Prefix-based Bounded-error Estimation with Intermittent Observations

Kwesi J. Rutledge, Sze Zheng Yong, Necmiye Ozay

***Abstract*— While observers with asymptotic convergence guarantees can be used to design output feedback controllers when considering control tasks like stability, if state constraints relevant to safety exist, it is crucial to bound the estimation error at all times. In this paper, we propose an optimization-based design technique for bounded-error state estimators for affine systems that provide estimation guarantees in the presence of intermittent measurements. We treat the affine system as a switched system where the measurement equation switches between two modes based on whether a measurement exists or is missing, and model potential intermittent measurement patterns with a finite language that constrains the feasible mode sequences. By utilizing Q-parametrization, we show that an optimal estimator can be constructed that simultaneously provides an estimate of the continuous-state and implicitly estimates the specific missing data pattern (i.e., mode sequence), within the given language, according to the prefix observed so far. We illustrate with numerical examples that this approach significantly improves the achievable estimation bounds compared to earlier work.**

I. Introduction

Control and decision-making loops in many applications may not have access to regularly sampled sensory data. One typical example of intermittent measurements is in networked control systems where sensory data is transmitted over unreliable communication networks, which are subject to packet drops [14], [12]. Another example is due to sensor glitches that may cause certain sensory readings to be unavailable occasionally, and it is desirable to have control and estimation algorithms that are robust to these temporary failures. Finally, for many autonomous systems where perception algorithms are used to provide information about the positions of external agents to the controlled agent, such position information might be missing temporarily due to classification errors in the perception algorithm or occlusion [6]. Therefore, there is a need for control and estimation algorithms robust to missing data.

When considering safety-critical applications, in addition to robustness to missing data, it is necessary to have a bound on the state estimation error when using an estimator so that safety of the state can be assessed even during the transient periods. Set-valued or set-membership estimation techniques have been proposed to obtain such non-asymptotic guarantees on the estimation error [2], [9], [11]. These techniques are also related to filters optimizing ℓ_∞-induced norm (i.e., ℓ_1-filters). A related notion is that of equalized performance ([4], [5]), which, roughly speaking, requires the estimation error at the next step to be within a given bound if the previous values of the error satisfy that bound. In our recent work [10], we proposed a generalization of this property, namely equalized recovery, that allows for a more relaxed bound on the estimation error within a time interval as long as the error gets back to its original bound at the end of the interval.

In this paper, we propose an optimization-based method to synthesize estimators that provide equalized recovery guarantees with intermittent measurements. Different frameworks exist to model intermittent measurements. Earlier works, especially those dealing with communication networks, consider probabilistic models of intermittent measurements [12], whereas recent works consider automaton-based [7] or language-based [10] models to represent feasible missing data patterns. Among these works only [10] deals with bounded-error estimation. Similar to [10], we also consider language-constraints to represent the missing data patterns. On the other hand, as opposed to the worst-case approach in [10], where robustness against the worst-case missing data pattern is considered, our proposed estimators adapt their filter gains based on the prefix of the missing data pattern observed thus far; hence, significantly improving achievable recovery levels. The design of these estimators are enabled by Q-parametrization, which is a technique used to recast the optimal control design for affine systems as a convex programming problem [13]. Although, in general, imposing additional structure on filter/controller gains in Q-parametrization-based design leads to non-convex problems, one of our main contributions is to show that the structure imposed by prefix dependency of the filter gains still leads to a convex problem. Therefore, the proposed filters not only provide significantly improved recovery levels, but also can be synthesized efficiently.

II. Notation and Preliminaries

The sets of real and binary numbers are denoted by $\mathbb{R}$ and $\mathbb{B}$, respectively. 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, $\mathbb{1}_k$ denotes a k dimensional vector of ones, and $\text{diag}(x)$ represents the $n\times n$ diagonal matrix with the elements of $x\in\mathbb{R}^n$ on the main diagonal. 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. For a (block) vector v, $(v)_k$ and $v_{i:j}$ denote its k^{th} entry, and its sub-vector consisting of entries from i^{th} to j^{th}, respectively.

KJR and NO are with the Dept. of Electrical Engineering and Computer Science, Univ. of Michigan, Ann Arbor, MI 48109, USA {krutledg,necmiye}@umich.edu. SZY is with the School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA szyong@asu.edu.

For a given set Σ, the symbol Σ^* denotes the set of all finite-length words and $\Sigma^{[T]}$ denotes the set of all words with length up to T that are formed by elements in Σ. For a word $\lambda \in \Sigma^*$, its length is denoted by $|\lambda|$. For $i \leq |\lambda|$, we use $\lambda^{[1:i]}$ to denote the length i prefix of λ. For example, if $\lambda = \sigma_0\sigma_1 \ldots \sigma_N$, then $\lambda^{[1:i]} = \sigma_0\sigma_1 \ldots \sigma_{i-1}$. Finally, the set of all non-empty prefixes of λ is denoted by $Pref(\lambda)$.

