Optimization-Based Design of Bounded-Error Estimators Robust to Missing Data ⋆

Kwesi J. Rutledge * **Sze Zheng Yong** ** **Necmiye Ozay** *

* *University of Michigan, Ann Arbor, MI 48105 USA (e-mail: {krutledg,necmiye}@ umich.edu).*
** *Arizona State University, Tempe, AZ 85287 USA (e-mail: szyong@asu.edu)*

Abstract: Non-asymptotic bounded-error state estimators that provide hard bounds on the estimation error are crucial for safety-critical applications. This paper proposes a class of optimal bounded-error affine estimators to achieve a novel property we are calling *Equalized Recovery* that can be computed by leveraging ideas from the dual problem of affine finite horizon optimal control design. In particular, by using Q-parametrization, the estimator design problem is reduced to a convex optimization problem. An extension of this estimator to handle missing data (e.g., due to package drops or sensor glitches) is also proposed. These ideas are illustrated with a numerical example motivated by vehicle safety systems.

Keywords: Robust estimators, bounded-error estimation, missing data

1. INTRODUCTION

State estimation is one of the key problems in control design. The celebrated Kalman filter (Kalman (1960)) and its various extensions for cases with missing and intermittent observations (Sinopoli et al. (2004)) have been widely used in control systems where the states of the system are not available for feedback so one may need to estimate the states from some output measurements. In safety-critical problems, non-asymptotic estimation techniques such as set-valued observers or ℓ_∞ filters (Milanese and Vicino (1991); Shamma and Tu (1999)) also become important as error bounds can be precomputed and used in control synthesis.

Equalized performance is an intuitive property for estimators to ensure that the estimation error does not increase at each step (Blanchini and Sznaier (2012)). This is shown to be useful in output feedback correct-by-construction control (Mickelin et al. (2014)). In this paper, we generalize equalized performance to allow for violation of the error bound for a limited horizon, within which a more relaxed error bound is satisfied. In the case of time horizons greater than one, this new condition can be thought of as ensuring a 'recovery' to some desired performance level after a lapse. We then show that optimal affine filters satisfying this generalized condition can be designed using convex optimization. In particular, we leverage results from affine finite horizon optimal control (Skaf and Boyd (2010)) and show how similar ideas can be employed for the dual problem of filter design.

As a further generalization of the new error bound condition, we consider the case of missing observations or intermittent data due to package drops or sensor glitches. In contrast to probabilistic packet drops in (Sinopoli et al. (2004)), we consider missing observations with no assumed probability distribution, but are instead described by fixed-length language specifications, e.g., 'every other measurement is missing in a time horizon of T time steps.' The system with missing observations can be seen as a hybrid system where the observation equation switches between normal observations and no observation. The optimal estimator design problem in this case can be reduced to a convex problem that allows us to analyze how many missing observations the estimator can tolerate within a given time horizon while meeting the error bound requirements.

Notation: We denote by $\|\cdot\|$ infinity norms of vectors and matrices. The symbol $\otimes$ represents the Kronecker product, I_k represents the identity matrix of size k, $0_{k\times m}$ represents the $k \times m$ zero matrix, and $\mathbb{1}_k$ denotes a k dimensional vector of ones. The subscripts are dropped when the dimension of the matrix is clear from the context. For matrices and vectors, the inequalities $\geq$ are always taken element-wise.

2. PROBLEM STATEMENT

In this work, the estimation of an affine dynamical system's state is discussed. To denote this system, its estimator, and several other quantities in the following sections, we present the following definitions.

⋆ This work is partially supported by the NSF Graduate Research Fellowship Grant Number DGE 1256260, an Early Career Faculty grant from NASA's Space Technology Research Grants Program, DARPA grant N66001-14-1-4045 and by Toyota Research Institute ("TRI"). TRI provided funds to assist the authors with their research but this article solely reflects the opinions and conclusions of its authors and not TRI or any other Toyota entity.

2.1 The Missing Data Model

In practice, a control system's measurements may be unpredictably, and unmanageably disturbed (e.g., by an unexpected flash of light that blinds a camera, or by a gliding leaf that suddenly blocks our LIDAR) during operation. Similarly, due to communication or sensor glitches, measurement data can be missing at sampling-time. We assert that the following 2-mode hybrid system represents the dynamics and measurement updates of a general affine system when missing data can occur:

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + f + w(t), \\ y(t) &= \begin{cases} Cx(t) + v(t), & q(t) = 1, \\ \emptyset, & q(t) = 0, \end{cases} \end{aligned} \tag{1}$$

where A, B, C, f are known system matrices, $q(t)$ is the discrete state/mode in the hybrid system, $u(t) \in \mathbb{R}^m$ is the input, $w(t) \in \mathbb{R}^n$ is the process noise and $v(t) \in \mathbb{R}^p$ is the measurement noise. The noise terms (or uncontrolled inputs) $w(t)$ and $v(t)$ are unknown but bounded, and their bounds are known ($\|w(t)\| \leq \eta_w$ and $\|v(t)\| \leq \eta_v$, $\forall t$).

