2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Robustification and Optimization of a Kalman Filter with Measurement Loss using Linear Precoding

Rainer Blind, Stefan Uhlich, Bin Yang and Frank Allgöwer

Abstract— **This paper considers a linear precoding for a Kalman filter which receives the measurements over an erasure channel. We show how to design the precoding matrix that either robustifies the Kalman filter with respect to packet losses or reduces the estimation error. The effectiveness of the proposed method is demonstrated by two examples.**

I. INTRODUCTION

Motivated by recent progress in microcontroller and network technology, more and more control loops are closed using a packet based digital network. Unfortunately, this approach also brings a fundamental problem: the delay or loss of packets during data transfer. Consequently, the control and estimation over unreliable networks has become a very interesting and active research area.

In this work, we will study the state estimation problem over a lossy network. Fig. 1 shows a block diagram of such a networked control system. The measurements are transmitted to the state estimator via a network where measurements may get lost. The goal of this paper is to robustify this transmission. We will consider a linear precoding of the measurements to combat the erasures that may occur. Therefore, we assume that all observations are available to the precoder before they are transmitted. The scheme as shown in Fig. 1 has received a lot of attention in the past and the following paragraph provides an overview of the literature.

One approach to estimate the state of a system over a lossy network is the usage of smart sensors. A smart sensor is composed of the sensor itself and a Kalman filter to estimate the plant state. Now, the estimate $\hat{\mathbf{x}}_t$ is sent instead of the measurements $\mathbf{y}_t$. This approach has the advantage that the current estimation error does not depend on previous losses. In [1], this is done by an encoder and decoder. In order to reduce the necessary bandwidth, it is suggested in [2] to add an open-loop estimator at the remote side and different scheduling schemes, when to send the estimate are presented. However, this scheme is only applicable if enough computational resources are available at the sensor side.

A series of contributions on Kalman filtering with intermittent measurements was started by [3]. In [3], the authors

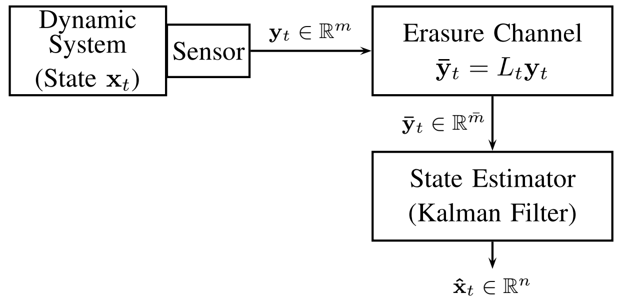

Fig. 1. State estimation over digital network

assume that either all the measurements are received, or they are completely lost. It turns out that the optimal Kalman filter is an ordinary time-varying Kalman filter. The excellent contribution in this work are convergence criteria for the covariance matrix. Often, only a part of the measurements will arrive. Thus, [3] was extended to the case of two independent channels in [4] and to the case of p channels in [5].

In order to minimize the distortion for a given bitrate, [6] used scalar quantizers for Multiple Description Coding (MDC) to code the measurements. This increases the robustness of the system as the number of available bits are split into two independent streams that are transmitted. If only one of the two streams is available, then the original measurement can still be approximately recovered. However, this approach has two drawbacks:

- Classical MDC is designed to minimize the mean squared error between the original sent signal $\mathbf{y}_t$ and the reconstructed signal $\bar{\mathbf{y}}_t$ at the receiver. However, in state estimation we are more interested in a small error between the state $\mathbf{x}_t$ and its estimate $\hat{\mathbf{x}}_t$ opposed to a good reconstruction of the measurements.
- The statistical properties of the output vector $\mathbf{y}_t$ change with time. In classical MDC, however, most methods assume that the signal statistics are time-invariant. Therefore, we can not apply the classical MDC methods directly. Often, this difficulty is circumvented by considering the steady state which is also done in this work.

In this work, we extend the approach of [3] and [4] by introducing a precoder such that either the Kalman filter can tolerate a higher packet loss rate or the estimation error is reduced.

A. Motivating Example

Before the main theorems are presented in Sec. III, we will motivate this approach with the following simple but

This work was supported in part by the Deutsche Forschungsgemeinschaft (Collaborative Research Center 732: Incremental specification in context and Priority Program 1305: Control Theory of Digitally Networked Dynamical Systems)

R. Blind and F. Allgöwer are with the Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany {blind,allgower}@ist.uni-stuttgart.de

S. Uhlich and B. Yang are with the Chair of System Theory and Signal Processing, University of Stuttgart, Germany {uhlich,yang}@lss.uni-stuttgart.de

978-1-4244-4524-0/09/$25.00 ©2009 AACC

illustrative discrete-time system: $\mathbf{x}_{t+1} = \left[\begin{smallmatrix} 2.5 & 0 \\ 0 & 2 \end{smallmatrix}\right] \mathbf{x}_t$. Moreover, we assume that both states are measured directly, i.e. $\mathbf{y}_t = \begin{bmatrix} y_{t,1} & y_{t,2} \end{bmatrix}' = \mathbf{x}_t$ and the two measurements are send over a network in individual packets. Obviously, the system is observable as long as both measurements arrive. Unfortunately, the system is no longer observable, if one of the measurements is lost. Note that the observability is a fundamental condition for the design of an observer.

