

State Estimation for Linear Systems with State Equality Constraints^{*}

Sangho Ko and Robert R. Bitmead,

*Department of Mechanical and Aerospace Engineering, University of California, San Diego,
9500 Gilman Drive, La Jolla, California, 92093-0411, USA.*

Abstract

This paper deals with state estimation problem for linear systems with state equality constraints. Using noisy measurements which are available from the observable system, we construct the optimal estimate which also satisfies linear equality constraints. For this purpose, after reviewing modeling problems in linear stochastic systems with state equality constraints, we formulate a projected system representation from a descriptor system form. By using the constrained Kalman filter for the projected system and comparing its filter Riccati Equation with those of the unconstrained and the projected Kalman filters, we reach the conclusion that the current constrained estimator outperforms other filters for estimating the constrained system. We extend the same procedures from discrete-time to the continuous-time case. Finally, a numerical example is presented, which demonstrates performance differences among those filters.

Key words: Estimation; Constraints; Kalman filters; Projection; Descriptor systems

1 Introduction

Control with constraints is increasingly applied in industry (Maciejowski, 2002) and the practicality of incorporating both state and input constraints into control problems via MPC (Model Predictive Control) methods is in large part responsible for the recent upsurge in interest in these latter methods. Using explicit constraints, in place of their implicit inclusion using penalty and barrier methods, simplifies the design specification to focus on the performance objective. Constraints in control, particularly optimal control, have a long history (Bryson and Ho, 1975) and have focused on full state feedback systems. The constraints can be of two basic types: *physical constraints* reflecting known limits to physical state variables, such as positivity of mass or pressure; and *design constraints* which represent desired operating limits which might otherwise be violated by the controlled system. Forbidden, as opposed to undesirable, state motions may be incorporated into system descriptions through descriptor system representations.

Estimation of systems having state constraints has drawn many practitioners' attention in diverse engineering disciplines: radio surgery (Altman and Tombropoulos, 1994) for maximizing dose to the tumor and minimizing dose to healthy tissues, attitude determination of spacecraft (Lefferts *et al.*, 1982; Peng *et al.*, 2000; Chiang *et al.*, 2002)

^{*} Corresponding author Robert R. Bitmead. Tel. +1-858-822-3477. Fax +1-858-822-3107.

Email addresses: sko@ucsd.edu (Sangho Ko), rbitmead@ucsd.edu (Robert R. Bitmead).

for constraining the unit norm of quaternion, robotic systems for multi-sensor data fusion (Wen and Durrant-Whyte, 1992) and locating objects (Geeter *et al.*, 1997), target tracking of air vehicles (Tahk and Speyer, 1990; Alouani and Blair, 1993) and land-vehicles (Simon and Chia, 2002) to implement kinematic constraints, and adaptive beam-forming (Chen and Chiang, 1993; Hayward, 1998) to avoid the problem of signal distortion along the look direction occurring in the unconstrained Kalman beam-former. Each of these works is an example of dealing with physical constraints in the form of *equality* state constraints in probabilistic framework. In the mainstream MPC literature, Rao (2000) and Rao *et al.* (2003) have developed a deterministic constrained state estimate which is based on a finite-horizon optimization. This deterministic approach does not yield an estimate quality measure of closeness to the actual state.

The problem which we consider here is the concomitant state estimation problem in constrained systems. Noisy measurements are available from the observable system and we desire to use these optimally to reconstruct a state estimate, which also is known to satisfy linear equality constraints. For this purpose, we formulate linear stochastic systems with linear equality constraints firstly in a stochastic descriptor system representation from which the projected system representation will be derived. Then, we construct the Kalman filter for the projected system and show that the current constrained estimator outperforms the other estimators available for constrained systems.

From the viewpoint of methods of implementing state equality constraints, regardless of whether the constraint is linear or nonlinear, they may be classified largely into the method of *soft* (or *weak*) and *hard* (or *strong*) constraints:

Soft Constraints

An equality constraint is termed “soft” or “weak” when a relation between states is only approximately known (Geeter *et al.*, 1997) and hence uncertainties in the physical adherence to the constraints are allowed (Tahk and Speyer, 1990). For example, suppose that we have static relations between some components of state \mathbf{x} described by

$$\mathbf{C}(\mathbf{x}) = \mathbf{0} \tag{1}$$

which is, however, not exactly known, and hence we may allow uncertainty in the constraint (1) by adding a noise $\boldsymbol{\mu}$ yielding

$$\mathbf{C}(\mathbf{x}) + \boldsymbol{\mu} = \mathbf{0}. \tag{2}$$

Noise $\boldsymbol{\mu}$ is generally assumed to be a zero-mean Gaussian random vector with appropriate covariance which reflects the level of uncertainty in the constraint. By considering $\mathbf{0}$ and $\boldsymbol{\mu}$ as a “measurement” and a “measurement noise”, respectively, we can modify the unconstrained Kalman filter estimate to which the constraint is not yet applied. Since the constraint is used as a measurement, this method is also termed the “pseudo-measurement” method. This was employed in applications by Altman and Tombropoulos (1994), Geeter *et al.* (1997), Tahk and Speyer (1990), Alouani and Blair (1993), Chen and Chiang (1993), and Hayward (1998), where, instead of the original nonlinear constraints, at each time step linearized (or linear in Chen and Chiang (1993), Hayward (1998)) constraints were implemented in the Kalman filter. Since this method increases the number of measurement equations, it causes some computational burden for inverting the covariance matrix when implementing the Kalman filter.

Hard Constraint

By comparison to the soft constraint, a “hard” or “strong” constraint is used when some relations between state variables are known exactly and hence (1) is used to describe the relation. This type of constraint also can be incorporated into the conventional Kalman filter through *perfect* pseudo-measurements similarly to the soft constraint case, but this case results in a singular measurement noise covariance increasing the possibility of numerical problems (Stengel, 1994). For linear equality constraints, it is always possible to reduce the system model parametrization and use the reduced state equation and the conventional Kalman filter. On the other hand, there are several good reasons for not using a reduced state space for treating the constrained system. Firstly, the dimension of the reduced state

space may vary between systems of different dimensionality, such as is the case for locomotion systems (Hemami and Wyman, 1979). Secondly, the reduction of the state equations makes their interpretation less natural and more difficult (Simon and Chia, 2002). The most recent hard constraints method with non-reduced form of the state equation is based on the projection method in which the constrained estimate is obtained from the unconstrained estimate of the conventional Kalman filter by projecting onto the subspace or the surface described by the hard constraint represented by (1). Recently, Mahata and Söderstrom (2004) used a similar approach in estimating deterministic parameters of viscoelastic materials. Here the additional linear constraints are imposed in the form of a partially known boundary condition to obtain better estimates.

Chia (1985), and Simon and Chia (2002) used this method with the Kalman prediction/filtering and proved that the state estimation error covariance of the projected estimate is smaller than that of the unconstrained estimate. Wen and Durrant-Whyte (1992) independently developed a similar method by firstly including the constraint as a perfect observation and then showing that their method is theoretically exactly the same as projecting the filtered estimate of the unconstrained Kalman filter onto the surface of the hard constraint represented by (1). In conventional linear stochastic models with additive white process noise, for a state vector to be constrained “strongly” in a proper subspace of the whole state space, the process noise must have a singular covariance consistent with a linear constraint on the state. However, because these authors first used a positive definite or a bigger process noise covariance for estimating the constrained system and then projected onto the constraint surface, their method is not optimal, which we show in later sections.

In the next section we consider possible models for linear stochastic systems with equality state constraints and then represent the system in a descriptor form and also in a projected form. In Section 3 we deal with the unconstrained Kalman filter and in Section 4 we consider two constrained estimators and compare these in terms of error covariance. In Section 5, we extend the same procedures applied to discrete systems for the continuous-time case. Finally a numerical example is presented in Section 6.

Notations: In this paper, matrices will be denoted by upper case boldface (e.g., \mathbf{A}), linear spaces are denoted by calligraphic uppercase (e.g., \mathcal{A}), column matrices (vectors) will be denoted by lower case boldface (e.g., \mathbf{x}), and scalars will be denoted by lower case (e.g., y) or upper case (e.g., Y). For a matrix \mathbf{A} , \mathbf{A}^T denotes its transpose and \mathbf{A}^\dagger represents the Moore-Penrose inverse of \mathbf{A} , and $\mathcal{N}(\mathbf{A})$ denote the null space of \mathbf{A} . For a symmetric matrices $\mathbf{P} > \mathbf{0}$ or $\mathbf{P} \geq \mathbf{0}$ denotes the fact that \mathbf{P} is positive definite or positive semi-definite, respectively. For a random vector \mathbf{x} , $\epsilon\{\mathbf{x}\}$ represents the mathematical expectation of \mathbf{x} .

2 Linear Stochastic Systems with Equality Constraints

We investigate a method of estimating the state of systems modeled by a linear stochastic difference equation of the form

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \quad (3)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \quad (4)$$

where the state $\mathbf{x}_k \in \mathbb{R}^n$ is known to be constrained in the null space of \mathbf{D}

$$\mathcal{N}(\mathbf{D}) \triangleq \{\mathbf{x} : \mathbf{D}\mathbf{x} = \mathbf{d} = \mathbf{0}\} \quad (5)$$

and $\mathbf{y} \in \mathbb{R}^p$ represents the measurement. For the $\mathbf{d} \neq \mathbf{0}$ case, translation of the state space such that $\mathbf{x}_k = \bar{\mathbf{x}}_k + \mathbf{D}^\dagger \mathbf{d}$ yields the state equation form (3) with an additional term in the state equation which can be treated as a deterministic

constant noise. Hence, without loss of generality, it suffices to consider $\mathbf{d} = \mathbf{0}$ case for the analysis of current state estimation problem. The noise sequences $\mathbf{w}_k \in \mathbb{R}^n$ and $\mathbf{v}_k \in \mathbb{R}^p$ are of zero-mean white gaussian distribution with

$$\mathcal{E} \left\{ \begin{bmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{bmatrix} \begin{bmatrix} \mathbf{w}_j \\ \mathbf{v}_j \end{bmatrix}^T \right\} = \begin{bmatrix} \mathbf{Q}^c & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \delta_{kj} \quad (6)$$

where δ_{kj} represents the Kronecker delta ($\delta_{kj} = 1$ if $k = j$, 0 otherwise). The matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$ is assumed to have a full row rank. If \mathbf{D} does not have a full row rank, there exist redundant state constraints. In that case we can simply remove linearly dependent rows from \mathbf{D} .