The discrete state/mode $q(t) = 1$ denotes that the measurement vector is available, while $q(t) = 0$ signifies "missing" data, when $y(t)$ is essentially empty.

Moreover, in line with the robust framework throughout our paper, we consider a novel missing data model with no assumed probability distribution for discrete state/mode switching, i.e., the discrete state/mode switching is not stochastic, but the missing data model is instead expressed by fixed-length language specifications, e.g., 'every other observation is missing' and 'at most m missing data in first M steps of time horizon with $T > M$ time steps.' More precisely, our missing data model is a fixed-length language $\mathcal{L} \subseteq \mathbb{B}^T$ that specifies the set of allowable mode sequences $\{q(t)\}_{t=t_0}^{t_0+T-1}$. Fixed-length prefixes of more general regular or omega regular language specifications can also be considered.

The goal is to obtain good estimation performance guarantees despite the worst case missing data scenario. Thus, the relatively well-established estimation approaches for probabilistic intermittent observations, e.g., in (Sinopoli et al. (2004)), that optimize the expected/average estimation performance do not apply.

Remark 1. In the system description (1), we assume that the noise terms w and v have an identity gain term to keep the notation in the proceeding derivations simpler. If a model contains non-identity gain terms on the process or measurement noise (i.e., $\bar{B}_w w(t)$, $\bar{C}_v v(t)$), this can be straightforwardly handled by the proposed methodology.

2.2 Finite Horizon Affine Estimator

We consider in this paper a finite horizon dynamic estimator design with the following update rules:

$$\begin{aligned} \hat{x}(t+1) &= A\hat{x}(t) + Bu(t) - u_e(t) + f, \\ \hat{y}(t) &= C\hat{x}(t), \end{aligned} \tag{2}$$

and the following *causal* output injection term:

$$u_e(t) = u_0(t) + \sum_{\tau=t_0}^{t} F_{(t,\tau)} y_\xi(t), \tag{3}$$

where $y_\xi(t) \triangleq \begin{cases} y(t) - \hat{y}(t) = y(t) - C\hat{x}(t), & q(t) = 1, \\ 0, & q(t) = 0, \end{cases}$

and t_0 is the initial time of the finite horizon.

It is assumed that the initial estimate at time t_0, $\hat{x}(t_0)$, is given with an initial estimation error, $\xi(t_0) \triangleq x(t_0) - \hat{x}(t_0)$, that satisfies $\|\xi(t_0)\| \leq M_1$. At each time step $t \in [t_0, t_0 + T - 1]$, measurements $y(t)$ according to (1) are available, if not missing, only for the interval $[t_0, t]$ (i.e., up to the current time step; thus, the estimator is real-time and causal), and the finite horizon affine estimator outputs state estimate for the following time step $\hat{x}(t+1)$. Note that this is different from the fixed-interval smoothing problem where state estimates over an interval are obtained in a "post mortem" manner.

2.3 Equalized Recovery

We would like to focus on designing state estimators that have a property we are calling "Equalized Recovery." This property is present if, when estimation error $\xi(t) \triangleq x(t) - \hat{x}(t)$ with $\hat{x}(t)$ being an estimate of $x(t)$, is bounded at a certain point in time, we can guarantee that at a certain time in the future, the same bound will hold again.

Definition 1. (Equalized Recovery). An estimator is said to achieve an equalized recovery level M_1 with recovery time T and intermediate level $M_2 \geq M_1$ at time t_0 if whenever $||\xi(t_0)|| \leq M_1$, we must have $\|\xi(t)\| \leq M_2$ for all $t \in [t_0, t_0 + T]$ and $\|\xi(t_0 + T)\| \leq M_1$.

As a special case, achieving equalized recovery level M with recovery horizon 1 and intermediate level M is equivalent to equalized performance (see Blanchini and Sznaier (2012); Mickelin et al. (2014)).

The objective of this work is to synthesize a finite horizon affine estimator given by (2) and (3) that achieves equalized recovery as defined in Definition 1 for the system in (1) where the mode signal satisfies a fixed-length language specification describing the missing data pattern, which is formally described in the following problem:

Problem 1. [Equalized Recovery Estimator Synthesis] Let the initial estimate at time t_0 be $\hat{x}(t_0)$ and the initial estimation error be $\xi(t_0) \triangleq x(t_0) - \hat{x}(t_0)$. Given that

- the dynamics of the system is (1),
- the recovery level is M_1 (i.e., with $\|\xi(t_0)\| \leq M_1$),
- the intermediate level is $M_2 \geq M_1$,
- the recovery time is T, and
- the mode signal $q(t)$, $t \in [t_0, t_0 + T - 1]$ satisfies a missing data model $\mathcal{L} \subseteq \mathbb{B}^T$,

find an estimator of the form (2) and (3) such that $\|\xi(t)\| \leq M_2$ $\forall t \in [t_0, t_0 + T]$ and $\|\xi(t_0 + T)\| \leq M_1$.

