Cubature Kalman Filtering for Continuous-Discrete Systems: Theory and Simulations

Ienkaran Arasaratnam, Simon Haykin, *Life Fellow, IEEE*, and Thomas R. Hurd

***Abstract*—In this paper, we extend the cubature Kalman filter (CKF) to deal with nonlinear state-space models of the continuous-discrete kind. To be consistent with the literature, the resulting nonlinear filter is referred to as the continuous-discrete cubature Kalman filter (CD-CKF). We use the Itô-Taylor expansion of order 1.5 to transform the process equation, modeled in the form of stochastic ordinary differential equations, into a set of stochastic difference equations. Building on this transformation and assuming that all conditional densities are Gaussian-distributed, the solution to the Bayesian filter reduces to the problem of how to compute Gaussian-weighted integrals. To numerically compute the integrals, we use the third-degree cubature rule. For a reliable implementation of the CD-CKF in a finite word-length machine, it is structurally modified to propagate the square-roots of the covariance matrices. The reliability and accuracy of the square-root version of the CD-CKF are tested in a case study that involves the use of a radar problem of practical significance; the problem considered herein is challenging in the context of radar in two respects-high dimensionality of the state and increasing degree of nonlinearity. The results, presented herein, indicate that the CD-CKF markedly outperforms existing continuous-discrete filters.**

***Index Terms*—Bayesian filters, cubature Kalman filter (CKF), Itô-Taylor expansion, nonlinear filtering, square-root filtering.**

I. Introduction

In [1], we described a new nonlinear filter named the cubature Kalman filter (CKF), for hidden state estimation based on nonlinear *discrete-time* state-space models. Like the celebrated Kalman filter for linear Gaussian models, an important virtue of the CKF is its mathematical rigor. This rigor is rooted in the *third-degree spherical-radial cubature rule* for numerically computing Gaussian-weighted integrals [3]. Although the idea of constructing cubature rules has been known for over four decades in the mathematical literature, to the best of our knowledge, there has been no reference made to it in the literature on nonlinear filtering except in [1] for the first time.

Manuscript received December 02, 2009; accepted June 21, 2010. Date of publication July 08, 2010; date of current version September 15, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. James Lam.

I. Arasaratnam is with the Center for Mechatronics and Hybrid Technologies, McMaster University, Hamilton, ON L8S 4K1 Canada (e-mail: haran@ieee.org).

S. Haykin is with the Cognitive Systems Laboratory, McMaster University, Hamilton, ON L8S 4K1 Canada (e-mail: haykin@mcmaster.ca).

T. R. Hurd is with the Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1 Canada (e-mail: hurdt@mcmaster.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2056923

A unique characteristic of the CKF is the fact that the cubature rule leads to an even number of equally weighted cubature points ($2n$ points, where n is the dimensionality of the state vector); these points are distributed uniformly on an ellipsoid centered on the origin. In a related context, the unscented Kalman filter (UKF) due to Julier *et al.* [26] has an odd number of sigma points ($(2n+1)$ points) also distributed on an ellipsoid but with a nonzero center point. Whereas the cubature points of the CKF follow rigorously from the cubature rule, the sigma points of the UKF are the result of the "unscented" transformation applied to inputs. In other words, there is a fundamental difference between the CKF and UKF:

- The CKF follows directly from the cubature rule, an important property of which is that it does not entail any free parameter;
- In contrast, the UKF purposely introduces a nonzero scaling parameter, commonly denoted by κ, which defines the nonzero center point that is often given more weighting than the remaining set of sigma points (see Section III for more details).

Indeed, it is the inclusion of the parameter κ that is responsible for the UKF to underperform compared to the UKF. What is truly interesting is that when κ of the "plain" UKF is set to zero (by "plain" we mean the UKF without using a scaled unscented transformation), the sigma point set boils down to the cubature point set and the algorithmic steps of the plain UKF become identical to that of the CKF. It is ironic that the observation for setting κ equal to zero has been largely overlooked in the literature on nonlinear filtering for the past many years.

To proceed on, the original derivation of the CKF is limited to a discrete-time domain, where the process and measurement equations are both described by stochastic difference equations. However, in many nonlinear filtering problems considered in practice, the state-space model is of the continuous-discrete kind. Specifically, the process equation is formulated from the underlying physics of a dynamic system and therefore expressed in the form of a set of stochastic ordinary differential equations (ODEs). As before, with digital devices as the method of choice, the measurements are naturally made in the discrete-time domain. Filtering problems in the continuous-discrete time domain often arise in numerous applications such as target tracking [8], [41], navigation [21], stochastic control [4], and finance [35].

The motivation of this paper is to extend the CKF to deal with state-space models of the continuous-discrete kind. The key to successful filtering in such continuous-discrete background lies in the effective extraction of useful information about the system's states from available measurements. A "good" process equation in discrete-time will certainly facilitate this

information extraction to a great extent. In this paper, we use the Itô-Taylor expansion of order 1.5 to transform the process equation in the stochastic ODE form into a more familiar stochastic difference equation for the first time in the nonlinear filtering literature. This transformation yields a more accurate yet approximate process equation in discrete time. Building on this transformation and assuming that all conditional densities are Gaussian-distributed, the filtering solution to the continuous-discrete state-space model reduces to the problem of how to compute Gaussian-weighted integrals. Given such a setting, we go on to use a third-degree cubature rule to numerically compute them. To be consistent with the literature, we have named the resulting algorithm the *Continuous-Discrete Cubature Kalman Filter (CD-CKF)*.

The rest of the paper is structured as follows: Section II defines the continuous-discrete filtering problem and presents the conceptual yet optimal Bayesian filtering solution. Section III briefly reviews a number of existing continuous-discrete filters. Section IV presents the Itô-Taylor expansion of order 1.5, a powerful numerical method used to approximate stochastic ODEs. In Section V, we derive the CD-CKF using the Itô-Taylor expansion of order 1.5. We go on to derive a square-root version of the CD-CKF for improved reliability in a limited-precision system in Section VI. Section VII presents a number of unique features of the proposed CD-CKF that distinguish it from existing continuous-discrete filters. Section VIII is devoted to a challenging experiment, which compares the square-root CD-CKF against existing continuous-discrete filters in tracking coordinated turns in an air-traffic-control environment. Section IX concludes the paper with some final remarks.

II. Bayesian Filtering in Nonlinear Continuous-Discrete State-Space Models

We consider a state-space model whose process equation is given by a stochastic differential equation [25]

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t)dt + \sqrt{\mathbf{Q}}d\boldsymbol{\beta}(t) \tag{1}$$

where $\mathbf{x}(t)$ denotes the n-dimensional state of the system at time t; $\boldsymbol{\beta}(t)$ denotes the n-dimensional standard Brownian motion with increment $d\boldsymbol{\beta}(t)$ that is independent of $\mathbf{x}(t)$; $\mathbf{f} : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ is a known nonlinear drift function having appropriate regularity properties; and $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is called the diffusion matrix, also known as the spectral density matrix or gain matrix of the process noise. The behavior of the system is observed through noisy measurements at sampled time instants $t_k = kT$, as shown by

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k, \quad k = 1, 2 \ldots \tag{2}$$

where $\mathbf{z}_k \in \mathbb{R}^d$; $\mathbf{h} : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^d$; the measurement noise $\mathbf{v}_k \in \mathbb{R}^n$ is assumed to be Gaussian with zero mean and known covariance matrix $\mathbb{E}[\mathbf{v}_i \mathbf{v}_j^T] = \mathbf{R}_i \delta(i-j)$, $(i, j = 1, 2, \ldots)$; and T is the measurement sampling interval. We use the symbols $\mathbb{E}$ and $\delta(i-j)$ for the expectation operator and the Kronecker delta, respectively. It is assumed that the initial condition and noise processes are all statistically independent.

A. Optimal Bayesian Filter

In the Bayesian filtering paradigm, the conditional density of the state given the measurements, also called the posterior density of the state, provides a complete statistical description of the state at that time [22]. The optimal continuous-discrete Bayesian filter consists of the following:

- Propagation of the 'old' posterior density between the measurement instants; this first equation is called the *time update*.
- Use of Bayes' rule to update this propagated posterior density at the measurement instants; this second equation is called the *measurement update*.

Let us examine the time update first. It is well known that for (1), the propagation or the temporal evolution $p(\mathbf{x}_t|\mathbf{z}_{1:k})$ of the old posterior density of the state from time t_k to t obeys the well-known Fokker-Planck equation, also called Kolmogorov's forward equation [19], [25], [37]

$$\frac{\partial p}{\partial t} = -\mathbf{f}^T \frac{\partial p}{\partial \mathbf{x}_t} - \mathrm{tr}\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}_t}\right)p + \frac{1}{2}\mathrm{tr}\left(\mathbf{Q}\frac{\partial^2 p}{\partial \mathbf{x}_t^2}\right) \tag{3}$$

with the initial condition $p(\mathbf{x}_k|\mathbf{z}_{1:k})$, where "tr" is used to denote the trace operator, $p = p(\mathbf{x}_t|\mathbf{z}_{1:k})$ with $t \leq t_{k+1}$ and the measurement history up to time t_k denoted by $\mathbf{z}_{1:k} = \{\mathbf{z}_i,\ i = 1, 2 \ldots k\}$.