Now, suppose we do not transmit the measurements directly but send the following virtual measurements $\tilde{y}_{t,1} = y_{t,1} + y_{t,2}$ and $\tilde{y}_{t,2} = y_{t,1} - y_{t,2}$ instead. Since this transformation, called correlating transform for obvious reasons, is invertible, there is no difference to the original system as long as both measurements arrive. However, the system is still observable if one of the virtual measurements is lost. As a consequence thereof, the observer will be more robust against packet losses.

B. Multiple Description Coding

Although, we motivated this approach by the loss of observability, such a precoding is very well known in the area of Multiple Description Coding (MDC). Multiple description coding aims at transmitting information over a channel which randomly erases elements. For an overview of MDC see [7]. One possibility to cope with such an erasure channel is the usage of a correlating transform, e.g. [8]–[10]. The correlating transform in classical MDC is chosen such that the average distortion is minimized. We will apply the same principle to a Kalman filter which receives its measurements over an erasure channel. However, in this work we optimize the correlating transform with respect to the estimation error of the Kalman filter and the robustness against measurement losses. The contribution of this paper is therefore the design of an optimal correlating transform which generates the virtual measurements that are transmitted. Note that the proposed approach can also be applied in connection to the design of a robust linear-quadratic regulator.

Motivated by the example above and the widespread use of such a precoder in MDC, we will proceed as follows: In Sec. II we recall the results on Kalman filtering with intermittent measurements and describe the channel model and precoding in detail. In Sec. III we combine these ideas and consider the Kalman filtering with linear precoding. Finally, Sec. IV provides simulation results for two examples and Sec. V concludes the work.

II. Preliminaries

In this section, we first repeat the results for Kalman filtering with intermittent measurements. Then we describe our channel model and the precoding.

A. Kalman Filtering with Intermittent Measurements

In this section, we briefly repeat the results of [3]. Although these results have been extended to the case of two and more channels in [4] and [5], we stick to [3] for simplicity.

We consider the following discrete time system

$$\begin{aligned} \mathbf{x}_{t+1} &= A\mathbf{x}_t + \mathbf{w}_t, \\ \mathbf{y}_t &= \gamma_t C\mathbf{x}_t + \mathbf{v}_t, \end{aligned} \tag{1}$$

where $\mathbf{x}_t \in \mathbb{R}^n$ is the system state and $\mathbf{y}_t \in \mathbb{R}^m$ the measurement output at time instance t. $\mathbf{w}_t \in \mathbb{R}^n$ and $\mathbf{v}_t \in \mathbb{R}^m$ are Gaussian white noise vectors with zero mean and covariance matrix $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$, respectively. Moreover, $\gamma_t \in \{0, 1\}$ is an independent and identically distributed (iid) random process with $\mathrm{E}[\gamma_t] = \lambda$, which indicates whether or not a measurement arrives.

First, we define

$$\begin{aligned} \hat{\mathbf{x}}_{t|t} &:= \mathrm{E}[\mathbf{x}_t|\mathcal{I}_t] \\ P_{t|t} &:= \mathrm{E}[(\mathbf{x}_t - \hat{\mathbf{x}}_t)(\mathbf{x}_t - \hat{\mathbf{x}}_t)'|\mathcal{I}_t] \\ \hat{\mathbf{x}}_{t+1|t} &:= \mathrm{E}[\mathbf{x}_{t+1}|\mathcal{I}_t] \\ P_{t+1|t} &:= \mathrm{E}[(\mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t})(\mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t})'|\mathcal{I}_t] \end{aligned}$$

where $\mathcal{I}_t$ is the information available at time t, i.e. $\mathbf{y}_0, \ldots, \mathbf{y}_t$ and $\gamma_0, \ldots, \gamma_t$. Note that $\mathrm{tr}\{P_{t+1|t}\} = \mathrm{E}\|\mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t}\|_2^2$, is the mean squared error between the true and the estimated state which we will use in Sec. III to find the optimal precoding.

