Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007



# State Estimation for Equality-Constrained Linear Systems

B. O. S. Teixeira*†, J. Chandrasekar†, L. A. B. Tôrres*, L. A. Aguirre*, and D. S. Bernstein†

***Abstract*— We address the state-estimation problem for linear systems in a context where prior knowledge, in addition to the model and the measurements, is available in the form of an equality constraint. First, we investigate from where an equality constraint arises in a dynamic system. Then, the equality-constrained Kalman filter (ECKF) is derived as the solution to the equality-constrained state-estimation problem and compared to alternative algorithms. These methods are investigated in an example. In addition to exactly satisfying an equality constraint on the system, ECKF produce more accurate and more informative estimates than the unconstrained estimates.**

## I. Introduction

The classical Kalman filter (KF) for linear systems provides optimal state estimates under standard noise and model assumptions [7]. In practice, however, additional information about the system may be available, and this information may be useful for improving state estimates. Technically speaking, it is not possible to improve estimates from KF since these are optimal. Instead, a scenario we have in mind is the case in which the dynamics and the disturbances are such that the state of the system satisfies an equality or inequality constraint. For example, in a chemical reaction, the species concentrations are nonnegative [13], whereas in a compartmental model with zero net inflow [4], mass is conserved. Likewise, in undamped mechanical systems, such as a system with Hamiltonian dynamics, conservation laws hold. In the quaternion-based attitude estimation problem, the attitude vector must have unit norm [5]. Additional examples arise in optimal control [6], parameter estimation [1], and navigation [2, 14]. In such cases, we wish to obtain state estimates that take advantage of prior knowledge of the states and use this information to obtain better estimates than would be provided by KF in the absence of such information.

Various algorithms have been developed for equality-constrained state estimation. One of the most popular techniques is the measurement-augmentation KF (MAKF), in which a perfect measurement of the constrained quantity is assumed to be available [2, 14]. In addition, estimate [12] and system [8] projection methods have been considered. Regarding inequality constraints, moving horizon techniques [9], Kalman-based algorithms [13], and probabilistic methods [10] have been developed.

This research was supported by the National Council for Scientific and Technological Development – CNPq, Brazil, and by the National Science Foundation Information Technology Research Initiative, through grants ATM-0325332 and CNS-0539053 to the University of Michigan, USA.

*Department of Electronic Engineering, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil, {brunoot,torres,aguirre}@cpdee.ufmg.br.

†Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, USA, {jchandra,dsbaero}@umich.edu.

In the context of equality-constrained systems, three contributions are presented in this paper. First, we investigate how a linear equality state constraint arises in a linear system and present necessary conditions on both the dynamics and process noise for the state to be equality constrained. In [8], this problem is stated in the opposite way, that is, given that a system satisfies an equality constraint, the goal is to characterize the process noise. In these cases, we show that an equality-constrained linear system is not controllable from the process noise and that additional information regarding the initial condition provided by the equality constraint is useful for improving the classical KF estimates.

Second, we derive the equality-constrained KF (ECKF) as the solution to the equality-constrained state-estimation problem.

Finally, we prove the equivalence between ECKF and MAKF and show the connections of the former with the estimate and system projection approaches mentioned above. We compare these four algorithms using a compartmental model in which the disturbances are constrained so that mass is conserved.

## II. State Estimation for Linear Systems

For the linear stochastic discrete-time dynamic system

$$x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + G_{k-1}w_{k-1}, \quad (2.1)$$

$$y_k = C_k x_k + v_k, \quad (2.2)$$

where $A_{k-1} \in \mathbb{R}^{n\times n}$, $B_{k-1} \in \mathbb{R}^{n\times p}$, $G_{k-1} \in \mathbb{R}^{n\times q}$, and $C_k \in \mathbb{R}^{m\times n}$ are known matrices, the state-estimation problem can be described as follows. Assume that, for all $k \geq 1$, the known data are the measurements $y_k \in \mathbb{R}^m$, the inputs $u_{k-1} \in \mathbb{R}^p$, and the statistical properties of $x_0$, $w_{k-1}$ and $v_k$. The initial state vector $x_0 \in \mathbb{R}^n$ is assumed to be Gaussian with mean $\hat{x}_0$ and error-covariance $P_0^{xx} \triangleq \mathrm{E}\left[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^{\mathrm{T}}\right]$. The process noise $w_{k-1} \in \mathbb{R}^q$, which represents unknown input disturbances, and the measurement noise $v_k \in \mathbb{R}^m$, concerning inaccuracies in the measurements, are assumed white, Gaussian, zero mean, and mutually independent with known covariance matrices $Q_{k-1}$ and $R_k$, respectively. Next, define the cost function

$$J(x_k) \triangleq \rho(x_k|(y_1, \dots, y_k)), \quad (2.3)$$

which is the conditional probability density function of the state vector $x_k \in \mathbb{R}^n$ given the past and present measured data $y_1, \dots, y_k$. Under the stated assumptions, the maximization of (2.3) is the state estimation problem.

The recursive solution $\hat{x}_k$ to the state-estimation problem is given by the Kalman filter (KF) [7], whose *forecast* step is given by

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1} + B_{k-1}u_{k-1}, \quad (2.4)$$
$$P^{xx}_{k|k-1} = A_{k-1}P^{xx}_{k-1}A^{\mathrm{T}}_{k-1} + G_{k-1}Q_{k-1}G^{\mathrm{T}}_{k-1}, \quad (2.5)$$
$$\hat{y}_{k|k-1} = C_k\hat{x}_{k|k-1}, \quad (2.6)$$
$$P^{yy}_{k|k-1} = C_kP^{xx}_{k|k-1}C^{\mathrm{T}}_k + R_k, \quad (2.7)$$
$$P^{xy}_{k|k-1} = P^{xx}_{k|k-1}C^{\mathrm{T}}_k, \quad (2.8)$$

where $P^{xx}_{k|k-1} \triangleq \mathrm{E}\left[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}\right]$, $P^{yy}_{k|k-1} \triangleq \mathrm{E}\left[(y_k - \hat{y}_{k|k-1})(y_k - \hat{y}_{k|k-1})^{\mathrm{T}}\right]$, and $P^{xy}_{k|k-1} \triangleq \mathrm{E}\left[(x_k - \hat{x}_{k|k-1})(y_k - \hat{y}_{k|k-1})^{\mathrm{T}}\right]$, and whose *data-assimilation* step is given by

$$K_k = P^{xy}_{k|k-1}(P^{yy}_{k|k-1})^{-1}, \quad (2.9)$$
$$\hat{x}_k = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}), \quad (2.10)$$
$$P^{xx}_k = P^{xx}_{k|k-1} - K_kP^{yy}_{k|k-1}K^{\mathrm{T}}_k, \quad (2.11)$$

where $P^{xx}_k \triangleq \mathrm{E}\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^{\mathrm{T}}\right]$ and $K_k \in \mathbb{R}^{n\times m}$ is the Kalman gain matrix. The notation $\hat{z}_{k|k-1}$ indicates an estimate of $z_k$ at time $k$ based on information available up to and including time $k-1$. Likewise, $\hat{z}_k$ indicates an estimate of $z$ at time $k$ using information available up to and including time $k$. Model information is used during the forecast step, while measurement data are injected into the estimates during the data-assimilation step.