Considering the measurement update, on the receipt of $\mathbf{z}_{k+1}$, the use of Bayes' rule yields the updated posterior density

$$p(\mathbf{x}_{k+1}|\mathbf{z}_{1:k+1}) = \frac{1}{c_{k+1}} p(\mathbf{x}_{k+1}|\mathbf{z}_{1:k}) p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}), \tag{4}$$

where the likelihood function

$$p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) = \mathcal{N}(\mathbf{z}_{k+1}; \mathbf{h}(\mathbf{x}_{k+1}, k+1), \mathbf{R}_{k+1}),$$

with $\mathcal{N}(.,.)$ being the conventional symbol for a Gaussian density; and the normalizing constant

$$c_{k+1} = \int_{\mathbb{R}^n} p(\mathbf{x}_{k+1}|\mathbf{z}_{1:k}) p(\mathbf{z}_{k+1}|\mathbf{x}_{k+1}) d\mathbf{x}_{k+1}. \tag{5}$$

The predictive density $p(\mathbf{x}_{k+1}|\mathbf{z}_{1:k})$ is obtained from (3).

The pair of (3) and (4) describes the continuous-discrete Bayesian filter only in conceptual terms. Apart from a few special cases, the new posterior will not remain within a closed family of distributions described by a finite summary statistic and so the Fokker-Planck equation has to be solved approximately. Known exact solutions are limited to the following cases. For a linear process equation, (3) reduces to the time-update of the Kalman-Bucy filter [28]. For a Beneš-type nonlinear process equation, (3) reduces to the time-update of the Beneš filter [10]. Finally, in [16], Daum has further extended the class of nonlinear dynamic systems that admit a sufficient statistic of a constant finite dimension. Apart from these restricted known cases, we have to be content with an approximate solution. The

next section reviews existing approximate continuous-discrete nonlinear filters.

III. Time Update of Existing Filters: A Brief Review

The time update of a continuous-discrete filter can be divided into two broadly defined types. Type-I filters compute the conditional density by explicitly solving the Fokker-Planck (3). For example, in the past, researchers have used a number of numerical methods such as the finite-difference method [29], adaptive finite-element methods [6, pp. 115–123]), the adjoint method [18], Galerkin's method [9], particle methods [13], [14], [36] and Markov chain Monte Carlo methods [11]. On the other hand, type-II filters compute merely a finite number of summary statistics in terms of conditional moments after discretizing the Itô-type process equation using the Euler or higher-order Runge-Kutta methods [23], [30]. The computational complexity of type-I filters increases exponentially with the dimension of the state vector, whereas the computational complexity of type-II filters varies only polynomially with the dimension of the state vector [17]. Because type-II filters are computationally cheaper than type-I filters and the proposed new filter, (the CD-CKF), fits into the type-II filter family, in what follows, we present a detailed review of the time-updates of two well-known type-II filters—continuous-discrete extended Kalman filter (CD-EKF) [25] and continuous-discrete unscented Kalman filter (CD-UKF) [38]. In so doing, we also facilitate comparisons presented in Section VII between the proposed CD-CKF and these two existing filters.

Before diving into the details of the CD-EKF and CD-UKF, we first set a common stage used for their derivations. Applying the Euler approximation to the stochastic differential (1) over the time interval $(t, t+\delta)$ yields

$$\mathbf{x}(t+\delta) = \mathbf{x}(t) + \delta \mathbf{f}(\mathbf{x}(t), t) + \sqrt{\mathbf{Q}}\mathbf{w} \tag{6}$$

where the n-dimensional Gaussian random variable $\mathbf{w}$ is related to the standard Gaussian random variable $\mathbf{u}$ via $\mathbf{w} = \sqrt{\delta}\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \delta \mathbf{I}_n)$ with $\mathbf{I}_n$ being used to denote the n-dimensional identity matrix. Taking the expectation yields

$$\mathbb{E}[\mathbf{x}(t+\delta)] = \mathbb{E}[\mathbf{x}(t)] + \delta \mathbb{E}[\mathbf{f}(\mathbf{x}(t), t)]. \tag{7}$$

Because $\mathbf{x}(t)$ is independent of $\mathbf{w}$, the associated error covariance matrix satisfies

$$\mathrm{var}[\mathbf{x}(t+\delta)] = \mathrm{var}[\mathbf{x}(t) + \delta \mathbf{f}(\mathbf{x}(t), t)] + \delta \mathbf{Q}. \tag{8}$$

CD-EKF: Historically, the first applicable continuous-discrete nonlinear filter was derived in Chapters 6 and 9 of [25] as generalization of the classical EKF. To approximately solve the moment equations given by (7) and (8) over the time interval (t_k, t_{k+1}), the CD-EKF uses the first-order Taylor series, which expands $\mathbf{f}(\mathbf{x}(t), t)$ around the latest known estimate $\hat{\mathbf{x}}_{k|k} = \mathbb{E}[\mathbf{x}_k | \mathbf{z}_{1:k}]$

$$\mathbf{f}(\mathbf{x}(t), t) = \mathbf{f}(\hat{\mathbf{x}}_{k|k}, k) + \mathbf{f}_{\mathbf{x}}(k)(\mathbf{x}(t) - \hat{\mathbf{x}}_{k|k}) + \mathrm{HOT} \tag{9}$$

where HOT refers to higher-order terms and the Jacobian of $\mathbf{f}(.,.)$ evaluated at $\hat{\mathbf{x}}_{k|k}$ is

$$\mathbf{f}_{\mathbf{x}}(k) = \left[\nabla_{\mathbf{x}} \mathbf{f}^T(\mathbf{x}, k)\right]^T_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}}.$$

Consider the case where $t = t_k$ and $(t+\delta) = t_{k+1}$, which also implies that $\delta = T$. Substituting (9) into (7) yields the predicted state estimate

$$\hat{\mathbf{x}}_{k+1|k} = \mathbb{E}[\mathbf{x}_{k+1} | \mathbf{z}_{1:k}] \approx \hat{\mathbf{x}}_{k|k} + T\mathbf{f}(\hat{\mathbf{x}}_{k|k}, k). \tag{10}$$

Similarly, substituting (9) into (8) yields the predicted state-error covariance matrix

$$\begin{aligned} \mathbf{P}_{k+1|k} &= \mathrm{var}[\mathbf{x}_{k+1} | \mathbf{z}_{1:k}] \\ &\approx \left(\mathbf{I}_n + T\mathbf{f}_{\mathbf{x}}(k)\right) \mathbf{P}_{k|k} \left(\mathbf{I}_n + T\mathbf{f}_{\mathbf{x}}(k)\right)^T \\ &\quad + T\mathbf{Q}. \end{aligned} \tag{11}$$

The predicted state estimate (10) and its covariance (11) collectively form the time-update of the CD-EKF.

CD-UKF: Next, we focus on the time-update of the CD-UKF [38], [39]. At the heart of the CD-UKF is the *unscented transformation*, which uses a set of $(2n+1)$ sigma points to approximate a symmetric density $\Pi(\mathbf{x})$ of an n-dimensional random variable $\mathbf{x}$ with mean μ and covariance Σ as shown by

$$\Pi(\mathbf{x}) \approx \sum_{i=0}^{2n} w_i \boldsymbol{\delta}(\mathbf{x} - \mathcal{X}_i), \tag{12}$$

where $\boldsymbol{\delta}(.)$ is the Dirac delta function and the sigma points $\{\mathcal{X}_i\}$ are chosen to match the first and second moments, such that $(i = 1, 2, \ldots n)$

$$\begin{aligned} \mathcal{X}_0 &= \mu, \quad w_0 = \frac{\kappa}{n+\kappa} \\ \mathcal{X}_i &= \mu + (\sqrt{(n+\kappa)\Sigma})_i, \quad w_i = \frac{1}{2(n+\kappa)} \\ \mathcal{X}_{n+i} &= \mu - (\sqrt{(n+\kappa)\Sigma})_i, \quad w_{n+i} = \frac{1}{2(n+\kappa)} \end{aligned}$$

where the ith column of a matrix $\mathbf{A}$ is denoted by $(\mathbf{A})_i$; and the parameter κ is chosen to be $\kappa = (3 - n)$. Hence, a nonlinear expectation is approximately computed using this sigma-point set

$$\begin{aligned} \mathbb{E}[\mathbf{f}(\mathbf{x})] &= \int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x}) \Pi(\mathbf{x}) d\mathbf{x} \\ &\approx \sum_{i=0}^{2n} w_i \mathbf{f}(\mathcal{X}_i) \end{aligned} \tag{13}$$

$$\begin{aligned} \mathrm{var}[\mathbf{f}(\mathbf{x})] &= \int_{\mathbb{R}^n} (\mathbf{f}(\mathbf{x}) - \mathbb{E}[\mathbf{f}(\mathbf{x})])(\mathbf{f}(\mathbf{x}) - \mathbb{E}[\mathbf{f}(\mathbf{x})])^T \Pi(\mathbf{x}) d\mathbf{x} \\ &\approx \sum_{i=0}^{2n} w_i (\mathbf{f}(\mathcal{X}_i) - \mathbb{E}[\mathbf{f}(\mathbf{x})])(\mathbf{f}(\mathcal{X}_i) - \mathbb{E}[\mathbf{f}(\mathbf{x})])^T \end{aligned} \tag{14}$$