The time update of the Kalman filter is

$$\hat{\mathbf{x}}_{t+1|t} = A\hat{\mathbf{x}}_{t|t}, \tag{2}$$

$$P_{t+1|t} = AP_{t|t}A' + Q, \tag{3}$$

which is identical to the case of no measurement losses. The measurement update becomes

$$\begin{aligned} \hat{\mathbf{x}}_{t+1|t+1} = \hat{\mathbf{x}}_{t+1|t} &+ \gamma_{t+1}P_{t+1|t}C' \left(CP_{t+1|t}C' + R\right)^{-1} \\ &\times (\mathbf{y}_{t+1} - C\hat{\mathbf{x}}_{t+1|t}), \end{aligned} \tag{4}$$

$$\begin{aligned} P_{t+1|t+1} = P_{t+1|t} &- \gamma_{t+1}P_{t+1|t}C' \left(CP_{t+1|t}C' + R\right)^{-1} \\ &\times CP_{t+1|t}. \end{aligned} \tag{5}$$

Both $\hat{\mathbf{x}}_{t+1|t+1}$ and $P_{t+1|t+1}$ are now random variables, depending on γ_{t+1}.

Using the shortcut $P_t := P_{t|t-1}$, (3) and (5) can be written as

$$P_{t+1} = AP_tA' + Q - \gamma_t AP_tC' \left(CP_tC' + R\right)^{-1} CP_tA'. \tag{6}$$

In order to derive an upper bound of $\mathrm{E}[P_t]$, the Modified Algebraic Riccati Equation (MARE) $g_\lambda(X)$ is defined as follows:

$$g_\lambda(X) = AXA' + Q - \lambda AXC' \left(CXC' + R\right)^{-1} CXA' \tag{7}$$

After these definitions, [3] showed that there exists a critical arrival rate λ_c which determines whether $\mathrm{E}[P_t]$ is bounded or not and give an upper bound for λ_c and $\mathrm{E}[P_t]$.

Theorem 1 ([3]): If $(A, Q^{\frac{1}{2}})$ is controllable, (A, C) is detectable, and A is unstable, then there exists a $\lambda_c \in [0, 1)$ such that

$$\lim_{t\to\infty} \mathrm{E}[P_t] = +\infty, \qquad \textit{for } 0 \le \lambda \le \lambda_c \textit{ and } \exists P_0 \ge 0 \tag{8}$$

$$\mathrm{E}[P_t] \le M_{P_0}\ \forall t, \qquad \textit{for } \lambda_c < \lambda \le 1 \textit{ and } \forall P_0 \ge 0 \tag{9}$$

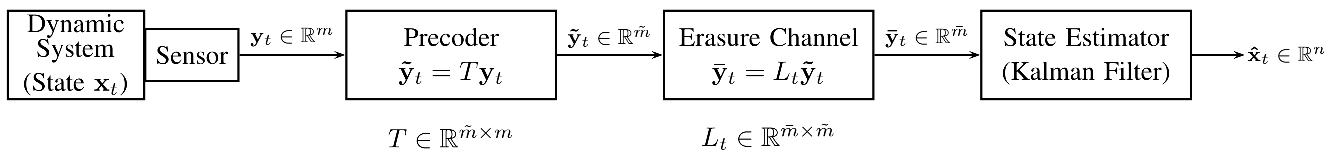

Fig. 2. State estimation over digital network with linear precoding

where $M_{P_0} > 0$ depends on the initial condition $P_0 \geq 0$.

Unfortunately, λ_c can not be calculated directly, but an upper bound $\bar{\lambda}$ can be found as follows:

Theorem 2 ([3]): The upper bound $\bar{\lambda}$ is given by the solution of the following optimization problem

$$\bar{\lambda} = \arg\min_{\lambda} \Psi_\lambda(Y, Z) > 0, \qquad 0 \leq Y \leq I, \tag{10}$$

where

$$\Psi(Y, Z) = \begin{bmatrix} Y & \sqrt{\lambda}(YA + ZC) & \sqrt{1-\lambda}YA \\ \star & Y & 0 \\ \star & \star & Y \end{bmatrix}. \tag{11}$$

The following theorem shows that there is an upper bound for $\mathrm{E}[P_t]$:

Theorem 3 ([3]): Assume that $(A, Q^{\frac{1}{2}})$ is controllable, (A, C) is detectable and $\lambda > \bar{\lambda}$. Then

$$\mathrm{E}[P_t] \leq V_t \qquad \forall\, \mathrm{E}[P_0] \geq 0 \tag{12}$$

where V_t is found by the sequence $V_{t+1} = g_\lambda(V_t)$, $V_0 = E[P_0]$. Moreover, $\lim_{t\to\infty} V_t = \bar{V}$ where $\bar{V}$ is the fixed point of (7), i.e. $\bar{V} = g_\lambda(\bar{V})$.

The last theorem states that the upper bound for $\lim_{t\to\infty} \mathrm{E}[P_t]$ can be found by a LMI-problem.