We remark that in the case of no missing data, if the system is detectable, there always exists a Luenberger filter that solves Problem 1 for some large enough M_1, M_2 and T, related to noise bounds and ℓ_1-gain of the error system. However, a non-asymptotic analysis and design technique that computes an estimator, whenever possible, for a given set of parameters $\{M_1, M_2, T\}$ is important as these parameters can be used as "contracts" if one proceeds with control design where hard safety constraints need to be enforced (Mickelin et al. (2014)). As we show in the following, searching for finite horizon affine solutions to Problem 1 with given parameters can be done efficiently.

3. APPROACH

To address Problem 1, we synthesize a finite horizon affine estimator that builds upon the idea of Q-parametrization in (Skaf and Boyd (2010)). Our key result shows that the feasibility of finite horizon affine estimators (2),(3) that solve Problem 1 is equivalent to the feasibility of a convex optimization problem. Thus, feasibility and design of this type of estimator can be performed using the efficient tools of convex optimization. First, we present a novel estimator design that achieves equalized recovery, as defined in Definition 1, for the case without missing measurements, i.e., the perfect output measurement case. Using this result, we then propose an extension that can deal with missing measurements that are expressed by fixed-length language specifications.

3.1 Perfect Output Measurement Case

We assume for the moment that no data is missing and all measurements are available for a given time horizon T, i.e., with $q(t) = 1$ for all $t \in [t_0, t_0 + T - 1]$. For this special case, the following theorem addresses the feasibility and synthesis problems when designing finite horizon affine estimators (2),(3).

Theorem 1. [Perfect Output Measurement Case] The existence of a causal finite horizon affine estimator that solves Problem 1 (i.e. achieves equalized recovery level M_1 with recovery time T and intermediate level M_2) when $q(t) = 1 \; \forall t \in [t_0, t_0 + T - 1]$ (i.e., with no missing data) is equivalent to the feasibility of the following problem:

$$\begin{array}{ll} \text{Find} & Q,\ r \\ \text{subject to} & Q \text{ is } m\text{-by-}n \text{ block lower triangular,} \\ & \forall (\|w\| \leq \eta_w, \|v\| \leq \eta_v, \|\xi(t_0)\| \leq M_1): \\ & \|R_{0:T}\xi\| \leq M_2 \text{ and } \|R_T\xi\| \leq M_1, \end{array} \tag{4}$$

where $R_{\tau_1:\tau_2} \in \mathbb{R}^{n(\tau_2-\tau_1+1)\times n(T+1)}$ and $\xi \in \mathbb{R}^{n(T+1)}$ are:

$$R_{\tau_1:\tau_2} = \left[\, 0_{n(\tau_2-\tau_1+1)\times n\tau_1} \;\; I_{n(\tau_2-\tau_1+1)} \;\; 0_{n(\tau_2-\tau_1+1)\times n(T-\tau_2)} \,\right],$$

$$\xi = (S + SQ\bar{C}S)w + SQv + (I + SQ\bar{C})J\xi(t_0) + Sr,$$

with

$$\bar{C} = [\, I_T \otimes C \;\; 0_{pT\times n} \,] \in \mathbb{R}^{pT\times n(T+1)},$$

$$J = \begin{bmatrix} I_n \\ A \\ \vdots \\ A^{T-1} \\ A^T \end{bmatrix},\; S = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I_n & 0 & 0 & \cdots & 0 \\ A & I_n & 0 & \cdots & 0 \\ A^2 & A & I_n & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ A^{T-1} & A^{T-2} & A^{T-3} & \cdots & I_n \end{bmatrix}. \tag{5}$$

Moreover, if (4) is feasible, then the finite horizon affine estimator is as given in (2),(3) with

$$\begin{aligned} F &\triangleq \begin{bmatrix} F_{(t_0,t_0)} & 0 & \cdots & 0 \\ F_{(t_0+1,t_0)} & F_{(t_0+1,t_0+1)} & & \vdots \\ \vdots & & \ddots & 0 \\ F_{(t_0+T-1,t_0)} & F_{(t_0+T-1,t_0+1)} & \cdots & F_{(t_0+T-1,t_0+T-1)} \end{bmatrix}, \\ &= (I + Q\bar{C}S)^{-1}Q, \\ u_0 &\triangleq (u_0(t_0), \cdots, u_0(t_0+T-1)) \in \mathbb{R}^{mT}, \\ &= (I + Q\bar{C}S)^{-1}r = (I + F\bar{C}S)r. \end{aligned} \tag{6}$$

