2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Unscented Filtering for Equality-Constrained Nonlinear Systems

Bruno O. Soares Teixeira*†, Jaganath Chandrasekar†,
Leonardo A. Borges Tôrres*, Luis A. Aguirre*, and Dennis S. Bernstein†

***Abstract*— This paper addresses the state-estimation problem for nonlinear systems in a context where prior knowledge, in addition to the model and the measurement data, is available in the form of an equality constraint. Three novel suboptimal algorithms based on the unscented Kalman filter are developed, namely, the equality-constrained unscented Kalman filter, the projected unscented Kalman filter, and the measurement-augmented unscented Kalman filter. These methods are compared on two examples: a quaternion-based attitude estimation problem and an idealized flow model involving conserved quantities.**

I. Introduction

Under Gaussian noise and linear dynamics assumptions, the equality-constrained Kalman filter (ECKF) [19] is the optimal solution to the equality-constrained state-estimation problem. ECKF takes advantage of prior knowledge of the state vector provided by an equality constraint and uses this information to obtain better estimates than would be provided by the classical Kalman filter (KF) [9] in the absence of such information.

Although it is difficult to make correspondingly precise statements in the case of nonlinear systems, the same principles and objectives apply. For example, 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 [4]. Additional examples arise in optimal control [6], parameter estimation [1, 22], and navigation [23].

However, the solution to the equality-constrained state-estimation problem for nonlinear systems is complicated by the fact that the random variables are not completely characterized by their first-order and second-order moments [5]. Thus, suboptimal solutions based on the extended KF (EKF) [9] are generally used. One of the most popular techniques is based on measurement augmentation, in which a perfect "measurement" of the constrained quantity is is appended to the physical measurements [22, 23]. In addition, the estimate-projection method [17] has also been considered. A two-step projection algorithm for handling nonlinear equality constraints is presented in [10].

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 this context, two contributions are presented in this paper. First, we present and compare three suboptimal algorithms for state estimation for equality-constrained nonlinear systems, namely, the equality-constrained unscented KF (ECUKF), the projected unscented KF (PUKF), and the measurement-augmented unscented KF (MAUKF). These methods, which extend algorithms for equality-constrained state estimation developed for linear systems [19], are based on the unscented Kalman filter (UKF) [11, 12], which is a specialized sigma-point Kalman filter (SPKF) [21]. Recent work [7, 12, 14, 21] reports the improved performance of SPKFs compared to EKF, which is prone to numerical problems such as sensitivity to initialization, divergence, and instability for strongly nonlinear systems [16]. A quaternion-based attitude estimation problem [4] is addressed to illustrate ECUKF, PUKF, MAUKF, and UKF. Although the state of the process model satisfies the unit norm constraint, this constraint is violated by the state estimates obtained from unconstrained Kalman filtering [4].

Second, we use equality-constrained Kalman filtering techniques to improve estimation when an approximate discretized model is used to represent a continuous-time process. According to [15], constraints can also be used to correct model error. The problem of using UKF with discrete-time models obtained from black-box identification to perform state estimation for continuous-time nonlinear systems is treated in [2]. We illustrate the application of equality-constrained unscented Kalman filter techniques to this problem through an example of a one-dimensional inviscid and compressible hydrodynamic model discretized by means of a finite-volume scheme [3]. The boundary conditions are chosen such that density and energy are conserved, and this knowledge is used to improve the state estimates. A detailed version of this paper appear as [18, 20].

II. State Estimation for Nonlinear Systems

For the stochastic nonlinear discrete-time dynamic system

$$x_k = f\left(x_{k-1},\ u_{k-1},\ w_{k-1},\ k-1\right), \tag{2.1}$$

$$y_k = h\left(x_k,\ k\right) + v_k, \tag{2.2}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{N} \to \mathbb{R}^n$ and $h : \mathbb{R}^n \times \mathbb{N} \to \mathbb{R}^m$ are, respectively, the process and observation models, 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 probability density functions (PDFs) $\rho(x_0)$, $\rho(w_{k-1})$ and $\rho(v_k)$, where $x_0 \in \mathbb{R}^n$ is the initial state vector, $w_{k-1} \in \mathbb{R}^q$ is the process noise, and

$v_k \in \mathbb{R}^m$ is the measurement noise. Next, define the profit function

$$J(x_k) \triangleq \rho(x_k|(y_1,\ldots,y_k)), \tag{2.3}$$

which is the conditional PDF of the state vector $x_k \in \mathbb{R}^n$ given the past and present measured data $y_1,\ldots,y_k$. Under the stated assumptions, the maximization of (2.3) is the state estimation problem, while the maximizer of J is the optimal state estimate.

However, the solution to this problem is complicated by the fact that, for nonlinear systems, $\rho(x_k|(y_1,\ldots,y_k))$ is not completely characterized by its mean $\hat{x}_{k|k}$ and covariance $P^{xx}_{k|k} \triangleq \mathrm{E}\left[(x_k-\hat{x}_{k|k})(x_k-\hat{x}_{k|k})^{\mathrm{T}}\right]$ [5]. We thus use an approximation based on the classical Kalman filter (KF) for linear systems [9] to provide a suboptimal solution to the nonlinear case, specifically, the unscented Kalman filter (UKF) [11]. To accomplish that, UKF propagates only approximations to $\hat{x}_{k|k}$ and $P^{xx}_{k|k}$ using the initial mean $\hat{x}_{0|0}$ and the covariance $P^{xx}_{0|0} \triangleq \mathrm{E}\left[(x_0-\hat{x}_0)(x_0-\hat{x}_0)^{\mathrm{T}}\right]$ of $\rho(x_0)$. We assume that the maximizer of J is $\hat{x}_{k|k}$. Furthermore, we assume that the mean and covariance of $\rho(w_{k-1})$ and $\rho(v_k)$ are known and equal to zero and Q_{k-1}, R_k, respectively. Also, w_{k-1} and v_k are assumed to be uncorrelated.