A. Properties of Block Triangular Matrices

Several important properties of block lower triangular matrices will be exploited in this work. To describe them, the following notion is introduced. Intuitively, it is very similar to the leading principal minors of a square matrix, but is applied to non-square block matrices.

Definition 1: The i^{th} *leading principal block minor* of a $l \times p$ block matrix $X \in \mathbb{R}^{al\times bp}$, written as $\mathcal{BM}_i(X)$, is the $l \times p$ block matrix:

$$\mathcal{BM}_i(X) = X([1 : il], [1 : ip])$$

for all $i \in [1, \min(a, b)]$.

Using this definition, several properties of the leading principal block matrix operator $\mathcal{BM}_i(\cdot)$ can be shown for lower block triangular matrices, as follows:

Lemma 1: Let $W, X \in \mathbb{R}^{ap\times bq}$, $Y \in \mathbb{R}^{bq\times cr}$ and $Z \in \mathbb{R}^{as\times as}$. The following properties hold:

1) $\mathcal{BM}_i(W + X) = \mathcal{BM}_i(W) + \mathcal{BM}_i(X)$;
2) If X and Y are $p\times q$ and $q\times r$ block lower triangular, respectively, then $\mathcal{BM}_i(XY) = \mathcal{BM}_i(X)\mathcal{BM}_i(Y)$;
3) If Z is nonsingular and $s \times s$ block lower triangular, then $\mathcal{BM}_i(Z^{-1}) = (\mathcal{BM}_i(Z))^{-1}$,

for all $i \in [1, \min(a, b, c)]$.

Proof: The first property is a trivial consequence of matrix addition, while the second property follows directly from multiplication of two block lower triangular matrices. Finally, the third property can be observed from the identity for partitioned matrix inversion of block lower triangular matrices. ■

Proposition 1: Let $\bar{C}^{(1)}$ and $\bar{C}^{(2)}$ be $p \times n$ block lower triangular matrices that share the same j^{th} leading principal block minor:

$$\mathcal{BM}_j(\bar{C}^{(1)}) = \mathcal{BM}_j(\bar{C}^{(2)}).$$

Also let $F^{(1)}$, $F^{(2)}$ be $n\times p$ block lower triangular matrices and let S be a $n \times n$ block lower triangular matrix, all with compatible block sizes. Define

$$Q^{(i)} \doteq F^{(i)}(I - \bar{C}^{(i)}SF^{(i)})^{-1} \tag{1}$$

for all $i \in \{1, 2\}$. Then,

$$\mathcal{BM}_j(F^{(1)}) = \mathcal{BM}_j(F^{(2)}) \in \mathbb{R}^{jn\times jp}$$

if and only if

$$\mathcal{BM}_j(Q^{(1)}) = \mathcal{BM}_j(Q^{(2)}) \in \mathbb{R}^{jn\times jp}.$$

Proof: Given $\mathcal{BM}_j(\bar{C}^{(1)}) = \mathcal{BM}_j(\bar{C}^{(2)})$, we prove both the sufficient and necessary directions:

Sufficiency: Suppose that $\mathcal{BM}_j(F^{(1)}) = \mathcal{BM}_j(F^{(2)})$. Using the fact that $\bar{C}^{(i)}$, S and $F^{(i)}$ are block lower triangular (and hence, $Q^{(i)}$ is also block lower triangular) for $i \in \{1, 2\}$ ax well as Lemma 1, we have:

$$\begin{aligned}
&\mathcal{BM}_j(Q^{(1)})\\
&= \mathcal{BM}_j[F^{(1)}(I - \bar{C}^{(1)}SF^{(1)})^{-1}]\\
&= \mathcal{BM}_j(F^{(1)})(\mathcal{BM}_j(I) - \mathcal{BM}_j(\bar{C}^{(1)})\mathcal{BM}_j(S)\mathcal{BM}_j(F^{(1)}))^{-1}\\
&= \mathcal{BM}_j(F^{(2)})(\mathcal{BM}_j(I) - \mathcal{BM}_j(\bar{C}^{(2)})\mathcal{BM}_j(S)\mathcal{BM}_j(F^{(2)}))^{-1}\\
&= \mathcal{BM}_j[F^{(2)}(I - \bar{C}^{(2)}SF^{(2)})^{-1}]\\
&= \mathcal{BM}_j(Q^{(2)}).
\end{aligned}$$