Proof. We begin by developing an expression for the estimation error, which according to (1) and the estimator definition (2) is:

$$\begin{aligned} \xi(t+1) &= A\xi(t) + u_e(t) + w(t), \\ y_\xi(t) &= C\xi(t) + v(t). \end{aligned} \tag{7}$$

Redefinition in terms of trajectories. Using a standard property of discrete time dynamics, we can write the entire trajectory of $\xi(t)$ as a matrix equation. Consider the following definitions:

$$\begin{aligned} \xi &= (\xi(t), \cdots, \xi(t+T)) \in \mathbb{R}^{n(T+1)}, \\ u_e &= (u_e(t), \cdots, u_e(t+T-1)) \in \mathbb{R}^{mT}, \\ w &= (w(t), \cdots, w(t+T-1)) \in \mathbb{R}^{nT}, \\ y_\xi &= (y_\xi(t), \cdots, y_\xi(t+T-1)) \in \mathbb{R}^{pT}, \\ v &= (v(t), \cdots, v(t+T-1)) \in \mathbb{R}^{pT}. \end{aligned}$$

Then the following matrix equations can be written:

$$\begin{aligned} \xi &= J\xi(t_0) + S(u_e + w), \\ y_\xi &= \bar{C}\xi + v, \\ u_e &= Fy_\xi + u_0, \end{aligned} \tag{8}$$

where $\bar{C}$, J and S are defined in (5).

Now, the three equations of (8) can be solved to create closed form expressions for the estimation error and input:

$$\begin{bmatrix} \xi \\ u_e \end{bmatrix} = P \begin{bmatrix} w \\ v \end{bmatrix} + \begin{bmatrix} \tilde{\xi} \\ \tilde{u}_e \end{bmatrix}, \tag{9}$$

where

$$P = \begin{bmatrix} P_{\xi w} & P_{\xi v} \\ P_{uw} & P_{uv} \end{bmatrix},$$

$$\begin{aligned} P_{\xi w} &= S + SF(I - \bar{C}SF)^{-1}\bar{C}S, & P_{\xi v} &= SF(I - \bar{C}SF)^{-1}, \\ P_{uw} &= F(I - \bar{C}SF)^{-1}\bar{C}S, & P_{uv} &= F(I - \bar{C}SF)^{-1}, \end{aligned}$$

and

$$\begin{aligned} \tilde{\xi} &= J\xi(t_0) + Su_0 + SF(I - \bar{C}SF)^{-1}(\bar{C}J\xi(t_0) + \bar{C}Su_0), \\ \tilde{u}_e &= F(I - \bar{C}SF)^{-1}(\bar{C}J\xi(t_0) + \bar{C}Su_0) + u_0. \end{aligned}$$

In the following, we make use of the Q-parametrization approach in (Skaf and Boyd (2010)) to obtain linear constraints that lead to the desired optimization problem.

Making the closed loop expression linear. The equations for ξ and u_e are nonlinear in the design variables (F, u_0). This would normally mean that searching for design variables that satisfy constraints like $\|\xi(t+T)\|_\infty \leq M_1$ would be a nonlinear and non-convex programming problem, but we simplify the search by taking inspiration from Q-parameterization and introducing the nonlinear mappings $Q \triangleq F(I - \bar{C}SF)^{-1}$ and $r \triangleq (I + Q\bar{C}S)u_0$. This transforms the P matrix into the following:

$$P_{\xi w} = S + SQ\bar{C}S, \quad P_{\xi v} = SQ, \quad P_{uw} = Q\bar{C}S, \quad P_{uv} = Q$$

and

$$\tilde{\xi} = (I + SQ\bar{C})J\xi(t_0) + Sr, \quad \tilde{u}_e = Q\bar{C}J\xi(t_0) + r.$$

With this change of variables the expression ξ is a linear function of variables Q and r. Thus, when we consider constraints on the estimation error over the entire time horizon $R_{0:T}\xi$ and for the final time step $R_T\xi$, the constraints will be linear in Q and r.

Deriving the Optimization Problem. Based on the requirements for equalized recovery in Definition 1, we must have $\|\xi(t)\| \leq M_2$ for all $t \in [t_0, t_0 + T]$ and $\|\xi(t_0 +$

$T)\| \leq M_1$, for the worst-case noise w, v and initial state estimation uncertainty $\xi(t_0)$. Using the aforementioned change of variables, it is understood that all constraints on estimation error are linear in our new design variables (Q, r). Furthermore, the sets that Q and r belong to are both convex ($Q \in \{Q \in \mathbb{R}^{mT \times pT} \mid Q$ is m-by-n block lower triangular$\}$ and $r \in \mathbb{R}^{mT}$). Thus, it follows that the feasibility problem in (4) is equivalent to the existence of a causal finite horizon affine estimator that solves Problem 1, similar to the robust optimization approach in Section III-C of (Skaf and Boyd (2010)).