A. Unscented Kalman Filter

Instead of analytically linearizing (2.1)-(2.2) and using the KF equations [9], UKF employs the unscented transform (UT) [12], which approximates the posterior mean $\hat{y} \in \mathbb{R}^m$ and covariance $P^{yy} \in \mathbb{R}^{m\times m}$ of a random vector y obtained from the nonlinear transformation $y = h(x)$, where x is a prior random vector whose mean $\hat{x} \in \mathbb{R}^n$ and covariance $P^{xx} \in \mathbb{R}^{n\times n}$ are assumed known. UT yields the actual mean $\hat{y}$ and the actual covariance P^{yy} if $h = h_1 + h_2$, where h_1 is linear and h_2 is quadratic [12]. Otherwise, $\hat{y}_k$ is a *pseudo mean* and P^{yy} is a *pseudo covariance*.

UT is based on a set of deterministically chosen vectors known as sigma points. To capture the mean $\hat{x}^{\mathrm{a}}_{k-1|k-1}$ of the augmented state vector $x^{\mathrm{a}}_{k-1} \triangleq \begin{bmatrix} x_{k-1} \\ w_{k-1}\end{bmatrix}$, where $x^{\mathrm{a}}_{k-1} \in \mathbb{R}^{n_{\mathrm{a}}}$ and $n_{\mathrm{a}} \triangleq n+q$, as well as the augmented error covariance $P^{xx\mathrm{a}}_{k-1|k-1} \triangleq \begin{bmatrix} P^{xx}_{k-1|k-1} & 0_{n\times q} \\ 0_{q\times n} & Q_{k-1}\end{bmatrix}$, the sigma-point matrix $\mathcal{X}_{k-1|k-1} \in \mathbb{R}^{n_{\mathrm{a}}\times(2n_{\mathrm{a}}+1)}$ is chosen as

$$\begin{aligned}\mathcal{X}_{k-1|k-1} \;&=\; \hat{x}^{\mathrm{a}}_{k-1|k-1}1_{1\times(2n_{\mathrm{a}}+1)} \;+\; \sqrt{(n_{\mathrm{a}}+\lambda)} \\ &\times \left[0_{n_{\mathrm{a}}\times 1} \;\; \left(P^{xx\mathrm{a}}_{k-1|k-1}\right)^{1/2} \;\; -\left(P^{xx\mathrm{a}}_{k-1|k-1}\right)^{1/2}\right],\end{aligned} \tag{2.4}$$

with weights

$$\begin{cases} \gamma_0^{(m)} \triangleq \dfrac{\lambda}{n_{\mathrm{a}}+\lambda}, \\ \gamma_0^{(c)} \triangleq \dfrac{\lambda}{n_{\mathrm{a}}+\lambda} + 1 - \alpha^2 + \beta, \\ \gamma_i^{(m)} \triangleq \gamma_i^{(c)} \triangleq \gamma_{i+n_{\mathrm{a}}}^{(m)} \triangleq \gamma_{i+n_{\mathrm{a}}}^{(c)} \triangleq \dfrac{1}{2(n_{\mathrm{a}}+\lambda)},\; i=1,\ldots,n_{\mathrm{a}}, \end{cases} \tag{2.5}$$

where $(\cdot)^{1/2}$ is the Cholesky square root, $0<\alpha\le 1$, $\beta \ge 0$, $\kappa \ge 0$, and $\lambda \triangleq \alpha^2(\kappa+n_{\mathrm{a}}) - n_{\mathrm{a}} > -n_{\mathrm{a}}$. We set $\alpha = 1$ and $\kappa = 0$ [7] such that $\lambda = 0$ [11] and set $\beta = 2$ [7]. Alternative schemes for choosing sigma points are given in [11]. The notation $\hat{x}_{k|k-1}$ indicates an estimate of x_k at time k based on information available up to and including time $k-1$. Likewise, $\hat{x}_{k|k}$ indicates an estimate of x_k at time k using information available up to and including time k.

The UKF *forecast* equations are given by (2.4)-(2.5) together with

$$\mathcal{X}^x_{i,k|k-1} = f(\mathcal{X}^x_{i,k-1|k-1}, u_{k-1}, \mathcal{X}^w_{i,k-1|k-1}, k-1),\; i=0,\ldots,2n_{\mathrm{a}}, \tag{2.6}$$

$$\hat{x}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(m)}\mathcal{X}^x_{i,k|k-1}, \tag{2.7}$$

$$P^{xx}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(c)}[\mathcal{X}^x_{i,k|k-1}-\hat{x}_{k|k-1}][\mathcal{X}^x_{i,k|k-1}-\hat{x}_{k|k-1}]^{\mathrm{T}}, \tag{2.8}$$

$$\mathcal{Y}_{i,k|k-1} = h(\mathcal{X}^x_{i,k|k-1},\; k),\; i=0,\ldots,2n_{\mathrm{a}}, \tag{2.9}$$

$$\hat{y}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(m)}\mathcal{Y}_{i,k|k-1}, \tag{2.10}$$

$$P^{yy}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(c)}[\mathcal{Y}_{i,k|k-1}-\hat{y}_{k|k-1}][\mathcal{Y}_{i,k|k-1}-\hat{y}_{k|k-1}]^{\mathrm{T}} + R_k, \tag{2.11}$$

$$P^{xy}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(c)}[\mathcal{X}^x_{i,k|k-1}-\hat{x}_{k|k-1}][\mathcal{Y}_{i,k|k-1}-\hat{y}_{k|k-1}]^{\mathrm{T}}, \tag{2.12}$$

where $\mathcal{X}_i$ is the ith column of $\mathcal{X}$, $\begin{bmatrix}\mathcal{X}^x_{k-1|k-1} \\ \mathcal{X}^w_{k-1|k-1}\end{bmatrix} \triangleq \mathcal{X}_{k-1|k-1}$, $\mathcal{X}^x_{k-1|k-1} \in \mathbb{R}^{n\times(2n_{\mathrm{a}}+1)}$, $\mathcal{X}^w_{k-1|k-1} \in \mathbb{R}^{q\times(2n_{\mathrm{a}}+1)}$, $P^{xx}_{k|k-1}$ is the forecast error covariance, $P^{yy}_{k|k-1}$ is the innovation covariance, $P^{xy}_{k|k-1}$ is the cross covariance, and $P^{xx}_{k|k}$ is the data-assimilation error-covariance.