Necessity: Suppose that $\mathcal{BM}_j(Q^{(1)}) = \mathcal{BM}_j(Q^{(2)})$. First, we note the (strictly) block lower triangular properties of $\bar{C}^{(i)}$, S, $Q^{(i)}$ and $F^{(i)}$. It was shown in [13] that we can solve for $F^{(i)}$, for $i \in \{1, 2\}$ from (1) as:

$$F^{(i)} = (I + Q^{(i)}\bar{C}^{(i)}S)^{-1}Q^{(i)}.$$

Then, using the fact that $\bar{C}^{(i)}$, S, $Q^{(i)}$ and $F^{(i)}$ are block lower triangular for $i \in \{1, 2\}$ and Lemma 1, we find that:

$$\begin{aligned}
&\mathcal{BM}_j(F^{(1)})\\
&= \mathcal{BM}_j[(I + Q^{(1)}\bar{C}^{(1)}S)^{-1}Q^{(1)}]\\
&= (\mathcal{BM}_j(I) + \mathcal{BM}_j(Q^{(1)})\mathcal{BM}_j(\bar{C}^{(1)})\mathcal{BM}_j(S))^{-1}\mathcal{BM}_j(Q^{(1)})\\
&= (\mathcal{BM}_j(I) + \mathcal{BM}_j(Q^{(2)})\mathcal{BM}_j(\bar{C}^{(2)})\mathcal{BM}_j(S))^{-1}\mathcal{BM}_j(Q^{(2)})\\
&= \mathcal{BM}_j[(I + Q^{(2)}\bar{C}^{(2)}S)^{-1}Q^{(2)}]\\
&= \mathcal{BM}_j(F^{(2)}).
\end{aligned}$$

■

Proposition 2: Consider the following pairs of matrices $(\bar{C}^{(1)}, \bar{C}^{(2)})$ and $(Q^{(1)}, Q^{(2)})$ that share the same j^{th} principal leading block minor amongst each pair

$$\mathcal{BM}_j(\bar{C}^{(1)}) = \mathcal{BM}_j(\bar{C}^{(2)}),$$
$$\mathcal{BM}_j(Q^{(1)}) = \mathcal{BM}_j(Q^{(2)})$$

and consider two vectors $u_0^{(1)}$ and $u_0^{(2)}$ and a block lower triangular matrix S. Define:

$$r^{(i)} = (I + Q^{(i)}\bar{C}^{(i)}S)u_0^{(i)} \tag{2}$$

for all $i \in \{1, 2\}$. Then, the vectors $u_0^{(1)}$ and $u_0^{(2)}$ satisfy:

$$(u_0^{(1)})_k = (u_0^{(2)})_k \quad \forall k \in [1, jn]$$

if and only if the first jn entries of the vector $r^{(1)}$ is identical to that of $r^{(2)}$:

$$(r^{(1)})_k = (r^{(2)})_k \quad \forall k \in [1, jn].$$

Proof: The proof is similar to Proposition 1:

$$\begin{aligned}
r^{(1)}_{1:jn} &= [(I + Q^{(1)}\bar{C}^{(1)}S)u_0^{(1)}]_{1:jn}\\
&= \mathcal{BM}_j(I + Q^{(1)}\bar{C}^{(1)}S)[u_0^{(1)}]_{1:jn}\\
&= (\mathcal{BM}_j(I) + \mathcal{BM}_j(Q^{(1)})\mathcal{BM}_j(\bar{C}^{(1)})\mathcal{BM}_j(S))[u_0^{(1)}]_{1:jn}\\
&= (\mathcal{BM}_j(I) + \mathcal{BM}_j(Q^{(2)})\mathcal{BM}_j(\bar{C}^{(2)})\mathcal{BM}_j(S))[u_0^{(2)}]_{1:jn}\\
&= \mathcal{BM}_j(I + Q^{(2)}\bar{C}^{(2)}S)[u_0^{(2)}]_{1:jn}\\
&= r^{(2)}_{1:jn},
\end{aligned}$$

where we again applied the fact that $\bar{C}^{(i)}$, S and $Q^{(i)}$ are block lower triangular for $i \in \{1, 2\}$ and Lemma 1. Opposite direction is similar. ■

III. Problem Setup

In this paper, we consider estimation problems for affine systems subject to missing measurements. We use the following two-mode switched system to represent the dynamics and measurement updates:

$$\begin{aligned} x(t+1) =& \ Ax(t)+Bu(t)+f+w(t), \\ y(t) =& \begin{cases} Cx(t)+v(t), & q(t)=1, \\ \emptyset, & q(t)=0, \end{cases} \end{aligned} \tag{3}$$

where A, B, C, f are known system matrices, $x(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ is the continuous state, $u(t) \in \mathcal{U} \subseteq \mathbb{R}^m$ is the input, $w(t) \in \mathcal{W} \subseteq \mathbb{R}^n$ is the process noise, $y(t) \in \mathcal{Y} \subseteq \mathbb{R}^p \cup \{\emptyset\}$ is the output measurements of the system, $q(t) \in \mathbb{B}$ is the discrete state/mode of the hybrid system, with $q(t) = 1$ denoting that the measurement vector is available and $q(t) = 0$ corresponding to "missing" data, and $v(t) \in \mathcal{V} \subseteq \mathbb{R}^p$ is the measurement noise. The noise terms $w(t)$ and $v(t)$ are unknown but bounded, and their bounds are known (i.e., $\mathcal{W} = \{w \in \mathbb{R}^n \mid \|w\| \leq \eta_w\}$ and $\mathcal{V} = \{v \in \mathbb{R}^p \mid \|v\| \leq \eta_v\}$).