Finally, the estimator gains F and u_0 can be recovered from the optimization variables Q and r using the equations (6). □

The feasibility problem in (4) contains semi-infinite constraints due to the "for all" quantifier on the uncertain terms. Nonetheless, it can be easily shown using techniques from robust optimization (Bertsimas et al. (2011); Ben-Tal et al. (2009)) that the feasibility problem can be robustified such that only finitely many linear constraints remain, as given in the following without proof for brevity.

Proposition 2. (Robustified Feasibility Problem). The feasibility problem in (4) is equivalent to the following:

$$\begin{aligned}
&\text{Find} && Q, r, \Pi_1, \Pi_2 \\
&\text{subject to} && Q \text{ is } m\text{-by-}n \text{ block lower triangular}, \\
& && \Pi_1 \geq 0, \Pi_2 \geq 0, \\
& && \Pi_1 \begin{bmatrix} \eta_w \mathbb{1} \\ \eta_v \mathbb{1} \\ M_1 \mathbb{1} \end{bmatrix} \leq M_2 \mathbb{1} - \begin{bmatrix} I \\ -I \end{bmatrix} R_{0:T} S r, \\
& && \Pi_2 \begin{bmatrix} \eta_w \mathbb{1} \\ \eta_v \mathbb{1} \\ M_1 \mathbb{1} \end{bmatrix} \leq M_1 \mathbb{1} - \begin{bmatrix} I \\ -I \end{bmatrix} R_T S r, \\
& && \Pi_1 \begin{bmatrix} I & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \\ 0 & 0 & -I \end{bmatrix} = \begin{bmatrix} I \\ -I \end{bmatrix} R_{0:T} G, \\
& && \Pi_2 \begin{bmatrix} I & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \\ 0 & 0 & -I \end{bmatrix} = \begin{bmatrix} I \\ -I \end{bmatrix} R_T G,
\end{aligned} \tag{10}$$

where $G \triangleq \left[(I + SQ\bar{C})S \;\; SQ \;\; (I + SQ\bar{C})J\right]$, while Π_1 and Π_2 are dual matrix variables of appropriate dimensions.

So far we have considered linear feasibility problems. One can also consider minimizing M_1, M_2 or T. The problem remains linear if M_2 is a variable to be minimized. For minimizing M_1 or T, one should resort to a combination of linear programs and a line search over M_1 or T.

3.2 Estimator Synthesis with Missing Data

Next, we extend the results in Theorem 1 to take into account the possibility of missing data. In particular, instead of probabilistic packet drops in (Sinopoli et al. (2004)), we consider missing observations with no assumed probability distribution, but are instead expressed by a fixed-length language, $\mathcal{L} \subseteq \mathbb{B}^T$, which we assume is given. For instance, a missing data specification of 'at most m missing data over a time horizon T' is given by $\mathcal{L} = \{\sigma \in \mathbb{B}^T \mid \sigma$ has at most m 0's in the first M time steps$\}$. This presents a difficulty because the nonlinear transformation that allowed us to linearize the problem in Theorem 1 was dependent on the fact that $|\mathcal{L}| = 1$. To accommodate the many different patterns that can exist in a language we introduce the following generalized inequality for any 2 words σ_1 and σ_2 in the space of words in $\mathbb{B}^T$:

$$\sigma_1 \preceq \sigma_2 \iff (\forall i \in [1, T])(\sigma_1[i] = 0 \implies \sigma_2[i] = 0).$$

Using this inequality, we derive a "worst-case" language $\mathcal{L}^* \triangleq \{\sigma^*\} \in \mathbb{B}^T$ where σ^* is the least upper bound of the set $\mathcal{L}$ according to $\preceq$. It can be computed by performing a bit-wise AND operation across all words in $\mathcal{L}$ and can be thought of as the pattern composed only of information that is always available to the designer $(\sigma^*(t) = 1 \iff (\forall \sigma \in \mathcal{L})(\sigma(t) = 1))$.

Theorem 3. [Estimator Synthesis with Missing Data] If a causal finite horizon affine estimator that solves Problem 1 (i.e. achieves equalized recovery level M_1 with recovery time T and intermediate level M_2) with single word language $\mathcal{L}^* = \{\sigma^*\}$ is feasible, then the following optimization problem has at least one feasible solution:

$$\begin{aligned}
&\text{Find} && Q, r \\
&\text{subject to} && Q \text{ is } m\text{-by-}n \text{ block lower triangular}, \\
& && \forall(\|w\| \leq \eta_w, \|v\| \leq \eta_v, \|\xi(t_0)\| \leq M_1): \\
& && \|R_{0:T}\xi^{\sigma^*}\| \leq M_2 \text{ and } \|R_T \xi^{\sigma^*}\| \leq M_1,
\end{aligned} \tag{11}$$

and the solution defines a feedback method that solves Problem 1 for the original language $\mathcal{L}$.