The *data-assimilation* equations are given by

$$K_k = P^{xy}_{k|k-1}(P^{yy}_{k|k-1})^{-1}, \tag{2.13}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}), \tag{2.14}$$

$$P^{xx}_{k|k} = P^{xx}_{k|k-1} - K_k P^{yy}_{k|k-1} K_k^{\mathrm{T}}, \tag{2.15}$$

where $K_k \in \mathbb{R}^{n\times m}$ is the Kalman gain matrix. 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 Nonlinear Systems

Assume that, for all $k \ge 1$, the state vector x_k satisfies the equality constraint

$$g(x_k,\; k-1) \;=\; d_{k-1}, \tag{3.1}$$

where $g : \mathbb{R}^n \times \mathbb{N} \to \mathbb{R}^s$ and $d_{k-1} \in \mathbb{R}^s$ are assumed known. Then, the objective of the equality-constrained state-estimation problem is to maximize (2.3) subject to (3.1). That is, we look for the maximizer of J that satisfies (3.1).

In addition to nonlinear dynamics, the solution to this problem is complicated due the inclusion of (3.1). We thus extend algorithms derived in the linear scenario to provide a suboptimal solution to the nonlinear case.

IV. Kalman Filters for Equality-Constrained Linear Systems

In this section, we briefly review three state-estimation algorithms for equality-constrained linear systems. For details the reader is referred to [19].

Assume that f, h, and g are linear functions, respectively, given by

$$x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + G_{k-1}w_{k-1}, \quad (4.1)$$

$$y_k = C_k x_k + v_k, \quad (4.2)$$

$$D_{k-1}x_k = d_{k-1}, \quad (4.3)$$

where $A_{k-1} \in \mathbb{R}^{n\times n}$, $B_{k-1} \in \mathbb{R}^{n\times p}$, $G_{k-1} \in \mathbb{R}^{n\times q}$, $C_k \in \mathbb{R}^{m\times n}$, and $D_{k-1} \in \mathbb{R}^{s\times n}$. Then, the recursive solution $\hat{x}^{\mathrm{p}}_{k|k}$ to the equality-constrained state-estimation problem is given by the equality-constrained Kalman filter (ECKF) [19], whose forecast step is given by

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}^{\mathrm{p}}_{k-1|k-1} + B_{k-1}u_{k-1}, \quad (4.4)$$

$$P^{xx}_{k|k-1} = A_{k-1}P^{xx\mathrm{p}}_{k-1|k-1}A^{\mathrm{T}}_{k-1} + G_{k-1}Q_{k-1}G^{\mathrm{T}}_{k-1}, \quad (4.5)$$

$$\hat{y}_{k|k-1} = C_k\hat{x}_{k|k-1}, \quad (4.6)$$

$$P^{yy}_{k|k-1} = C_kP^{xx}_{k|k-1}C^{\mathrm{T}}_k + R_k, \quad (4.7)$$

$$P^{xy}_{k|k-1} = P^{xx}_{k|k-1}C^{\mathrm{T}}_k, \quad (4.8)$$

where $P^{xx\mathrm{p}}_{k|k} \triangleq \mathrm{E}\left[(x_k - \hat{x}^{\mathrm{p}}_{k|k})(x_k - \hat{x}^{\mathrm{p}}_{k|k})^{\mathrm{T}}\right]$, whose data-assimilation step is given by (2.13)-(2.15), and whose *projection* step is given by

$$\hat{d}_{k-1|k} = D_{k-1}\hat{x}_{k|k}, \quad (4.9)$$

$$P^{dd}_{k|k} = D_{k-1}P^{xx}_{k|k}D^{\mathrm{T}}_{k-1}, \quad (4.10)$$

$$P^{xd}_{k|k} = P^{xx}_{k|k}D^{\mathrm{T}}_{k-1}, \quad (4.11)$$

$$K^{\mathrm{p}}_k = P^{xd}_{k|k}(P^{dd}_{k|k})^{-1}, \quad (4.12)$$

$$\hat{x}^{\mathrm{p}}_{k|k} = \hat{x}_{k|k} + K^{\mathrm{p}}_k(d_{k-1} - \hat{d}_{k-1|k}), \quad (4.13)$$

$$P^{xx\mathrm{p}}_{k|k} = P^{xx}_{k|k} - K^{\mathrm{p}}_kP^{dd}_{k|k}K^{\mathrm{p}\,\mathrm{T}}_k. \quad (4.14)$$

Note that the projection step is absent in the KF algorithm.