Remark 1: The assumption that the noise terms w and v in (3) have an identity gain in front of them is just 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., $\bar{B}_w w(t)$, $\bar{C}_v v(t)$), it is straightforward to incorporate this in the proposed methodology.

We further impose a constraint on the allowable missing data patterns. This constraint is modeled by a fixed-length language $\mathcal{L} \subseteq \mathbb{B}^T$ that specifies the set of allowable mode sequences $\{q(t)\}_{t=t_0}^{t_0+T-1}$.

Our goal is to design an estimator

$$\mathcal{O} : (\mathcal{U} \times \mathcal{Y})^* \to \mathcal{X} \tag{4}$$

that given the input output data so far, generates an estimate of the state. In particular, we are interested in bounded error estimators that satisfy the following equalized recovery condition proposed in [10].

Definition 2 (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 for any estimation error $\xi(t) \doteq x(t) - \hat{x}(t)$ with $\|\xi(t_0)\| \leq M_1$, we have $\|\xi(t)\| \leq M_2$ for all $t \in [t_0, t_0+T]$ and $\|\xi(t_0+T)\| \leq M_1$.

If we consider the special case of achieving equalized recovery level M with recovery horizon one and intermediate level M, this is equivalent to equalized performance, proposed earlier in the literature [5], [8].

Now we are ready to state the problem of interest.

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 (3),
- 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 $\mathcal{O} : (\mathcal{U} \times \mathcal{Y})^{[T]} \to \mathcal{X}$ such that $\|\xi(t)\| \leq M_2$ for all $t \in [t_0, t_0+T]$ and $\|\xi(t_0+T)\| \leq M_1$.

We consider a finite horizon dynamic estimator with the following update rules:

$$\hat{x}(t+1) = A\hat{x}(t) + Bu(t) - u_e(t) + f, \tag{5}$$

where $u_e(t)$ is an output injection term to be designed. In our earlier work [10], for a single missing data pattern (i.e., $|\mathcal{L}| = 1$), we consider the following causal output injection term:

$$u_e(t) = u_0(t) + \sum_{\tau=t_0}^{t} F_{(t,\tau)} y_\xi(\tau), \tag{6}$$

where $y_\xi(\tau) \triangleq \begin{cases} y(\tau) - C\hat{x}(\tau), & q(\tau)=1, \\ 0, & q(\tau)=0, \end{cases}$
and t_0 is the initial time of the finite horizon. A necessary and sufficient condition for solving Problem 1 with a singleton language is given for this class of estimators in terms of a convex feasibility problem in [10]. In case the problem has a solution, the convex program returns the set of estimator gains $u_0(t)$ and $F_{(t,\tau)}$ for $t \in [t_0, t_0+T]$ and $\tau \in [t_0, t]$. The approach is generalized to missing data patterns represented by $\mathcal{L}'$ with $|\mathcal{L}'| > 1$ by considering a worst-case word $\lambda^* \in \mathbb{B}^T$ such that a solution for Problem 1 with $\mathcal{L}^* = \{\lambda^*\}$ is also a solution for the same problem with $\mathcal{L}'$.

We now illustrate the limitations of this approach from [10] with an example language and propose a more general output injection term. Let $\mathcal{L}' = \{\lambda_1, \lambda_2\}$ with $\lambda_1 = 1011$ and $\lambda_2 = 1101$. The worst-case word for $\mathcal{L}'$ is $\lambda^* = 1001$. Obviously, if we can find estimator gains that achieve equalized recovery with levels M_1 and M_2 when missing both the second and third measurements (i.e., $\mathcal{L}^* = \{\lambda^*\}$), this same estimator achieves the same levels both for λ_1 and λ_2. On the other hand, assume there exists a set of estimator gains that achieve equalized recovery with levels M_1 and M_2 for λ_1 and another set of gains that achieves the same levels for σ_2. While this does not imply existence of an estimator for λ^* (conservativeness of [10]), it does not imply the feasibility of Problem 1 either. To see the latter, observe that at time t_0 with $q(t_0) = 1$, we have no way to know if λ_1 or λ_2 will be the upcoming pattern. Therefore, if the gain pairs "$u_0(t_0)$, $F_{(t_0,t_0)}$" corresponding to λ_1 and λ_2 are not the same, we have no way to compute $u_e(t_0)$ in (6) that will guarantee equalized recovery with the desired level. In other words, existence of estimator gains for λ_1 and λ_2 separately does not guarantee that there is a causally implementable filter.

