an even function remains even after application of the Bellman operator. As $J_N \equiv 0$ is an even function, it follows by induction that every value function J_k is even for $k \in \{0, \ldots, N-1\}$. This implies that the f_k^0 resulting in the first iteration step from Algorithm 1 is an even mapping of e_k, if $\alpha^0 \equiv 0$.
Next, we calculate α^1 assuming f_k^0 being an even function of e_k for $k \in \{0, \ldots, N-1\}$. Let $\phi_{e_k|\tau}$ be defined as the density function of the conditional probability distribution of e_k given τ_k and $\delta_k = 0$, when using event-trigger f^0. The definition of $\phi_{e_k|\tau}$ yields the following calculation of $\alpha^1_{k,\tau}$

$$\alpha^1_{k,\tau} = \int_{e\in\mathbb{R}} e \cdot \phi_{e_k|\tau}(e) de$$

For $k = 0$, $\phi_{e_k|\tau}$ is determined by truncating the density function ϕ_{e_0} of the initial state e_0 at all (e, τ), where f_0^1 takes a value of 1 and by normalizing the resulting function, i.e.

$$\phi_{e_0|\tau}(e) = \frac{\phi_{e_0}(e) \cdot (1 - f_0^0(e, \tau))}{\int_{e\in\mathbb{R}} \phi_{e_0}(e) \cdot (1 - f_0^0(e, \tau)) de}. \tag{12}$$

Since ϕ_{e_0} and f_0^0 are even functions, we conclude that $\phi_{e_0|\tau}$ is even and therefore we have $\alpha^1_{0,-1} = 0$. Along the same lines, we can show that $\phi_{e_k|k-1}$ is even and $\alpha^1_{k,k-1} = 0$ for $k \in \{1, \ldots, N-1\}$ by replacing ϕ_{e_0} with ϕ_w in (12) . For a constant τ, the conditional density function $\phi_{e_k|\tau}$ evolves by the recursion

$$\phi_{e_{k+1}|\tau}(e) = \frac{(\frac{1}{|a|}\phi_{e_k|\tau}(\frac{(\cdot)}{a}) * \phi_w)(e) \cdot (1 - f_k^0(e, \tau))}{\int_{e\in\mathbb{R}} (\frac{1}{|a|}\phi_{e_k|\tau}(\frac{(\cdot)}{a}) * \phi_w)(e) \cdot (1 - f_k^0(e, \tau)) de}.$$

It can be observed that this recursion preserves symmetry of the conditional density function $\phi_{e_k|\tau}$, as f_k^0 is an even function. Therefore, we have shown that $\alpha^* \equiv 0$ is a fixpoint of Algorithm 1, which completes the proof. □

In above lemma, the distributions need not to be unimodal, but only symmetry properties are required. A natural question arising from Lemma 4 is whether the fixpoint at 0 is a stable and unique fixpoint. This question is partly answered in the following Theorem by adding the assumption that the distributions are unimodal.

Theorem 5. Let the initial state e_0 and the noise process $\{w_k\}$ have symmetric and unimodal distributions. Then, $\alpha^* \equiv 0$ is a globally asymptotically stable fixpoint of Algorithm 1.

Proof. By considering the bias α^i as a state variable that evolves according Algorithm 1 with increasing iteration index i, and using the supremum norm of α^i as a Lyapunov candidate, it can be shown that the resulting system is globally asymptotically stable with a unique equilibrium at $\alpha^* \equiv 0$. A detailed proof can be found in Molin and Hirche [2012].

Remark 2. As the iterative Algorithm 1 produces a sequence of pairs (f^i, α^i) whose costs are non-increasing with increasing i, we conclude that 0 is the optimal choice for α, when noise distributions are symmetric and unimodal according to Theorem 5. The optimal state estimator of x_k is then given by the linear predictor in (6) and is therefore independent of the choice of the event-trigger f. The distribution of the initial state, x_0, must be also symmetric and unimodal, but its mean $\bar{x}_0$ can be chosen arbitrarily. Hence, the symmetry axis of the distribution of x_0 need not to be at zero. In order to determine the optimal f^*, dynamic programming must only be applied once with $\alpha \equiv 0$. Therefore, the joint design approach in the case of symmetric densities can be considered as an independent design of event-trigger and estimator.

Remark 3. The result is in accordance with Lipsa and Martins [2011] and constitutes an alternative way by analyzing the asymptotic behavior of Algorithm 1 to prove that symmetric event-triggering laws are optimal in the presence of symmetric unimodal distributions. Moreover, the iterative algorithm may be applied to arbitrary distributions. Although $\alpha \equiv 0$ is a fixpoint of the Algorithm 1 by Lemma 4 assuming symmetric density functions, the next section shows that an independent approach given by $\alpha \equiv 0$ can be outperformed by Algorithm 1 by almost 50%. Hence, we can conclude that symmetry of the densities is not sufficient to show that the independent design is optimal. Therefore, additional assumptions are required to show that the independent design is optimal. In the case of Theorem 5 such requirement is given by the unimodality assumption of the density functions.