For all $k \geq 1$, the projected Kalman filter by estimate projection (PKF-EP) [17] projects $\hat{x}_{k|k}$ given by (2.14) onto the hyperplane (4.3) yielding the projected estimate $\hat{x}^{\mathrm{p}}_{k|k}$. However, unlike ECKF, $\hat{x}^{\mathrm{p}}_{k|k}$ of PKF-EP is not recursively fed back in the next iteration. That is, PKF-EP equations are equal to ECKF equations except for (4.4)-(4.5), which are replaced by

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1}, \quad (4.15)$$

$$P^{xx}_{k|k-1} = A_{k-1}P^{xx}_{k-1|k-1}A^{\mathrm{T}}_{k-1} + G_{k-1}Q_{k-1}G^{\mathrm{T}}_{k-1}. \quad (4.16)$$

Finally, the measurement–augmentation Kalman filter (MAKF) [23] treats (4.3) as perfect measurements. That is, we replace (4.2) by the augmented observation

$$\tilde{y}_k \triangleq \tilde{C}_kx_k + \tilde{v}_k, \quad (4.17)$$

where $\tilde{y}_k \triangleq \begin{bmatrix} y_k \\ d_{k-1} \end{bmatrix}$, $\tilde{C}_k \triangleq \begin{bmatrix} C_k \\ D_{k-1} \end{bmatrix}$, and $\tilde{v}_k \triangleq \begin{bmatrix} v_k \\ 0_{s\times 1} \end{bmatrix}$. With (4.17), MAKF uses (4.15)-(4.16) together with the augmented forecast equations

$$\hat{\tilde{y}}_{k|k-1} = \tilde{C}_k\hat{x}_{k|k-1}, \quad (4.18)$$

$$\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1} = \tilde{C}_kP^{xx}_{k|k-1}\tilde{C}^{\mathrm{T}}_k + \tilde{R}_k, \quad (4.19)$$

$$\tilde{P}^{x\tilde{y}}_{k|k-1} = P^{xx}_{k|k-1}\tilde{C}^{\mathrm{T}}_k, \quad (4.20)$$

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

$$\tilde{K}_k = \tilde{P}^{x\tilde{y}}_{k|k-1}(\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1})^{-1}, \quad (4.21)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \tilde{K}_k(\tilde{y}_k - \hat{\tilde{y}}_{k|k-1}), \quad (4.22)$$

$$P^{xx}_{k|k} = P^{xx}_{k|k-1} - \tilde{K}_k\tilde{P}^{x\tilde{y}}_{k|k-1}\tilde{K}^{\mathrm{T}}_k. \quad (4.23)$$

For the linear case, MAKF and ECKF estimates are equal [19, 20].

V. Equality-Constrained UKFs

In this section, by using UT, we extend PKF-EP, ECKF, and MAKF to the nonlinear case. Unlike the linear case (Section IV), these approaches do not guarantee that the nonlinear equality constraint (3.1) is exactly satisfied by the state estimates, but they provide approximate solutions.

A. Equality-Constrained Unscented Kalman Filter

The projection step of ECKF given by (4.9)-(4.14) is now extended to the nonlinear case by means of UT.

Using (2.4)-(2.5) to choose sigma points and associated weights, we have

$$\mathcal{X}^x_{k|k} = \hat{x}_{k|k}1_{1\times(2n+1)} + \sqrt{(n+\lambda)}\left[0_{n\times 1}\ \left(P^{xx}_{k|k}\right)^{1/2}\ -\left(P^{xx}_{k|k}\right)^{1/2}\right], \quad (5.1)$$

where $\hat{x}_{k|k}$ and $P^{xx}_{k|k}$ are given by (2.14) and (2.15), and with weights given by (2.5), replacing n_{a} by n. Then the sigma points $\mathcal{X}^x_{i,k|k} \in \mathbb{R}^n$, $i = 0, \ldots, 2n$, are propagated through (3.1) yielding

$$\mathcal{D}_{i,k|k} = g(\mathcal{X}^x_{i,k|k}, k-1), \quad i = 0, \ldots, 2n, \quad (5.2)$$

such that $\hat{d}_{k-1|k}$, $P^{dd}_{k|k}$, and $P^{xd}_{k|k}$ are given by

$$\hat{d}_{k-1|k} = \sum_{i=0}^{2n}\gamma^{(m)}_i\mathcal{D}_{i,k|k}, \quad (5.3)$$

$$P^{dd}_{k|k} = \sum_{i=0}^{2n}\gamma^{(c)}_i[\mathcal{D}_{i,k|k} - \hat{d}_{k-1|k}][\mathcal{D}_{i,k|k} - \hat{d}_{k-1|k}]^{\mathrm{T}}, \quad (5.4)$$

$$P^{xd}_{k|k} = \sum_{i=0}^{2n}\gamma^{(c)}_i[\mathcal{X}^{x|k}_{i,k} - \hat{x}_{k|k}][\mathcal{D}_{i,k|k} - \hat{d}_{k-1|k}]^{\mathrm{T}}, \quad (5.5)$$

and K^{p}_k, $\hat{x}^{\mathrm{p}}_{k|k}$, and $P^{xx\mathrm{p}}_{k|k}$ are respectively given by (4.12), (4.13), and (4.14).

Now define $x^{\mathrm{ap}}_{k-1} \triangleq \begin{bmatrix} x^{\mathrm{p}}_{k-1} \\ w_{k-1} \end{bmatrix}$, and $P^{xx\mathrm{ap}}_{k-1|k-1} \triangleq \begin{bmatrix} P^{xx\mathrm{p}}_{k-1|k-1} & 0_{n\times q} \\ 0_{q\times n} & Q_{k-1} \end{bmatrix}$, such that the sigma points

$$\mathcal{X}_{k-1|k-1} = \hat{x}^{\mathrm{ap}}_{k-1|k-1}1_{1\times(2n_{\mathrm{a}}+1)} + \sqrt{(n_{\mathrm{a}}+\lambda)} \times\left[0_{n_{\mathrm{a}}\times 1}\ \left(P^{xx\mathrm{ap}}_{k-1|k-1}\right)^{1/2}\ -\left(P^{xx\mathrm{ap}}_{k-1|k-1}\right)^{1/2}\right] \quad (5.6)$$

are chosen based on $\hat{x}^{\mathrm{p}}_{k-1|k-1}$ (4.13). Then, by appending the projection step given by (5.1)-(5.5), (4.12)-(4.14) to (5.6), (2.5)-(2.15), we obtain the *equality-constrained unscented Kalman filter* (ECUKF).

B. Projected Unscented Kalman Filter

If, unlike ECUKF, we append the projection step given by (5.1)-(5.5),(4.12)-(4.14) to UKF equations given by (2.4)-(2.15) without feedback recursion, we obtain the *projected unscented Kalman filter* (PUKF). PUKF is the nonlinear extension of PKF-EP.