In order to reduce conservativeness while preserving causality, in this paper we propose prefix-based estimators that use the following output injection term at time t:

$$u_e(q_{t_0:t}) = u_0(t, q_{t_0:t}) + \sum_{\tau=t_0}^{t} F_{(t,\tau,q_{t_0:t})} y_\xi(\tau), \tag{7}$$

where $y_\xi(t)$ is defined as before. Since we know that $q_{t_0:t_0+T-1} \in \mathcal{L}$, u_e should be defined for all prefixes in $\bigcup_{\lambda \in \mathcal{L}} Pref(\lambda)$. For our example in the previous paragraph, we need filter gains for prefixes in $\{1, 10, 11, 100, 110, 1011, 1101\}$. Our main result is to show that prefix-based estimators, i.e., the filter gains in (7), which achieve a given equalized recovery specification can be computed efficiently using convex programming.

We call the estimator (5) with output injection mechanism (6), a *time-based estimator* and the new estimator (5) with (7), a *prefix-based estimator*. While the time-based estimators use the available (non-missing) output history for feedback, prefix-based estimators use both the output history and the discrete-state history. By its definition, it essentially also performs estimation at the discrete-level (or online model detection) to detect which missing data patterns in $\mathcal{L}$ may be active and adapts the filter gains accordingly. Whereas, the time-based estimator is agnostic to the missing data pattern and tries to be robust rather than adaptive. The next proposition formally captures the fact that prefix-based estimators are more general than time-based estimators.

Proposition 3: For any time-based estimator for the dynamical system in (3) with missing data pattern given by a fixed-length language $\mathcal{L}$, identical performance can be obtained using a prefix-based estimator.

Proof: Let the output injection term for the time-based estimator be

$$u_e(t) = \bar{u}_0(t) + \sum_{\tau=t_0}^{t} \bar{F}_{(t,\tau)} y_\xi(\tau). \tag{8}$$

Define the filter gains of the prefix based estimator's output injection term in (7) as $u_0(t,\lambda) \doteq \bar{u}_0(t)$, $F_{(t,\tau,\lambda)} \doteq \bar{F}_{(t,\tau)}$ for all $t \in [t_0, t_0+T]$, $\tau \in [t_0, t]$ and for all $\bar{\lambda} \in \bigcup_{\lambda\in\mathcal{L}} Pref(\lambda)$. Then the two estimators are equivalent. ■

IV. Synthesis of Prefix-based Estimators

In this section, we discuss how to synthesize a prefix-based estimator using robust linear programming and present a necessary and sufficient condition for the existence of an estimator that uses prefix-based feedback of the form in (7) and solves Problem 1.

Theorem 1: Given a prefix-based estimator with the output injection term (7), we associate with it block matrices $\{(F^{(i)}, u_0^{(i)})\}_{i=1}^{|\mathcal{L}|}$ formed from the filter gains, where for all $\lambda_i \in \mathcal{L}$, the (j,k) block entry $(F^{(i)})_{jk}$ of $F^{(i)}$ is defined as

$$(F^{(i)})_{jk} \doteq F_{(t_0+j-1, t_0+j-k-1, \lambda_i^{[1:j]})} \tag{9}$$

$\forall k \in [1,j]$,$\forall j \in [1,T]$, and $(F^{(i)})_{jk} = 0$ otherwise; and the jth block entry of the feedforward term $u_0^{(i)}$ is defined as

$$(u_0^{(i)})_j \doteq u_0(t_0+j-1, \lambda_i^{[1:j]})$$

$\forall j \in [1,T]$. Let

$$\bar{C}^{(i)} \doteq \left[\text{diag}(\lambda_i) \otimes C \;\; 0_{|\lambda_i|T\times n}\right],$$
$$S \doteq \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ I_n & 0 & 0 & \cdots & 0 \\ A & I_n & 0 & \cdots & 0 \\ A^2 & A & I_n & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ A^{T-1} & A^{T-2} & A^{T-3} & \cdots & I_n \end{bmatrix}. \tag{10}$$