Theorem 4 ([3]): If $\lambda > \bar{\lambda}$, then the matrix $\bar{V} = g_\lambda(\bar{V})$ is given by

a) $\bar{V} = \lim_{t\to\infty} V_t$; $V_{t+1} = g_\lambda(V_t)$ where $V_0 \geq 0$.

b) $\arg\max_V \mathrm{tr}\{V\}$

$$\textit{subject to } \begin{bmatrix} AVA' - V + Q & \sqrt{\lambda}AVC' \\ \star & CVC' + R \end{bmatrix} \geq 0, \; V \geq 0.$$

B. Precoding

To robustify the Kalman filter against random measurement losses, we introduce a linear precoding. Instead of transmitting $\mathbf{y}_t$ directly, we add another block to the sender as shown in Fig. 2. Thus, the new measurement equation is

$$\tilde{\mathbf{y}}_t = T\mathbf{y}_t = TC\mathbf{x}_t + T\mathbf{v}_t, \tag{13}$$

where the precoding is done by the left-multiplication of the original measurement $\mathbf{y}_t$ with the matrix $T \in \mathbb{R}^{\tilde{m}\times m}$.

As can be seen in Fig. 2 and by Eq. (13), we require the precoder to receive all measurement packets. After the precoding, the virtual measurements can be send over independent channels.

Although there is no restriction on the dimensions of T, we assume T to be a square matrix in this work. This has the advantage that the necessary bandwidth is not changed since the number of original and virtual measurements is identical. All results, however, can be easily extended to the non-square case. Note that a non-square matrix means compression or adding redundancy, two interesting topics in the field of channel coding.

There are two special cases of T where we do not expect good results:

- T is singular. In this case, the precoding introduces some redundant measurements which do not improve the estimation if more than one of these is available to the state estimator, i.e. there is no difference whether they are lost or arrive if at least one was received.
- T is such that the system is not observable if one virtual measurement is lost. As shown in the motivating example, this might be the case for $T = I$. It might also be possible to choose T such that the system is not observable even if all virtual measurements arrive.

Obviously, both transforms should be avoided. Fortunately, they will be automatically avoided while searching for an optimal transformation.

C. Channel Model

In [3], the lossy network is modeled by the random process γ_t. A proper model for the multichannel case is slightly more complicated. We use a channel model in the spirit of [9], which allows us to present the results in a compact form.

We model the lossy network by a left-multiplication with an erasure matrix $L_t \in \mathbb{R}^{\bar{m}\times\tilde{m}}$ ($\bar{m} \leq \tilde{m}$) which is the identity matrix where a row is removed if the corresponding measurement is lost. Hence, the dimension of L_t depends on the loss of measurements. Note that we can not define L_t as an identity matrix where a row is replaced by a row with only zeros if the corresponding measurement is lost. This is due to the fact that there is a subtle difference whether $y_{t,i}$ is lost or $y_{t,i} = 0$. Another reason will be clarified in Sec. III.

The following two examples will give the basic idea of this notation. If all measurements arrive, then L_t is obviously the identity matrix. If there are 3 measurements and the first and second arrive, then we have $\tilde{m} = 3$, $\bar{m} = 2$, and $L_t = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right]$.

Finally, we define the set of all possible erasure constellations $\mathcal{L}$ as $\mathcal{L} = \{\tilde{L}_1, \ldots, \tilde{L}_E\}$ where $E = 2^{\tilde{m}}$ is the total number of error constellations. At each time instance t, L_t is randomly chosen from $\mathcal{L}$.

In the following, we will assume that the erasure process is independent and identically distributed for each measurement. Moreover, for simplicity of notation, we assume that it does not depend on the channel. The probability that a measurement arrives safely is given by the arrival rate λ. The probability $w_e(\lambda)$ that a particular $\tilde{L}_e$ is chosen is thus

$$
\begin{aligned}
&\Psi(Y, Z_1, \ldots, Z_E, T) \\
&= \begin{bmatrix}
Y & \sqrt{w_1(\lambda)}(YA + Z_1\tilde{L}_1TC) & \sqrt{w_2(\lambda)}(YA + Z_2\tilde{L}_2TC) & \cdots & \sqrt{w_E(\lambda)}(YA + Z_E\tilde{L}_ETC) \\
\star & Y & 0 & \cdots & 0 \\
\star & \star & Y & & 0 \\
\vdots & \vdots & \vdots & \ddots & \\
\star & \star & \star & & Y
\end{bmatrix}
\end{aligned} \quad (20)
$$

$$
\begin{aligned}
&\Gamma(V, T) \\
&= \begin{bmatrix}
AVA' + Q - V & \sqrt{w_1(\lambda)}AVC'T'\tilde{L}_1' & \sqrt{w_2(\lambda)}AVC'T'\tilde{L}_2' & \cdots & \sqrt{w_E(\lambda)}AVC'T'\tilde{L}_E' \\
\star & \Xi_1 & 0 & \cdots & 0 \\
\star & \star & \Xi_2 & & 0 \\
\vdots & \vdots & \vdots & \ddots & \\
\star & \star & \star & & \Xi_E
\end{bmatrix}
\end{aligned} \quad (21)
$$

where $\Xi_e := \tilde{L}_eTCVC'T'\tilde{L}_e' + \tilde{L}_eTRT'\tilde{L}_e'$.