C. Measurement-Augmentation Unscented Kalman Filter

To extend the MAKF to the nonlinear case, we replace (2.2) by the augmented observation

$$\tilde{y}_k \triangleq \tilde{h}(x_k, k) + \tilde{v}_k, \tag{5.7}$$

where $\tilde{h}(x_k, k) \triangleq \begin{bmatrix} h(x_k, k) \\ g(x_k, k-1) \end{bmatrix}$. With (5.7), the *measurement-augmented unscented Kalman filter* (MAUKF) combines (2.4)-(2.8) with the augmented forecast equations

$$\tilde{\mathcal{Y}}_{i,k|k-1} = \tilde{h}(\mathcal{X}^x_{i,k|k-1}, k), \quad i = 0, \ldots, 2n_{\mathrm{a}}, \tag{5.8}$$

$$\hat{\tilde{y}}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(m)} \tilde{\mathcal{Y}}_{i,k|k-1}, \tag{5.9}$$

$$\tilde{P}^{\tilde{y}\tilde{y}}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(c)} [\tilde{\mathcal{Y}}_{i,k|k-1} - \hat{\tilde{y}}_{k|k-1}][\tilde{\mathcal{Y}}_{i,k|k-1} - \hat{\tilde{y}}_{k|k-1}]^{\mathrm{T}} + \tilde{R}_k, \tag{5.10}$$

$$\tilde{P}^{x\tilde{y}}_{k|k-1} = \sum_{i=0}^{2n_{\mathrm{a}}} \gamma_i^{(c)} [\mathcal{X}^x_{i,k|k-1} - \hat{x}_{k|k-1}][\tilde{\mathcal{Y}}_{i,k|k-1} - \hat{\tilde{y}}_{k|k-1}]^{\mathrm{T}}, \tag{5.11}$$

and the data-assimilation equations given by (4.21)-(4.23).

VI. Numerical Examples

A. Attitude Estimation

Consider an attitude estimation problem [4], whose kinematics is modeled as

$$\dot{e}(t) = -\frac{1}{2}\Omega(t)e(t), \tag{6.1}$$

where the state vector is the quaternion vector $e(t) \triangleq [e_0(t)\ e_1(t)\ e_2(t)\ e_3(t)]^{\mathrm{T}}$, the matrix $\Omega(t)$ is given by

$$\Omega(t) = \begin{bmatrix} 0 & r(t) & -q(t) & p(t) \\ -r(t) & 0 & p(t) & q(t) \\ q(t) & -p(t) & 0 & r(t) \\ -p(t) & -q(t) & -r(t) & 0 \end{bmatrix}, \tag{6.2}$$

and the angular velocity vector $u(t) = [p(t)\ q(t)\ r(t)]^{\mathrm{T}}$ is a known input. Since $||e(0)||_2 = 1$ and $\Omega(t)$ is a skew-symmetric matrix, it follows that, for all $t > 0$,

$$||e(t)||_2 = 1. \tag{6.3}$$

We set $e(0) = [0.9603\ \ 0.1387\ \ 0.1981\ \ 0.1387]^{\mathrm{T}}$, and
$u(t) = \left[\, 0.03\sin\left(\frac{2\pi}{600}t\right)\ \ 0.03\sin\left(\frac{2\pi}{600}t - 300\right)\ \ 0.03\sin\left(\frac{2\pi}{600}t - 600\right) \right]^{\mathrm{T}}$.

To perform attitude estimation, we assume that

$$\breve{u}_{k-1} = u((k-1)T) + \beta_{k-1} + w^u_{k-1} \tag{6.4}$$

is measured by rate gyros, where T is the discretization step, $w^u_{k-1} \in \mathbb{R}^3$ is zero mean, Gaussian noise, and $\beta_{k-1} \in \mathbb{R}^3$ is drift error. The discrete-time equivalent of (6.1) augmented by the gyro drift random-walk model [4] is given by

$$\begin{bmatrix} e_k \\ \beta_k \end{bmatrix} = \begin{bmatrix} A_{k-1} & 0_{4\times 3} \\ 0_{3\times 4} & I_{3\times 3} \end{bmatrix} \begin{bmatrix} e_{k-1} \\ \beta_{k-1} \end{bmatrix} + \begin{bmatrix} 0_{4\times 1} \\ w^\beta_{k-1} \end{bmatrix}, \tag{6.5}$$

where $e_k \triangleq e(kT)$, $w^\beta_{k-1} \in \mathbb{R}^3$ is process noise associated to drift-error model, $x_k \triangleq \begin{bmatrix} e_k \\ \beta_k \end{bmatrix} \in \mathbb{R}^7$ is the state vector, $w_{k-1} \triangleq \begin{bmatrix} w^u_{k-1} \\ w^\beta_{k-1} \end{bmatrix} \in \mathbb{R}^6$ is the process noise, and

$$A_{k-1} \triangleq \cos(s_{k-1})I_{4\times 4} - \frac{1}{2}\frac{T\sin(s_{k-1})}{s_{k-1}}\Omega_{k-1}, \tag{6.6}$$

$$s_{k-1} \triangleq \frac{T}{2}\,||\breve{u}_{k-1} - \beta_{k-1} - w^u_{k-1}||_2, \tag{6.7}$$

$$\Omega_{k-1} \triangleq \Omega((k-1)T). \tag{6.8}$$

The constraint (6.3) also holds for $t = kT$, that is,

$$x^2_{1,k} + x^2_{2,k} + x^2_{3,k} + x^2_{4,k} = 1. \tag{6.9}$$

We set $T = 0.1$ s, $\beta_{k-1} = [\, 0.001\ \ -0.001\ \ 0.0005 \,]^{\mathrm{T}}$ rad/s, and $Q_{k-1} = \mathrm{diag}\left(10^{-5}I_{3\times 3},\ 10^{-10}I_{3\times 3}\right)$. Attitude observations $y^{[i]}_k \in \mathbb{R}^3$ for a direction sensor are given by [4]