Since the allowable space of \mathbf{x}_k with constraint is $\mathcal{N}(\mathbf{D})$, \mathbf{x}_{k+1} given by (3) also must satisfy the constraint, that is, $\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \in \mathcal{N}(\mathbf{D})$, for which we identify the following possible cases:

Case 1: $(\mathbf{A}\mathbf{x}_k, \mathbf{B}\mathbf{u}_k, \mathbf{w}_k) \notin \mathcal{N}(\mathbf{D})$

Since the sum of the three elements must satisfy $\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \in \mathcal{N}(\mathbf{D})$ for any $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, this case causes a correlated noise \mathbf{w}_k with current input \mathbf{u}_k and state \mathbf{x}_k and causes theoretical as well as practical problems (Rao, 2000). Hence it is not a proper Markovian system model.

Case 2: $(\mathbf{A}\mathbf{x}_k, \mathbf{B}\mathbf{u}_k) \notin \mathcal{N}(\mathbf{D})$ but $\mathbf{w}_k \in \mathcal{N}(\mathbf{D})$

Similarly, the sum of the first two elements must satisfy $\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \in \mathcal{N}(\mathbf{D})$. This case allows uncorrelated noise sequences \mathbf{w}_k with the input \mathbf{u}_k or the state \mathbf{x}_k , but the system cannot maintain the state constraint without corrective action of the input \mathbf{u}_k . Hence, this model is suitable for modeling systems having *design constraints*.

Case 3: $(\mathbf{A}\mathbf{x}_k, \mathbf{B}\mathbf{u}_k, \mathbf{w}_k) \in \mathcal{N}(\mathbf{D})$

This case also allows uncorrelated noise sequences \mathbf{w}_k with the input \mathbf{u}_k or the state \mathbf{x}_k and has a proper form for modeling systems with *physical constraints*, since regardless of the corrective input \mathbf{u}_k , the state stays within the constraint surface $\mathcal{N}(\mathbf{D})$. Since for all $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, it is required that $\mathbf{A}\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, $\mathcal{N}(\mathbf{D})$ is \mathbf{A} -invariant (Wonham, 1979).

As in the Case 2 and 3, the process noise sequence \mathbf{w}_k being constrained in the null space of \mathbf{D} is a reasonable choice from the viewpoint of usual system modeling practice: we model a system without considering noise or model uncertainty and then add a noise sequence to compensate for uncertainty of the system. Since the noise is in a lower dimensional space than the whole state space \mathbb{R}^n , the support of the probability density function must have a lower dimension than the whole state space, which means the covariance matrix of the noise must be singular.

2.1 Descriptor System Representation of Constrained System

The linear system (3) with the constraint (5) can be represented in the form of a descriptor system. To do this, first we define the following:

Definition 1 (Skelton et al. (1998)) For a given matrix $\mathbf{N} \in \mathbb{R}^{n \times m}$ with rank r , $\mathbf{N}^\perp \in \mathbb{R}^{(n-r) \times n}$ is defined as any matrix such that $\mathbf{N}^\perp \mathbf{N} = \mathbf{0}$ and $\mathbf{N}^\perp \mathbf{N}^{\perp T} > \mathbf{0}$.

Remark 1 Note that the matrix \mathbf{N}^\perp defined in the Definition 1 exists if and only if \mathbf{N} has linearly dependent rows ($n > r$), and the set of all such matrices can be captured by $\mathbf{N}^\perp = \mathbf{T}\mathbf{U}_2^T$, where \mathbf{T} is an arbitrary nonsingular matrix

and \mathbf{U}_2 is from the singular value decomposition (SVD)

$$\mathbf{N} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}. \quad (7)$$

In this paper, we consider only $\mathbf{T} = \mathbf{I}$ case and hence $\mathbf{N}^\perp = \mathbf{U}_2^T$.

Since \mathbf{D} is of full row rank, \mathbf{D}^T has dependent rows and therefore \mathbf{D}^{T^\perp} can be defined. Then, by multiplying both sides of the state equation (3) by \mathbf{D}^{T^\perp} and using the constrained equation (5) we obtain the following descriptor system representation

$$\mathbf{E}\mathbf{x}_{k+1} = \bar{\mathbf{A}}\mathbf{x}_k + \bar{\mathbf{B}}\mathbf{u}_k + \mathbf{E}\mathbf{w}_k \quad (8)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \quad (9)$$

where

$$\mathbf{E} = \begin{bmatrix} \mathbf{D}^{T^\perp} \\ \mathbf{0} \end{bmatrix}, \quad \bar{\mathbf{A}} = \begin{bmatrix} \mathbf{D}^{T^\perp} \mathbf{A} \\ \mathbf{D} \end{bmatrix}, \quad \bar{\mathbf{B}} = \begin{bmatrix} \mathbf{D}^{T^\perp} \mathbf{B} \\ \mathbf{0} \end{bmatrix}. \quad (10)$$

2.2 Projected System Representation of Constrained System

In a descriptor system representation such as (8), \mathbf{x}_k is termed a *descriptor* vector. Luenberger (1977) defined a *state* as a vector $\mathbf{z}_k = \mathbf{\Gamma}\mathbf{x}_k$, whose dimension is smaller than the descriptor vector \mathbf{x}_k with a matrix $\mathbf{\Gamma}$ having a specific property, for a set of dynamic equations if knowledge of its value and the value of input and the noise sequence are sufficient to uniquely determine the descriptor vector \mathbf{x}_k . Furthermore, an equivalent condition was obtained for both a time-invariant descriptor system to be *regular*¹ and for a vector $\mathbf{z}_k = \mathbf{\Gamma}\mathbf{x}_k$ to be a *state* for the descriptor system.

The following Lemma 1 can be easily verified.

Lemma 1 $\begin{bmatrix} \mathbf{D}^{T^\perp} \\ \mathbf{D} \end{bmatrix}$ is invertible and $\begin{bmatrix} \mathbf{D}^{T^\perp} \\ \mathbf{D} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{D}^{T^\perp T} & \mathbf{D}^\dagger \end{bmatrix}$.

By Lemma 1 above and Theorem 6 and 7 in Luenberger (1977), the following theorem is obtained.

Theorem 1 *The following statements are equivalent:*

- (i) *The descriptor system (8) is regular and the vector $\mathbf{z}_k = \mathbf{D}^{T^\perp} \mathbf{x}_k$ is a state.*
- (ii) *The matrix $\begin{bmatrix} \mathbf{D}^{T^\perp} \\ \mathbf{D} \end{bmatrix}$ is square and nonsingular.*

By Lemma 1 and Theorem 1, the descriptor system (8) describing the constrained system (3) and (5) is regular and $\mathbf{z}_k = \mathbf{D}^{T^\perp} \mathbf{x}_k$ is a state, from which the descriptor system (8) can be further simplified through combining the state

¹ A set of dynamic equation is said to be *regular* if there is an initial condition vector which when propagated forward serves as a state vector for every time period.

expression $\mathbf{z}_k = \mathbf{D}^{T^\perp} \mathbf{x}_k$ with the state constraint $\mathbf{D}\mathbf{x}_k = \mathbf{0}$, leading to

$$\begin{bmatrix} \mathbf{D}^{T^\perp} \\ \mathbf{D} \end{bmatrix} \mathbf{x}_k = \begin{bmatrix} \mathbf{z}_k \\ \mathbf{0} \end{bmatrix}. \quad (11)$$

Applying Lemma 1 to (11) yields

$$\mathbf{x}_k = \mathbf{D}^{T^\perp T} \mathbf{z}_k \quad \text{or} \quad \mathbf{x}_{k+1} = \mathbf{D}^{T^\perp T} \mathbf{z}_{k+1}. \quad (12)$$

From the definition of \mathbf{z}_k and (3),

$$\mathbf{z}_{k+1} = \mathbf{D}^{T^\perp} \mathbf{x}_{k+1} = \mathbf{D}^{T^\perp} (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k). \quad (13)$$

Substituting (13) into (12), we have

$$\mathbf{x}_{k+1} = \mathbf{D}^{T^\perp T} \mathbf{D}^{T^\perp} (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k) = \mathbf{P}_{\mathcal{N}(\mathbf{D})} (\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k) \quad (14)$$

where $\mathbf{P}_{\mathcal{N}(\mathbf{D})} \triangleq \mathbf{D}^{T^\perp T} \mathbf{D}^{T^\perp} = \mathbf{U}_2 \mathbf{U}_2^T$ is the orthogonal projector onto the null space of \mathbf{D} since \mathbf{U}_2 spans $\mathcal{N}(\mathbf{D})$ and, therefore, owing to the projector $\mathbf{P}_{\mathcal{N}(\mathbf{D})}$ being applied to $\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k$, the system representation (14) is termed a *projected system*. If the three elements $(\mathbf{A}\mathbf{x}_k, \mathbf{B}\mathbf{u}_k, \mathbf{w}_k)$ satisfy one of the three cases which are defined in Section 2, the projected system can be written as a ordinary state equation (3), and the system matrices \mathbf{A} and/or \mathbf{B} must satisfy some special properties. For instance, for Case 3-constrained systems, one important consequence can be drawn. From the fact that, for any $\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$, $\mathbf{A}\mathbf{x}_k \in \mathcal{N}(\mathbf{D})$ also holds,

$$\mathbf{A}\mathbf{x}_k = \mathbf{P}\mathbf{A}\mathbf{x}_k = \mathbf{A}\mathbf{P}\mathbf{x}_k \quad (15)$$

where \mathbf{P} is any projection matrix onto the null space of \mathbf{D} . By taking a conditional expectation for given any measurements \mathcal{Y} on both sides of (15), we obtain also

$$\mathbf{P}\mathbf{A} \mathcal{E}\{\mathbf{x}_k|\mathcal{Y}\} = \mathbf{A}\mathbf{P} \mathcal{E}\{\mathbf{x}_k|\mathcal{Y}\}. \quad (16)$$

Combining (15) with (16) yields

$$\begin{aligned} \mathbf{A}\mathbf{P}\boldsymbol{\Sigma} &= \mathbf{P}\mathbf{A}\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Sigma} \\ \mathbf{A}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^T \mathbf{A}^T &= \mathbf{P}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T \mathbf{P}^T = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T \end{aligned} \quad (17)$$

where $\boldsymbol{\Sigma} = \mathcal{E}\left\{[\mathbf{x}_k - \mathcal{E}(\mathbf{x}_k|\mathcal{Y})][\mathbf{x}_k - \mathcal{E}(\mathbf{x}_k|\mathcal{Y})]^T\right\}$. The equations in (17) are crucial relations in subsequent analysis for comparing performance between different estimators which can be used for estimating the constrained system.