$w_e(\lambda) := \lambda^i(1-\lambda)^{\tilde{m}-i}$, where i is the number of arrived measurements and $\tilde{m} - i$ the number of lost measurements. Note that the results of this work can be easily extended to the case of non-identical arrival rates.

III. Kalman Filtering with Linear Precoding

In this section, we combine the ideas of the previous section in order to achieve one of the following goals:

Goal 1: Make the Kalman filter more robust against packet loss.

Goal 2: Reduce the estimation error.

Before we show how to design the precoding matrix T, we first derive some equations and theorems.

Combining the precoding and the channel model, we get

$$\bar{\mathbf{y}}_t = L_t\tilde{\mathbf{y}}_t = L_tTC\mathbf{x}_t + L_tT\mathbf{v}_t \quad (14)$$

as input for the Kalman filter. Since the time update of the Kalman filter does not depend on the measurements, it is still

$$\hat{\mathbf{x}}_{t+1|t} = A\hat{\mathbf{x}}_{t|t}, \quad (15)$$

$$P_{t+1|t} = AP_{t|t}A' + Q. \quad (16)$$

By replacing C with L_tTC and R with $L_tTRT'L_t'$, (4) and (5) for the measurement update become

$$
\begin{aligned}
\hat{\mathbf{x}}_{t+1|t+1} = \hat{\mathbf{x}}_{t+1|t} &+ P_{t+1|t}C'T'L_t' \\
&\times \left(L_tTCP_{t+1|t}C'T'L_t' + L_tTRT'L_t'\right)^{-1} \\
&\times (\bar{\mathbf{y}}_{t+1} - L_tTC\hat{\mathbf{x}}_{t+1|t}), \quad (17)\\
P_{t+1|t+1} = P_{t+1|t} &- P_{t+1|t}C'T'L_t' \\
&\times \left(L_tTCP_{t+1|t}C'T'L_t' + L_tTRT'L_t'\right)^{-1} \\
&\times L_tTCP_{t+1|t}. \quad (18)
\end{aligned}
$$

Obviously, $\left(L_tTCP_{t+1|t}C'T'L_t' + L_tTRT'L_t'\right)^{-1}$ can only be calculated if L_t does not contain rows with only zeros.

The MARE for this case can be derived from (7) by replacing λ with $\sum_{e=1}^{E} w_e(\lambda)$, C with $\tilde{L}_eTC$, and R with $\tilde{L}_eTRT'\tilde{L}_e'$:

$$
\begin{aligned}
g(X) = AXA' + Q - \sum_{e=1}^{E} w_e(\lambda)AXC'T'\tilde{L}_e' \\
\times \left(\tilde{L}_eTCXC'T'\tilde{L}_e' + \tilde{L}_eTRT'\tilde{L}_e'\right)^{-1}\tilde{L}_eTCXA' \quad (19)
\end{aligned}
$$

where E is the total number of error constellations and $w_e(\lambda)$ is the probability that a particular $\tilde{L}_e$ will occur. In order to get the upper bounds $\bar{\lambda}$ and $\bar{V}$, Theorem 2 and 4 become now:

Theorem 5: The upper bound $\bar{\lambda}$ is given by the solution of the following optimization problem

$$\bar{\lambda} = \arg\min_{\lambda} \Psi(Y, Z_1, \ldots, Z_E, T) > 0 \qquad 0 \le Y \le I,$$

where $\Psi(.)$ is given in (20) at the top of the page.

Theorem 6: If $\lambda > \bar{\lambda}$, then the matrix $\bar{V} = g_\lambda(\bar{V})$ is given by

(a) $\bar{V} = \lim_{t\to\infty} V_t$; $V_{t+1} = g_\lambda(V_t)$ where $V_0 \ge 0$.

(b) $\arg\max_V \operatorname{tr}\{V\}$ subject to $\Gamma(V,T) \ge 0, \qquad V \ge 0$, where $\Gamma(V,T)$ is given in (21) at the top of the page.

The proof of these theorems follows the same line as in [4] and is omitted here.

Now we can choose T according to the two goals:

Goal 1: Make the Kalman filter more robust against packet loss. Here we choose T in such a way that the upper bound $\bar{\lambda}$ is minimized. Consequently, we use Theorem 5 and search for $Y, Z_1, \ldots, Z_E$ and T such that $\arg\min_\lambda \Psi(Y, Z_1, \ldots, Z_E, T) > 0$ is minimized. Note that we have to use the upper bound $\bar{\lambda}$ to obtain the optimal T because we cannot calculate λ_c exactly.