## III. State Estimation for Equality-Constrained Linear Systems

In (2.1), assume that $\mathrm{rank}(G_{k-1}) = q < n$, and define $r \triangleq n - q$, where $1 \le r \le n$. The case $r = n$ indicates that $G_{k-1}w_{k-1}$ is absent. Therefore, there exists $E_{k-1} \in \mathbb{R}^{r\times n}$ such that $\mathrm{rank}(E_{k-1}) = r$ and

$$E_{k-1}G_{k-1} = 0_{r\times q}. \quad (3.1)$$

Let $T_{1,k-1} \in \mathbb{R}^{(n-r)\times n}$ be such that $T_{k-1} \triangleq \begin{bmatrix} T_{1,k-1} \\ E_{k-1} \end{bmatrix} \in \mathbb{R}^{n\times n}$ is invertible. For example, we can choose $T_{1,k-1} \triangleq G^{\mathrm{T}}_{k-1}$. Multiplying (2.1) by $T$ yields

$$\begin{bmatrix} T_{1,k-1} \\ E_{k-1} \end{bmatrix} x_k = \begin{bmatrix} T_{1,k-1}A_{k-1} \\ E_{k-1}A_{k-1} \end{bmatrix} x_{k-1} + \begin{bmatrix} T_{1,k-1}B_{k-1} \\ E_{k-1}B_{k-1} \end{bmatrix} u_{k-1} + \begin{bmatrix} T_{1,k-1}G_{k-1} \\ 0_{r\times q} \end{bmatrix} w_{k-1}. \quad (3.2)$$

Hence, for all $k \ge 1$,

$$E_{k-1}x_k = e_{k-1}, \quad (3.3)$$

where $e_{k-1} \triangleq E_{k-1}(A_{k-1}x_{k-1} + B_{k-1}u_{k-1})$. Note that $e_{k-1}$ is not constant even if system (2.1)-(2.2) is time-invariant. Since $\mathrm{rank}(G_{k-1}) < n$, $G_{k-1}w_{k-1}$ has singular covariance $G_{k-1}Q_{k-1}G^{\mathrm{T}}_{k-1}$ [6, 8].

Let $s$ be an integer satisfying $1 \le s \le r$, and let $E_{k-1}$ be partitioned as $E_{k-1} \triangleq \begin{bmatrix} E_{1,k-1} \\ D_{k-1} \end{bmatrix}$, where $E_{1,k-1} \in \mathbb{R}^{(r-s)\times n}$ and $D_{k-1} \in \mathbb{R}^{s\times n}$. It thus follows from (3.1) that

$$D_{k-1}G_{k-1} = 0_{s\times q}. \quad (3.4)$$

Note that (3.4) holds, for all $D_{k-1} \in \mathbb{R}^{s\times n}$, if $r = n$.

*Proposition 3.1:* Assume that

$$D_{k-1}A_{k-1} = D_{k-1}, \quad (3.5)$$
$$D_{k-1}B_{k-1}u_{k-1} = 0_{s\times 1}, \quad \textit{for all} \quad k \ge 1. \quad (3.6)$$

Then, for all $k \ge 1$,

$$D_{k-1}x_k = d_{k-1}, \quad (3.7)$$

where

$$d_{k-1} \triangleq D_{k-1}x_{k-1}. \quad (3.8)$$

**Proof.** It follows from (3.3) that $D_{k-1}x_k = D_{k-1}(A_{k-1}x_{k-1} + B_{k-1}u_{k-1}) = D_{k-1}x_{k-1} = d_{k-1}$. □

*Corollary 3.1:* If (2.1)-(2.2) is time-invariant and (3.4)-(3.6) hold, then, for all $k \ge 1$,

$$Dx_k = d, \quad (3.9)$$

where $d \triangleq Dx_0$.

Note that the case $s = r = n$ is not of practical interest because it indicates $x_k = D^{-1}_{k-1}d_{k-1}$.

The next result shows that, if (2.1) is equality constrained, then it is not controllable from the process noise.

*Proposition 3.2:* Assume that (3.4)-(3.6) hold, then $(A_{k-1}, G_{k-1})$ is not controllable.

**Proof.** Multiplying the controllability matrix

$$\mathcal{K}(A_{k-1}, G_{k-1}) \triangleq \begin{bmatrix} G_{k-1} & A_{k-1}G_{k-1} & \ldots & A^{n-1}_{k-1}G_{k-1} \end{bmatrix} \quad (3.10)$$

by $D_{k-1}$ yields

$$\begin{aligned} D_{k-1}\mathcal{K}(A_{k-1}, G_{k-1}) &= \begin{bmatrix} D_{k-1}G_{k-1} & D_{k-1}A_{k-1}G_{k-1} & \cdots & D_{k-1}A^{n-1}_{k-1}G_{k-1} \end{bmatrix} \\ &= \begin{bmatrix} 0_{s\times q} & D_{k-1}G_{k-1} & \cdots & D_{k-1}A^{n-2}_{k-1}G \end{bmatrix} \\ &= 0_{s\times nq}. \end{aligned}$$ □

Assuming that, for all $k \ge 1$, $D_{k-1}$ satisfies (3.4)-(3.6) and $d_{k-1}$ is known, the objective of the *equality-constrained state-estimation* problem is to maximize (2.3) subject to (3.7).

## IV. Equality-Constrained Kalman Filter

In this section we solve the equality-constrained state-estimation problem to obtain the equality-constrained Kalman filter (ECKF).

*Lemma 4.1:* $\hat{x}_k$ maximizes $J$ given by (2.3) if and only if $\hat{x}_k$ minimizes

$$\mathcal{J}(x_k) = \Big[(x_k - \hat{x}_{k|k-1})^{\mathrm{T}}(P^{xx}_{k|k-1})^{-1}(x_k - \hat{x}_{k|k-1}) + (y_k - C_kx_k)^{\mathrm{T}}(R_k)^{-1}(y_k - C_kx_k)\Big], \quad (4.1)$$

where $\hat{x}_{k|k-1}$ and $P^{xx}_{k|k-1}$ are given by (2.4) and (2.5).