3 Unconstrained Kalman Filter

As discussed in Section 2, for the constrained system given by (3) and (5) to have an uncorrelated process noise sequence \mathbf{w}_k with both current state \mathbf{x}_k and input \mathbf{u}_k , it is required that the process noise be constrained in the null space $\mathcal{N}(\mathbf{D})$ of the constraint matrix \mathbf{D} , which means the covariance matrix of \mathbf{w}_k must be singular. Hence, in order to design the ‘‘correct’’ Kalman filter for such a constrained system, the exact (singular) covariance matrix of the constrained noise must be used. When one does not know accurately the noise covariance, it is common practice (Anderson and Moore, 1979) to use the upper bound of the noise covariance or simply any bigger noise covariance than the (expected) correct value, causing, in some sense, a worst case design. Instead of the true singular (positive semi-definite) covariance matrix \mathbf{Q}^c , if a positive definite process noise covariance $\mathbf{Q}^u (\geq \mathbf{Q}^c)$ is used for the constrained system (3), then bigger or *unconstrained* estimation error covariances will be resulted. The corresponding

Kalman predictor and filter are given by the following equations:

Unconstrained Kalman predictor

$$\hat{\mathbf{x}}_{k+1|k}^u = (\mathbf{A} - \mathbf{M}_k^u \mathbf{C}) \hat{\mathbf{x}}_{k|k-1}^u + \mathbf{B} \mathbf{u}_k + \mathbf{M}_k^u \mathbf{y}_k \quad (18)$$

$$\mathbf{M}_k^u = \mathbf{A} \Sigma_{k|k-1}^u \mathbf{C}^T (\mathbf{C} \Sigma_{k|k-1}^u \mathbf{C}^T + \mathbf{R})^{-1} \quad (19)$$

$$\Sigma_{k+1|k}^u = \mathbf{A} \Sigma_{k|k-1}^u \mathbf{A}^T + \mathbf{Q}^u - \mathbf{A} \Sigma_{k|k-1}^u \mathbf{C}^T (\mathbf{C} \Sigma_{k|k-1}^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_{k|k-1}^u \mathbf{A}^T \quad (20)$$

Unconstrained Kalman filter

$$\hat{\mathbf{x}}_{k|k}^u = (\mathbf{I} - \mathbf{L}_k^u \mathbf{C}) \mathbf{A} \hat{\mathbf{x}}_{k-1|k-1}^u + (\mathbf{I} - \mathbf{L}_k^u \mathbf{C}) \mathbf{B} \mathbf{u}_{k-1} + \mathbf{L}_k^u \mathbf{y}_k \quad (21)$$

$$\mathbf{L}_k^u = \Sigma_{k|k-1}^u \mathbf{C}^T (\mathbf{C} \Sigma_{k|k-1}^u \mathbf{C}^T + \mathbf{R})^{-1} \quad (22)$$

$$\Sigma_{k|k}^u = \Sigma_{k|k-1}^u - \Sigma_{k|k-1}^u \mathbf{C}^T (\mathbf{C} \Sigma_{k|k-1}^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_{k|k-1}^u \quad (23)$$

$$\Sigma_{k|k-1}^u = \mathbf{A} \Sigma_{k-1|k-1}^u \mathbf{A}^T + \mathbf{Q}^u \quad (24)$$

Clearly, with the use of $\mathbf{Q}^u \geq \mathbf{Q}^c$ or the unconstrained process noise, the resulting state estimates need not satisfy the state constraint, and hence they are termed *unconstrained* estimates. The conditions of the stabilizable pair $[\mathbf{A}, \mathbf{Q}^{u^{1/2}}]$ and the detectable pair $[\mathbf{A}, \mathbf{C}]$ guarantee convergence of the above Riccati Equations, which is summarized by Theorem 2.

Theorem 2 (Anderson and Moore (1979)) *Assume that $[\mathbf{A}, \mathbf{Q}^{u^{1/2}}]$ is stabilizable and the pair $[\mathbf{A}, \mathbf{C}]$ is detectable. Then, for $\mathbf{Q}^u \geq \mathbf{0}$ and $\mathbf{R} > \mathbf{0}$, following hold:*

(i) *For any nonnegative symmetric initial condition $\Sigma_{k_0|k_0-1}$, one has*

$$\lim_{k \rightarrow \infty} \Sigma_{k+1|k}^u = \Sigma_{\infty}^u \quad (25)$$

with Σ_{∞}^u independent of $\Sigma_{k_0|k_0-1}$ and satisfying the Algebraic Riccati Equation (ARE):

$$\Sigma_{\infty}^u = \mathbf{A} \Sigma_{\infty}^u \mathbf{A}^T - \mathbf{A} \Sigma_{\infty}^u \mathbf{C}^T (\mathbf{C} \Sigma_{\infty}^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_{\infty}^u \mathbf{A}^T + \mathbf{Q}^u \quad (26)$$

(ii) $|\lambda_i(\mathbf{A} - \mathbf{M}_{\infty}^u \mathbf{C})| < 1$, where from (19)

$$\mathbf{M}_{\infty}^u = \mathbf{A} \Sigma_{\infty}^u \mathbf{C}^T (\mathbf{C} \Sigma_{\infty}^u \mathbf{C}^T + \mathbf{R})^{-1}. \quad (27)$$

4 Constrained Kalman Filter

In this section, we consider and compare two different constrained predictors (filters) that can be used for the Case 3 constrained system (physically constrained system) described in Section 2. Performance comparison between estimators is made by comparing the relative magnitude of covariance matrices for the Kalman predictor and filter. For a simple notation, the subscripts $(\cdot)_k$ will be used for denoting the Kalman predictor instead of $(\cdot)_{k|k-1}$.

4.1 Projected Kalman Predictor

Chia (1985), and Simon and Chia (2002) derived a constrained Kalman predictor by directly projecting the unconstrained state estimate $\hat{\mathbf{x}}_k^u$ onto the constraint subspace $\mathcal{N}(\mathbf{D})$. Let us name it *projected* estimate which will be

denoted by the superscript $(\cdot)^p$. They solved the problem, for any symmetric positive definite weighting matrix \mathbf{W} ,

$$\min_{\hat{\mathbf{x}}_k^p} (\hat{\mathbf{x}}_k^p - \hat{\mathbf{x}}_k^u)^T \mathbf{W} (\hat{\mathbf{x}}_k^p - \hat{\mathbf{x}}_k^u) \quad \text{subject to } \mathbf{D}\hat{\mathbf{x}}_k^p = \mathbf{0} \quad (28)$$

and obtained

$$\hat{\mathbf{x}}_k^p = P_{\mathcal{N}(\mathbf{D})}^W \hat{\mathbf{x}}_k^u \quad (29)$$

where

$$P_{\mathcal{N}(\mathbf{D})}^W \triangleq \mathbf{I} - \mathbf{W}^{-1} \mathbf{D}^T (\mathbf{D} \mathbf{W}^{-1} \mathbf{D}^T)^{-1} \mathbf{D} \quad (30)$$

which is a projector to the constraint subspace $\mathcal{N}(\mathbf{D})$ with a weighting matrix \mathbf{W} . The property of the projected Kalman predictor is summarized in the following theorem.

Theorem 3 (Chia (1985) and Simon and Chia (2002), Projected Kalman predictor)

- (i) *The projected state estimate $\hat{\mathbf{x}}_k^p$ given by (29) with $\mathbf{W} = (\boldsymbol{\Sigma}_k^u)^{-1}$ has a smaller state error covariance than that of the unconstrained state estimate $\hat{\mathbf{x}}_k^u$. That is*

$$\boldsymbol{\Sigma}_k^p \triangleq \mathbf{Cov}(\mathbf{x}_k - \hat{\mathbf{x}}_k^p) \leq \mathbf{Cov}(\mathbf{x}_k - \hat{\mathbf{x}}_k^u) = \boldsymbol{\Sigma}_k^u \quad (31)$$

and the covariance of the projected estimator is given by

$$\boldsymbol{\Sigma}_k^p = \mathbf{P}_k \boldsymbol{\Sigma}_k^u \mathbf{P}_k^T = \mathbf{P}_k \boldsymbol{\Sigma}_k^u, \quad (32)$$

where $\mathbf{P}_k = \mathbf{I} - \boldsymbol{\Sigma}_k^u \mathbf{D}^T (\mathbf{D} \boldsymbol{\Sigma}_k^u \mathbf{D}^T)^{-1} \mathbf{D}$ is a projection matrix onto the null space of \mathbf{D} .