Goal 2: Minimize the estimation error. Here we choose T in such a way that the estimation error $(\mathbf{x}_t - \hat{\mathbf{x}}_t)$ is minimized for a given λ. This means, we use Theorem 6 and search for T such that $\operatorname{tr}\{\bar{V}\}$ is minimized. This is especially interesting, if the network is given and the packet arrival rates are known. Again, we have to use the upper bound $\bar{V}$ because $\lim_{t\to\infty} \mathrm{E}[P_t]$ is not known exactly.

Unfortunately, both problems are nonconvex and hence very difficult to solve since there does not exist any algorithm, which can guarantee to find the global optimal solution of a nonconvex problem. Reformulating these problems as convex ones would be very valuable.

Finally, we show that the rows of T can be normalized without affecting the precoding. This will simplify the analysis of the examples in Sec. IV.

Lemma 1: Let T be an arbitrary precoding matrix. The solution of the MARE $X = g(X)$ in (19) *and the upper bounds on the critical arrival rate $\bar{\lambda}$ are invariant with respect to a scaling of the rows of T, i.e. by replacing T with NT where $N = \mathrm{diag}(n_{11}, \ldots, n_{\tilde{m}\tilde{m}})$ and $n_{ii} \neq 0, \forall i$.*

Proof: A scaling of the rows of the precoding matrix T corresponds to the replacement of $L_t T$ in (14) by $L_t N T$. Since L_t contains different row vectors of the identity matrix, a scaling of the columns of L_t by $L_t N$ can also be written as a scaling of the rows of L_t, i.e. $\tilde{N} L_t$, where $\tilde{N} = L_t N L_t'$ is a $\tilde{m} \times \tilde{m}$ invertible diagonal matrix containing $\tilde{m}$ diagonal elements of N. Hence, the net effect of a scaling of the rows of T is to scale the measurements $\bar{\mathbf{y}}_t$ which is an invertible process. ■

IV. Examples

In this section, we present two examples to show the benefits of the proposed method. For the sake of clarity, we consider two relatively simple examples with two measurements ($m = 2$). We parametrize the precoding transform as $T := \left[\begin{smallmatrix} 1 & t_1 \\ t_2 & 1 \end{smallmatrix}\right]$ according to Lemma 1. Now we can plot $\bar{\lambda}$ or $\mathrm{tr}\{\bar{V}\}$ over t_1 and t_2 and easily see the influence of T on these performance measures. Note that $t_1 = t_2 = 0$ corresponds to the no precoding case.

Example 1: First, we reconsider the motivating example of Sec. I-A with

$$A = \begin{bmatrix} 2.5 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Moreover, we set $Q = 10I$ and $R = 2.5I$.

Without the precoding transform, we obtain $\bar{\lambda} = 0.84$. Hence, if more than 84% of the measurements arrive, then the expected value of the covariance matrix P_t is guaranteed to be bounded.

In order to achieve Goal 1, we search for a transform which minimizes $\bar{\lambda}$. To easily see the influence of T on $\bar{\lambda}$, Fig. 3 shows $\bar{\lambda}$ over T. We see that $\bar{\lambda}$ can be significantly reduced by a proper precoding transform and it would be easy to find a transform such that the covariance matrix converges even for an arrival rate less than 75%. Note that there is a dramatic increase in $\bar{\lambda}$ if $t_2 = 0$. In this case, the more unstable first mode is no longer observable if the first virtual measurement is lost. Moreover, there are peaks in $\bar{\lambda}$ along the lines where T is singular, i.e. $t_1 = 1/t_2$.

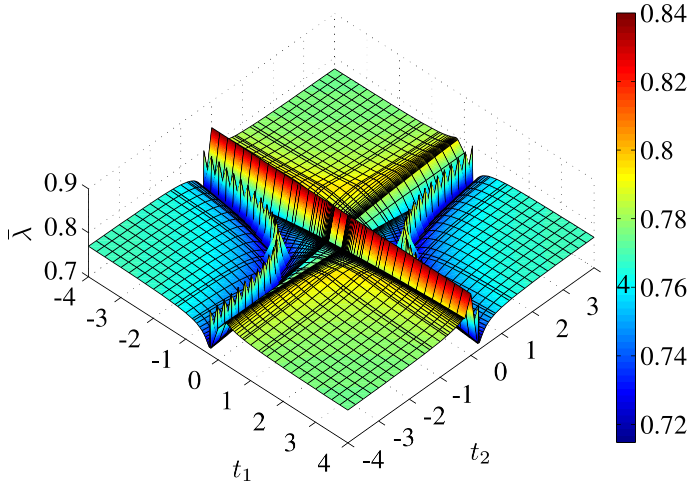

Fig. 3. Upper bound $\bar{\lambda}$ of the critical arrival rate over the precoding matrix T for Example 1