By neglecting terms of the order δ^2, the predicted state-error covariance matrix (8) is simplified to obtain the following

approximate result (see, for example, [38, Appendix B])

$$\mathrm{var}[\mathbf{x}(t+\delta)] \approx \mathrm{var}[\mathbf{x}(t+\delta)] + \delta\big(\mathrm{cov}[\mathbf{x}(t), \mathbf{f}(\mathbf{x}(t),t)] + \mathrm{cov}[\mathbf{f}(\mathbf{x}(t),t), \mathbf{x}(t)] + \mathbf{Q}\big). \quad (15)$$

As δ tends to zero, the moments equations in (7) and (15) can be expressed in vector and matrix differential equations, respectively, as follows:

$$\frac{d}{dt}\mathbb{E}[\mathbf{x}(t)] = \mathbb{E}[\mathbf{f}(\mathbf{x}(t),t)] \quad (16)$$

$$\frac{d}{dt}\mathrm{var}[\mathbf{x}(t)] = \mathrm{cov}[\mathbf{f}(\mathbf{x}(t),t), \mathbf{x}(t)] + \mathrm{cov}[\mathbf{x}(t), \mathbf{f}(\mathbf{x}(t),t)] + \mathbf{Q}. \quad (17)$$

Given that $\mathbf{x}_k$ at time t_k is Gaussian distributed with the mean $\mathbb{E}[\mathbf{x}(t_k)|\mathbf{z}_{1:k}] = \hat{\mathbf{x}}_{k|k}$ and covariance $\mathrm{var}[\mathbf{x}(t_k)|\mathbf{z}_{1:k}] = \mathbf{P}_{k|k}$, $\hat{\mathbf{x}}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ are obtained as follows: First, the operations $\mathbb{E}[.]$ and $\mathrm{cov}[.]$ present on the right-hand side (RHS) of (16) and (17) are approximated by the unscented transformation as discussed above. Then, the results are numerically integrated up to time t_{k+1}; the Runge-Kutta method is used as a numerical integration tool in [38].

In this section, so far, we have considered only a single-step time update. To compute the predicted state and its error covariance more accurately at time t_{k+1}, an m-step prediction over the "small" time intervals of length δ, where $\delta = (t_{k+1} - t_k)/m = T/m$ and the integer $m > 1$, is usually performed in a recursive manner.

On the receipt of a new measurement $\mathbf{z}_{k+1}$, the posterior density is obtained by fusing $\mathbf{z}_{k+1}$ with the predictive density using Bayes' rule. Because the measurement-update relies only on the measurement equation, which is modeled in discrete time for a continuous-discrete state-space model case, the measurement-updates of the CD-EKF and CD-UKF reduce to that of the EKF and UKF, respectively. In the next section, we describe the Itô-Taylor Expansion that directly transforms the process equation into a stochastic difference equation. As it captures process-model nonlinearity and process noise more accurately, this new approach is well suited to continuous-discrete filtering problems. In fact, it provides increased stability and improved state estimation.

IV. Itô-Taylor Expansion of Order 1.5 for Solving Stochastic ODEs

Itô-Taylor expansions, including the stochastic versions of the Euler and Milstein schemes, are a general method for discretizing stochastic differential equations [30]. An Itô-Taylor expansion for a stochastic differential equation is said to be strongly convergent with order β if, for any $m = 1, 2, \ldots$ and time interval $[t_k, t_{k+1}]$, the error of the m-step approximation of $\hat{\mathbf{x}}^{(m)}(t)$ to $\mathbf{x}(t)$ representing the exact solution given the initial condition $\mathbf{x}(t_k) = \mathbf{x}_k$, satisfies

$$\mathbb{E}\left[\sup_{t\in[t_k,t_{k+1}]} |\mathbf{x}(t) - \hat{\mathbf{x}}^{(m)}(t)|\right] \le \lambda(\delta^{(m)})^{\beta}.$$

Here, $\delta^{(m)} = (t_{k+1} - t_k)/m$ and λ is a constant uniform in m. Hence, the expansion becomes more accurate as m increases. According to [30, Sec. 10.2], the Euler approximation (6) is equivalent to the Itô-Taylor expansion of order $\beta = 0.5$.

Because an Itô-Taylor expansion of higher order is theoretically more accurate than that of lower order, the Itô-Taylor expansion of order 1.5, which we denote by IT-1.5 in the rest of the paper, is ideal for our purpose. According to [30, Sec. 10.4], applying the IT-1.5 to the stochastic differential (1) over the time interval $(t, t+\delta)$ yields

$$\mathbf{x}(t+\delta) = \underbrace{\mathbf{x}(t) + \delta\mathbf{f}(\mathbf{x}(t),t) + \frac{1}{2}\delta^2(\mathbb{L}_0\mathbf{f}(\mathbf{x}(t),t))}_{\mathbf{f}_d(\mathbf{x}_t,t)} + \sqrt{\mathbf{Q}}\mathbf{w} + (\mathbb{L}\mathbf{f}(\mathbf{x}(t),t))\mathbf{y} \quad (18)$$

This equation entails the following notations:

- Two differential operators, $\mathbb{L}_0$ and $\mathbb{L}_j$ $(j = 1, 2 \ldots n)$, which are, respectively, defined by

$$\mathbb{L}_0 = \frac{\partial}{\partial t} + \sum_{i=1}^{n} \mathbf{f}_i \frac{\partial}{\partial \mathbf{x}_i} + \frac{1}{2}\sum_{j,p,q=1}^{n} \sqrt{\mathbf{Q}}_{p,j}\sqrt{\mathbf{Q}}_{q,j}\frac{\partial^2}{\partial \mathbf{x}_p \partial \mathbf{x}_q}$$

$$\mathbb{L}_j = \sum_{i=1}^{n} \sqrt{\mathbf{Q}}_{i,j}\frac{\partial}{\partial \mathbf{x}_i};$$

- The term $\mathbb{L}\mathbf{f}$ denotes a square matrix with its (i,j)th element being $\mathbb{L}_j\mathbf{f}_i$, $(i, j = 1, \ldots, n)$;
- Noise-free process function, which, in discrete time, is defined by

$$\mathbf{f}_d(\mathbf{x}(t),t) = \mathbf{x}(t) + \delta\mathbf{f}(\mathbf{x}(t),t) + \frac{1}{2}\delta^2(\mathbb{L}_0\mathbf{f}(\mathbf{x}(t),t)); \quad (19)$$

- Pair of correlated n-dimensional Gaussian random variables $(\mathbf{w}, \mathbf{y})$, which can be generated from a pair of independent n-dimensional standard Gaussian random variables $(\mathbf{u}_1, \mathbf{u}_2)$ as follows:

$$\mathbf{w} = \sqrt{\delta}\mathbf{u}_1$$
$$\mathbf{y} = \frac{1}{2}\delta^{3/2}\left(\mathbf{u}_1 + \frac{\mathbf{u}_2}{\sqrt{3}}\right).$$

Accordingly, we find the following three covariance matrices:

$$\mathbb{E}[\mathbf{w}\mathbf{w}^T] = \delta\mathbf{I}_n$$
$$\mathbb{E}[\mathbf{w}\mathbf{y}^T] = \frac{1}{2}\delta^2\mathbf{I}_n$$
$$\mathbb{E}[\mathbf{y}\mathbf{y}^T] = \frac{1}{3}\delta^3\mathbf{I}_n.$$

A. Motivating Example Using the IT-1.5

For a target moving with a 'nearly' constant velocity in a single physical dimension, we assume that the velocity follows a random walk with a constant diffusion gain and position is the integral of velocity. Hence, the target dynamics evolve according to the following equation:

$$d\mathbf{x}(t) = \mathbf{F}\mathbf{x}(t)dt + \sqrt{\mathbf{Q}}d\boldsymbol{\beta}(t), \quad (20)$$

where the two-dimensional state of the target $\mathbf{x} = [\epsilon,\ \dot{\epsilon}]^T$ with ϵ and $\dot{\epsilon}$ denoting the position and velocity of the target, respectively; the drift function $\mathbf{F} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$; and the gain matrix $\mathbf{Q} = \mathrm{diag}([0,\ \sigma^2])$. Because of the linear structure of the process equation, this stochastic ODE can be analytically solved over the interval $[0, T]$, obtaining [8, Sec. 4.3.1]:

$$\mathbf{x}(T) = \exp(\mathbf{F}T)\mathbf{x}(0) + \int_0^T \exp(\mathbf{F}(T-t))\sqrt{\mathbf{Q}}\, d\boldsymbol{\beta}(t)$$

where $\exp(\mathbf{F}T) = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}$ and the Itô integral can be computed to be a two-dimensional zero-mean Gaussian with covariance matrix $\begin{pmatrix} T^3/3 & T^2/2 \\ T^2/2 & T \end{pmatrix}\sigma^2$.