- (ii) *Among all the projected Kalman predictors of (29), the predictor that uses $\mathbf{W} = (\boldsymbol{\Sigma}_k^u)^{-1}$ has the smallest estimation error covariance.*

Therefore, at $(k+1)$ th-stage the state estimation error covariance of the projected Kalman predictor is, from (20) and (32), given by

$$\begin{aligned} \boldsymbol{\Sigma}_{k+1}^p &= \mathbf{P}_{k+1} \boldsymbol{\Sigma}_{k+1}^u \mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1} \left\{ \mathbf{A} \boldsymbol{\Sigma}_k^u \mathbf{A}^T + \mathbf{Q}^u - \mathbf{A} \boldsymbol{\Sigma}_k^u \mathbf{C}^T (\mathbf{C} \boldsymbol{\Sigma}_k^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \boldsymbol{\Sigma}_k^u \mathbf{A}^T \right\} \mathbf{P}_{k+1}^T. \end{aligned} \quad (33)$$

4.2 Constrained Kalman Filter by Projected System

In Section 2.2, we have proved that the original state equation (3) with the constraint (5) can be reduced to the projected system (14). Therefore, for the Case 3 constrained system, the Kalman predictor is given by

$$\hat{\mathbf{x}}_{k+1}^c = (\mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{A} - \mathbf{M}_k^c \mathbf{C}) \hat{\mathbf{x}}_k^c + \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{B} \mathbf{u}_k + \mathbf{M}_k^c \mathbf{y}_k = (\mathbf{A} - \mathbf{M}_k^c \mathbf{C}) \hat{\mathbf{x}}_k^c + \mathbf{B} \mathbf{u}_k + \mathbf{M}_k^c \mathbf{y}_k \quad (34)$$

$$\mathbf{M}_k^c = \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{C}^T (\mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{C}^T + \mathbf{R})^{-1} = \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{C}^T (\mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{C}^T + \mathbf{R})^{-1} \quad (35)$$

$$\boldsymbol{\Sigma}_{k+1}^c = \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{A}^T \mathbf{P}_{\mathcal{N}(\mathbf{D})} + \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{Q}^u \mathbf{P}_{\mathcal{N}(\mathbf{D})} - \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{C}^T (\mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{A}^T \mathbf{P}_{\mathcal{N}(\mathbf{D})} \quad (36)$$

$$= \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{A}^T + \mathbf{Q}^c - \mathbf{A} \boldsymbol{\Sigma}_k^c \mathbf{C}^T (\mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \boldsymbol{\Sigma}_k^c \mathbf{A}^T \quad (37)$$

where it is assumed that $\mathbf{Q}^c = \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{Q}^u \mathbf{P}_{\mathcal{N}(\mathbf{D})}$.

Remark 2 In Section 2.2, a reduced state $\mathbf{z}_k = \mathbf{D}^{T\perp} \mathbf{x}_k$ was used in the middle of deriving the projected system (14). We can also construct the Kalman predictor for estimating this reduced state \mathbf{z}_k and use the relation $\hat{\mathbf{x}}_k = \mathbf{D}^{T\perp T} \hat{\mathbf{z}}_k$ from (12) which will yield the Kalman predictor given in (34 - 37).

4.3 Comparison of Constrained Kalman Filters

From optimality of the Kalman predictor/filter, we expect that the Kalman predictor derived from the projected state equation (14) is the optimal filter for the original constrained system modeled by (3) and (5). But, we need to know whether the projected estimator of Section 4.1 is also optimal or not. To compare the error covariances given by the two equations (33) and (37), let us express (33) as

$$\Sigma_{k+1}^p = \mathbf{P}_k \left\{ \mathbf{A} \Sigma_k^p \mathbf{A}^T + \mathbf{Q}^u - \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T \right\} \mathbf{P}_k^T + \Delta \mathbf{Q}_k^d, \quad (38)$$

where

$$\begin{aligned} \Delta \mathbf{Q}_k^d \triangleq & \mathbf{P}_{k+1} \left\{ \mathbf{A} \Sigma_k^u \mathbf{A}^T + \mathbf{Q}^u - \mathbf{A} \Sigma_k^u \mathbf{C}^T (\mathbf{C} \Sigma_k^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^u \mathbf{A}^T \right\} \mathbf{P}_{k+1}^T \\ & - \mathbf{P}_k \left\{ \mathbf{A} \Sigma_k^p \mathbf{A}^T + \mathbf{Q}^u - \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T \right\} \mathbf{P}_k^T. \end{aligned} \quad (39)$$

Since Σ_k^p is constrained and then using (17), we have

$$\begin{aligned} \mathbf{P}_k \mathbf{A} \Sigma_k^p \mathbf{A}^T \mathbf{P}_k^T &= \mathbf{P}_{k+1} \mathbf{P}_k \mathbf{A} \Sigma_k^p \mathbf{A}^T \mathbf{P}_k^T \mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1} \mathbf{A} \mathbf{P}_k \Sigma_k^p \mathbf{P}_k^T \mathbf{A}^T \mathbf{P}_{k+1}^T \\ &= \mathbf{P}_{k+1} \mathbf{A} \Sigma_k^p \mathbf{A}^T \mathbf{P}_{k+1}^T \end{aligned} \quad (40)$$

and also

$$\mathbf{P}_k \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T \mathbf{P}_k^T = \mathbf{P}_{k+1} \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T \mathbf{P}_{k+1}^T. \quad (41)$$

Therefore, we have

$$\Delta \mathbf{Q}_k^d = \mathbf{P}_{k+1} \Delta_k^d \mathbf{P}_{k+1}^T + \mathbf{P}_{k+1} \mathbf{Q}^u \mathbf{P}_{k+1}^T - \mathbf{P}_k \mathbf{Q}^u \mathbf{P}_k^T, \quad (42)$$

where

$$\Delta_k^d = \mathbf{A} \Sigma_k^u \mathbf{A}^T - \mathbf{A} \Sigma_k^u \mathbf{C}^T (\mathbf{C} \Sigma_k^u \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^u \mathbf{A}^T - \mathbf{A} \Sigma_k^p \mathbf{A}^T + \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T. \quad (43)$$

To prove that $\Delta_k^d \geq \mathbf{0}$, we need the following lemma about a monotonicity property of the Riccati Difference Equation.

Lemma 2 (De Souza (1989), Bitmead and Gevers (1991)) *Consider two Riccati Difference Equations with the same \mathbf{A} , \mathbf{B} and \mathbf{R} matrices but possibly different \mathbf{Q}^1 and \mathbf{Q}^2 . Denote their solution matrices Σ_k^1 and Σ_k^2 respectively. Suppose that $\mathbf{Q}^1 \geq \mathbf{Q}^2$, and, for some k we have $\Sigma_k^1 \geq \Sigma_k^2$, then for all $i > 0$*

$$\Sigma_{k+i}^1 \geq \Sigma_{k+i}^2. \quad (44)$$

Since Δ_k^d in (43) can be considered as the difference $\Sigma_{k+1}^u - \Sigma_{k+1}^p$ of the two Riccati Difference Equations with the same \mathbf{Q}^u and \mathbf{R} , using the fact that $\Sigma_k^u \geq \Sigma_k^p$ from Theorem 3 together with Lemma 2, we can deduce that $\Delta_k^d \geq \mathbf{0}$ for all k . Hence, (38) can be written as

$$\begin{aligned} \Sigma_{k+1}^p &= \mathbf{P}_k \left\{ \mathbf{A} \Sigma_k^p \mathbf{A}^T - \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T \right\} \mathbf{P}_k^T + \mathbf{P}_{k+1} \Delta_k \mathbf{P}_{k+1}^T + \mathbf{P}_{k+1} \mathbf{Q}^u \mathbf{P}_{k+1}^T \\ &= \mathbf{A} \Sigma_k^p \mathbf{A}^T - \mathbf{A} \Sigma_k^p \mathbf{C}^T (\mathbf{C} \Sigma_k^p \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^p \mathbf{A}^T + \mathbf{P}_{k+1} \Delta_k \mathbf{P}_{k+1}^T + \mathbf{P}_{k+1} \mathbf{Q}^u \mathbf{P}_{k+1}^T. \end{aligned} \quad (45)$$

Recall that from Section 4.2 and using (17) we have, with $\mathbf{Q}^c = \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{Q}^u \mathbf{P}_{\mathcal{N}(\mathbf{D})}$,

$$\Sigma_{k+1}^c = \mathbf{A} \Sigma_k^c \mathbf{A}^T + \mathbf{Q}^c - \mathbf{A} \Sigma_k^c \mathbf{C}^T (\mathbf{C} \Sigma_k^c \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \Sigma_k^c \mathbf{A}^T. \quad (46)$$

The following Theorem 4 summarizes what we have shown regarding the performance of the three versions of Kalman predictors that can be used for estimating constrained state.

Theorem 4 *For Case 3 constrained systems with the assumptions that the pair $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ is stabilizable and $[\mathbf{A}, \mathbf{C}]$ is detectable, the following hold:*

(i) *If $\mathbf{Q}^u > \mathbf{Q}^c$, with the initial conditions such that $\Sigma_{0|-1}^c \leq \Sigma_{0|-1}^p \leq \Sigma_{0|-1}^u$, we have*

$$\Sigma_{k+1|k}^c \leq \Sigma_{k+1|k}^p \leq \Sigma_{k+1|k}^u \quad (47a)$$

$$\lim_{k \rightarrow \infty} \Sigma_{k+1|k}^c \leq \lim_{k \rightarrow \infty} \Sigma_{k+1|k}^p \leq \lim_{k \rightarrow \infty} \Sigma_{k+1|k}^u. \quad (47b)$$

(ii) *If $\mathbf{Q}^u = \mathbf{Q}^c$, with the initial conditions such that $\Sigma_{0|-1}^c = \Sigma_{0|-1}^p = \mathbf{P}_0 \Sigma_{0|-1}^u \mathbf{P}_0^T$, we have*

$$\Sigma_{k+1|k}^c \leq \Sigma_{k+1|k}^p \leq \Sigma_{k+1|k}^u \quad (48a)$$

$$\lim_{k \rightarrow \infty} \Sigma_{k+1|k}^c = \lim_{k \rightarrow \infty} \Sigma_{k+1|k}^p = \lim_{k \rightarrow \infty} \Sigma_{k+1|k}^u. \quad (48b)$$

PROOF.

(i) Using $\mathbf{P}_{k+1} \mathbf{Q}^u \mathbf{P}_{k+1}^T \geq \mathbf{P}_{\mathcal{N}(\mathbf{D})} \mathbf{Q}^u \mathbf{P}_{\mathcal{N}(\mathbf{D})} = \mathbf{Q}^c$ from the fact that $\mathbf{P}_{\mathcal{N}(\mathbf{D})}$ is the orthogonal projector, and $\mathbf{P}_{k+1} \Delta_k \mathbf{P}_{k+1}^T \geq \mathbf{0}$, it can be shown that, again through Lemma 2 and Theorem 3, we obtain (47a) and also (47b). The existence of limit matrices is guaranteed from the assumptions of stabilizability of $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ (which implies also stabilizable $[\mathbf{A}, \mathbf{Q}^{u^{1/2}}]$, since $\mathbf{Q}^u \geq \mathbf{Q}^c$) and detectability of $[\mathbf{A}, \mathbf{C}]$.