4. NUMERICAL ILLUSTRATION

This section intends to outline the benefits of the proposed iterative algorithm by numerical examples. These also show that $\alpha \equiv 0$ is optimal for unimodal symmetric noise distributions. We compare the iterative algorithm with the optimal symmetric event-trigger having a linear predictor, i.e. assuming $\alpha \equiv 0$. Suppose the system process to be defined by (1) with $a = 1$, a communication penalty $\lambda = 0.5$, and the distribution of the initial state and the system noise to be identical and defined by the density function ϕ_w

$$\phi_w(\mu,\sigma) = \frac{1}{2}\phi_{\mathcal{N}}(\mu,\sigma) + \frac{1}{2}\phi_{\mathcal{N}}(-\mu,\sigma)$$

with

$$\phi_{\mathcal{N}}(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

In the special case of $\mu = 0$, we retrieve the normal distribution. In order to facilitate comparability between different distributions, we choose $\mu \in [0, 1)$ and set

$$\sigma = \sqrt{1-\mu^2}$$

that yields an identical variance of 1 for all $\mu \in [0, 1)$. In the limit $\mu \to 1$, the noise process degrades to a Bernoulli process taking discrete values $\{-1, 1\}$ with probability $\frac{1}{2}$. Various density functions for different μ are sketched in Fig. 3.

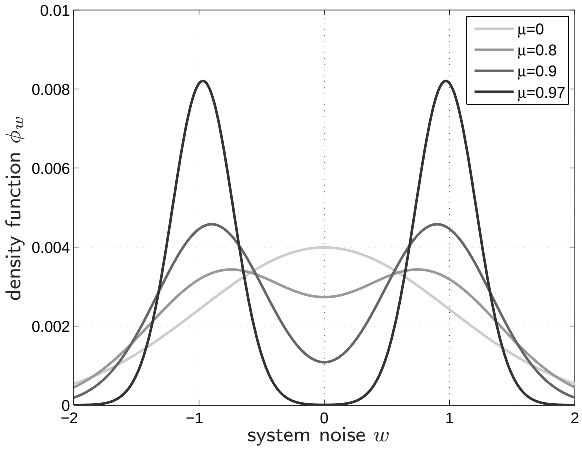

Fig. 3. Various bimodal/unimodal density functions with zero-mean and identical variance of 1 composed of two Gaussian kernels shifted by $\pm\mu$.

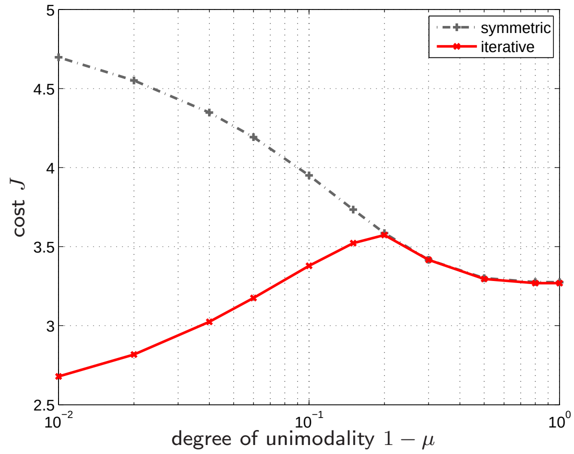

Fig. 4. Performance comparison for a horizon $N = 10$. The degree of unimodality $1-\mu$ (1 for zero-mean Gaussian and 0 for Bernoulli process with discrete parameters in {-1,1}) is drawn on a logarithmic scale.

We observe that for $\mu < 0.8$ the peaks of the bimodal density function are less distinctive. Therefore, we can not expect that large gains of the iterative procedure can be attained compared with the optimal symmetric solution for $\mu < 0.8$. A performance comparison of the iterative procedure and the optimal symmetric event-trigger is drawn in Fig. 4 for a horizon $N = 10$ and various μ. The initialization for the iterative procedure is chosen to be $\alpha^0 \equiv 0.1$. As expected the costs are almost identical for $\mu \in [0, 0.8]$. This also validates Theorem 5, since ϕ_w is unimodal for sufficient small choice of μ. For $\mu > 0.8$ a rapid performance improvement can be observed. In the limit $\mu \to 1$, the costs are reduced by a factor of 45% by the iterative procedure compared with the optimal symmetric event-trigger. This may seem surprising, because the cost function as well as the noise distribution are all even functions. Fig. 5 gives an illustrative explanation of such significant performance improvement for $N = 1$ and $\mu = 0.95$. With