Applying the IT-1.5 to (20) over the same time interval yields

$$\mathbf{x}(T) = \underbrace{\mathbf{x}(0) + T\begin{pmatrix} \dot{\epsilon}(0) \\ 0 \end{pmatrix}}_{\mathbf{F}_d\mathbf{x}(0)} + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix}\mathbf{w} + \begin{pmatrix} 0 & \sigma \\ 0 & 0 \end{pmatrix}\mathbf{y}}_{\mathbf{v}}$$

where

$$\mathbf{F}_d = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}$$

and $\mathbf{v}$ is a white noise sample from zero-mean Gaussian process with covariance matrix

$$\boldsymbol{\Sigma} = \mathbb{E}[\mathbf{v}\mathbf{v}^T] = \begin{pmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ \frac{1}{2}T^2 & T \end{pmatrix}\sigma^2.$$

We thus see that in this model, the IT-1.5 is exactly the same as the above analytical solution.

V. CD-CKF

In this section, we extend the CKF with the state-space models of the continuous-discrete kind using the Itô-Taylor expansion for discretizing the process equation. We refer to the resulting algorithm as the CD-CKF. Given that the process equation is approximated by a stochastic difference equation, the key assumption taken to develop this new algorithm is that the state and measurement processes at time t_{k+1}, given the past measurement history up to time t_k, are jointly approximated to Gaussian. Under this assumption, the Bayesian filter reduces to the problem of how to compute integrals whose integrands are of the form

$$\text{nonlinear function} \times \text{Gaussian}.$$

To numerically compute these integrals, we adopt a similar technique presented in [1]. For completeness, the technique is briefly described next.

A. Review of Third-Degree Cubature Rule

Cubature rules approximate a multidimensional weighted integral of the form

$$I(\mathbf{f}) = \int_{\mathsf{D}} \mathbf{f}(\mathbf{x})w(\mathbf{x})d\mathbf{x} \tag{21}$$

where $\mathbf{f}(.)$ is some arbitrary function, $\mathsf{D} \subseteq \mathbb{R}^n$ is the region of integration, and the known weighting function $w(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathsf{D}$, using a set of weighted points:

$$I(\mathbf{f}) \approx \sum_{i=1}^{m} \omega_i \mathbf{f}(\mathbf{x}_i). \tag{22}$$

For example, a third-degree cubature rule can be constructed to approximate an n-dimensional Gaussian weighted integral as follows [40]:

$$\int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x})\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})d\mathbf{x} \approx \frac{1}{2n}\sum_{i=1}^{2n} \mathbf{f}(\boldsymbol{\mu} + \sqrt{\boldsymbol{\Sigma}}\xi_i),$$

where $\boldsymbol{\Sigma} = \sqrt{\boldsymbol{\Sigma}}\sqrt{\boldsymbol{\Sigma}}^T$ and the cubature points

$$\xi_i = \begin{cases} \sqrt{n}\mathbf{e}_i, & i = 1, 2 \ldots n \\ -\sqrt{n}\mathbf{e}_{i-n}, & i = n+1, n+2 \ldots 2n. \end{cases}$$

Thus, the cubature rule uses an even set of $2n$ equally weighted symmetric cubature points. The cubature rule is exact for nonlinear functions belonging to monomials of degree three or less [3, Proposition 3.1]. Next, we derive the time update of the CD-CKF using the third-degree cubature rule.

B. Time-Update of the CD-CKF

Equipped with the third-degree cubature rule, we can now use the IT-1.5 to update the old posterior density of the state before receiving a new measurement, and obtain the first and second moments of the predictive density. In what follows, we use the notation $\mathbf{x}_k^j$ to denote $\mathbf{x}(t)$ at time $t = kT + j\delta$, where $1 \leq j \leq m$ and $\delta = T/m$. Given the statistic of $\mathbf{x}_k$ such that $\mathbf{x}_k \sim \mathcal{N}(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$, for $j = 1$ we may thus write the predicted state estimate

$$\begin{aligned} \hat{\mathbf{x}}_{k|k}^1 &= \mathbb{E}[\mathbf{x}_k^1 | \mathbf{z}_{1:k}] \\ &\approx \mathbb{E}[\mathbf{f}_d(\mathbf{x}_k, kT) + \sqrt{\mathbf{Q}}\mathbf{w} + (\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT))\mathbf{y} | \mathbf{z}_{1:k}]. \end{aligned}$$

where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \delta\mathbf{I}_n)$; $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, (\delta^3/3)\mathbf{I}_n)$; and $\mathbb{E}[\mathbf{w}\mathbf{y}^T] = (\delta^2/2)\mathbf{I}_n$. Because the noise terms are independent of the state vector and zero-mean Gaussian, we may further simplify matters by writing

$$\begin{aligned} \hat{\mathbf{x}}_{k|k}^1 &= \mathbb{E}[\mathbf{f}_d(\mathbf{x}_k, kT) | \mathbf{z}_{1:k}] \\ &= \int_{\mathbb{R}^n} \mathbf{f}_d(\mathbf{x}_k, kT)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k. \end{aligned} \tag{23}$$

Similarly, we may write the predicted state-error covariance matrix [see (24) at the bottom of the next page]. Note that the CD-CKF not only propagates the state estimate (23) but also the predicted state error covariance matrix (24) for each δ during the

time-update. As long as the nonlinearity of $\mathbb{L}\mathbf{f}(.,.)$ is not "severe," we have found that a simpler approximation will be as effective as (24), namely to replace $\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT)$ by $\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)$

$$\begin{aligned}\mathbf{P}_{k|k}^1 \approx & \int_{\mathbb{R}^n} \mathbf{f}_d(\mathbf{x}_k, kT)\mathbf{f}_d^T(\mathbf{x}_k, kT)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k \\ & + \frac{\delta^3}{3}\big(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)\big)\big(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)\big)^T \\ & + \frac{\delta^2}{2}\Big[\sqrt{\mathbf{Q}}\big(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)\big)^T + \big(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)\big)\sqrt{\mathbf{Q}}^T\Big] \\ & - (\hat{\mathbf{x}}_{k|k}^1)(\hat{\mathbf{x}}_{k|k}^1)^T + \delta\mathbf{Q}. \end{aligned} \tag{25}$$

The alternative expression (25) allows us to compute the predicted state-error covariance matrix efficiently. Because (25) admits a simple expression as a sum of squared matrices, it also helps us formulate a square-root version of the CD-CKF, which is discussed in Section VI.

To numerically compute the integral present on the RHS of (23), we now use the third-degree cubature rule, obtaining

$$\hat{\mathbf{x}}_{k|k}^1 \approx \frac{1}{2n}\sum_{i=1}^{2n} X_{i,k|k}^{*(1)}, \tag{26}$$

where

$$X_{i,k|k}^{*(1)} = \mathbf{f}_d\left(\hat{\mathbf{x}}_{k|k} + \sqrt{\mathbf{P}_{k|k}}\xi_i, kT\right) \quad i = 1, 2 \ldots 2n.$$

Similarly, we numerically compute (25) and obtain

$$\begin{aligned}\mathbf{P}_{k|k}^1 \approx & (\mathsf{X}_{k|k}^{*(1)})(\mathsf{X}_{k|k}^{*(1)})^T \\ & + \frac{\delta^3}{3}(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT))(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT))^T \\ & + \frac{\delta^2}{2}\left[\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT)\mathbf{Q}^T + \mathbf{Q}(\mathbb{L}\mathbf{f}(\hat{\mathbf{x}}_{k|k}, kT))^T\right] \\ & + \delta\mathbf{Q} \end{aligned} \tag{27}$$

where

$$\begin{aligned}\mathsf{X}_{k|k}^{*(1)} = \frac{1}{\sqrt{2n}}[X_{1,k|k}^{*(1)} - \hat{\mathbf{x}}_{k|k}^1 \; X_{2,k|k}^{*(1)} - \hat{\mathbf{x}}_{k|k}^1 \ldots \\ X_{2n,k|k}^{*(1)} - \hat{\mathbf{x}}_{k|k}^1]. \end{aligned} \tag{28}$$

The RHS of (28) appears as a scaled version of a set of terms, from each of which the prior mean is subtracted off; hereafter, we refer to such a matrix 'a weighted-centered matrix'. To compute the predicted state statistics $(\hat{\mathbf{x}}_{k|k}^m, \mathbf{P}_{k|k}^m)$ at time $t_{k+1} = (kT + m\delta)$, the integration steps (26) and (27) are performed in a successive manner. The measurement update of the CD-CKF is exactly the same as that of the CKF, the equations of which can be found in [1]. For completeness and coding purposes, all the equations involved in computing the CD-CKF are summarized in Appendix A.

VI. Square-Root Continuous-Discrete Cubature Kalman Filtering for Improved Reliability

The CD-CKF performs numerically sensitive operations such as matrix inversion and subtraction of two positive definite matrices. Unfortunately, these operations can lead to meaningless covariance matrices that fail to be symmetric and positive (semi-)definite. Specifically, the loss of positive definiteness becomes unacceptable as it stops the CD-CKF from running continuously.

In order to preserve the properties of symmetry and positive (semi-)definiteness and to improve numerical accuracy, various ad hoc methods have been introduced in the literature on Bayesian filtering. Some of them include measurement-update with a sequence of scalar measurements in a preferred order, decoupled or quasi-decoupled covariances, symmetrization of covariances, Joseph's covariance update and use of large process and measurement noise covariance matrices. These ad hoc methods were developed in an era of limited computing power. Equipped with the power of currently available computers, to find a systematic solution that preserves the properties of a covariance matrix and improves numerical accuracy, we look to square-root filtering algorithms, which propagate the square-roots of various error covariance matrices [2], [27]. Following this line of thinking, we may also structurally reformulate the CD-CKF developed for a continuous-discrete state-space model.