Proof. The proof follows nearly identical reasoning to the proof in Theorem 1, with the primary difference being that the $\bar{C}$ matrix is replaced with $\bar{C}^{\sigma^*} = \text{diag}(\sigma^*) \otimes C$. The sufficiency of $\mathcal{L}^*$ follows from the fact that, by construction, the solution (F^*, u_0^*) extracted from (11) sets $F^*_{(t,\tau)} = 0, \forall t \in [\tau, t_0 + T - 1]$, whenever $\sigma^*(\tau) = 0$, and thus, whenever $\sigma(\tau) = 0$ for any $\sigma \in \mathcal{L}$. □

Note that the optimization may not be feasible for some combinations of parameters $(M_1, M_2, T, \mathcal{L})$. This can be due to (i) infeasibility of the problem (e.g., a language constraint of $L = \{00 \cdots 0\}$ corresponds to the system running in an open loop setting and thus it may be impossible to recover for unstable systems) or (ii) the conservative nature of the "worst-case" language $\mathcal{L}^*$.

3.3 Implementation of the Estimator

Assuming that the optimization is feasible, there are two different ways to implement this estimator. First, if the missing data pattern periodically repeats itself (every T time-step), then, the same gain matrices F as in (3) can be applied with period T since the estimator guarantees that the estimation error bound returns to the equalized recovery level M_1 at the end of the period.

Second, these estimators can also be used in conjunction with a filter that guarantees equalized performance. Recall that such a filter ensures a uniform bound M on the estimation error at all times when there is no missing data. If we consider languages $\mathcal{L}$ where each word starts with a $q(t) = 0$, then one can switch from the equalized performance filter to equalized recovery filter whenever a missing measurement occurs and revert back to the equalized performance filter after the recovery time T. One can also investigate for a given language $\mathcal{L}$, the minimum time T required for recovery using the proposed optimization problems. Such T, in general, can be longer than the longest word in $\mathcal{L}$. Such analysis is helpful in understanding what type of missing data patterns can be accommodated or how much time is required for recovery, which in turn is useful for designing controllers in a compositional manner.

4. EXAMPLES

In this section we demonstrate the utility of the proposed estimators. We present motivating examples that show the benefits of the proposed synthesis method when compared to previous methods as well as show what the novel method can achieve on an adaptive cruise control system.

4.1 Conservativeness in Earlier Estimator Designs

It is shown in Mickelin et al. (2014) that if the gain L of the Luenberger observer:

$$\begin{aligned}\hat{x}(t+1) &= A\hat{x}(t) + B\hat{u}(t) + L(y(t) - \hat{y}(t)),\\ \hat{y}(t) &= C\hat{x}(t)\end{aligned} \tag{12}$$

satisfies

$$||A - LC|| + ||L||\frac{\eta_v}{M} + ||B_w||\frac{\eta_w}{M} \le 1, \tag{13}$$

the observer achieves equalized performance level M.

However, the condition (13) can be infeasible, when a Luenberger observer that achieves equalized performance *is feasible*, as shown below. This implies that the above condition is only a *sufficient* condition for the existence of observers that achieve equalized performance.

Consider a discrete time system as follows:

$$x(t+1) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -0.5 & -3.5 & -0.25 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} (u(t) + w(t)),$$

$$y(t) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} x(t)$$

with the following bounds in effect: $\|\xi(t)\| \le M$ and $\|w(t)\| \le 0.6M = \eta_w$ (with $\eta_v = 0$).

Here, the choice of L^* that minimizes the left hand side of (13) and the lack of measurement noise implies:

$$\|(A - L^*C)\| + \|L^*\|\frac{\eta_v}{M} + \frac{\eta_w}{M} = 0.5 + 0 + 0.6 = 1.1,$$

which means that (13) cannot be satisfied, even as we know that L^* creates the following estimation error update:

$$\xi(t+1) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -0.5 & 0 & 0 \end{bmatrix} \xi(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} w(t) = \begin{bmatrix} w(t) \\ 0 \\ -0.5\xi_1(t) \end{bmatrix}$$

and clearly this vector's norm should not exceed M. In this situation, our proposed synthesis method correctly identifies the feasibility of an estimator because it does not rely on an over-approximation and correctly identifies the minimum estimation error $||\xi(t+1)|| = 0.6M$.