**Proof.** See [7, pp. 207–208]. □

*Theorem 4.1:* Let $\hat{x}^{\mathrm{p}}_k$ denote the solution of the equality-constrained state estimation problem and define the error covariance $P^{xx\mathrm{p}}_k \triangleq \mathrm{E}\left[(x_k - \hat{x}^{\mathrm{p}}_k)(x_k - \hat{x}^{\mathrm{p}}_k)^{\mathrm{T}}\right]$. Also let $\hat{x}_{k|k-1}$ and $P^{xx}_{k|k-1}$ be given by

$$\hat{x}_{k|k-1} \triangleq A_{k-1}\hat{x}^{\mathrm{p}}_{k-1} + B_{k-1}u_{k-1}, \quad (4.2)$$
$$P^{xx}_{k|k-1} \triangleq A_{k-1}P^{xx\mathrm{p}}_{k-1}A^{\mathrm{T}}_{k-1} + G_{k-1}Q_{k-1}G^{\mathrm{T}}_{k-1}. \quad (4.3)$$

Define

$$\hat{d}_{k-1} \triangleq D_{k-1}\hat{x}_k, \tag{4.4}$$
$$P_k^{dd} \triangleq D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}}, \tag{4.5}$$
$$P_k^{xd} \triangleq P_k^{xx}D_{k-1}^{\mathrm{T}}, \tag{4.6}$$
$$K_k^{\mathrm{p}} \triangleq P_k^{xd}(P_k^{dd})^{-1}, \tag{4.7}$$

where $\hat{x}_k$ is given by (2.10) and $P_k^{xx}$ is given by (2.11). Then $\hat{x}_k^{\mathrm{p}}$ and $P_k^{xx\mathrm{p}}$ are given by

$$\hat{x}_k^{\mathrm{p}} = \hat{x}_k + K_k^{\mathrm{p}}(d_{k-1} - \hat{d}_{k-1}), \tag{4.8}$$
$$P_k^{xx\mathrm{p}} = P_k^{xx} - K_k^{\mathrm{p}}P_k^{dd}K_k^{\mathrm{p}\,\mathrm{T}}. \tag{4.9}$$

**Proof.** Using Lemma 4.1, let $\lambda \in \mathbb{R}^s$ and define the Lagrangian $L \triangleq \mathcal{J}(x_k) + 2\lambda^{\mathrm{T}}(D_{k-1}x_k - d_{k-1})$. The necessary conditions for a minimizer $\hat{x}_k^{\mathrm{p}}$ are given by

$$\frac{\partial L}{\partial x_k} = (P_{k|k-1}^{xx})^{-1}(\hat{x}_k^{\mathrm{p}} - \hat{x}_{k|k-1}) - C_k^{\mathrm{T}}R_k^{-1}(y_k - C_k\hat{x}_k^{\mathrm{p}}) + D_{k-1}^{\mathrm{T}}\lambda = 0_{n\times 1}, \tag{4.10}$$
$$\frac{\partial L}{\partial \lambda} = D_{k-1}\hat{x}_k^{\mathrm{p}} - d_{k-1} = 0_{s\times 1}. \tag{4.11}$$

It follows from (4.10) that

$$((P_{k|k-1}^{xx})^{-1} + C_k^{\mathrm{T}}R_k^{-1}C_k)(\hat{x}_k^{\mathrm{p}} - \hat{x}_{k|k-1}) = C_k^{\mathrm{T}}R_k^{-1}(y_k - C_k\hat{x}_{k|k-1}) - D_{k-1}^{\mathrm{T}}\lambda. \tag{4.12}$$

From (2.11), using (2.7)-(2.9) and the matrix inversion lemma [3], we have

$$\begin{aligned} P_k^{xx} &= P_{k|k-1}^{xx} - K_k P_{k|k-1}^{yy}K_k^{\mathrm{T}} \\ &= P_{k|k-1}^{xx} - P_{k|k-1}^{xx}C_k^{\mathrm{T}}(C_kP_{k-1}^{xx}C_k^{\mathrm{T}} + R_k)^{-1}C_kP_{k|k-1}^{xx} \\ &= ((P_{k|k-1}^{xx})^{-1} + C_k^{\mathrm{T}}R_k^{-1}C_k))^{-1}. \end{aligned} \tag{4.13}$$

Furthermore, from (2.9), using (2.7)-(2.8), we have

$$\begin{aligned} K_k &= P_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1} = P_{k|k-1}^{xx}C_k^{\mathrm{T}}(C_kP_{k|k-1}^{xx}C_k^{\mathrm{T}} + R_k)^{-1} \\ &= P_k^{xx}(P_k^{xx})^{-1}P_{k|k-1}^{xx}C_k^{\mathrm{T}}(C_kP_{k|k-1}^{xx}C_k^{\mathrm{T}} + R_k)^{-1} \\ &= P_k^{xx}(C_k^{\mathrm{T}}R_k^{-1}C_k + (P_{k|k-1}^{xx})^{-1})P_{k|k-1}^{xx}C_k^{\mathrm{T}}(C_kP_{k|k-1}^{xx}C_k^{\mathrm{T}} + R_k)^{-1} \\ &= P_k^{xx}C_k^{\mathrm{T}}R_k^{-1}(C_kP_{k|k-1}^{xx}C_k^{\mathrm{T}} + R_k)(C_kP_{k|k-1}^{xx}C_k^{\mathrm{T}} + R_k)^{-1} \\ &= P_k^{xx}C_k^{\mathrm{T}}R_k^{-1}. \end{aligned} \tag{4.14}$$

Substituting (4.13) and (4.14) into (4.12) and multiplying by $P_k^{xx}$ yields

$$\hat{x}_k^{\mathrm{p}} = \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}) - P_k^{xx}D_{k-1}^{\mathrm{T}}\lambda. \tag{4.15}$$

Substituting (4.15) into (4.11) yields

$$d_{k-1} = D_{k-1}\hat{x}_{k|k-1} + D_{k-1}P_k^{xx}C_k^{\mathrm{T}}R_k^{-1}(y_k - C_k\hat{x}_{k|k-1}) - D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}}\lambda,$$

which implies

$$\begin{aligned} \lambda &= (D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1}(D_{k-1}\hat{x}_{k|k-1} - d_{k-1}) + \\ &\quad (D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1}D_{k-1}K_k(y_k - C_k\hat{x}_{k|k-1}). \end{aligned} \tag{4.16}$$

Likewise, substituting (4.16) into (4.15) yields

$$\begin{aligned} \hat{x}_k^{\mathrm{p}} &= \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}) - \\ &\quad P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1}(D_{k-1}\hat{x}_{k|k-1} - d_{k-1}) - \\ &\quad P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1}D_{k-1}K_k(y_k - C_k\hat{x}_{k|k-1}) \\ &= \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}) - P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1} \times \\ &\quad (D_{k-1}\hat{x}_{k|k-1} - d_{k-1} + D_{k-1}K_ky_k - D_{k-1}K_kC_k\hat{x}_{k|k-1}) \\ &= \hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}) + P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1} \times \\ &\quad [d_{k-1} - D_{k-1}(\hat{x}_{k|k-1} + K_k(y_k - C_k\hat{x}_{k|k-1}))]. \end{aligned} \tag{4.17}$$

Now using (4.4)-(4.7), (2.9)-(2.11), we obtain

$$\begin{aligned} \hat{x}_k^{\mathrm{p}} &= \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) + P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1} \\ &\quad [d_{k-1} - D_{k-1}(\hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}))] \\ &= \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) + K_k^{\mathrm{p}}(d_{k-1} - \hat{d}_{k-1}) \\ &= \hat{x}_k + K_k^{\mathrm{p}}(d_{k-1} - \hat{d}_{k-1}), \end{aligned}$$

which proves (4.8).