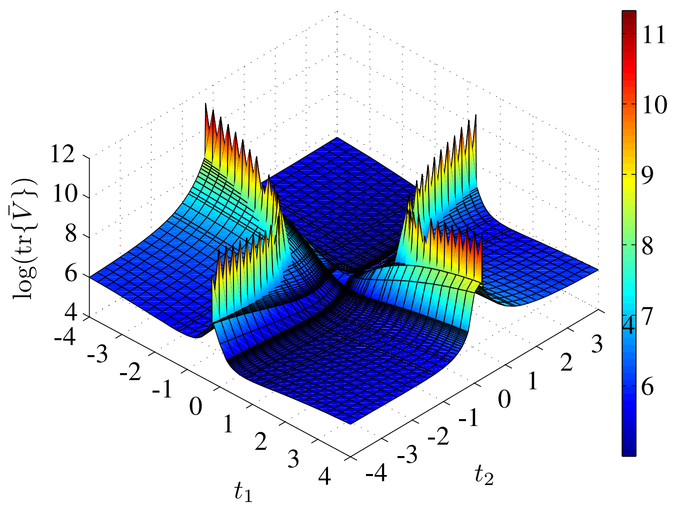

Fig. 4. Upper bound $\mathrm{tr}\{\bar{V}\}$ of the expected estimation error over the precoding matrix T for Example 1

However, changes in $\bar{\lambda}$ are not only due to observability and singularity. Even if the system is observable with one virtual measurement and T is not singular, $\bar{\lambda}$ depends on T.

In order to achieve Goal 2, we ask for a transform which minimizes $\mathrm{tr}\{\bar{V}\}$ for a fixed λ. Fig. 4 shows $\log(\mathrm{tr}\{\bar{V}\})$ over T for $\lambda = 0.85$. Here, we see that $\mathrm{tr}\{\bar{V}\}$ is getting huge when T is chosen such that the system is not observable if one virtual measurement is lost. Moreover, $\mathrm{tr}\{\bar{V}\}$ is increased along the line $t_1 = 1/t_2$, where T is singular. For $T = I$, we get $\mathrm{tr}\{\bar{V}\} = 415.94$. Using MATLAB's `fminsearch` function, we obtain the optimal[1] transform $T_{opt} = \left[\begin{smallmatrix} 1 & 0.6803 \\ -1.4813 & 1 \end{smallmatrix}\right]$ and $\mathrm{tr}\{\bar{V}\} = 148.69$, which improves the quality of the Kalman filter considerably.

Note that the peaks in Fig. 3 and 4 occur when T is singular and therefore both transmitted measurements are redundant. The key of the precoder T is that each row

[1] Obviously, we can not guarantee that this, and the following transforms, are globally optimal.

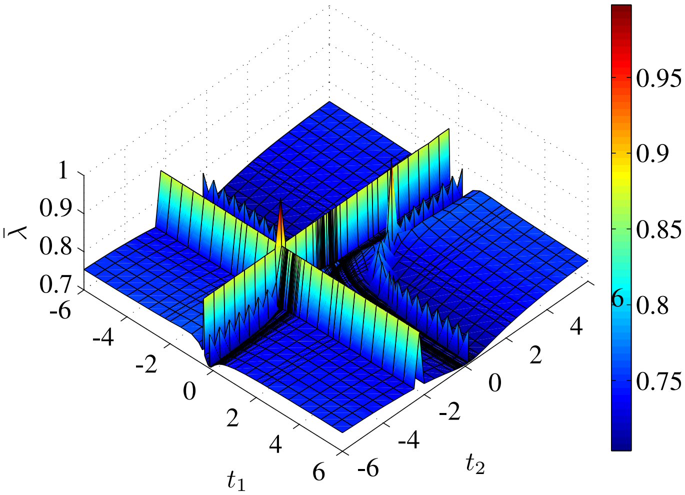

Fig. 5. Upper bound $\bar{\lambda}$ of the critical arrival rate over the precoding matrix T for Example 2