$$y^{[i]}_k = C_k r^{[i]} + v^{[i]}_k, \tag{6.10}$$

where $r^{[i]} \in \mathbb{R}^3$ is a reference direction vector to a known point, and C_k is the rotation matrix from the reference to the body-fixed frame

$$C_k = \begin{bmatrix} (x^2_{1,k} - x^2_{2,k} - x^2_{3,k} + x^2_{4,k}) & 2(x_{1,k}x_{2,k} - x_{3,k}x_{4,k}) & 2(x_{1,k}x_{3,k} - x_{2,k}x_{4,k}) \\ 2(x_{1,k}x_{2,k} - x_{3,k}x_{4,k}) & (-x^2_{1,k} + x^2_{2,k} - x^2_{3,k} + x^2_{4,k}) & 2(x_{1,k}x_{4,k} + x_{2,k}x_{3,k}) \\ 2(x_{1,k}x_{3,k} + x_{2,k}x_{4,k}) & 2(-x_{1,k}x_{4,k} + x_{2,k}x_{3,k}) & (-x^2_{1,k} - x^2_{2,k} + x^2_{3,k} + x^2_{4,k}) \end{bmatrix}. \tag{6.11}$$

We assume that two directions are available [4, 13], which can be obtained using either a star tracker or a combined three-axis accelerometer/three-axis magnetometer. We set $r^{[1]} = [\, 1\ \ 0\ \ 0 \,]^{\mathrm{T}}$, $r^{[2]} = [\, 0\ \ 1\ \ 0 \,]^{\mathrm{T}}$, and $R_k = 10^{-4}I_{6\times 6}$. These direction measurements are assumed to be provided at a lower rate, specifically, at 1 Hz, which corresponds to a sample interval of $10T$ s.

We implement Kalman filtering using UKF, ECUKF, PUKF, and MAUKF with (6.5), (6.10), and constraint (6.9). We initialize these algorithms with $\hat{x}_0 = [\, 1\ 0\ 0\ 0\ 0\ 0\ 0 \,]^{\mathrm{T}}$ and $P^{xx}_0 = \mathrm{diag}\left(0.5I_{4\times 4}\ \ 0.01I_{3\times 3}\right)$. Table I shows the results obtained from a 100-run Monte Carlo simulation. Note that, though the state vector of the model (6.5) satisfes (6.9), UKF estimates do not satisfy (6.9). Nevertheless, with the usage of prior knowledge, more informative (smaller trace of error covariance) estimates are produced compared to the unconstrained estimates given by UKF. However, a slight increase in the root-mean-square (RMS) error is

TABLE I: Average of percent RMS constraint error, trace of error covariance matrix, and RMS estimation error for 100-run Monte Carlo simulation for attitude estimation using UKF, ECUKF, PUKF, and MAUKF.

UKF	ECUKF	PUKF	MAUKF
Percent RMS constraint error ($\times 10^{-4}$)			
382.1	6.23	8.34	6.26
Trace of error covariance matrix for attitude ($\times 10^{-5}$)			
8.21	5.42	8.17	5.43
Trace of error covariance matrix for drift ($\times 10^{-8}$)			
9.214	9.204	9.213	9.209
RMS estimation error for e_0, e_1, e_2, e_3, β_1, β_2, β_3 ($\times 10^{-3}$)			
1.388	8.523	1.384	8.517
1.387	3.069	1.385	3.067
1.367	5.548	1.365	5.544
1.289	2.283	1.284	2.282
0.126	0.375	0.126	0.375
0.127	0.380	0.127	0.380
0.125	0.302	0.126	0.302

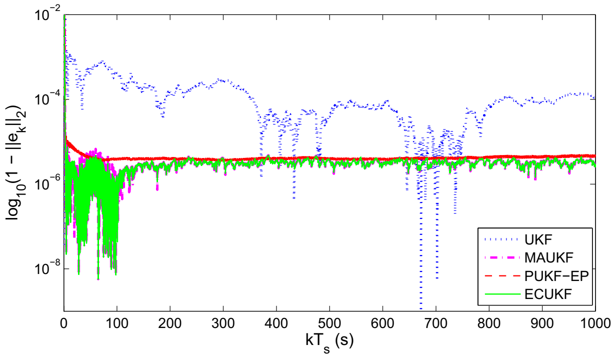

Fig. 1. Estimation errors of the quaternion vector norm using UKF, ECUKF, PUKF-EP, and MAUKF algorithms. ECUKF and MAUKF estimates almost coincide.

observed for algorithms ECUKF and MAUKF implying loss of accuracy around 5% compared to UKF. The equality constraint is more closely tracked whenever a constrained filter is employed; see Figure 1. Percent errors around 10^{-4} are obtained and ECUKF exhibits the smallest error.

B. One-Dimensional Hydrodynamics

We consider state estimation for one-dimensional hydrodynamic flow. The flow of an inviscid, compressible fluid along a one-dimensional channel is governed by Euler's equations [8]

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial z}\rho v, \tag{6.12}$$

$$\frac{d}{dt}\left(\frac{p}{\rho^{\gamma}}\right) = 0, \tag{6.13}$$

$$\rho\frac{\partial v}{\partial t} = -\rho v\frac{\partial v}{\partial z} - \frac{\partial p}{\partial z}, \tag{6.14}$$

where $\rho \in \mathbb{R}$ is the density, $v \in \mathbb{R}$ is the velocity, $p \in \mathbb{R}$ is the pressure of the fluid, $z \in \mathbb{R}$ is the spatial coordinate, and $\gamma = \frac{5}{3}$ is the ratio of specific heats of the fluid. A discrete-time model of hydrodynamic flow is obtained by using finite-volume-based spatial and temporal discretization [3].