4.2 Adaptive Cruise Control

An adaptive cruise controller (ACC) is a driver assistance system that aims to maintain a safe headway (the distance between an ego vehicle and the lead vehicle) in the existence of a lead vehicle and, if possible, drive at a set speed during operation. Consider an ACC system where the acceleration of the lead car is considered to be an uncontrolled disturbance and the controller applies force inputs to the ego vehicle. This can be written in the affine, discrete-time form:

Table 1. Constants used in the Automatic Cruise Control (ACC) Example.

m	1370 kg	T_s	0.5 s
$\bar{k}_0$	7.58 N	η_w	0.1
$\bar{k}_1$	9.9407 Ns/m	η_v	0.05

$$\begin{cases} x(t+1) &= A_d x(t) + B_d u(t) + f_d + E_d w(t) \\ y(t) &= Cx(t) + v(t) \end{cases},$$

where the state $x(t) = [v_e(t), h(t), v_L(t)]^T$ consists of the speed v_e of the ego vehicle, headway h, and speed v_L of the lead vehicle. The system matrices (A_d, B_d, f_d, E_d) are:

$$A_d = \begin{bmatrix} e^{-\kappa T_s} & 0 & 0 \\ \dfrac{e^{-\kappa T_s} - 1}{\kappa} & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix}, B_d = \frac{1}{\bar{k}_1^2}\begin{bmatrix} (1-e^{-\kappa T_s})\bar{k}_1 \\ m(1-e^{-\kappa T_s}) - \bar{k}_1 T_s \\ 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, E_d = \begin{bmatrix} 0 \\ \dfrac{T_s^2}{2} \\ T_s \end{bmatrix}, f_d = \begin{bmatrix} -\dfrac{\bar{k}_0}{\bar{k}_1}(1 - e^{-\kappa T}) \\ -\dfrac{\bar{k}_0}{\bar{k}_1^2}\left(m(1-e^{-\kappa T}) - \bar{k}_1 T\right) \\ 0 \end{bmatrix},$$

where the constant m is the mass of the vehicle, the constants $\bar{k}_0$ and $\bar{k}_1$ are coefficients related to friction and drag (with $\kappa \triangleq \bar{k}_1/m$), and T_s is the sampling time. The values of these parameters are given in Table 1.

For this problem, a reasonable assumption on the lead car (or another driver on the road) is that they limit their acceleration to a certain range for their own comfort or safety among other things. Another reasonable assumption is that our sensors have documented or known quantities such as sensitivity and discretization error (typically detailed in a component's data sheet). Assume that the maximum magnitude of acceleration that the lead car uses is 0.1 m/s^2 ($\eta_w = 0.1$) and that the maximum sensor error (consider a speedometer rated to have an upper bound of 0.01 m/s of error during operation and a radar rated with 50 cm of error) is 0.05 ($\eta_v = 0.05$).

We start with demonstrating the proposed filter for the no missing data case. We let the recovery level M_1 be 1, and find the minimal intermediate level to be $M_2 = 1.05$ with recovery time $T = 6$ (or 3 seconds). Note that this indicates the system does not admit a filter that achieves equalized performance of $M = 1$ (indeed it can be shown that it does not admit a full-order filter that achieves equalized performance of any level). Figure 1 shows the trajectories of the estimation error for random initializations and disturbances. As seen there, the estimator guarantees the prescribed bounds.

Next, consider a missing data pattern within the language $\mathcal{L}_1 = \{111111, 101111, 110111, 111011, 111101\}$, with recovery time $T = 6$ (or 3 seconds). For this example, $\mathcal{L}^* = \{100001\}$. Taking equalized recovery level of $M_1 = 1$, the minimal intermediate level is found to be $M_2 = 2.9864$ by solving (11). Assuming that the missing data pattern takes an arbitrary word from within $\mathcal{L}_1$ at every 3 seconds,

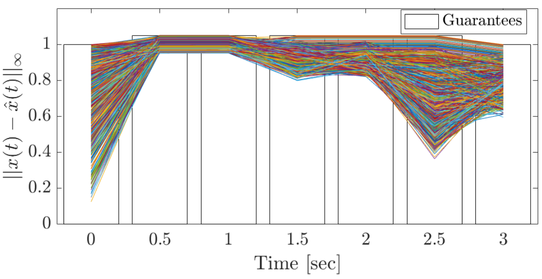

Fig. 1. A filter is synthesized that guarantees equalized recovery of the ACC system's estimate when equalized performance cannot be guaranteed.

the synthesized estimator can be used periodically. A few estimation error trajectories with the resulting estimator are depicted in Figure 2.

Finally, we consider the use of a hybrid filter in the ACC setting. This filter switches between an equalized performance filter during normal operation to an equalized recovery filter when missing data occurs. While a full state estimator cannot achieve equalized performance for the ACC system, a reduced order observer can [Mickelin and Ozay (2018)] and we use it as the baseline for the final example of this synthesis method. When estimation is restricted to the not directly measured state v_L, an equalized performance level of $M = 0.4102$ is achieved.