Then, equations (1) and (2) define a bijection such that any estimator $\{(F^{(i)}, u_0^{(i)})\}_{i=1}^{|\mathcal{L}|}$ is paired with one and only one element in the polyhedral set:

$$\mathcal{Q}(\mathcal{L}) \doteq \left\{ \begin{array}{c|l} \{(Q^{(i)}, & Q^{(i)} \text{ is block lower diagonal } \forall i \\ r^{(i)})\}_{i=1}^{|\mathcal{L}|} & (p \in Pref(\lambda_i) \wedge p \in Pref(\lambda_j)) \implies \\ & \quad (\mathcal{BM}_{|p|}(Q^{(i)}) = \mathcal{BM}_{|p|}(Q^{(j)})) \wedge \\ & \quad ((r^{(i)})_{1:|p|m} = (r^{(j)})_{1:|p|m}) \\ & \quad \forall \lambda_i, \lambda_j \in \mathcal{L} \end{array} \right\}. \tag{11}$$

Proof: This follows directly from Propositions 1 and 2 as well as the invertibility of the mappings (1) and (2). ■

Theorem 2: There exists a prefix-based estimator (i.e., $\{(F^{(i)}, u_0^{(i)})\}_{i=1}^{|\mathcal{L}|}$) that satisfies equalized recovery with parameters $(M_1, M_2, \mathcal{L})$ if and only if the following robust linear programming problem is feasible:

$$\text{Find} \quad \left\{(Q^{(i)}, r^{(i)})\right\}_{i=1}^{|\mathcal{L}|} \in \mathcal{Q}(\mathcal{L}) \tag{12a}$$

$$\text{subject to} \quad \begin{array}{c} \forall(\|w\| \le \eta_w, \|v\| \le \eta_v, \|\xi(t_0)\| \le M_1): \\ \|\xi^{(i)}\| \le M_2 \text{ and } \|\left[0_{n\times nT} \;\; I_n\right]\xi^{(i)}\| \le M_1, \end{array} \quad \forall i \in [1, |\mathcal{L}|], \tag{12b}$$

where

$$\begin{aligned} \xi^{(i)} &= (S + SQ^{(i)}\bar{C}^{(i)}S)w + SQ^{(i)}N^{(i)}v \\ &\quad + (I + SQ^{(i)}\bar{C}^{(i)})J\xi^{(i)}(t_0) + Sr^{(i)}, \end{aligned} \tag{13}$$

$\bar{C}^{(i)}$ and S are defined in (10), $N^{(i)} \doteq \text{diag}(\lambda_i) \otimes I$, (14)

$$J \doteq \begin{bmatrix} I_n \\ A \\ \vdots \\ A^{T-1} \\ A^T \end{bmatrix}. \tag{15}$$

Proof: For a given prefix-based feedback law $\{(F^{(i)}, u_0^{(i)})\}_{i=1}^{|\mathcal{L}|}$, the trajectory $\xi^{(i)} = [\xi^{(i)}(t_0)^\intercal, \ldots, \xi^{(i)}(t_0+T)^\intercal]^\intercal$ of the estimation error under the i^{th} missing data pattern can be written as a nonlinear function of $\{(F^{(i)}, u_0^{(i)})\}_{i=1}^{|\mathcal{L}|}$ just by plugging in the output injection term (7) in (5) and computing the error. After applying a change of variables via the mapping in Theorem 1, we can express $\xi^{(i)}$ as a linear function of $\{(Q^{(i)}, r^{(i)})\}_{i=1}^{|\mathcal{L}|} \in \mathcal{Q}(\mathcal{L})$ as in (13). Since the equalized recovery condition can also be written as linear constraints in $\xi^{(i)}$ that should hold for all initial estimation errors satisfying M_1 bound and for all possible noise values, problem (12) is a robust linear program, whose feasibility is equivalent to the existence of the desired estimator. Finally, the gains of the prefix-based feedback law are obtained by applying the inverse of the mapping in Theoreom 1. ■

Remark 2: Per a similar argument to Theorem 2's proof, finding an estimator that minimizes the intermediate level M_2 subject to a given equalized recovery level M_1 and given missing data language $\mathcal{L}$ can be posed as a robust linear program over the decision variables $\left\{(Q^{(i)}, r^{(i)})\right\}_{i=1}^{|\mathcal{L}|}$ and M_2.

Since the feasibility problem in (12) contains semi-infinite constraints due to the "for all" quantifier on the uncertain terms, the problem is not readily solvable. However, as in [10], techniques from robust optimization and duality [1], [3] can be applied to obtain a linear programming (LP) problem with only finitely many linear constraints. In particular, we have the following theorem:

Theorem 3: There exists a prefix-based finite horizon affine estimator of the form (7) that satisfies equalized recovery with parameters $(M_1, M_2, \mathcal{L})$ if and only if the following linear program is feasible:

$$
\begin{aligned}
&\text{Find} \quad \{(Q^{(i)}, r^{(i)})\}_{i=1}^{|\mathcal{L}|} \in \mathcal{Q}(\mathcal{L}),\ \left\{(\Pi_1^{(i)}, \Pi_2^{(i)})\right\}_{i=1}^{|\mathcal{L}|} \\
&\text{subject to } \forall i \in [1, |\mathcal{L}|], \\
&\quad \Pi_1^{(i)} \geq 0, \Pi_2^{(i)} \geq 0, \\
&\quad \Pi_1^{(i)} \begin{bmatrix} \eta_w \mathbb{1} \\ \eta_v \mathbb{1} \\ M_1 \mathbb{1} \end{bmatrix} \leq M_2 \mathbb{1} - \begin{bmatrix} I \\ -I \end{bmatrix} S r^{(i)}, \\
&\quad \Pi_2^{(i)} \begin{bmatrix} \eta_w \mathbb{1} \\ \eta_v \mathbb{1} \\ M_1 \mathbb{1} \end{bmatrix} \leq M_1 \mathbb{1} - \begin{bmatrix} I \\ -I \end{bmatrix} R_T S r^{(i)}, \\
&\quad \Pi_1^{(i)} \Omega = \begin{bmatrix} I \\ -I \end{bmatrix} G^{(i)},\ \Pi_2^{(i)} \Omega = \begin{bmatrix} I \\ -I \end{bmatrix} R_T G^{(i)},
\end{aligned} \tag{16}
$$

where $\bar{C}^{(i)}$, $N^{(i)}$, J, and S are as defined in (14) and (15), and

$$
G^{(i)} = \left[(I + SQ^{(i)}\bar{C}^{(i)})S \ \ SQ^{(i)}N^{(i)} \ \ (I + SQ^{(i)}\bar{C}^{(i)})J\right],
$$

$$
\Omega \doteq \begin{bmatrix} I & 0 & 0 \\ -I & 0 & 0 \\ 0 & I & 0 \\ 0 & -I & 0 \\ 0 & 0 & I \\ 0 & 0 & -I \end{bmatrix}.
$$

V. Discussion and Extensions

Assuming that the optimization problem above is feasible, there are multiple scenarios in which this prefix-based estimator can be applied. First, if the estimation problem under consideration is one that is for a finite horizon, the estimators apply directly. Second, if the missing data pattern repeats itself with a period of T time-steps, then, the same estimator can similarly be used with period T since the estimator guarantees that the estimation error bound always recovers at the end of the period to the equalized recovery level M_1.

Moreover, as in [10], the prefix-based estimators can be used in conjunction with a filter that guarantees equalized performance. In particular, if we consider languages $\mathcal{L}$ with words that start with a $q(t) = 0$, then we can switch from the equalized performance estimator to equalized recovery estimator whenever a missing measurement occurs and revert back to the equalized performance estimator after the recovery time T.

It is easy to extend the algorithm to any finite-length language $\mathcal{L}$ (as opposed to fixed-length). This just requires defining $\xi^{(i)}$ in the optimization problem (12) to be of length $|\lambda_i|$ and M_1 bound is enforced at time $|\lambda_i|$ instead of at time T. Furthermore, for a given language $\mathcal{L}$, we can find the minimum time T, which can be longer than the longest word in $\mathcal{L}$ assuming no missing measurements will occur in the extension, that is required for equalized recovery when using the proposed prefix-based optimization formulation. This analysis is helpful for understanding how much time is required for recovery with different types of missing data patterns, which in turn is useful for designing controllers in a compositional manner.

TABLE I

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

VI. Example: Adaptive Cruise Control

In this section, we demonstrate the superiority of the proposed prefix-based estimator on an automotive problem that was considered in previous works. This motivating example shows that the smallest level of estimation error that we can guarantee using these methods is drastically improved for a realistic scenario when using the prefix-based estimator instead of a time-based one.

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 presence of a lead vehicle and, if possible, drive at a set speed during operation. Let the acceleration of the lead car in the ACC system be an uncontrolled disturbance and let the inputs that the controller manipulates be force inputs to the ego vehicle. This can be written in the affine, discrete-time form:

$$
\begin{aligned}
x(t+1) &= Ax(t) + Bu(t) + f + Ew(t), \\
y(t) &= Cx(t) + v(t),
\end{aligned}
$$

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, B, C, E, f) are:

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

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

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 I.

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 ($\eta_v = 0.05$).