(ii) We observe that Riccati Difference Equations (20) and (46) are the same except for the different initial conditions such that $\Sigma_{0|-1}^u \geq \Sigma_{0|-1}^c = \mathbf{P}_0 \Sigma_{0|-1}^u \mathbf{P}_0^T$, from which, in combination with Lemma 2, we have $\Sigma_k^p = \mathbf{P}_k \Sigma_k^u \mathbf{P}_k^T \geq \mathbf{P}_k \Sigma_k^c \mathbf{P}_k^T = \Sigma_k^c$. Therefore, we obtain (48a). With the assumptions of stabilizable $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ and detectable $[\mathbf{A}, \mathbf{C}]$, the initial condition effects fade away as $k \rightarrow \infty$ from the Theorem 2 and thus we obtain (48b). \square

As in Chia (1985) and Wen and Durrant-Whyte (1992), we can construct a projected filter from the unconstrained Kalman filter onto the constraint subspace, obtaining

$$\hat{\mathbf{x}}_{k|k}^p = \mathbf{P}_{k|k} \hat{\mathbf{x}}_{k|k}^u \quad (49)$$

where

$$\mathbf{P}_{k|k} \triangleq \mathbf{I} - \Sigma_{k|k}^u \mathbf{D}^T (\mathbf{D} \Sigma_{k|k}^u \mathbf{D}^T)^{-1} \mathbf{D} \quad (50)$$

is a projection matrix onto the null space of \mathbf{D} , and

$$\Sigma_{k|k}^p = \mathbf{P}_{k|k} \Sigma_{k|k}^u \mathbf{P}_{k|k}^T = \mathbf{P}_{k|k} \Sigma_{k|k}^u. \quad (51)$$

For Case 3 constrained systems, by using the same technique as used in the predictor case, we can show that the filter version of Theorem 4 also holds.

Corollary 1 *For Case 3 constrained systems with the assumption that the pair $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ is stabilizable and $[\mathbf{A}, \mathbf{C}]$ is detectable, the following hold:*

(i) If $\mathbf{Q}^u > \mathbf{Q}^c$, with the initial conditions such that $\Sigma_{0|0}^c \leq \Sigma_{0|0}^p \leq \Sigma_{0|0}^u$, we have

$$\Sigma_{k|k}^c \leq \Sigma_{k|k}^p \leq \Sigma_{k|k}^u \quad (52a)$$

$$\lim_{k \rightarrow \infty} \Sigma_{k|k}^c \leq \lim_{k \rightarrow \infty} \Sigma_{k|k}^p \leq \lim_{k \rightarrow \infty} \Sigma_{k|k}^u. \quad (52b)$$

(ii) If $\mathbf{Q}^u = \mathbf{Q}^c$, with the initial conditions such that $\Sigma_{0|0}^c = \Sigma_{0|0}^p = \mathbf{P}_0 \Sigma_{0|0}^u \mathbf{P}_0^T$, we have

$$\Sigma_{k|k}^c \leq \Sigma_{k|k}^p \leq \Sigma_{k|k}^u \quad (53a)$$

$$\lim_{k \rightarrow \infty} \Sigma_{k|k}^c = \lim_{k \rightarrow \infty} \Sigma_{k|k}^p = \lim_{k \rightarrow \infty} \Sigma_{k|k}^u. \quad (53b)$$

PROOF. See Appendix A. \square

5 Continuous-Time Case

In this section, we develop the continuous-time counterpart of the state estimation results presented in the previous sections. Consider a continuous time linear stochastic system described with the Itô calculus

$$d\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t)dt + \mathbf{B}\mathbf{u}(t)dt + d\mathbf{w}(t) \quad (54)$$

$$d\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t)dt + d\mathbf{v}(t) \quad (55)$$

where the state $\mathbf{x}(t) \in \mathbb{R}^n$ is known to be constrained to

$$\mathcal{N}(\mathbf{D}) \triangleq \{\mathbf{x} : \mathbf{D}\mathbf{x} = \mathbf{0}\} \quad (56)$$

where $\dim(\mathcal{N}(\mathbf{D})) < n$. Here $\mathbf{w}(t)$ and $\mathbf{v}(t)$ are mutually independent Brownian motions with intensities

$$\mathcal{E} \left\{ \begin{bmatrix} \mathbf{w}(t) \\ \mathbf{v}(t) \end{bmatrix} \begin{bmatrix} \mathbf{w}(\tau) \\ \mathbf{v}(\tau) \end{bmatrix}^T \right\} = \begin{bmatrix} \mathbf{Q}^c & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \delta(t - \tau) \quad (57)$$

where $\delta(t - \tau)$ denotes the Dirac delta function. Then, the continuous time (*unconstrained*) Kalman filter without considering the state equality constraint (56) is given by the following equations

$$d\hat{\mathbf{x}}^u(t) = [\mathbf{A} - \mathbf{M}^u(t)\mathbf{C}] \hat{\mathbf{x}}^u(t)dt + \mathbf{B}\mathbf{u}(t) + \mathbf{M}^u(t)d\mathbf{y}(t) \quad (58a)$$

$$\mathbf{M}^u(t) = \Sigma^u(t)\mathbf{C}^T\mathbf{R}^{-1} \quad (58b)$$

$$\dot{\Sigma}^u(t) = \mathbf{A}\Sigma^u(t) + \Sigma^u(t)\mathbf{A}^T - \Sigma^u(t)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma^u(t) + \mathbf{Q}^u. \quad (58c)$$

With appropriate stabilizability and detectability properties of the projected system in continuous time, we obtain the following convergence property of the unconstrained Kalman filter.

Theorem 5 Assume that $[\mathbf{A}, \mathbf{Q}^{u/2}]$ is stabilizable and $[\mathbf{A}, \mathbf{C}]$ is detectable. Then, the following hold

(i) For any nonnegative symmetric initial condition $\Sigma_{t_0}^u$, one has

$$\lim_{t \rightarrow \infty} \Sigma^u(t) = \Sigma_\infty^u \quad (59)$$

with Σ_∞^u independent of $\Sigma_{t_0}^u$ and satisfying the ARE

$$\mathbf{0} = \mathbf{A}\Sigma_\infty^u + \Sigma_\infty^u\mathbf{A}^T - \Sigma_\infty^u\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma_\infty^u + \mathbf{Q}^u. \quad (60)$$

(ii) $\text{Re}\left\{\lambda_i(\mathbf{A} - \mathbf{M}_\infty^u\mathbf{C})\right\} < 0$, where

$$\mathbf{M}_\infty^u = \Sigma_\infty^u\mathbf{C}^T\mathbf{R}^{-1}. \quad (61)$$

By following the same steps used in Section 4.1, we can establish the continuous-time projected Kalman filter for the system (54) such that

$$\hat{\mathbf{x}}^p(t) = P_{\mathcal{N}(\mathbf{D})}^W \hat{\mathbf{x}}^u(t), \quad (62)$$

where

$$P_{\mathcal{N}(\mathbf{D})}^W \triangleq \mathbf{I} - \mathbf{W}^{-1}\mathbf{D}^T(\mathbf{D}\mathbf{W}^{-1}\mathbf{D}^T)^{-1}\mathbf{D} \quad (63)$$

which is a projector to the constraint subspace $\mathcal{N}(\mathbf{D})$ with a weighting matrix \mathbf{W} . The property of the projected Kalman predictor is summarized in the following theorem.

Theorem 6 (Continuous-time projected Kalman predictor)

(i) The projected state estimate $\hat{\mathbf{x}}^p(t)$ given by (62) with $\mathbf{W} = [\Sigma^u(t)]^{-1}$ has a smaller state error covariance than that of the unconstrained state estimate. That is

$$\Sigma^p(t) \triangleq \text{Cov}[\mathbf{x}(t) - \hat{\mathbf{x}}^p(t)] \leq \text{Cov}[\mathbf{x}(t) - \hat{\mathbf{x}}^u(t)] = \Sigma^u(t) \quad (64)$$

and the covariance of the projected estimator is given by

$$\Sigma^p(t) = \mathbf{P}(t)\Sigma^u(t)\mathbf{P}^T(t) = \mathbf{P}(t)\Sigma^u(t) \quad (65)$$

where

$$\mathbf{P}(t) \triangleq \mathbf{I} - \Sigma^u(t)\mathbf{D}^T[\mathbf{D}\Sigma^u(t)\mathbf{D}^T]^{-1}\mathbf{D} \quad (66)$$

is a projection matrix onto the null space of \mathbf{D} .

- (ii) Among all the projected Kalman filters of (62), the filter that uses $\mathbf{W} = [\Sigma^u(t)]^{-1}$ has the smallest estimation error covariance.
- (iii) The state estimation error covariance of the continuous-time projected Kalman filter satisfies the following Riccati Differential Equation

$$\dot{\Sigma}^p(t) = \mathbf{A}\Sigma^p(t) + \Sigma^p(t)\mathbf{A}^T - \Sigma^p(t)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma^p(t) + \mathbf{P}(t)\mathbf{Q}^u\mathbf{P}^T(t) + \mathbf{P}(t)\Delta^c(t)\mathbf{P}^T(t) \quad (67)$$

where

$$\Delta^c(t) = \mathbf{A}\Sigma^u(t) + \Sigma^u(t)\mathbf{A}^T - \Sigma^u(t)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma^u(t) - [\mathbf{A}\Sigma^p(t) + \Sigma^p(t)\mathbf{A}^T - \Sigma^p(t)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma^p(t)]. \quad (68)$$

PROOF. See Appendix B. \square

Similarly to the discrete-time case, in order to prove that $\Delta^c(t) \geq \mathbf{0}$ we need the following Lemma 3 which provides a means to compare two Riccati Differential Equations having different initial conditions and driving noise intensities.

Lemma 3 (Poubelle et al. (1988)) Consider two Riccati Differential Equations with the same \mathbf{A} , \mathbf{B} and \mathbf{R} matrices but possibly different \mathbf{Q}^1 and \mathbf{Q}^2 , and possibly different initial conditions, Σ_0^1 and Σ_0^2 . Denote their solution matrices $\Sigma^1(t)$ and $\Sigma^2(t)$, respectively. Then

$$\Sigma_0^2 \leq \Sigma_0^1 \text{ and } \mathbf{Q}^2 \leq \mathbf{Q}^1 \quad (69)$$

implies

$$\Sigma^2(t) \leq \Sigma^1(t), \text{ for all } t \geq 0. \quad (70)$$

Furthermore, if for some t_0 , $\dot{\Sigma}(t_0) \geq \mathbf{0}$, then $\dot{\Sigma}(t) \geq \mathbf{0}$ for all $t \geq t_0$.