To this end, a covariance matrix $\mathbf{P}$ of the CD-CKF is written in the form

$$\mathbf{P} = \mathbf{S_o}\mathbf{S_o}^T, \tag{29}$$

where $\mathbf{P} \in \mathbb{R}^{n\times n}$, $\mathbf{S_o} \in \mathbb{R}^{n\times m}$, $m > n$, is a 'fat' matrix. Though $\mathbf{S_o}$ in (29) can be considered as a square-root of $\mathbf{P}$, we prefer to keep the square-root as an $n \times n$ triangular matrix for computational reasons. The transformation of $\mathbf{S_o}$ into a

---

$$\begin{aligned}\mathbf{P}_{k|k}^1 = & \mathbb{E}\big[(\mathbf{x}_k^1 - \hat{\mathbf{x}}_{k|k}^1)(\mathbf{x}_k^1 - \hat{\mathbf{x}}_{k|k}^1)^T|\mathbf{z}_{1:k}\big] \\ = & \int_{\mathbb{R}^n} \mathbf{f}_d(\mathbf{x}_k, kT)\mathbf{f}_d^T(\mathbf{x}_k, kT)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k + \frac{\delta^3}{3}\int_{\mathbb{R}^n}(\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT))(\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT))^T\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k \\ & + \frac{\delta^2}{2}\left[\sqrt{\mathbf{Q}}\left(\int_{\mathbb{R}^n}\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k\right)^T + \left(\int_{\mathbb{R}^n}\mathbb{L}\mathbf{f}(\mathbf{x}_k, kT)\mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})d\mathbf{x}_k\right)\sqrt{\mathbf{Q}}^T\right] \\ & - (\hat{\mathbf{x}}_{k|k}^1)(\hat{\mathbf{x}}_{k|k}^1)^T + \delta\mathbf{Q}. \end{aligned} \tag{24}$$

triangular matrix $\mathbf{S_n} \in \mathbb{R}^{n\times n}$ is performed by a triangularization procedure (e.g., Gram-Schmidt based QR-decomposition). When the matrix $\mathbf{S_o}^T$ is decomposed into an orthogonal matrix $\mathbf{V} \in \mathbb{R}^{m\times n}$ and an upper triangular matrix $\mathbf{D} \in \mathbb{R}^{n\times n}$ such that $\mathbf{S_o}^T = \mathbf{VD}$, we get

$$\mathbf{P} = \mathbf{S_o}\mathbf{S_o}^T = \mathbf{D}^T\mathbf{V}^T\mathbf{V}\mathbf{D} = \mathbf{D}^T\mathbf{D} = \mathbf{S_n}\mathbf{S_n}^T$$

where the 'new' square-root of $\mathbf{P}$, $\mathbf{S_n} = \mathbf{D}^T$. In this paper, we simplify matters by using the notation

$$\mathbf{S_n} = \mathbf{Tria}(\mathbf{S_o})$$

where $\mathbf{S_o}$ is referred to as the 'old' square-root of $\mathbf{P}$.

In what follows, we illustrate how this triangularization algorithm is fitted into the two-step square-root continuous-discrete cubature filter involving the time and measurement updates. Indeed, we may construct a square-root factor of the predicted state-error covariance matrix in a straightforward way (see the time-update of the square-root CD-CKF in Appendix B). Because the measurement update of the square-root CD-CKF is not straightforward, it is derived in the following subsection.

A. Measurement-Update of the Square-Root CD-CKF

From the measurement-update of the CD-CKF, we recall that the following three covariance matrices may be expressed in squared-matrix forms (see Appendix A):

$$\mathbf{P}_{k+1|k} = \mathsf{X}_{k+1|k}\mathsf{X}_{k+1|k}^T \tag{30}$$

$$\mathbf{P}_{zz,k+1|k} = \mathsf{Z}_{k+1|k}\mathsf{Z}_{k+1|k}^T + \mathbf{S}_{R,k+1}^T\mathbf{S}_{R,k+1} \tag{31}$$

$$\mathbf{P}_{xz,k+1|k} = \mathsf{X}_{k+1|k}\mathsf{Z}_{k+1|k}^T, \tag{32}$$

where the weighted-centered matrices

$$\mathsf{X}_{k+1|k} = \frac{1}{\sqrt{2n}}[X_{1,k+1|k} - \hat{\mathbf{x}}_{k+1|k}\ X_{2,k+1|k} - \hat{\mathbf{x}}_{k+1|k} \cdots X_{2n,k+1|k} - \hat{\mathbf{x}}_{k+1|k}] \tag{33}$$

$$\mathsf{Z}_{k+1|k} = \frac{1}{\sqrt{2n}}[\mathsf{Z}_{1,k+1|k} - \hat{\mathbf{z}}_{k+1|k}\ \mathsf{Z}_{2,k+1|k} - \hat{\mathbf{z}}_{k+1|k} \cdots \mathsf{Z}_{2n,k+1|k} - \hat{\mathbf{z}}_{k+1|k}]. \tag{34}$$

We combine (30)–(32) together and rewrite them in a squared-matrix form, as shown by

$$\begin{pmatrix} \mathbf{P}_{zz,k+1|k} & \mathbf{P}_{zx,k+1|k} \\ \mathbf{P}_{xz,k+1|k} & \mathbf{P}_{k+1|k} \end{pmatrix} = \begin{pmatrix} \mathsf{Z}_{k+1|k} & \mathbf{S}_{R,k+1} \\ \mathsf{X}_{k+1|k} & \mathbf{O} \end{pmatrix} \begin{pmatrix} \mathsf{Z}_{k+1|k} & \mathbf{S}_{R,k+1} \\ \mathsf{X}_{k+1|k} & \mathbf{O} \end{pmatrix}^T \tag{35}$$

where $\mathbf{O} \in \mathbb{R}^{n\times d}$ is the zero matrix. Applying the triangularization procedure to the square-root factor available on the RHS of (35) yields

$$\mathbf{Tria}\begin{pmatrix} \mathsf{Z}_{k+1|k} & \mathbf{S}_{R,k+1} \\ \mathsf{X}_{k+1|k} & \mathbf{O} \end{pmatrix} = \begin{pmatrix} \mathbf{T}_{11} & \mathbf{O} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} \tag{36}$$

where $\mathbf{T}_{11} \in \mathbb{R}^{d\times d}$, and $\mathbf{T}_{22} \in \mathbb{R}^{n\times n}$ are lower-triangular matrices, and $\mathbf{T}_{21} \in \mathbb{R}^{n\times d}$. Hence, we may rewrite (35) in the "new" squared-matrix form:

$$\begin{aligned} &\begin{pmatrix} \mathbf{P}_{zz,k+1|k} & \mathbf{P}_{zx,k+1|k} \\ \mathbf{P}_{xz,k+1|k} & \mathbf{P}_{k+1|k} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{T}_{11} & \mathbf{O} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{T}_{11} & \mathbf{O} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix}^T. \\ &= \begin{pmatrix} \mathbf{T}_{11}\mathbf{T}_{11}^T & \mathbf{T}_{11}\mathbf{T}_{21}^T \\ \mathbf{T}_{21}\mathbf{T}_{11}^T & \mathbf{T}_{21}\mathbf{T}_{21}^T + \mathbf{T}_{22}\mathbf{T}_{22}^T \end{pmatrix}. \end{aligned} \tag{37}$$

Recalling from the cubature filter, the filter gain is defined by

$$\mathbf{W}_{k+1} = \mathbf{P}_{xz,k+1|k}\mathbf{P}_{zz,k+1|k}^{-1}. \tag{38}$$

Substituting the results obtained in (37) into (38) yields

$$\begin{aligned} \mathbf{W}_{k+1} &= \mathbf{T}_{21}\mathbf{T}_{11}^T(\mathbf{T}_{11}\mathbf{T}_{11}^T)^{-1} \\ &= \mathbf{T}_{21}\mathbf{T}_{11}^{-1}. \end{aligned} \tag{39}$$

Because $\mathbf{T}_{11}$ is a lower-triangular matrix, we may efficiently compute $\mathbf{W}_{k+1}$. Using the new symbol '/' to represent the matrix right-divide operator, we may write $(\mathbf{T}_{21}/\mathbf{T}_{11}^{-1})$ to be $\mathbf{T}_{21}/\mathbf{T}_{11}$ and thus essentially apply the *forward substitution* algorithm to compute $\mathbf{W}_{k+1}$. We therefore write the updated state estimate

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + (\mathbf{T}_{21}/\mathbf{T}_{11})(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}). \tag{40}$$

Next, we consider how the associated updated state error covariance matrix $\mathbf{P}_{k+1|k+1}$ is written in matrix-squared form. We first write

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{W}_{k+1}\mathbf{P}_{zz,k+1|k}\mathbf{W}_{k+1}^T. \tag{41}$$

Then, substituting the results obtained in (37) into (41) yields

$$\begin{aligned} \mathbf{P}_{k+1|k+1} &= (\mathbf{T}_{21}\mathbf{T}_{21}^T + \mathbf{T}_{22}\mathbf{T}_{22}^T) \\ &\quad - \mathbf{T}_{21}\mathbf{T}_{11}^{-1}(\mathbf{T}_{11}\mathbf{T}_{11}^T)(\mathbf{T}_{21}\mathbf{T}_{11}^{-1})^T \\ &= \mathbf{T}_{22}\mathbf{T}_{22}^T. \end{aligned}$$

Hence, the 'new' square-root factor of $\mathbf{P}_{k+1|k+1}$ is $\mathbf{T}_{22}$.