Given (2.11) and the symmetry between (4.8) and (2.10), it follows that $P_k^{xx\mathrm{p}}$ is given by (4.9). □

Note that ECKF is expressed in three steps, namely, the *forecast* step (4.2)-(4.3), (2.6)-(2.8), the *data-assimilation* step (2.9)-(2.11), and the *projection* step (4.4)-(4.9), where the updated estimates are projected onto the hyperplane defined by the equality constraint (3.7).

*Lemma 4.2:* Let $\mathcal{N}(D_{k-1})$ denote the null space of $D_{k-1}$, let $W \in \mathbb{R}^{n\times n}$ be positive definite, and define

$$\mathcal{P}_{\mathcal{N}(D_{k-1})} \triangleq I_{n\times n} - WD_{k-1}^{\mathrm{T}}(D_{k-1}WD_{k-1}^{\mathrm{T}})^{-1}D_{k-1}. \tag{4.18}$$

Then $\mathcal{P}_{\mathcal{N}(D_{k-1})}$ is an oblique projector with range $\mathcal{N}(D_{k-1})$.

For the following two results, let $\hat{x}_k$ given by (2.10) and $P_k^{xx}$ given by (2.11) denote the updated estimate and updated error covariance of ECKF. Also, let $\hat{x}_k^{\mathrm{p}}$ given by (4.8) and $P_k^{xx\mathrm{p}}$ given by (4.9) denote the projected estimate and projected error covariance of ECKF.

*Proposition 4.1:* Set $W = P_k^{xx}$ in (4.18). Then, the projection step (4.4)-(4.9) is equivalent to

$$\hat{x}_k^{\mathrm{p}} = \mathcal{P}_{\mathcal{N}(D_{k-1})}\hat{x}_k + \overline{d}_{k-1}, \tag{4.19}$$
$$P_k^{xx\mathrm{p}} = \mathcal{P}_{\mathcal{N}(D_{k-1})}P_k^{xx}, \tag{4.20}$$

where $\overline{d}_{k-1} \triangleq P_k^{xx}D_{k-1}^{\mathrm{T}}(D_{k-1}P_k^{xx}D_{k-1}^{\mathrm{T}})^{-1}d_{k-1}$.

**Proof.** Using Lemma 4.2 and substituting (4.4)-(4.7) into (4.8) and (4.9) yields (4.19)-(4.20). □

*Proposition 4.2:* Assume that (2.1)-(2.2) is time invariant. Also, assume that $D$ in (3.9) satisfies (3.4)-(3.6). Furthermore, assume that, for a given $k-1$, $D\hat{x}_{k-1}^{\mathrm{p}} = d$ and $DP_{k-1}^{xx\mathrm{p}} = 0_{s\times n}$. Then $D\hat{x}_k = d$, $DP_k^{xx} = 0_{s\times n}$, $\hat{x}_k^{\mathrm{p}} = \hat{x}_k$, and $P_k^{xx\mathrm{p}} = P_k^{xx}$.

**Proof.** Multiplying (4.2)-(4.3) by $D$ yields

$$D\hat{x}_{k|k-1} = DA\hat{x}_{k-1}^{\mathrm{p}} + DBu_{k-1} = D\hat{x}_{k-1}^{\mathrm{p}} + 0_{s\times 1} = d, \tag{4.21}$$
$$\begin{aligned} DP_{k|k-1}^{xx} &= DAP_{k-1}^{xx\mathrm{p}}A^{\mathrm{T}} + DGQ_{k-1}G^{\mathrm{T}} \\ &= DP_{k-1}^{xx\mathrm{p}}A^{\mathrm{T}} + 0_{s\times q}Q_{k-1}G^{\mathrm{T}} = 0_{s\times n}A^{\mathrm{T}} = 0_{s\times n}. \end{aligned} \tag{4.22}$$

With (2.8) and (4.22), multiplying (2.9) by $D$ yields

$$DK_k = DP_{k|k-1}^{xy}(P_{k|k-1}^{yy})^{-1} = DP_{k|k-1}^{xx}C^{\mathrm{T}}(P_{k|k-1}^{yy})^{-1} = 0_{s\times m}. \tag{4.23}$$

With (4.21) and (4.23), multiplying (2.10)-(2.11) by $D$ yields

$$D\hat{x}_k = D\hat{x}_{k|k-1} + DK_k(y_k - \hat{y}_{k|k-1}) = d, \tag{4.24}$$
$$DP_k^{xx} = DP_{k|k-1}^{xx} - DK_kP_{k|k-1}^{yy}K_k^{\mathrm{T}} = 0_{s\times n}. \tag{4.25}$$

Given (4.24)-(4.25), from (4.19)-(4.20), we have $\hat{x}_k^{\mathrm{p}} = \hat{x}_k$ and $P_k^{xx\mathrm{p}} = P_k^{xx}$. □

*Corollary 4.1:*. Assume $D\hat{x}_1^{\mathrm{p}} = d$ and $DP_1^{xx\mathrm{p}} = 0_{s\times n}$. Then, for all $k \geq 2$, $D\hat{x}_k = d$ and $DP_k^{xx} = 0_{s\times n}$.

Therefore, for time-invariant systems, whenever (3.4)-(3.6) hold, the projection step of ECKF given by (4.4)-(4.9) is required only at $k = 1$, so that, for all $k \geq 2$, the updated estimate $\hat{x}_k$ given by (2.10) satisfies $D\hat{x}_k = d$.

## V. Connections of ECKF to Other Approaches

We now compare ECKF to three Kalman filtering algorithms whose state estimates satisfy an equality constraint.

First we consider the *measurement–augmentation Kalman filter* (MAKF) [2, 14], which treats (3.7) as perfect measurements. In Appendix I, we present the MAKF equations and prove that MAKF and ECKF estimates are equal.

In Appendix II, in the context of time-invariant systems, we show the connection between ECKF and the *projected Kalman filter by system projection* (PKF-SP) [8], which, assuming that (3.4)-(3.6) hold, incorporates the information provided by (3.7) only in filter initialization, that is, $k = 0$.