Theorem 6 tells us that $\Sigma^p(t) = \mathbf{P}(t)\Sigma^u(t)$ and by using (B.2) we can easily show that $\dot{\Sigma}^p(t) = \mathbf{P}(t)\dot{\Sigma}^u(t)\mathbf{P}^T(t)$ which means, at $t = 0$,

$$\dot{\Sigma}^u(0) \geq \dot{\Sigma}^p(0) = \mathbf{P}(0)\dot{\Sigma}^u(0)\mathbf{P}^T(0). \quad (71)$$

Here, $\dot{\Sigma}^u(0)$ is determined by (58c) for given initial condition $\Sigma^u(0)$. Now, direct applying Lemma 3 together with (71) to (68) yields $\Delta^c(t) \geq \mathbf{0}$ for all $t \geq 0$. Now, to incorporate the known state constraint (56) into the filter, consider a vector $\mathbf{z}(t) \triangleq \mathbf{D}^{T^\perp} \mathbf{x}(t)$, similarly to the discrete-time counterpart. Following the same procedures yields

$$\mathbf{x}(t) = \mathbf{D}^{T^\perp} \mathbf{z}(t) \quad (72)$$

and then

$$\begin{aligned} d\mathbf{x}(t) &= \mathbf{D}^{T^\perp} dz(t) \\ &= \mathbf{P}_{\mathcal{N}(\mathbf{D})} \left[\mathbf{A}\mathbf{x}(t)dt + \mathbf{B}\mathbf{u}(t)dt + d\mathbf{w}(t) \right] \end{aligned} \quad (73)$$

which states that for the state $\mathbf{x}(t)$ to be constrained to the constraint subspace $\mathcal{N}(\mathbf{D})$, the time variation of the state must remain within $\mathcal{N}(\mathbf{D})$, the null space of \mathbf{D} . Let us assume that all the three elements of the time variation in (73) are in $\mathcal{N}(\mathbf{D})$, enabling to have uncorrelated process noise $\mathbf{w}(t)$ with the state $\mathbf{x}(t)$ and the input $\mathbf{u}(t)$. Therefore, the Kalman filter for the continuous-time projected system (73) is given by

$$d\hat{\mathbf{x}}^c(t) = [\mathbf{A} - \mathbf{M}^c(t)\mathbf{C}]\hat{\mathbf{x}}^c(t)dt + \mathbf{B}\mathbf{u}(t)dt + \mathbf{M}^c(t)d\mathbf{y}(t) \quad (74a)$$

$$\mathbf{M}^c(t) = \Sigma^c(t)\mathbf{C}^T\mathbf{R}^{-1} \quad (74b)$$

$$\dot{\Sigma}^c(t) = \mathbf{A}\Sigma^c(t) + \Sigma^c(t)\mathbf{A}^T - \Sigma^c(t)\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\Sigma^c(t) + \mathbf{Q}^c. \quad (74c)$$

Direct application of Lemma 3 to (67) and (74) yields the following theorem.

Theorem 7 For Case 3 constrained systems with the assumption that the pair $[\mathbf{A}, \mathbf{Q}^{c^{1/2}}]$ is stabilizable and $[\mathbf{A}, \mathbf{C}]$ is detectable, the following hold:

(i) If $\mathbf{Q}^u > \mathbf{Q}^c$, with the initial conditions such that $\Sigma^c(0) \leq \Sigma^p(0) \leq \Sigma^u(0)$, we have

$$\Sigma^c(t) \leq \Sigma^p(t) \leq \Sigma^u(t) \quad (75a)$$

$$\lim_{t \rightarrow \infty} \Sigma^c(t) \leq \lim_{t \rightarrow \infty} \Sigma^p(t) \leq \lim_{t \rightarrow \infty} \Sigma^u(t). \quad (75b)$$

(ii) If $\mathbf{Q}^u = \mathbf{Q}^c$, with the initial conditions such that $\Sigma^c(0) = \Sigma^p(0) = \mathbf{P}(0)\Sigma^u(0)\mathbf{P}^T(0)$, we have

$$\Sigma^c(t) \leq \Sigma^p(t) \leq \Sigma^u(t) \quad (76a)$$

$$\lim_{t \rightarrow \infty} \Sigma^c(t) = \lim_{t \rightarrow \infty} \Sigma^p(t) = \lim_{t \rightarrow \infty} \Sigma^u(t). \quad (76b)$$

6 Numerical Example: Tracking of a Land-Based Vehicle

To show numerically the performance differences of the three estimators, consider the following linear systems and measurements describing movement of a land-based vehicle.

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 0 \\ 0 \\ T \sin \theta \\ T \cos \theta \end{bmatrix} u_k + \mathbf{w}_k \quad (77)$$

$$\mathbf{y}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{v}_k$$

This example was used in Simon and Chia (2002) to compare the performance between the unconstrained and the projected Kalman predictor, where a nonlinear measurement equation was used. The first two state components of \mathbf{x}_k are the northerly and easterly positions and the last two are the northerly and easterly velocities. Vector \mathbf{w}_k and \mathbf{v}_k , represent the process and measurements noise sequences respectively, and are zero-mean gaussian noise sequences with covariances \mathbf{Q}^c and \mathbf{R} , respectively. Furthermore, it is known that the vehicle is moving on a road with a heading of θ which is described by the constraint equation

$$\mathbf{D}\mathbf{x}_k = [0, 0, 1, -\tan \theta] \mathbf{x}_k = 0. \quad (78)$$

For a simulation study, 2 seconds sample period T and 60 deg heading θ are used. The commanded acceleration is alternately set to ± 1 m/sec² (which is assumed to be known exactly) as if the vehicle was alternately accelerating and decelerating in traffic. The initial conditions are set to

$$\begin{aligned} \hat{\mathbf{x}}_0 &= \text{Diag}[0, 0, 17, 10], \\ \Sigma_{0|-1}^u &= \text{Diag}[400, 400, 10, 10], \text{ for the unconstrained estimator, and} \\ \Sigma_{0|-1}^c &= \Sigma_{0|-1}^p = \mathbf{P}_{\mathcal{N}(\mathbf{D})} \Sigma_{0|-1}^u \mathbf{P}_{\mathcal{N}(\mathbf{D})}, \text{ for the constrained estimators.} \end{aligned} \quad (79)$$

As for the noise covariances, the following are used:

$$\mathbf{Q}^c = \begin{bmatrix} 30 & 30/\tan \theta & 0 & 0 \\ 30/\tan \theta & 30/\tan^2 \theta & 0 & 0 \\ 0 & 0 & 10 & 10/\tan \theta \\ 0 & 0 & 10/\tan \theta & 10/\tan^2 \theta \end{bmatrix} \leq \mathbf{Q}^u = \begin{bmatrix} 35 & 15 & 0 & 0 \\ 15 & 20 & 0 & 0 \\ 0 & 0 & 15 & 5 \\ 0 & 0 & 5 & 10 \end{bmatrix} \quad (80)$$

$$\mathbf{R} = \text{Diag}[400, 10]$$

Here \mathbf{Q}^u is the process noise covariance used for the unconstrained estimator.

Tables 1 and 2 summarize the result of a simulation conducted for 100 seconds to see performance difference between the three estimators studied in the previous sections. As proved earlier, the constrained Kalman predictor has the smallest variances for all components of state estimator (Table 2). Although the differences among the three are small but still the least is the constrained estimate (Table 1). These differences are also shown in Fig. 1 through Fig.

Estimators	RMS Estimation Error [m]			
	$\hat{\mathbf{x}}_k(1)$	$\hat{\mathbf{x}}_k(2)$	$\hat{\mathbf{x}}_k(3)$	$\hat{\mathbf{x}}_k(4)$
Unconstrained Kalman Filter	43.44	22.46	4.36	2.06
Projected Kalman Filter	41.84	22.79	3.53	2.03
Constrained Kalman Filter	41.82	21.23	3.31	1.91

Table 1
RMS Estimation Error

Estimators	Variance of Estimate [m^2]			
	$\hat{\mathbf{x}}_k(1)$	$\hat{\mathbf{x}}_k(2)$	$\hat{\mathbf{x}}_k(3)$	$\hat{\mathbf{x}}_k(4)$
Unconstrained Kalman Filter	591.08	181.68	49.20	16.09
Projected Kalman Filter	442.10	170.97	31.09	10.36
Constrained Kalman Filter	255.74	90.47	21.86	7.29

Table 2
Variance of estimates

4 where Fig. 1 and 2 compare the time history and the estimation errors of the northerly position and velocity, and Fig. 3 and 4 show the variances of each estimate for the northerly position and velocity.

In the case of $\mathbf{Q}^u = \mathbf{Q}^c$, that is when the process noise covariance for the unconstrained estimator is the same as that of the constrained one, Fig. 5 and Fig. 6 show that the variances of the unconstrained estimator approach to those of the constrained estimator, as proven in Theorem 4 (ii). In this case, it is observed that the variances of the projected and the constrained estimators are the exactly same.