Assume that the channel consists of l identical cells (see Figure 2). For $i = 1, \ldots, l$, let $\rho^{[i]}$, $v^{[i]}$, and $p^{[i]}$ be the density, velocity, and pressure in the ith cell, and define $U^{[i]} \triangleq \begin{bmatrix} \rho^{[i]} & m^{[i]} & \mathcal{E}^{[i]} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^3$, where $m^{[i]} = \rho^{[i]}v^{[i]}$ is the momentum and $\mathcal{E}^{[i]} = \frac{1}{2}\rho^{[i]}(v^{[i]})^2 + \frac{p^{[i]}}{\gamma-1}$ is the energy in the ith cell. We use a second-order Rusanov scheme [8] to discretize (6.12)-(6.14) and obtain a discrete-time model that updates the flow variables at the center of each cell.

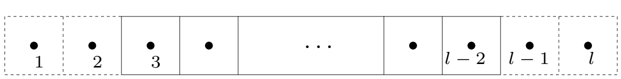

Fig. 2. One-dimensional grid used in the finite volume scheme.

The discrete-time state update equation [8] is given by

$$U_k^{[i]} = U_{k-1}^{[i]} - \frac{T}{\Delta z}\left[\overline{F}_{\mathrm{Rus},k-1}^{[i]} - \overline{F}_{\mathrm{Rus},k-1}^{[i-1]}\right], \tag{6.15}$$

where $T > 0$ is the sampling time, Δz is the width of each cell, and $\overline{F}_{\mathrm{Rus},k-1}^{[i]}$ is a nonlinear function of $U_{k-1}^{[i-1]}, \ldots, U_{k-1}^{[i+2]}$ whose equation is in [8]. Hence, $U_k^{[i]}$ depends on $U_{k-1}^{[i-2]}, \ldots, U_{k-1}^{[i+2]}$, as expected for a spatially second-order scheme.

Next, define the state vector $x_k \triangleq \begin{bmatrix} (U_k^{[3]})^{\mathrm{T}} & \cdots & (U_k^{[l-2]})^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{3(l-4)}$. Furthermore, we assume reflective boundary conditions at cells 1, 2, $l-1$, and l so that these mimic walls reflecting incident waves. Let $l = 54$ so that $n = 3(l-4) = 150$. It follows from (6.15) that the second-order Rusanov scheme yields a nonlinear discrete-time update model of the form $x_k = f(x_{k-1})$, where $f : \mathbb{R}^n \times \mathbb{R}^n$. To account for disturbances, we assume that the truth model is given by $x_k = f(x_{k-1}) + Gw_{k-1}$, which is a special case of (2.1), where $u_{k-1} = 0_{p\times 1}$, $w_{k-1} \in \mathbb{R}$, $Q_{k-1} = I_{6\times 6}$, and $G \in \mathbb{R}^{150\times 1}$, where

$$\mathrm{row}_j(G) = \begin{cases} 0.7, & \text{if } j = 7, 22, 37, 67, 97, 127, \\ 0, & \text{otherwise.} \end{cases} \tag{6.16}$$

It follows from (2.1) and (6.16) that although the flow variables in cells 5, 10, 15, 25, 35, and 45 are directly affected by w_{k-1}, the total density and the energy of cells $3, \ldots, l-2$, is not altered by w_{k-1}, that is,

$$\sum_{i=3}^{l-2} \rho_k^{[i]} = \sum_{i=3}^{l-2} \rho_0^{[i]}, \quad \sum_{i=3}^{l-2} \mathcal{E}_k^{[i]} = \sum_{i=3}^{l-2} \mathcal{E}_0^{[i]}. \tag{6.17}$$

Consequently, the linear equality constraint in (6.17) can be expressed as (4.3) with

$$D = \begin{bmatrix} 1 & 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 1 \end{bmatrix}, \quad d = \begin{bmatrix} \sum_{i=3}^{l-2} \rho_0^{[i]} \\ \sum_{i=3}^{l-2} \mathcal{E}_0^{[i]} \end{bmatrix}. \tag{6.18}$$

Next, for $i = 3, \ldots, l-2$, define $C^{[i]} \triangleq \begin{bmatrix} 0_{3\times 3(l-4-i)} & I_{3\times 3} & 0_{3\times 3(i-1)} \end{bmatrix} \in \mathbb{R}^{3\times 3(l-4)}$, so that the linear measurement model given by (4.2), for $y_k \in \mathbb{R}^m$, $m = 6$, corresponding to density, momentum and energy at cells 24 and 26, has $C = \begin{bmatrix} (C^{[24]})^{\mathrm{T}} & (C^{[26]})^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ and $R = 0.01I_{6\times 6}$. That is, we have a nonlinear process model with additive noise, a linear observation model, and a linear equality constraint.

Next, we compare the performance of UKF, ECUKF, PUKF, and MAUKF. Note that since the constraint in (6.18)

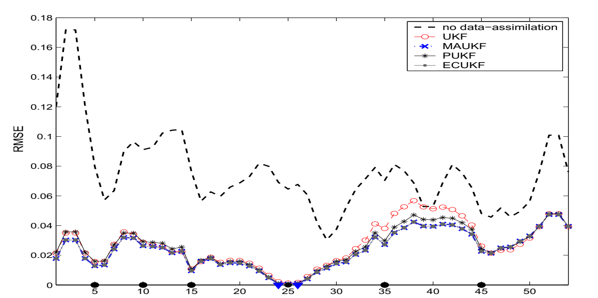

Fig. 3. RMS energy-estimate error at each cell i. The error when no data assimilation is performed is also shown for comparison. The performance of ECUKF, PUKF, and MAUKF is better than the performance of UKF. The performance of MAKF and ECUKF is very similar because the constraints are linear with respect to the state. The cells where disturbance enters the system is indicated by '•' and the cells where measurements are available is indicated by '▼'.