Finally, in Appendix III, we briefly review the *projected Kalman filter by estimate projection* (PKF-EP) [11, 12], which projects $\hat{x}_k$ onto the hyperplane (3.7) for all $k \geq 1$. Unlike ECKF, the projected estimate of PKF-EP is not recursively fed back in the next iteration. Fig. 1 illustrates how the forecast, data-assimilation, and projection steps are connected for ECKF, PKF-SP, and PKF-EP.

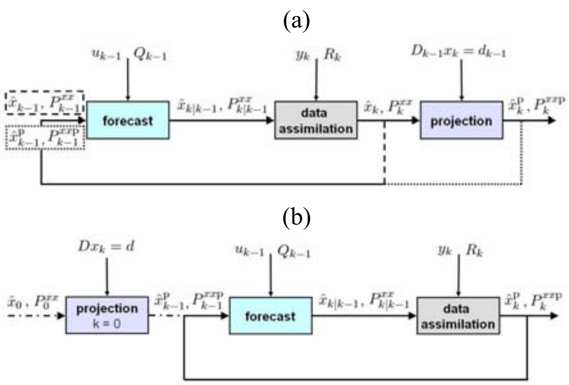


Fig. 1. Comparative diagram of (a) the equality-constrained Kalman filter (ECKF) (. . .) and the projected Kalman filter by estimate projection (PKF-EP) (−−) and (b) the projected Kalman filter by system projection (PKF-SP) (−·−). In ECKF, the projection step is connected by feedback recursion. In PKF-SP, the initial state estimate and the associated error covariance carry the information provided by the equality constraint.

## VI. Compartmental System Example

Consider the linear discrete-time compartmental model (2.1)-(2.2) [4] whose parameters are given by

$$A = \begin{bmatrix} 0.94 & 0.028 & 0.019 \\ 0.038 & 0.95 & 0.001 \\ 0.022 & 0.022 & 0.98 \end{bmatrix}, \quad B = 0_{3\times 1},$$

$$G = \begin{bmatrix} 0.05 & -0.03 \\ -0.02 & 0.01 \\ -0.03 & 0.02 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \tag{6.1}$$

with initial condition $x_0 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$ and noise covariance matrices $Q_{k-1} = \sigma_w^2 I_{3\times 3}$ and $R_k = \sigma_v^2 I_{2\times 2}$. The free-run simulation of this system is shown in Fig. 2ab for $\sigma_w = 1.0$ and $\sigma_v = 0.01$. Note that (3.4)-(3.6) hold for (6.1) such that the trajectory of $x_k \in \mathbb{R}^3$ lies on a plane (3.7), whose parameters are given by $D = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$, and $d = 3$. that is, conservation of mass is verified.

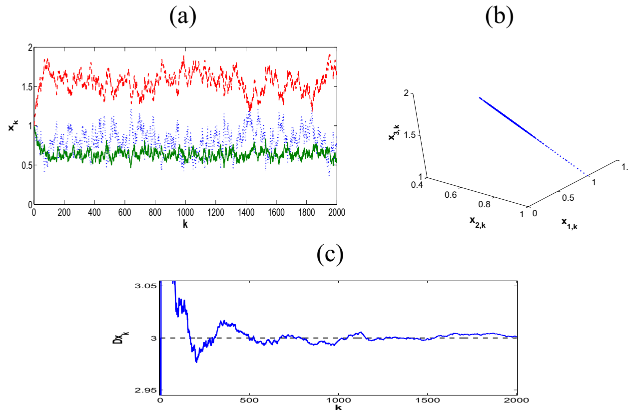


Fig. 2. Free-run simulation of the compartmental model. In (a), the state components are shown evolving with time and, in (b), in state space. In (c), it is shown the estimate of the total mass (constraint) $Dx_k$ using KF (——) in comparison with the true value (−−).

TABLE I: Average of percent RMS constraint error, trace of error covariance matrix, and RMS estimation error for 100-run Monte Carlo simulation for compartmental system, concerning different levels of process noise $\sigma_w = 0,\ 0.1,\ 0.5,$ and $1.0$, and algorithms, namely, KF, ECKF, MAKF, PKF-EP, and PKF-SP.

| $\sigma_w$ | KF | ECKF | MAKF | PKF-EP | PKF-SP |
|---|---|---|---|---|---|
| | Percent RMS constraint error | | | | |
| 0 | 0.12 | $4.52\times10^{-15}$ | $4.24\times10^{-11}$ | $4.53\times10^{-15}$ | $8.19\times10^{-12}$ |
| 0.1 | 0.22 | $4.52\times10^{-15}$ | $2.01\times10^{-11}$ | $4.52\times10^{-15}$ | $4.05\times10^{-12}$ |
| 0.5 | 0.40 | $4.50\times10^{-15}$ | $0.88\times10^{-11}$ | $4.51\times10^{-15}$ | $3.92\times10^{-12}$ |
| 1.0 | 0.62 | $4.53\times10^{-15}$ | $0.50\times10^{-11}$ | $4.51\times10^{-15}$ | $3.98\times10^{-12}$ |
| | Trace of error covariance matrix ($\times10^{-4}$) | | | | |
| 0 | 0.0996 | 0.0012 | | | |
| 0.1 | 1.0515 | 0.6352 | | | |
| 0.5 | 2.8057 | 1.4722 | | | |
| 1.0 | 5.4646 | 1.8387 | | | |
| | RMS estimation error for $x_1$, $x_2$, and $x_3$ ($\times10^{-3}$) | | | | |
| 0 | 0.57, 0.36, 2.93 | 0.10, 0.16, 0.21 | | | |
| 0.1 | 6.26, 2.60, 7.34 | 6.25, 2.54, 4.19 | | | |
| 0.5 | 9.01, 4.58, 13.2 | 9.01, 4.55, 6.75 | | | |
| 1.0 | 9.35, 5.58, 19.7 | 9.35, 5.56, 8.07 | | | |

For state estimation, the KF algorithm is initialized with

$$\hat{x}_0 = \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}^{\mathrm{T}}, \quad P_0^{xx} = I_{3\times 3}. \tag{6.2}$$

Fig. 2c shows that KF estimates do not lie on the plane (3.7). Even if $\hat{x}_0 = x_0$ or $\sigma_w = 0$, KF does not produce estimates satisfying (3.7). Next, we implement the ECKF algorithm. From a 100-run Monte Carlo simulation for each one of these process noise levels, namely, $\sigma_w = 0,\ 0.1,\ 0.5,\ 1.0,$ and $\sigma_v = 0.01$, Table I shows that the ECKF estimates satisfy the equality constraint. In addition, these estimates are both more accurate (smaller root-mean-square (RMS) errors) and more informative (smaller trace of error covariance) than the KF estimates.