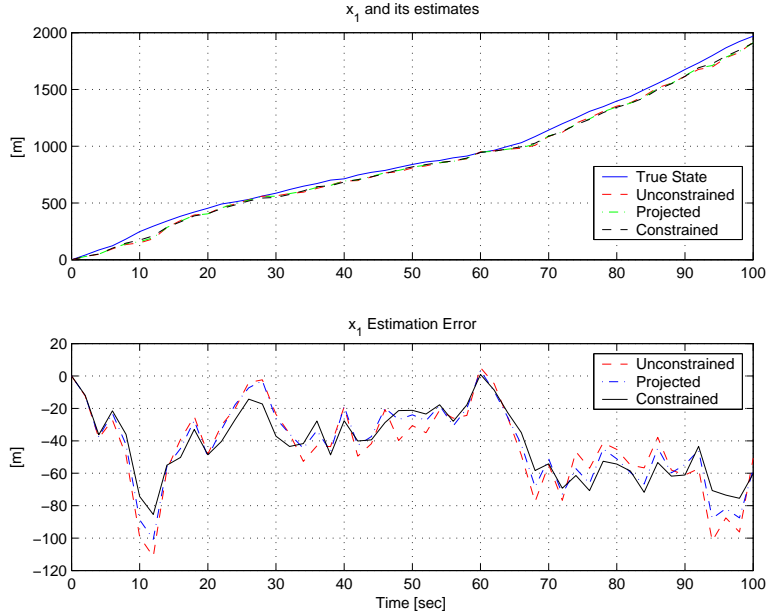


Fig. 1. Estimation result of the northerly position $\mathbf{x}_k(1)$ for a land-based vehicle

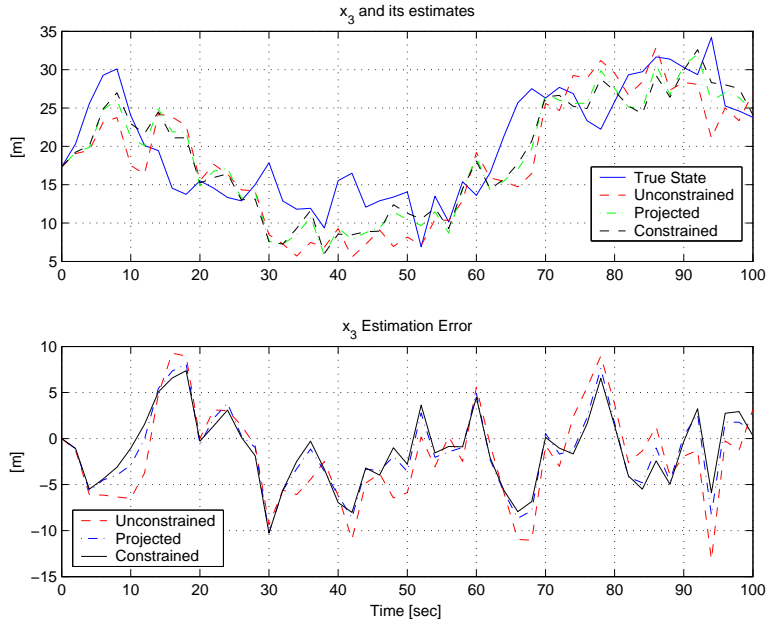


Fig. 2. Estimation result of the northerly velocity $\mathbf{x}_k(3)$ for a land-based vehicle

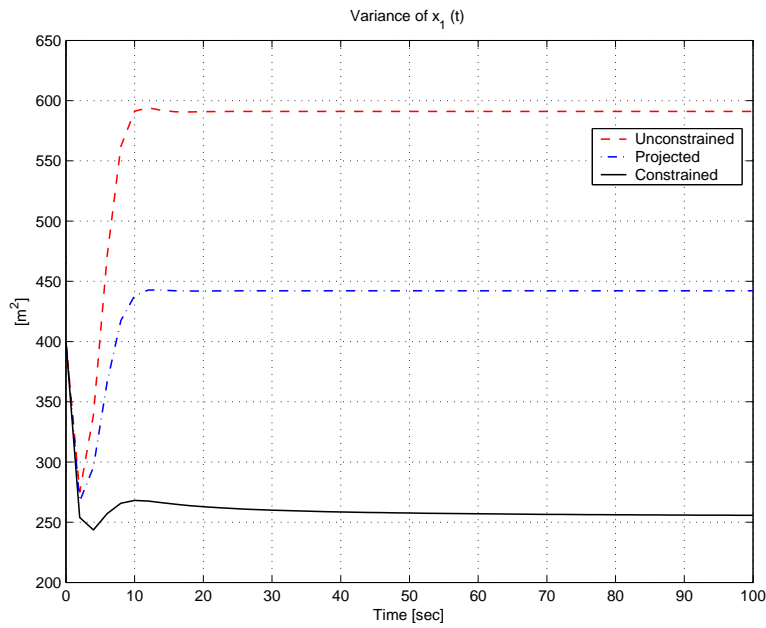


Fig. 3. Estimation error variance of the northerly position $\mathbf{x}_k(1)$ for a land-based vehicle

7 Concluding Remarks

In this paper, we have analyzed the three estimators that can be used for estimating linear systems with known state equality constraints. Among them, it was proved that the current constrained estimator is optimal and thus outperforms the unconstrained and the projected estimators. This fact was demonstrated through a numerical example. We have also extended the procedures used for discrete-time system to continuous-time system with state equality constraints.

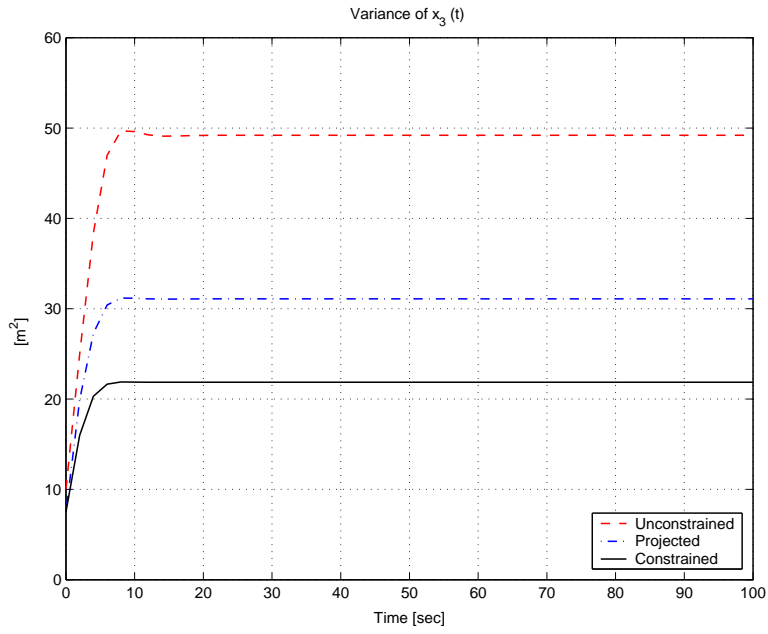


Fig. 4. Estimation error variance of the northerly velocity $\mathbf{x}_k(3)$ for a land-based vehicle

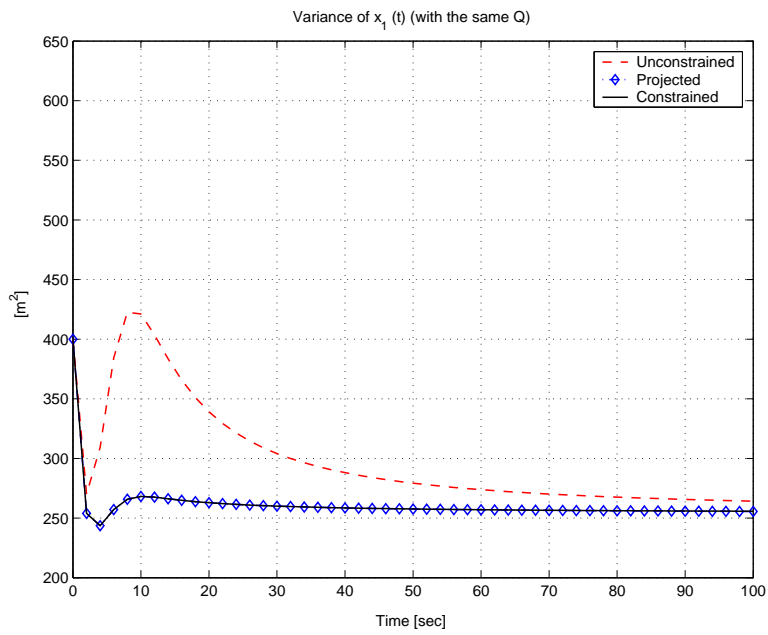


Fig. 5. Estimation error variance of the northerly position $\mathbf{x}_k(1)$ for a land-based vehicle (with the same $\mathbf{Q}^u = \mathbf{Q}^c$)

Acknowledgements

The research was supported by USA National Science Foundation Grants ECS-0200449 and ECS-0225530.

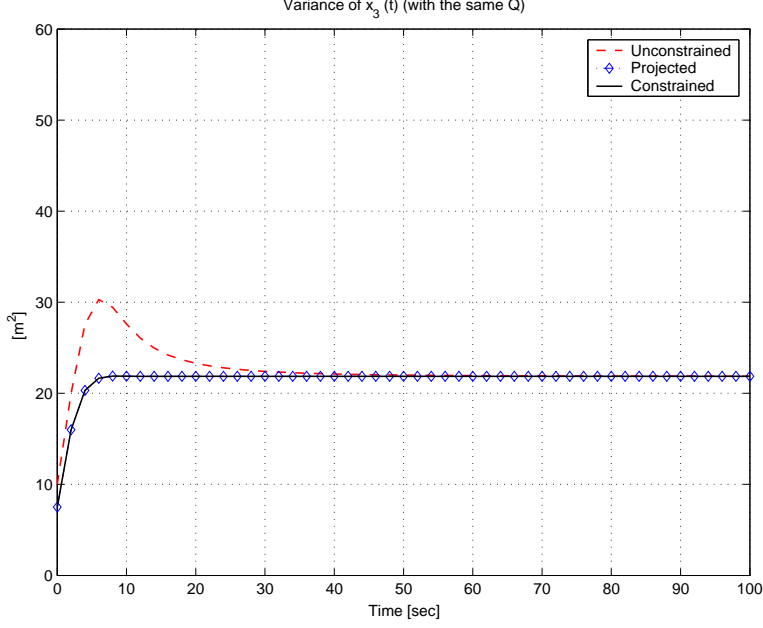


Fig. 6. Estimation error variance of the northerly velocity $\mathbf{x}_k(3)$ for a land-based vehicle (with the same $\mathbf{Q}^u = \mathbf{Q}^c$)

Appendices

A Proof of Corollary 1

(i) From (23) and (51), we have

$$\Sigma_{k+1|k+1}^p = \mathbf{P}_{k+1|k+1} \left\{ \Sigma_{k+1|k}^u - \Sigma_{k+1|k}^u \mathbf{C}^T \left(\mathbf{C} \Sigma_{k+1|k}^u \mathbf{C}^T + \mathbf{R} \right)^{-1} \mathbf{C} \Sigma_{k+1|k}^u \right\} \mathbf{P}_{k+1|k+1}, \quad (\text{A.1})$$

and for the constrained estimator

$$\Sigma_{k+1|k+1}^c = \Sigma_{k+1|k}^c - \Sigma_{k+1|k}^c \mathbf{C}^T \left(\mathbf{C} \Sigma_{k+1|k}^c \mathbf{C}^T + \mathbf{R} \right)^{-1} \mathbf{C} \Sigma_{k+1|k}^c. \quad (\text{A.2})$$