The set of feasible missing data patterns that we consider for this problem is $\mathcal{L}_1 = \{101111, 110111, 111011, 111101\}$. This may be written in plain english as an 'only 1 piece of data can be missing in a given time window' constraint and would be translated, according to the worst case language approach, to a

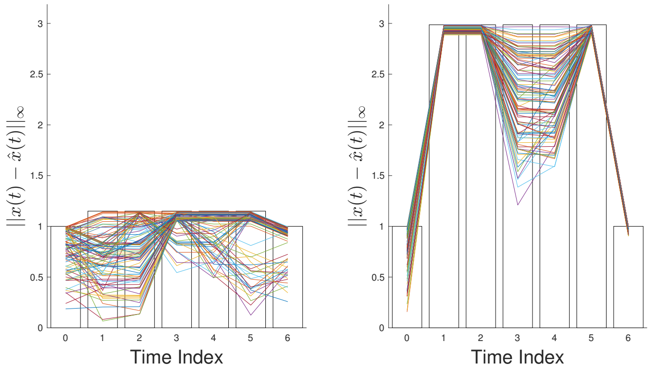

Fig. 1. Estimation error trajectories for the adaptive cruise control system with prefix-based (left) and time-based (right) estimators for 100 different initial conditions and with measurements that experience missing data events according to $\mathcal{L} = \mathcal{L}_1$. For both cases, an estimator that minimizes the intermediate level M_2 is synthesized given the partial equalized recovery specification with $M_1 = 1$ and $\mathcal{L} = \mathcal{L}_1$. The guaranteed estimation errors for each estimator are shown with the clear bars in each plot. Observe that the prefix-based estimator's intermediate level M_2 is much lower than that of the time-based estimator's.

constraint $\mathcal{L}_1^* = \{100001\}$. For a given set of parameters, or guarantees that the system would like to satisfy, it is obviously harder to guarantee them for $\mathcal{L}_1^*$ than any word in $\mathcal{L}_1$ and we would like for the optimal M_2 value that we find to reflect that.

The two approaches presented can be used to find the optimal guarantee for M_2 through a straight forward convex optimization problem and an optimal M_1 through line search. One natural question about the two methods that we can ask is which provides "better," in this case meaning smaller, bounds on the estimation error. The intuitive answer is that it is the estimator with the prefix-based feedback law and we illustrate this with Figure 1. The optimal M_2 that can be guaranteed for the language $\mathcal{L}_1$ is $M_2^* = 1.1490$ when using a prefix-based observer, nearly a third of what can be done with the time-based observer (recall that that number was $M_2^* = 2.9864$ in [10]), which can only consider the worst-case language $\mathcal{L}_1^*$. The two synthesis problems were completed using YALMIP with the Gurobi solver on an iMac with 3.4 GHz Intel Core i5 processor. The prefix-based problem was solved in 0.11s (773 iterations) while the time-based problem was solved in 0.05s (122 iterations). The designs that each optimization created were then applied to 100 sets of random disturbance trajectories and initial conditions $(w, v, \xi(0))$ that satisfied the assumptions of the problem and the resulting estimation error trajectories are shown to lie within the domain of the bars (guarantees) in Figure 1.

VII. Conclusions

In this work, we presented a method for synthesizing bounded-error estimators for affine systems that provide equalized recovery guarantees even in the presence of missing data, where the missing data patterns are constrained by a finite-length language. Our proposed optimal estimator leveraged Q-parametrization as well as some additional structure in our problem to provide an estimate of the continuous-state while implicitly estimating the specific missing data pattern (i.e., mode sequence), within the given language, based on the observed history of the missing data pattern. Using numerical examples, we demonstrated that this approach significantly improves the achievable estimation bounds compared to our earlier work that did not implicitly estimate the mode sequence.

Our future work includes correct-by-construction control synthesis with output feedback for safety applications, where state estimators satisfying equalized recovery are implemented to convert the control synthesis problem into one with state feedback subject to additional bounded disturbance/noise due to the estimation errors. We are also interested in extending the framework to detect and handle outliers and/or corrupted measurements.

Acklowledgments: 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, 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.

References

[1] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[2] D. Bertsekas and I. Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Transactions on Automatic Control*, 16(2):117–128, 1971.

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

[4] F. Blanchini and M. Sznaier. A convex optimization approach to fixed-order controller design for disturbance rejection in siso systems. *IEEE Transactions on Automatic Control*, 45(4):784–789, April 2000.

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

[6] K. Habib. ODI Resume Report on Investigation PE 16-007 concerning Tesla automatic vehicle control systems. Technical report, NHTSA Office of Defects Investigation, 2016.

[7] R. M. Jungers, A. Kundu, and W. P. M. H. Heemels. Observability and controllability analysis of linear systems subject to data losses. *IEEE Transactions on Automatic Control*, 2018.

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

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

[10] K. J. Rutledge, S. Z. Yong, and N. Ozay. Optimization-based design of bounded-error estimators robust to missing data. *IFAC-PapersOnLine*, 51(16):157 – 162, 2018. 6th IFAC Conference on Analysis and Design of Hybrid Systems ADHS 2018.

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

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

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

[14] W. Zhang, M. S. Branicky, and S. M. Phillips. Stability of networked control systems. *IEEE Control Systems*, 21(1):84–99, 2001.