For MAKF, PKF-EP, and PKF-SP, initialization is given by (6.2), except for PKF-SP (see (7.30) in Appendix II). Table I summarizes the results. ECKF, MAKF, PKF-SP, and PKF-EP guarantee that (3.9) is satisfied and yield improved estimates compared to KF. All equality-constrained methods produce similar results concerning RMS error and trace of error covariance for this time-invariant system. This is in accordance with [8, Theorem 2] regarding PKF-SP and PKF-EP. However, though not shown in Table I, it is relevant to mention that PKF-EP produces less accurate and less informative forecast estimates $\hat{x}_{k|k-1}$ compared to the other constrained algorithms. This is expected because PKF-EP do not use $\hat{x}^{\mathrm{p}}_{k-1}$ to calculate $\hat{x}_{k|k-1}$.

## VII. Concluding Remarks

We have shown that the problem of equality-constrained state estimation for linear systems arises from both process noise and dynamic equations with special properties (3.4)-(3.6), such that the system is not controllable from the process noise. In this case the optimal estimates of the classical Kalman filter (KF) do not match the equality constraint (3.7).

Then we have presented the equality-constrained KF (ECKF) as the solution to this problem. Moreover, we have proved its equivalence to the measurement-augmentation KF (MAKF) and have pointed its connections to the projection KF by system-projection (PKF-SP) and the projection KF by estimate-projection (PKF-EP).

We have compared these four methods by means of an example: a compartmental model with mass conservation. Numerical results suggest that, in addition to exactly satisfying the equality constraint, ECKF produce more accurate and more informative estimates than KF. For the time-invariant linear scenario, ECKF, MAKF, PKF-SP, and PKF-EP have produced similar results.

## Appendix I: Equivalence of ECKF and MAKF

Define the augmented observation

$$\tilde{y}_k \triangleq \begin{bmatrix} y_k \\ d_{k-1} \end{bmatrix} = \tilde{C}_k x_k + \begin{bmatrix} v_k \\ 0_{s\times 1} \end{bmatrix}, \tag{7.1}$$

where

$$\tilde{C}_k \triangleq \begin{bmatrix} C_k \\ D_{k-1} \end{bmatrix}. \tag{7.2}$$

With (7.1), MAKF uses (2.4)-(2.5) together with the augmented forecast equations

$$\hat{\tilde{y}}_{k|k-1} = \tilde{C}_k \hat{x}_{k|k-1}, \tag{7.3}$$

$$\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1} = \tilde{C}_k P^{xx}_{k|k-1} \tilde{C}_k^{\mathrm{T}} + \tilde{R}_k, \tag{7.4}$$

$$\tilde{P}^{x\tilde{y}}_{k|k-1} = P^{xx}_{k|k-1} \tilde{C}_k^{\mathrm{T}}, \tag{7.5}$$

where $\tilde{R}_k \triangleq \begin{bmatrix} R_k & 0_{m\times s} \\ 0_{s\times m} & 0_{s\times s} \end{bmatrix}$, and the augmented data-assimilation equations given by

$$\tilde{K}_k = \tilde{P}^{x\tilde{y}}_{k|k-1} (\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1})^{-1}, \tag{7.6}$$

$$\hat{x}_k = \hat{x}_{k|k-1} + \tilde{K}_k (\tilde{y}_k - \hat{\tilde{y}}_{k|k-1}), \tag{7.7}$$

$$P^{xx}_k = P^{xx}_{k|k-1} - \tilde{K}_k \tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1} \tilde{K}_k^{\mathrm{T}}. \tag{7.8}$$

Let $\tilde{x}_{k|k-1} \triangleq \hat{x}_{k|k-1}$ (2.4) denote the forecast estimate provided by MAKF. Furthermore, let $\tilde{P}^{xx}_{k|k-1} \triangleq P^{xx}_{k|k-1}$ (2.5) be the associated forecast error covariance of MAKF. Also let $\hat{x}_{k|k-1}$ (4.2) and $P^{xx}_{k|k-1}$ (4.3) denote the forecast estimate and the associated error covariance of ECKF.

*Proposition 7.1:* Assume that $\tilde{x}_{k|k-1} = \hat{x}_{k|k-1}$ and $\tilde{P}^{xx}_{k|k-1} = P^{xx}_{k|k-1}$. Then $\tilde{x}_{k+1|k} = \hat{x}_{k+1|k}$ and $\tilde{P}^{xx}_{k+1|k} = P^{xx}_{k+1|k}$.

**Proof.** $\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1}$ (7.4) is equivalent to

$$\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1} = \begin{bmatrix} P^{yy}_{k|k-1} & C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} \\ D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} & P^{dd}_{k|k-1} \end{bmatrix}. \tag{7.9}$$

It follows from [3] that $\tilde{P}^{\tilde{y}\tilde{y}\ -1}_{k|k-1}$ has entries

$$\tilde{P}^{\tilde{y}\tilde{y}\ -1}_{k|k-1} = \begin{bmatrix} (\tilde{P}^{-1}_{k|k-1})_1 & (\tilde{P}^{-1}_{k|k-1})_{12} \\ \left((\tilde{P}^{-1}_{k|k-1})_{12}\right)^{\mathrm{T}} & (\tilde{P}^{-1}_{k|k-1})_2 \end{bmatrix}, \tag{7.10}$$

where

$$\begin{aligned} (\tilde{P}^{-1}_{k|k-1})_1 &\triangleq \left(P^{yy}_{k|k-1} - C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} P^{dd\ -1}_{k|k-1} D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}}\right)^{-1}, \\ (\tilde{P}^{-1}_{k|k-1})_{12} &\triangleq -\left(P^{yy}_{k|k-1} - C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} P^{dd\ -1}_{k|k-1} D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}}\right)^{-1} \\ &\qquad C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} P^{dd\ -1}_{k|k-1}, \\ (\tilde{P}^{-1}_{k|k-1})_2 &\triangleq \left(P^{dd}_{k|k-1} - D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} P^{yy\ -1}_{k|k-1} C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1}\right)^{-1} \end{aligned} \tag{7.11}$$

Furthermore, it can be shown that

$$\begin{aligned} (\tilde{P}^{-1}_{k|k-1})_1 &= (P^{yy}_{k|k-1})^{-1} + (P^{yy}_{k|k-1})^{-1} C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} \\ &\qquad (\tilde{P}^{-1}_{k|k-1})_2 D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} (P^{yy}_{k|k-1})^{-1}, \end{aligned} \tag{7.12}$$

$$(\tilde{P}^{-1}_{k|k-1})_{12} = -(P^{yy}_{k|k-1})^{-1} C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} (\tilde{P}^{-1}_{k|k-1})_2. \tag{7.13}$$