Assuming that the occurrence of a missing data can be detected and that the maximum duration of the missing data pattern is known to be 1.5 seconds, we synthesize one mode of the hybrid filter to achieve equalized performance and another mode which achieves equalized recovery. The minimum time required for recovery, T, is unknown, and so a set of fixed-length languages is considered that are parameterized by T, $\mathcal{L}_2(T) = \{q \in \mathbb{B}^T \mid q(t_0+1) = 0, q(\tau) = 1 \; \forall \tau \in \{t_0, t_0+4, t_0+5, ..., t_0+T-1\}\}$.[1] The synthesis methods derived in this paper are used to identify that the minimum time for recovery was 8 discrete time steps (or 4 seconds) and that the minimum intermediate level of the recovery filter is $M_2 = 0.5249$. This reduced order hybrid estimator is applied with random initializations and disturbances and resulting estimates are shown in Figure 3. In all cases, a missing data pattern from $\mathcal{L}_2(8)$ begins at time step 10 (or 5 seconds) and our guarantees hold.

5. CONCLUSIONS

In this work, we present a method for synthesizing filters for affine systems that are robust to missing data. We define the notion of equalized recovery that generalizes equalized performance. An optimization-based necessary and sufficient condition is provided to synthesize finite horizon affine filters that satisfy the equalized recovery condition for given parameters using Q-parameterization.

[1] Note that the use of a reduced order observer changes the information available to an estimator. For reduced order observers, the estimate $\hat{x}(t+1)$ at time $t+1$ depends on $y(t+1)$ and the past measurements, whereas for the full order filter in (2), $\hat{x}(t+1)$ depends on $y(t)$ and earlier measurements. This requires slightly different information (with one step look-ahead on the mode) to be available to the estimator in order to switch between the equalized performance and recovery filters. Such look-ahead is not required when this hybrid scheme is used with full-order observers in the presence of missing data.

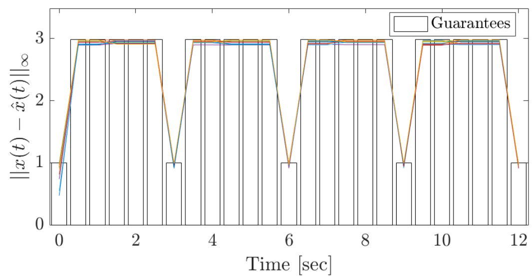

Fig. 2. A filter for the ACC system that achieves equalized recovery is synthesized to be robust against missing data within the language $\mathcal{L}_1$ and applied periodically.

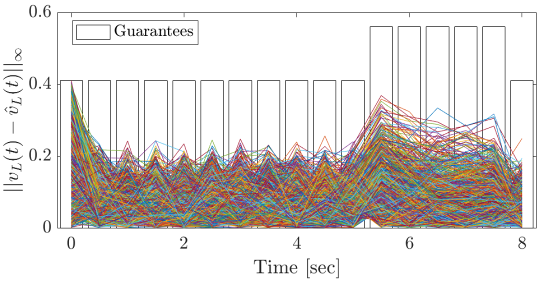

Fig. 3. A reduced order filter for ACC is synthesized to achieve equalized performance level $M_1 = 0.4102$. When a missing data sequence within $\mathcal{L}_2$ occurs, the recovery filter kicks-in and reduces the estimation error enough to then switch back to the equalized performance filter.

Our future work includes closing the loop with correct-by-construction control synthesis that guarantees safety when implemented together with filters satisfying equalized recovery condition. We are also interested in extending the framework to handle outliers or corrupted measurements with the main difficulty being separating such measurements from uncorrupted ones.

REFERENCES

Ben-Tal, A., El Ghaoui, L., and Nemirovski, A. (2009). *Robust optimization*. Princeton University Press.

Bertsimas, D., Brown, D., and Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM review*, 53(3), 464–501.

Blanchini, F. and Sznaier, M. (2012). A convex optimization approach to synthesizing bounded complexity ℓ_∞ filters. *IEEE Trans. on Automatic Control*, 57(1), 216–221.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35–45.

Mickelin, O. and Ozay, N. (2018). Correct-by-construction output feedback control synthesis for hybrid systems. (in preparation).

Mickelin, O., Ozay, N., and Murray, R.M. (2014). Synthesis of correct-by-construction control protocols for hybrid systems using partial state information. In *IEEE American Control Conference*, 2305–2311.

Milanese, M. and Vicino, A. (1991). Optimal estimation theory for dynamic systems with set membership uncertainty: an overview. *Automatica*, 27(6), 997–1009.

Shamma, J. and Tu, K.Y. (1999). Set-valued observers and optimal disturbance rejection. *IEEE Trans. on Automatic Control*, 44(2), 253–264.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., and Sastry, S.S. (2004). Kalman filtering with intermittent observations. *IEEE Trans. on Automatic Control*, 49(9), 1453–1464.

Skaf, J. and Boyd, S.P. (2010). Design of affine controllers via convex optimization. *IEEE Trans. on Automatic Control*, 55(11), 2476–2487.