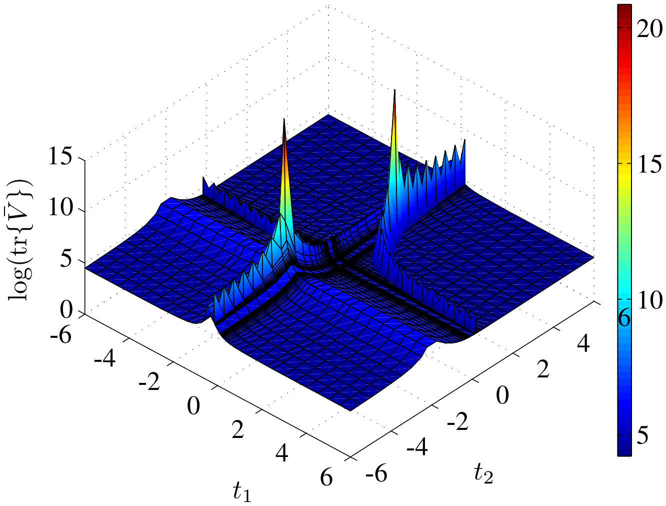

Fig. 6. Upper bound $\mathrm{tr}\{\bar{V}\}$ of the expected estimation error over the precoding matrix T for Example 2

has to point into the most important direction for the state estimation task but also to make each row a little bit different so that the combination of both allows a better estimation of the state. This is analog to MDC where the most important direction is the direction of maximum scatter of the data vectors which is called the principal component, see [7].

Example 2: The previous example is composed of two unstable systems which become unobservable if one of the measurements is lost. Now, we look at a system which is observable if any of the two packets arrive:

$$A = \begin{bmatrix} 2.5 & 0.25 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Note that this system is not observable if $t_1 = \pm 1/\sqrt{5} \approx \pm 0.447$ or $t_2 = \pm\sqrt{5} \approx \pm 2.236$ and the corresponding measurement is lost.

Again, we set $Q = 10I$ and $R = 2.5I$. Here, we get $\bar{\lambda} = 0.7618$ without the correlating transform. Fig. 5 shows $\bar{\lambda}$ over T and we see similar effects as in the previous example for the two special cases of T. As in the previous example, $\bar{\lambda}$ depends not only on these two effects. In this example, we used `fminsearch` to find the transform which minimizes $\bar{\lambda}$ and obtained $T_{opt} = \left[\begin{smallmatrix} 1 & 17.2102 \\ -0.053 & 1 \end{smallmatrix}\right]$. Using this correlating transform, we get $\bar{\lambda} = 0.7039$. Hence, even if the original system is observable if one measurement is lost, we still can reduce $\bar{\lambda}$ by a correlating transform.

As in the previous example, we now fix λ to 0.88 and search for that transform which minimizes $\mathrm{tr}\{\bar{V}\}$. Fig. 6 shows $\log(\mathrm{tr}\{\bar{V}\})$ over T. Again, we see a significant influence of the two special cases. In this example, we get $\mathrm{tr}\{\bar{V}\} = 103.27$ for the original system and $\mathrm{tr}\{\bar{V}\} = 82.76$ for the optimal transform $T_{opt} = \left[\begin{smallmatrix} 1 & 4.1512 \\ 0.5653 & 1 \end{smallmatrix}\right]$. Hence, the correlating transform improves the quality of the Kalman filter although the original system is observable if one measurement is lost.

V. Conclusion

This paper shows how to design a linear precoding matrix for a Kalman filter in the presence of lossy transfer channels. We showed that the precoding matrix can be chosen such that the Kalman filter will be more robust against measurement loss or to reduce the estimation error. Unfortunately, this is a nonconvex optimization problem and reformulating it as a convex one would be very valuable. The results indicate that the observer and the channel coder have to be considered together to design a good observer in a networked control system.

References

[1] V. Gupta, B. Hassibi, and R. M. Murray, "Optimal LQG control across packet-dropping links," *System and Control Letters*, vol. 56, no. 6, pp. 439–446, 2007.

[2] Y. Xu and J. P. Hespanha, "Estimation under uncontrolled and controlled communications in networked control systems," in *Proceedings of the Conference on Decision and Control, and the European Control Conference*, 2005, pp. 842–847.

[3] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automatic Control*, vol. 49, no. 9, pp. 1453– 1464, Sept. 2004.

[4] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," *IEEE Conf. on Decision and Control*, pp. 4180–4186, 2004.

[5] E. Garone, B. Sinopoli, A. Goldsmith, and A. Casavola, "LQR control for distributed systems over TCP-like erasure channels," in *Proceedings of the Conference on Decision and Control*, 2007, pp. 44–49.

[6] Z. Jin, V. Gupta, and R. M. Murray, "State estimation over packet dropping networks using multiple description coding," *Automatica*, vol. 42, no. 9, pp. 1441–1452, 2006.

[7] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Magazine*, vol. 18, no. 5, 2001.

[8] V. K. Goyal and J. Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Information Theory*, vol. 47, no. 6, pp. 2199–2224, 2001.

[9] G. Romano, P. Salvo Rossi, and F. Palmieri, "Optimal correlating transform for erasure channels," *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 677–680, 2005.

[10] S. Uhlich and B. Yang, "A generalized optimal correlating transform for multiple description coding and its theoretical analysis," *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.