Subtracting (A.2) from (A.1), and then using the property (17) yields

$$\Sigma_{k+1|k+1}^p - \Sigma_{k+1|k+1}^c = \mathbf{P}_{k+1|k+1} \Delta_{k+1}^f \mathbf{P}_{k+1|k+1} \quad (\text{A.3})$$

where

$$\begin{aligned} \Delta_{k+1}^f \triangleq & \Sigma_{k+1|k}^u - \Sigma_{k+1|k}^u \mathbf{C}^T \left(\mathbf{C} \Sigma_{k+1|k}^u \mathbf{C}^T + \mathbf{R} \right)^{-1} \mathbf{C} \Sigma_{k+1|k}^u \\ & - \Sigma_{k+1|k}^c + \Sigma_{k+1|k}^c \mathbf{C}^T \left(\mathbf{C} \Sigma_{k+1|k}^c \mathbf{C}^T + \mathbf{R} \right)^{-1} \mathbf{C} \Sigma_{k+1|k}^c. \end{aligned} \quad (\text{A.4})$$

Since Δ_{k+1}^f can be considered as the difference $\Sigma_{k+2|k+1}^u - \Sigma_{k+2|k+1}^c$ for a system of the pair $[\mathbf{I}, \mathbf{C}]$ with the same process noise and measurement noise covariances, we deduce that $\Delta_{k+1}^f \geq \mathbf{0}$, by using the monotonicity of the Riccati Difference Equation for the given fact $\Sigma_{k+1|k}^u \geq \Sigma_{k+1|k}^c$. Therefore, together with (51), we obtain (52a) and (52b).

(ii) can be proved similarly to (ii) of Theorem 4. \square

B Proof of Theorem 6

Since (i) and (ii) can be proved similarly to the discrete-time case given in Theorem 3, we prove only (iii). Before we derive (67), the following lemma is needed.

Lemma 4 (Skelton (1988)) *For a nonsingular matrix $\mathbf{M}(t)$, the following holds.*

$$\frac{d}{dt}[\mathbf{M}^{-1}(t)] = -\mathbf{M}^{-1}(t) \left[\frac{d}{dt} \mathbf{M}(t) \right] \mathbf{M}^{-1}(t) \quad (\text{B.1})$$

Then, by using Lemma 4, the time-derivative of the projection matrix $\mathbf{P}(t)$ in (66) is given by

$$\begin{aligned} \dot{\mathbf{P}}(t) &= -\dot{\Sigma}^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} - \Sigma^u(t) \mathbf{D}^T \frac{d}{dt} \left\{ [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \right\} \mathbf{D} \\ &= -\dot{\Sigma}^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} + \Sigma^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} [\mathbf{D} \dot{\Sigma}^u(t) \mathbf{D}^T] [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \\ &= -\left\{ \mathbf{I} - \Sigma^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \right\} \dot{\Sigma}^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \\ &= -\mathbf{P}(t) \dot{\Sigma}^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D}. \end{aligned} \quad (\text{B.2})$$

By differentiating (65) with respect to t , and using (58) and (B.2), we obtain

$$\begin{aligned} \dot{\Sigma}^p(t) &= \mathbf{P}(t) \dot{\Sigma}^u(t) \mathbf{P}^T(t) + \dot{\mathbf{P}}(t) \Sigma^u(t) \mathbf{P}^T(t) + \mathbf{P}(t) \Sigma^u(t) \dot{\mathbf{P}}^T(t) \\ &= \mathbf{P}(t) \left\{ \mathbf{A} \Sigma^u(t) + \Sigma^u(t) \mathbf{A}^T - \Sigma^u(t) \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \Sigma^u(t) + \mathbf{Q}^u \right\} \mathbf{P}^T(t) \\ &\quad - \mathbf{P}(t) \dot{\Sigma}^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \Sigma^u(t) \mathbf{P}^T(t) - \mathbf{P}(t) \Sigma^u(t) \mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \dot{\Sigma}^u(t) \mathbf{P}^T(t) \\ &= \mathbf{P}(t) \left\{ \mathbf{A} \Sigma^p(t) + \Sigma^p(t) \mathbf{A}^T - \Sigma^p(t) \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C} \Sigma^p(t) + \mathbf{Q}^u \right\} \mathbf{P}^T(t) + \mathbf{P}(t) \Delta^c(t) \mathbf{P}^T(t) \end{aligned} \quad (\text{B.3})$$

where $\Delta^c(t)$ is given by (68), and for the last equality the following relation was used

$$\mathbf{D}^T [\mathbf{D} \Sigma^u(t) \mathbf{D}^T]^{-1} \mathbf{D} \Sigma^u(t) \mathbf{P}^T(t) = [\mathbf{I} - \mathbf{P}^T(t)] \mathbf{P}^T(t) = \mathbf{0}, \quad (\text{B.4})$$

and finally using the continuous-time version of the property (17) yields (67). \square

References

- Alouani, A. T. and W. D. Blair (1993). Use of a kinematic constraint in tracking constant speed, maneuvering targets. *IEEE Trans. Automat. Contr.* **38**, 1107–1111.
- Altman, Russ B. and Rhea Tombropoulos (1994). Probabilistic constraint satisfaction. Reports SMI-94-0531. Stanford Medical Informatics.
- Anderson, Brian D. O. and John B. Moore (1979). *Optimal Filtering*. Prentice-Hall. Englewood Cliffs, NY.
- Bitmead, Robert R. and Michel Gevers (1991). Riccati difference and differential equations: Convergence, monotonicity and stability. In: *The Riccati Equation* (Sergio Bittanti, Alan J. Laub and Jan C. Willems, Eds.). pp. 263–291. Springer-Verlag. Berlin, Germany.
- Bryson, Arthur E. Jr. and Yu-Chi Ho (1975). *Applied optimal control : optimization, estimation, and control*. Hemisphere Pub. Corp.. New York, NY.
- Chen, Yuan-Hwang and Ching-Tai Chiang (1993). Adaptive beamforming using the constrained Kalman filter. *IEEE Trans. Antennas Propagat.* **41**, 1576–1580.
- Chia, Tien-Li (1985). Parameter Identification and State Estimation of Constrained Systems. PhD thesis. Case Western Reserve Univ.. Cleveland, OH.

- Chiang, Y. T., L. S. Wang, F. R. Chang and H. M. Peng (2002). Constrained filtering method for attitude determination using GPS and Gyro. *IEE Proc.-Radar Sonar Navig.* **149**, 258–264.
- De Souza, Carlos E. (1989). On stabilizing properties of solutions of the riccati difference equation. *IEEE Trans. Automat. Contr.* **34**, 1313–1316.
- Geeter, Jan De, Hendrik Van Brussel and Joris De Schutter (1997). A smoothly constrained Kalman filter. *IEEE Trans. Pattern Anal. Machine Intell.* **19**(10), 1171–1177.
- Hayward, S. D. (1998). Constrained Kalman filter for least-squared estimation of time-varying beamforming weights. In: *Mathematics in Signal Processing IV* (J. G. McWhirter and I. K. Proudler, Eds.). pp. 113–125. Oxford University Press Inc.. New York.
- Hemami, Hooshang and Bostwick F. Wyman (1979). Modeling and control of constrained dynamic systems with application to biped locomotion in the frontal plane. *IEEE Trans. Automat. Contr.* **24**(4), 526–535.
- Lefferts, E. J., F. L. Markley and M. D. Shuster (1982). Kalman filtering for spacecraft attitude estimation. *J. Guid. Control Dyn.* **5**, 417–429.
- Luenberger, David G. (1977). Dynamic equations in descriptor form. *IEEE Trans. Automat. Contr.* **22**(3), 312–321.
- Maciejowski, Jan Marian (2002). *Predictive Control with Constraints*. Pearson Education Ltd.. Essex, England.
- Mahata, Kaushik and Törsen Söderstrom (2004). Improved estimation performance using known linear constraints. *Automatica* **40**, 1307–1318.
- Peng, H. M., Y. T. Chiang, F. R. Chang and L. S. Wang (2000). Maximum-likelihood-based filtering for attitude determination via GPS carrier phase. In: *Proc. of IEEE Position Location and Navigation Symposium*. San Diego, California. pp. 480–487.
- Poubelle, Marie-Antoinette, Robert R. Bitmead and Michel R. Gevers (1988). Fake algebraic riccati techniques and stability. *IEEE Trans. Automat. Contr.* **33**(4), 379–381.
- Rao, Christopher V. (2000). Moving Horizon Strategies for the Constrained Monitoring and Control of Nonlinear Discrete-Time Systems. PhD thesis. Univ. of Wisconsin-Madison. Madison, WI.
- Rao, Christopher V., James B. Rawlings and David Q. Mayne (2003). Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Trans. Automat. Contr.* **48**, 246–258.
- Simon, Dan and Tien Li Chia (2002). Kalman filtering with state equality constraints. *IEEE Trans. Aerosp. Electron. Syst.* **38**, 128–136.
- Skelton, R. E. (1988). *Dynamic systems control*. John Wiley & Sons. New York, NY.
- Skelton, R. E., T. Iwasaki and K. Grigoriadis (1998). *A Unified Algebraic Approach to Linear Control Design*. Taylor and Francis. London, UK.
- Stengel, Robert F. (1994). *Optimal Control and Estimation*. Dover Publications, Inc.. Mineola, NY.
- Tahk, Minjea and Jason L. Speyer (1990). Target tracking problems subject to kinematic constraints. *IEEE Trans. Automat. Contr.* **35**, 324–326.
- Wen, W. and H. F. Durrant-Whyte (1992). Model-based multi-sensor data fusion. In: *Proc. 1992 IEEE International Conference on Robotics and Automation*. Nice, France. pp. 1720–1726.
- Wonham, W. Murray (1979). *Linear multivariable control: a Geometric Approach*. Springer-Verlag. New York, NY.