It follows from (2.9) that

$$K_k = P^{xx}_{k|k-1} C_k^{\mathrm{T}} (P^{yy}_{k|k-1})^{-1}, \tag{7.14}$$

Furthermore substituting (7.14) into (2.11) yields

$$P^{xx}_k = P^{xx}_{k|k-1} - P^{xx}_{k|k-1} C_k^{\mathrm{T}} (P^{yy}_{k|k-1})^{-1} C_k P^{xx}_{k|k-1}. \tag{7.15}$$

Hence,

$$\begin{aligned} (D_{k-1} P^{xx}_k D^{\mathrm{T}}_{k-1})^{-1} &= \Big(D_{k-1} P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} - \\ &\qquad D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} (P^{yy}_{k|k-1})^{-1} C_k P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1}\Big)^{-1} \\ &= (\tilde{P}^{-1}_{k|k-1})_2. \end{aligned} \tag{7.16}$$

Substituting (7.14) into (2.10) yields (4.8). Substituting (4.4) into (4.8) yields

$$\hat{x}^{\mathrm{p}}_k = \hat{x}_{k|k-1} + \begin{bmatrix} K_k - K^{\mathrm{p}}_k D_{k-1} K_k & K^{\mathrm{p}}_k \end{bmatrix} \left(\tilde{y}_k - \tilde{C}_k \hat{x}_{k|k-1}\right). \tag{7.17}$$

It follows from (4.5), (7.11), (7.13) and (7.16) that

$$K^{\mathrm{p}}_k = P^{xx}_{k|k-1} \tilde{C}_k^{\mathrm{T}} \begin{bmatrix} (\tilde{P}^{-1}_{k|k-1})_{12} \\ (\tilde{P}^{-1}_{k|k-1})_2 \end{bmatrix}. \tag{7.18}$$

Substituting (7.13) into (7.18) and substituting the resulting expression into $K_k - K^{\mathrm{p}}_k D_{k-1} K_k$ yields

$$\begin{aligned} K_k - K^{\mathrm{p}}_k D_{k-1} K_k &= P^{xx}_{k|k-1} C_k^{\mathrm{T}} \Big[P^{yy\ -1}_{k|k-1} + P^{yy\ -1}_{k|k-1} C_k P^{xx}_{k|k-1} \\ &\qquad D^{\mathrm{T}}_{k-1} (\tilde{P}^{-1}_{k|k-1})_2 D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} P^{yy\ -1}_{k|k-1}\Big] - \\ &\qquad P^{xx}_{k|k-1} D^{\mathrm{T}}_{k-1} (\tilde{P}^{-1}_{k|k-1})_2 D_{k-1} P^{xx}_{k|k-1} C_k^{\mathrm{T}} P^{yy\ -1}_{k|k-1}. \end{aligned} \tag{7.19}$$

Hence, (7.12) and (7.13) imply that

$$K_k - K^{\mathrm{p}}_k D_{k-1} K_k = P^{xx}_{k|k-1} \tilde{C}_k^{\mathrm{T}} \begin{bmatrix} (\tilde{P}^{-1}_{k|k-1})_1 \\ (\tilde{P}^{-1}_{k|k-1})^{\mathrm{T}}_{12} \end{bmatrix}. \tag{7.20}$$

Therefore, it follows from (7.18) and (7.20) that

$$\begin{bmatrix} K_k - K^{\mathrm{p}}_k D_{k-1} K_k & K^{\mathrm{p}}_k \end{bmatrix} = P^{xx}_{k|k-1} \tilde{C}_k^{\mathrm{T}} \tilde{P}^{\tilde{y}\tilde{y}\ -1}_{k|k-1}. \tag{7.21}$$

Since the estimate $\tilde{x}_k$ of MAKF is given by

$$\tilde{x}_k = \tilde{x}_{k|k-1} + \tilde{K}_k(\tilde{y}_k - \tilde{C}_k\tilde{x}_{k|k-1}), \qquad (7.22)$$

where $\tilde{K}_k = \tilde{P}^{xx}_{k|k-1}\tilde{C}_k^{\mathrm{T}}(\tilde{C}_k\tilde{P}^{xx}_{k|k-1}\tilde{C}_k^{\mathrm{T}} + \tilde{R}_k)^{-1}$, it follows from (7.21) that

$$\tilde{K}_k = \left[\begin{array}{cc} K_k - K_k^{\mathrm{p}}D_{k-1}K_k & K_k^{\mathrm{p}} \end{array}\right]. \qquad (7.23)$$

Therefore (7.17) and (7.22) imply that $\tilde{x}_k = \hat{x}_k^{\mathrm{p}}$ and (2.4) and (4.2) imply that $\tilde{x}_{k+1|k} = \hat{x}_{k+1|k}$.

Note that (2.11) and (4.9) can be expressed as

$$P_k^{xx} = (I_{n\times n} - K_kC_k)P^{xx}_{k|k-1}(I_{n\times n} - K_kC_k)^{\mathrm{T}} + K_kR_kK_k^{\mathrm{T}}, \qquad (7.24)$$

$$P_k^{xx\mathrm{p}} = (I_{n\times n} - K_k^{\mathrm{p}}D_{k-1})P_k^{xx}(I_{n\times n} - K_k^{\mathrm{p}}D_{k-1})^{\mathrm{T}}. \qquad (7.25)$$

Substituting (7.24) into (7.25) yields

$$\begin{aligned} P_k^{xx\mathrm{p}} = {} & (I_{n\times n} - K_k^{\mathrm{p}}D_{k-1})(I_{n\times n} - K_kC_k)P^{xx}_{k|k-1} \\ & (I_{n\times n} - K_kC_k)^{\mathrm{T}}(I_{n\times n} - K_k^{\mathrm{p}}D_{k-1})^{\mathrm{T}} + \\ & (K_k - K_k^{\mathrm{p}}D_{k-1}K_k)R_k(K_k - K_k^{\mathrm{p}}D_{k-1}K_k)^{\mathrm{T}}. \end{aligned} \qquad (7.26)$$

Substituting (7.2) and (7.23) into (7.26) yields

$$P_k^{xx\mathrm{p}} = (I_{n\times n} - \tilde{K}\tilde{C}_k)P^{xx}_{k|k-1}(I_{n\times n} - \tilde{K}\tilde{C}_k)^{\mathrm{T}} + \tilde{K}\tilde{R}_k\tilde{K}^{\mathrm{T}}. \qquad (7.27)$$

Since, (2.11) implies that

$$\tilde{P}_k^{xx} = (I_{n\times n} - \tilde{K}\tilde{C}_k)\tilde{P}^{xx}_{k|k-1}(I_{n\times n} - \tilde{K}\tilde{C}_k)^{\mathrm{T}} + \tilde{K}\tilde{R}_k\tilde{K}^{\mathrm{T}}, \qquad (7.28)$$

it follows from (7.27) and (7.28) that $\tilde{P}_k^{xx} = P_k^{xx\mathrm{p}}$. Hence, (2.4) and (4.3) imply that $\tilde{P}^{xx}_{k+1|k} = P^{xx}_{k+1|k}$. □