To sum up, given the weighted-centered matrices $\mathsf{X}_{k+1|k}$, and $\mathsf{Z}_{k+1|k}$, and the square-root factor of the measurement noise covariance matrix $\mathbf{S}_{R,k+1}$, the heart of the measurement-update of square-root continuous-discrete cubature filtering resides in computing the matrices $\mathbf{T}_{11}, \mathbf{T}_{21}$ and $\mathbf{T}_{22}$ as defined in (36). Subsequently, we compute the updated state estimate $\hat{\mathbf{x}}_{k+1|k+1}$ as shown in (40) and the square-root of the corresponding error covariance matrix that is simply given by $\mathbf{T}_{22}$.

Remarks:

- The CD-CKF derived in this paper shares a number of common features with its predecessor CKF such as the approximate formulation under Gaussian assumption and the use of the third-degree cubature rule for Gaussian weighted integrals. For a continuous-discrete setting, we have used the IT-1.5 method to discretize the Itô-type process equation which in turn translates the continuous-discrete filtering problem into a discrete-time

filtering problem. In this context, we could raise a question: After this time-domain transformation using the IT-1.5 method, what aspects of the CD-CKF make it distinct from the CKF and its square-root version? Indeed, to make the CD-CKF more accurate, stable and computationally efficient, we have done a number of clever things:
— An m-step $(m > 1)$ iterative time-update for improved perdition. This idea has been in the literature for many years.
— The way we simplify the time-update for efficient computations [see (25)].
— The way we restructure the time-update for a stable square-root formulation using *Tria*.
— The way we elegantly reformulate the square-root measurement-update by combining submatrices into a big one.

- Comparison of the measurement-update of the existing square-root CKF derived in [1] with the square-root CD-CKF derived in this paper reveals that both algorithms approximately require the same computational cost of $(6n^3 + 10n^2d)$ flop counts (or simply flops) per update cycle (see Appendix C). In computing these costs, note that we do not account for flops associated with problem-specific function evaluations, which are common to both algorithms. However, the derivation of the square-root CD-CKF is more elegant in that it applies the triangularization procedure to the array of matrices only once; in so doing, the square-root CD-CKF avoids an explicit computation of the cross-covariance matrix and repeated use of the forward substitution algorithm in computing the cubature Kalman gain.
- The square-root CD-CKF closely resembles the square-root formulation of the linear Kalman filter proposed by Kaminski *et al.* [27]. In this regard, a key exception is that in the square-root CD-CKF, various covariance matrices are expressed in the form of outer products of weighted-centered matrices.

VII. What Makes the CD-CKF Distinct From the CD-EKF and the CD-UKF?

In this section, we summarize three unique features that distinguish the proposed CD-CKF from the CD-EKF [25] and CD-UKF [38]:

- *Use of the IT-1.5*: The CD-CKF uses the IT-1.5 whereas current versions of both the CD-EKF and the CD-UKF use the Euler approximation, which is equivalent to the IT-0.5; they are therefore less accurate than the IT-1.5. To the best of the authors' knowledge, this is the first time that the IT-1.5 is utilized for Bayesian filtering in continuous-discrete systems. Because the IT-1.5 captures the state evolution more accurately, the proposed CD-CKF provides a more accurate state estimate and increased filter stability than existing continuous-discrete filters. Moreover, the IT-1.5 is almost as easy to compute as lower-order methods even in high-dimensions. Though it may be tempting to go for an Itô-Taylor expansion of higher-orders (higher than 1.5), our experience has revealed that the use of such an expansion in the CD-CKF does not help improve estimation accuracy significantly.
- *Continuous-Discrete Filtering from the Integration Perspective*: After transforming a continuous-discrete filtering problem into a more familiar stochastic difference equation using the IT-1.5, the Bayesian filter under the Gaussian assumption boils down to the problem of how to compute Gaussian-weighted integrals. As such, the CD-CKF uses the cubature rule to approximate these integrals. On the other hand, to solve the moment (7) and (8) for the first and second moments, the CD-EKF and CD-UKF use adhoc methods- As described in Section III, the CD-EKF uses the first-order Taylor expansion around its latest estimate whereas the CD-UKF is built on the unscented transformation idea.
- *Square-root Continuous-Discrete Filtering*: The CD-CKF and the CD-EKF naturally inherit the square-root filtering feature of classical Kalman filter theory whereas the CD-UKF is not always guaranteed to preserve this capability. In this paper, the square-root version of the CD-CKF is obtained by clever matrix manipulations and ideas from linear algebra such as the least-squares method and matrix triangularization.

VIII. Case Study: Radar Tracker for Coordinated Turns

To illustrate the information processing power of the CD-CKF, we now consider a typical air-traffic-control scenario where the objective is to track the trajectory of an aircraft that executes a maneuver at (nearly) constant speed and turn rate in the horizontal plane. Specifically, the motion in the horizontal plane and the motion in the vertical plane are considered to be decoupled from each other. In the aviation language, this kind of motion is commonly refereed to as (nearly) coordinated turn [8, Sec. 4.2]). Hence, we may write the coordinated turn in the three-dimensional space, subject to fairly "small" noise modeled by independent Brownian motions as shown by

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t))dt + \sqrt{\mathbf{Q}}d\boldsymbol{\beta}(t) \tag{42}$$

where the seven-dimensional state of the aircraft $\mathbf{x} = [\epsilon\ \dot{\epsilon}\ \eta\ \dot{\eta}\ \zeta\ \dot{\zeta}\ \omega]^T$ with ϵ, η and ζ denoting positions and $\dot{\epsilon}$, $\dot{\eta}$ and $\dot{\zeta}$ denoting velocities in the x, y and z Cartesian coordinates, respectively; ω denotes the turn rate; the drift function $\mathbf{f}(\mathbf{x}) = [\dot{\epsilon},\ (-\omega\dot{\eta}),\ \dot{\eta},\ \omega\dot{\epsilon},\ \dot{\zeta},\ 0,\ 0]^T$; the noise term $\boldsymbol{\beta}(t) = [\beta_1(t), \beta_2(t) \ldots \beta_7(t)]^T$ with $\{\beta_i(t)\}, i = 1, 2 \ldots 7$, being all mutually independent standard Brownian motions, accounts for unpredictable modeling errors due to turbulence, wind force, etc.; and finally the diffusion matrix $\mathbf{Q} = \mathrm{diag}([0,\ \sigma_1^2,\ 0,\ \sigma_1^2,\ 0,\ \sigma_2^2,\ \sigma_2^2])$. For the experiment at hand, the radar is located at the origin and equipped to measure the range, r, azimuth angle, θ, and elevation angle, ϕ, at measurement sampling time T. Hence, we write the measurement equation

$$\begin{pmatrix} r_k \\ \theta_k \\ \phi_k \end{pmatrix} = \begin{pmatrix} \sqrt{\epsilon_k^2 + \eta_k^2 + \zeta_k^2} \\ \tan^{-1}\left(\frac{\eta_k}{\epsilon_k}\right) \\ \tan^{-1}\left(\frac{\zeta_k}{\sqrt{\epsilon_k^2+\eta_k^2}}\right) \end{pmatrix} + \mathbf{w}_k$$

where the measurement noise $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ with $\mathbf{R} = \text{diag}([\sigma_r^2,\ \sigma_\theta^2,\ \sigma_\phi^2])$.

Why Is This Tracking Problem Challenging? The considered radar problem of tracking coordinated turns entails a number of desirable features:

- The problem is nonlinear in process and measurement models.
- It is a problem of practical significance that has not been successfully solved in the past; it is our belief that the CD-CKF is capable of solving this tracking problem in the 3-D space for the first time ever.
- From a practical perspective, the importance of the CD-CKF becomes highly significant as the degree of nonlinearity or the dimensionality of the problem at hand is relatively high; in this line of thinking, the considered seven-dimensional estimation problem may well be the highest dimensional problem tackled in the context of target tracking.
- We have a better control over the degree of nonlinearity via the turn rate parameter, ω—as ω increases, the aircraft maneuvers more quickly as illustrated in Fig. 1, which, in turn, makes the tracking job more difficult. Note that motion trajectories of shapes other than circles can be obtained by varying the magnitude and sign of ω.
- Finally, in the later part of this section, we demonstrate that the CD-CKF is indeed far superior to the traditional CD-EKF and the recently published CD-UKF [38].

Data Description: $\sigma_1 = \sqrt{0.2}$; $\sigma_2 = 7 \times 10^{-3}$; $\sigma_r = 50\,\text{m}$; $\sigma_\theta = 0.1^\circ$; $\sigma_\phi = 0.1^\circ$; and the true initial state $\mathbf{x}_0 = [1000\,\text{m},\, 0\,\text{ms}^{-1},\, 2650\,\text{m},\, 150\,\text{ms}^{-1},\, 200\,\text{m},\, 0\,\text{ms}^{-1},\, \omega^\circ/\text{s}]^T$.