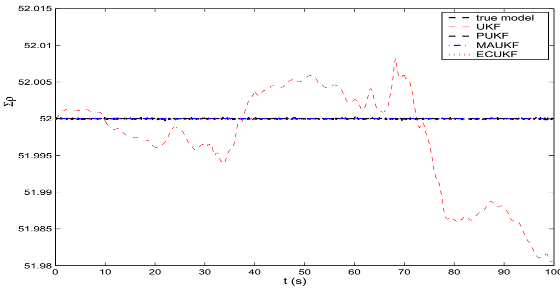

Fig. 4. Total density $\Sigma\rho$ obtained using the estimates from UKF, ECUKF, PUKF, and MAUKF. ECUKF, PUKF, and MAUKF provide estimates that ensure that the estimates satisfy the constraint. However, since no constraints are enforced in UKF, the total density or the total energy is not conserved.

and the observation model are linear, ECUKF and MAUKF estimates are equal [19]. The RMS estimation error at cells $1, \ldots, 54$ is shown in Figure 3. The error in the energy estimates from data-free simulation is shown for comparison. Note that the performance of ECUKF, PUKF, and MAUKF is better than the performance of UKF because of the enforcement of the constant total density and energy constraint. Figure 4 shows the total density estimated by UKF, ECUKF, PUKF, and MAUKF. The actual total density of the truth model is also plotted for comparison. The total density of the estimates of ECUKF, PUKF, and MAUKF is very close to the truth model, but UKF does not conserve total density or the total energy.

VII. Concluding Remarks

We have addressed the equality-constrained state-estimation problem for nonlinear systems. Three novel nonlinear extensions of equality-constrained linear state-estimation algorithms based on the unscented Kalman filter (UKF) were presented, namely, the equality-constrained UKF (ECUKF), the projected UKF (PUKF), and the measurement-augmented UKF (MAUKF). These methods were compared on two examples, including a quaternion-based attitude estimation problem, as well as an idealized flow model involving conserved quantities.

Numerical results suggest that, in addition to exactly (for linear constraints) or very closely (nonlinear constraints) satisfying a known equality constraint of the system, the proposed methods can be used to produce more accurate and more informative estimates. Considering the examples investigated, we recommend the user to test ECUKF, MAUKF, and PUKF in this order for a given equality-constrained state-estimation application. Recall that, since these methods are approximate, their performance depends on the application. Also, all equality-constrained approaches have required similar processing time, which was competitive to UKF processing time.

Finally, we have also addressed the case where an approximate discretized model is used to represent a continuous-time process in state estimation. Improved estimates were obtained when equality-constrained Kalman filtering algorithms were employed to enforce conserved quantities of the original continuous-time model, but without the higher computational burden required by more accurate integration schemes.

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] L. A. Aguirre, B. O. S. Teixeira, and L. A. B. Tôrres. Using data-driven discrete-time models and the unscented Kalman filter to estimate unobserved variables of nonlinear systems. *Physical Review E*, 72(2):026226, 2005.

[3] J. Chandrasekar, O. Barrero, A. J. Ridley, D. S. Bernstein, and D. De Moor. State estimation for linearized MHD flow. In *Proceedings of the IEEE Conference on Decision and Control*, pages 2584–2589, Paradise Island, The Bahamas, 2004.

[4] 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.

[5] F. E. Daum. Nonlinear filters: Beyond the Kalman filter. *IEEE Aerospace and Electronics Systems Magazine*, 20(8):57–69, 2005.

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

[7] S. Haykin, editor. *Kalman Filtering and Neural Networks*. Wiley Publishing, New York – NY, USA, 2001.

[8] C. Hirsch. *Numerical Computation of Internal and External Flows*. John Wiley & Sons, 1990.

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

[10] S. J. Julier and J. J. LaViola Jr.. On Kalman filtering with nonlinear equality constraints. *IEEE Transactions on Signal Processing*, 55(6), 2774–2784, 2007.

[11] S. J. Julier and J. K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92:401–422, 2004.

[12] S. J. Julier, J. K. Uhlmann, and Hugh F. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482, 2000.

[13] T. Lee, M. Leok, N. H. McClamroch, and A. Sanyal. Global attitude estimation using single direction measurements. In *Proceedings of the American Control Conference*, New York City – NY, USA, pp. 3659–3664, 2007.

[14] T. Lefebvre, H. Bruyninckx, and J. De Schutter. Kalman filters for nonlinear systems: A comparison of performance. *International Journal of Control*, 77(7):639–653, May 2004.

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

[16] K. Reif, S. Günther, E. Yaz, and R. Unbehauen. Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic Control*, 44(4):714–728, 1999.

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

[18] B. O. S. Teixeira. Constrained State Estimation for Linear and Nonlinear Dynamic Systems. *PhD Thesis*. Graduate Program in Electrical Engineering, Federal University of Minas Gerais. Belo Horizonte – MG, Brazil, 2008.

[19] B. O. S. Teixeira, J. Chandrasekar, L. A. B. Tôrres, L. A. Aguirre, and D. S. Berstein. State estimation for equality-constrained linear systems. In *Proceedings of the 46th IEEE Conference on Decision and Control*, New Orleans, LA, USA, pp. 6220–6225, 2007.

[20] B. O. S. Teixeira, J. Chandrasekar, L. A. B. Tôrres, L. A. Aguirre, and D. S. Bernstein. State estimation for linear and nonlinear equality-constrained systems. *International Journal of Control*, submitted.

[21] R. van der Merwe, E. A. Wan, and S. J. Julier. Sigma-point Kalman filters for nonlinear estimation and sensor-fusion – applications to integrated navigation. In *Proceedings of the AIAA Guidance, Navigation & Control Conference*, Providence – RI, USA, AIAA-2004-5120, 2004.

[22] D. M. Walker. Parameter estimation using Kalman filters with constraints. *Int. J. of Bifurcation and Chaos*, 4(16):1067–1078, 2006.

[23] 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.