## Appendix II: Connection of ECKF and PKF-SP

Assume that system (2.1)-(2.2) is time-invariant and that (3.4)-(3.6) hold. Then, consider PKF-SP which uses KF equations (2.4)-(2.11), but initialized with

$$\hat{x}_0^{\mathrm{p}} = (D^{\mathrm{T}}D)^{-1}D^{\mathrm{T}}d, \qquad (7.29)$$

$$P_0^{xx\mathrm{p}} = \mathcal{P}_{\mathcal{N}(D)}P_0^{xx}, \qquad (7.30)$$

where $P_0^{xx\mathrm{p}}$ is singular and the projector $\mathcal{P}_{\mathcal{N}(D)} \in \mathbb{R}^{n\times n}$ is obtained by the singular value decomposition

$$D^{\mathrm{T}} = \left[\begin{array}{cc} U_1 & U_2 \end{array}\right]\left[\begin{array}{c} S_s \\ 0_{(n-s)\times s} \end{array}\right]\left[\begin{array}{c} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \end{array}\right],$$

where $U_2 \in \mathbb{R}^{n\times(n-s)}$ such that $\mathcal{P}_{\mathcal{N}(D)} = U_2U_2^{\mathrm{T}}$. Also, note that, since (3.4) holds, $Gw_{k-1}$ is constrained in $\mathcal{P}_{\mathcal{N}(D)}$ and $GQ_{k-1}G^{\mathrm{T}}$ is a "constrained" covariance [8].

With Corollary 4.1 and comparing (7.29)-(7.30) to (4.19)-(4.20), we see that, similar to ECKF, which performs projection only at $k = 1$ to guarantee constraint satisfaction for all $k \geq 1$, PKF-SP performs projection in initialization, that is, only at $k = 0$, providing that (3.4)-(3.6) hold.

## Appendix III: Connection of ECKF and PKF-EP

PKF-EP projects the updated estimate $\hat{x}_k$ (2.10) onto the hyperplane defined by (3.7) by minimizing the cost function $\mathrm{J}(x_k) \triangleq (x_k - \hat{x}_k)^{\mathrm{T}} W^{-1}(x_k - \hat{x}_k)$ subject to (3.7), where $W \in \mathbb{R}^{n\times n}$ is positive definite. The solution $\hat{x}_k^{\mathrm{p}}$ to $\mathrm{J}(x_k)$ is given by

$$\hat{x}_k^{\mathrm{p}} = \hat{x}_k + K_k^{\mathrm{p}}(d_{k-1} - D_{k-1}\hat{x}_k), \qquad (7.31)$$

where

$$K_k^{\mathrm{p}} \triangleq WD_{k-1}^{\mathrm{T}}(D_{k-1}WD_{k-1}^{\mathrm{T}})^{-1}. \qquad (7.32)$$

The projected error covariance $P_k^{xx\mathrm{p}}$ associated with $\hat{x}_k^{\mathrm{p}}$ is given by (4.18) and (4.20).

PKF-EP is formed by forecast (2.4)-(2.8), data-assimilation (2.9)-(2.11), and projection (7.31)-(7.32), (4.18), (4.20) steps.

We set $W = P_k^{xx}$ in (7.32), where $P_k^{xx}$ is given by (2.11), such that $\hat{x}_k^{\mathrm{p}}$ (7.31) is optimal according to the maximum a posteriori and minimum variance criteria [12]. In this case, note that the projection equations (7.31)-(7.32), (4.18), (4.20) of PKF-EP are equal to the projection equations (4.4)-(4.7), (4.8), (4.9) of ECKF.

However, unlike ECKF, PKF-EP does not recursively feed the projected estimate $\hat{x}_k^{\mathrm{p}}$ (7.31) and the error covariance $P_k^{xx\mathrm{p}}$ given by (4.18), (4.20) back in forecast (2.4)-(2.5). Therefore, the PKF-EP forecast estimate $\hat{x}_{k|k-1}$ (2.4) is different from its ECKF counterpart (4.2).

## References


[1] L. A. Aguirre, M. F. S. Barroso, R. R. Saldanha, and E. M. A. M. Mendes. Imposing steady-state performance on identified nonlinear polynomial models by means of constrained parameter estimation. *IEE Proc. in Control Theory and Applications*, 151(2):174–179, 2004.

[2] A. T. Alouani and W. D. Blair. Use of a Kinematic Constraint in Target Tracking Constant Speed, Maneuvering Targets. *IEEE Trans. on Automatic Control*, 38:1107–1111, 1993.

[3] D. S. Bernstein. *Matrix Mathematics*. Princeton University Press, Princeton – NJ, USA, 2005.

[4] D. S. Bernstein and D. C. Hyland. Compartmental modelling and second-moment analysis of state space systems. *SIAM Journal on Matrix Analysis and Applications*, 14(3):880–901, 1993.

[5] J. L. Crassidis and F. L. Markley. Unscented Filtering for Spacecraft Attitude Estimation. *AIAA Journal of Guidance, Control, and Dynamics*, 26(4):536–542, July–August 2003.

[6] G. C. Goodwin, M. M. Seron, and J. A. de Doná. *Constrained Control and Estimation: An Optimisation Approach*. Springer, 2005.

[7] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, Inc., New York – NY, USA, 1970.

[8] S. Ko and R. R. Bitmead. State estimation for linear systems with state equality constraints. *Automatica*, 43(8):1363–1368, 2007.

[9] C. V. Rao, J. B. Rawlings, and D. Q. Mayne. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations. *IEEE Trans. on Automatic Control*, 48(2):246–258, 2003.

[10] M. Rotea and C. Lana. State Estimation with Probability Constraints. In *44th IEEE Conference on Decision and Control*, pages 380–385, Seville, Spain, 2005.

[11] D. Simon. *Optimal State Estimation: Kalman, H-Infinity and Nonlinear Approaches*. Wiley-Interscience, 2006.

[12] D. Simon and T. Chia. Kalman filtering with state equality constraints. *IEEE Trans. on Aerospace and Electronic Systems*, 38(1):128–136, 2002.

[13] P. Vachhani, S. Narasimhan, and R. Rengaswamy. Robust and reliable estimation via Unscented Recursive Nonlinear Dynamic Data Reconciliation. *Journal of Process Control*, 16:1075–1086, 2006.

[14] W. Wen and H. F. Durrant-Whyte. Model-based multi-sensor data fusion. In *Proceedings of the 1992 IEEE International Conference on Robotics and Automation*, pages 1720–1726, Nice, France, May 1992.