As aforementioned, we use the notation $\mathbf{x}_k^j$ to denote $\mathbf{x}(t)$ at time $t = t_k + j\delta$, where $1 \le j \le m$ and $\delta = T/m$. Applying the IT-1.5 to (42), we get the stochastic difference equation

$$\mathbf{x}_k^{(j+1)} = \mathbf{f}_d(\mathbf{x}_k^j) + \sqrt{\mathbf{Q}}\mathbf{w} + (\mathbb{L}\mathbf{f}(\mathbf{x}_k^j))\mathbf{y} \tag{43}$$

where

$$\mathbf{f}_d(\mathbf{x}) = \begin{pmatrix} \epsilon + \delta\dot{\epsilon} - \frac{\delta^2}{2}\omega\dot{\eta} \\ \dot{\epsilon} - \delta\omega\dot{\eta} - \frac{\delta^2}{2}\omega^2\dot{\epsilon} \\ \eta + \delta\dot{\eta} + \frac{\delta^2}{2}\omega\dot{\epsilon} \\ \dot{\eta} + \delta\omega\dot{\epsilon} - \frac{\delta^2}{2}\omega^2\dot{\eta} \\ \zeta + \delta\dot{\zeta} \\ \dot{\zeta} \\ \omega \end{pmatrix}$$

$$\mathbb{L}\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 0 & \sigma_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sigma_1\omega & 0 & 0 & -\sigma_2\dot{\eta} \\ 0 & 0 & 0 & \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_1\omega & 0 & 0 & 0 & 0 & -\sigma_2\dot{\epsilon} \\ 0 & 0 & 0 & 0 & 0 & \sigma_1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For the purpose of generating independent aircraft trajectories, we used the IT-1.5 with $m = 1000$ time-steps/sampling interval.

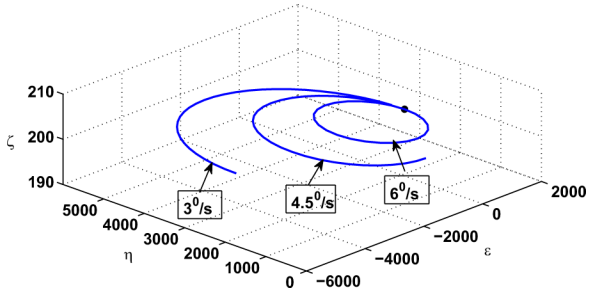


Fig. 1. Change of motion trajectories with varying turn rate, ω, when simulated for 60 s (The filled circle denotes the starting point).

We are now ready to compare the performance of the CD-CKF against the CD-EKF and the CD-UKF. All three approximate Bayesian filters were initialized with the same initial condition for each run. The initial state density was assumed to be Gaussian and the two-point differencing method, which uses the first two measurements to estimate the states' statistics, was adopted for initialization [8, Sec. 5.5.3].

Performance Metric. To compare the three filter performances, we use the accumulative root-mean square error (RMSE) of the position, velocity and turn rate. For example, we define the accumulative RMSE in position as

$$\sqrt{\frac{1}{NK}\sum_{n=1}^{N}\sum_{k=1}^{K}\left((\epsilon_k^n - \hat{\epsilon}_k^n)^2 + (\eta_k^n - \hat{\eta}_k^n)^2 + (\zeta_k^n - \hat{\zeta}_k^n)^2\right)}$$

where $(\epsilon_k^n, \eta_k^n, \zeta_k^n)$ and $(\hat{\epsilon}_k^n, \hat{\eta}_k^n, \hat{\zeta}_k^n)$ are the true and estimated positions at time index k in the nth Monte Carlo run. In a similar manner, we may also write formulas of the accumulative RMSE in velocity and turn rate. For a fair comparison, a total of $N = 100$ independent Monte Carlo runs were made with measurements collected for time intervals of 210s. However, the accumulative RMSEs were computed for a period of 60s-210s.

In addition to turn rate, the accumulative RMSE is dependant on the measurement sampling interval (T) and the number of time update iterations/sampling interval (m, where $m = 2^0, 2^1, \ldots 2^6$ in our experiment). For example, as T increases, the measurements that contain information about the state are available only after long 'waiting' periods; if the time update of a tracking filter is capable of capturing the state evolution more accurately, it will continue to satisfactorily run even for a substantially longer T without divergence. In fact, to alleviate the detrimental performance due to increased T, we may increase m. In our experiment, we also wanted to see how m helps reduce estimation errors and divergence rates of a tracking filter.

Observations. Referring to the results displayed in Figs. 2–5, we now make the following observations:

- Fig. 2 was obtained by fixing sampling interval T at $T = 2\,\text{s}$ and varying turn rate ω from $\omega = 3^\circ/\text{s}$ to $\omega = 6^\circ/\text{s}$ as depicted in Fig. 1. When m increases, the performances of the CD-EKF and the CD-UKF tend to improve as expected. On the other hand, from Fig. 2 we see that the CD-CKF has already approached a steady state and therefore, unlike the CD-EKF and CD-UKF, it does not require any further time-update iterations. The CD-CKF uses the approximate discretized process equation, which, with small m, seems

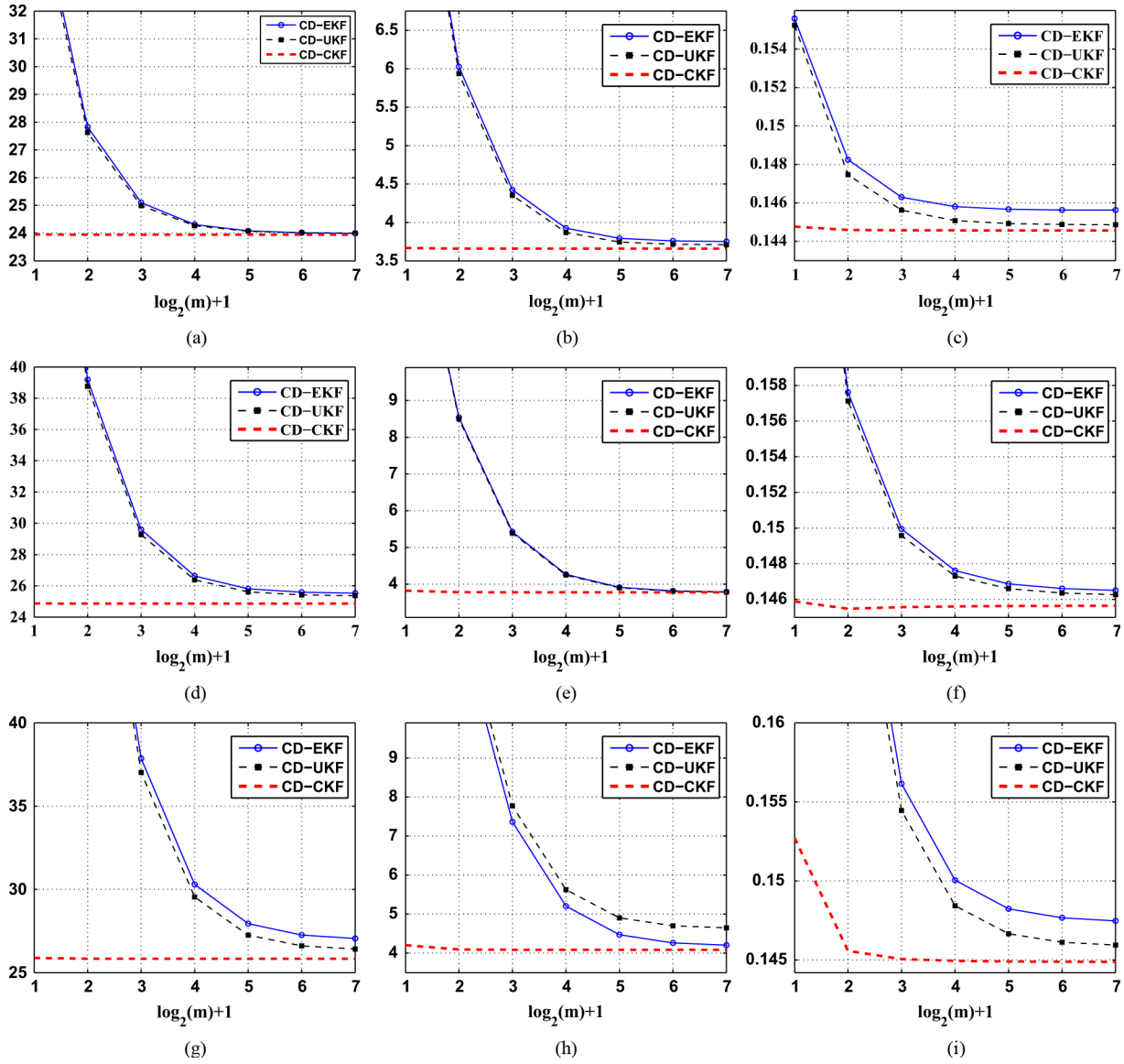


Fig. 2. Accumulative RMSE $\forall$s. $(\log_2(m)+1)$, where m is the number of time-update iterations/sampling interval, for a fixed sampling interval $T = 2$ s and varying turn rates. (a)–(c) $\omega = 3^\circ/\text{s}$; (d)–(f) $\omega = 4.5^\circ/\text{s}$; (g)–(i) $\omega = 6^\circ/\text{s}$. (a) RMSE in Position. (b) RMSE in Velocity. (c) RMSE in Turn Rate. (d) RMSE in Position. (e) RMSE in Velocity. (f) RMSE in Turn Rate. (g) RMSE in Position. (h) RMSE in Velocity. (i) RMSE in Turn Rate.

sufficient to closely capture the underlying state process responsible for the trajectory generation. For insight, consider the case where ω increases and T is fixed; for this scenario, we see that unlike the CD- EKF, the CD-UKF survives in many cases but at the expense of a large number of steps m to get closer to the CD-CKF as shown in Fig. 2(a), (d), and (g) (see also Table I).