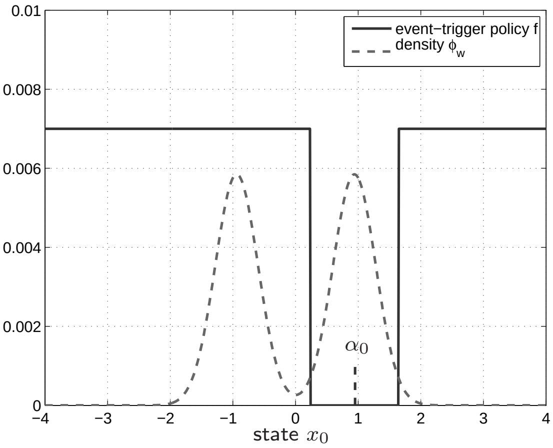

Fig. 5. Event-trigger policy f (scaled by 0.007) resulting from the iterative Algorithm 1 with initial noise distribution ϕ_w, $\mu = 0.95$, horizon $N = 1$ and initial choice $\alpha_0^0 = 0.1$. The algorithm converges to $\alpha_0 = 0.95$ and an asymmetric event-trigger $f(x_0) = \mathbb{1}_{\{[0.25,0.65]\}^c}(x_0)$.

an initial value $\alpha_0^0 = 0.1$, the iterative algorithm converges to $\alpha_0 = 0.95$ and an asymmetric event-trigger policy $f(x_0) = \mathbb{1}_{\{[0.25,1.65]\}^c}(x_0)$, whereas the optimal symmetric event-trigger is given by $f(x_0) = \mathbb{1}_{\{[-0.7,0.7]\}^c}(x_0)$. By asymmetric, we mean that the triggering thresholds are not evenly spaced from the origin. The event-trigger and estimator resulting from the iterative procedure have therefore an implicit agreement, if no state update is sent over the resource-constrained channel. In that case, no transmission indicates the estimator that the state x_0 is situated at the right peak resulting in the estimate α_0. In contrast to that, the linear predictor defined in (6), which is optimal for the symmetric event-trigger, is independent to the choice of the threshold of the symmetric event-trigger and the noise-distribution.

5. CONCLUSIONS

By considering the joint optimal design of state estimator and event-trigger as a two-player problem, an efficient iterative algorithm is developed, which alternates between optimizing the estimator while fixing the event-trigger and vice versa. The iterative method reveals certain structural characteristics of the optimal event-triggered estimator in the case of unimodal and symmetric distributions of the uncertainty. In this situation it is shown that the optimal event-triggered estimator can be obtained by a separate design and is given by a linear predictor and a symmetric threshold policy. This result is along previous results and offers an alternative line of proof for showing that such separate design is optimal in case of symmetric unimodal distributions.

In the case of symmetric and bimodal distributions, the iterative procedure offers a systematic method, which leads surprisingly to asymmetric event-triggers and biased estimators that outperform symmetric threshold policies. Similar properties of the iterative method are likely to hold as well in the case of multidimensional systems and are a subject of current investigations. Further research also investigates to extend the proposed iterative procedure to a sensor network setting, where various spatially distributed sensors shall find a common state estimate through exchanging information through a common digital network.

REFERENCES

K.J. Åström and B.M. Bernhardsson. Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, pages 2011–2016, Las Vegas, Nevada, 2002.

D. P. Bertsekas. *Dynamic programming and optimal control. Vol. I.* 3rd ed., Athena Scientific, Belmont, MA, 2007.

B. Hajek, K. Mitzel, and S. Yang. Paging and registration in cellular networks: Jointly optimal policies and an iterative algorithm. *Information Theory, IEEE Transactions on*, 54(2):608–622, 2008.

S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss. Network traffic reduction in haptic telepresence systems by deadband control. In *Proceedings IFAC World Congress, International Federation of Automatic Control*, 2005.

J. Karlsson, A. Gattami, T. J. Oechtering, and M. Skoglund. Iterative source-channel coding approach to Witsenhausen's counterexample. In *American Control Conference (ACC'11)*, pages 5348 –5353, 2011.

G.M. Lipsa and N.C. Martins. Remote state estimation with communication costs for first-order LTI systems. *Automatic Control, IEEE Transactions on*, 56(9):2013–2025, 2011.

A. Molin and S. Hirche. On LQG joint optimal scheduling and control under communication constraints. In *Decision and Control, 2009, Proceedings of the 48th IEEE Conference on*, pages 5832–5838, Shanghai, China, 2009.

A. Molin and S. Hirche. Structural characterization of optimal event-based controllers for linear stochastic systems. In *Decision and Control, 2010, Proceedings of the 49th IEEE Conference on*, pages 3227–3233, Atlanta, USA, December 2010a.

A. Molin and S. Hirche. Optimal event-triggered control under costly observations. In *Proc. 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS'10)*, Budapest, Hungary, 2010b.

A. Molin and S. Hirche. Event-triggered estimation of linear systems: An iterative algorithm and optimality properties. *Arxiv preprint arXiv:1203.4980*, 2012.

P.G. Otanez, J.R. Moyne, and D.M. Tilbury. Using deadbands to reduce communication in networked control systems. In *American Control Conference, 2002. Proceedings of the 2002*, volume 4, pages 3015–3020, 2002.

Maben Rabi, George V. Moustakides, and John S. Baras. Adaptive sampling for linear state estimation. *SIAM Journal on Control and Optimization*, 50(2):672–702, 2012.

Y. Xu and J.P. Hespanha. Optimal communication logics in networked control systems. In *Decision and Control, 2004, Proceedings of the 43rd IEEE Conference on*, pages 3527–3532, Nassau, Bahamas, 2004.