- In Figs. 3–5, we repeated the experiments for different sampling intervals, namely, $T = 4$ s, $T = 6$ s, and $T = 8$ s, respectively, while ω was varied from $\omega = 3^\circ/\text{s}$ to $\omega = 6^\circ/\text{s}$. In all these cases, the CD-CKF outperforms the CD-EKF and CD-UKF. Its importance becomes more and more obvious as T increases (compare, for example, Figs. 2(g)–(i) and 5(g)–(i)). Also, as can be seen from Fig. 5(c), (f), and (i), the CD-CKF improves its performance in tracking the turn rate as m increases and saturates after $m = 8$. However, if we are interested only in accurate position and velocity estimates, it will be sufficient to use the CD-CKF with at most two time update iterations/sampling interval (i.e., $m \leq 4$).
- Divergence is declared when the positional error $\sqrt{\left((\epsilon_k^n - \hat{\epsilon}_k^n)^2 + (\eta_k^n - \hat{\eta}_k^n)^2 + (\zeta_k^n - \hat{\zeta}_k^n)^2\right)}$ is greater than 500 m. The results are shown in Tables I–III. All three filters do not diverge when the tracking scenario assumes

TABLE I
DIVERGENCE TABLE FOR A FIXED $T = 2$ s: THREE ARGUMENTS WITHIN EACH PARENTHESIS GOING FROM LEFT TO RIGHT DENOTE THE NUMBER OF DIVERGENCES OUT OF 100 EXPERIMENTS FOR THE CD-EKF, CD-UKF, AND CD-CKF, RESPECTIVELY

$T = 2$ s	$m = 1$	$m = 2$	$m = 4$	$m = 8$	$m = 16$	$m = 32$	$m = 64$
$\omega = 6^\circ/\text{s}$	(11,0,0)	(9,0,0)	(9,0,0)	(9,0,0)	(9,0,0)	(8,0,0)	(8,0,0)

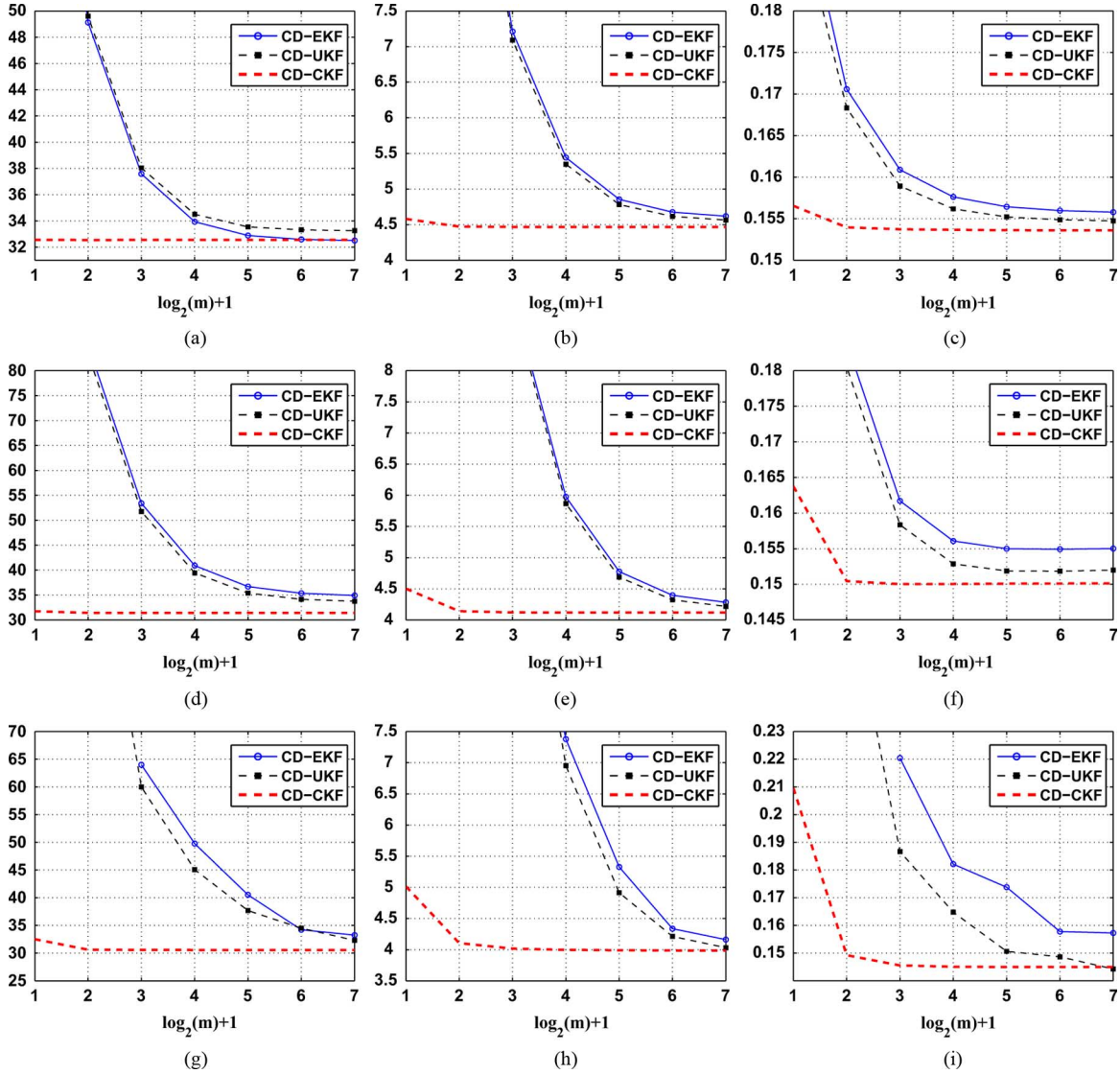


Fig. 3. Accumulative RMSE $\forall$s. $(\log_2(m)+1)$, where m is the number of time-update iterations/sampling interval, for a fixed sampling interval $T = 4$ s and varying turn rates. (a)–(c) $\omega = 3^\circ/\mathrm{s}$; (d)–(f) $\omega = 4.5^\circ\mathrm{s}$; (g)–(i) $\omega = 6^\circ\mathrm{s}$. (a) RMSE in Position. (b) RMSE in Velocity. (c) RMSE in Turn Rate. (d) RMSE in Position. (e) RMSE in Velocity. (f) RMSE in Turn Rate. (g) RMSE in Position. (h) RMSE in Velocity. (i) RMSE in Turn Rate.

ω to be $3^\circ/\mathrm{s}$ and T equals $2s$ or $4s$; similarly, they do not diverge when ω and T are set at $4.5^\circ/\mathrm{s}$ and 2s. For this reason, we do not present them in the tables. Again, we consider the case where T is fixed, and ω increases- in this case, as ω increases, the CD-EKF and CD-UKF diverge more frequently. The CD-EKF and CD-UKF completely break down when ω and T are fixed at $6^\circ/\mathrm{s}$ and 6s, respectively, irrespective of m (see Table III). One of the striking features of the CD-CKF lies in its reliability despite the fact that its computational complexity is comparable to that of the CD-EKF and CD-UKF for fixed m. From Table IV, we see that the CD-CKF has reached its own breakdown condition- it diverges in more than half of Monte Carlo trials when ω and T are fixed at $6^\circ/\mathrm{s}$ and 8s, respectively. However, it is important to note that the underlying target parameters for this condition are far more severe than those experienced with the CD-EKF and CD-UKF.

More specifically, of the three approximate Bayesian filters, the CD-CKF is the most accurate and reliable followed by the CD-UKF and then the CD-EKF. We may therefore say that the CD-CKF is the method of choice for challenging radar tracking problems, exemplified by the case study chosen herein.

Remark: The considered tracking problem falls under the simple class of single-target single-sensor target tracking without considering complications such as false alarms and misdetections. To handle such complications in a more complicated multitarget multisensor environment, the joint probabilistic data association (JPDA) filter [7], [12], the multihypothesis tracking (MHT) filter [7], [12] and the probability hypothesis density (PHD) filter [34], [42], [43] are often used. For tracking problems defined by a nonlinear continuous-discrete state-space model, the continuous-discrete version of these filters can be easily derived by adopting ideas from the CD-CKF. For example, consider the JPDA filter, a classical filter for tracking a known number of targets in clutter. It uses

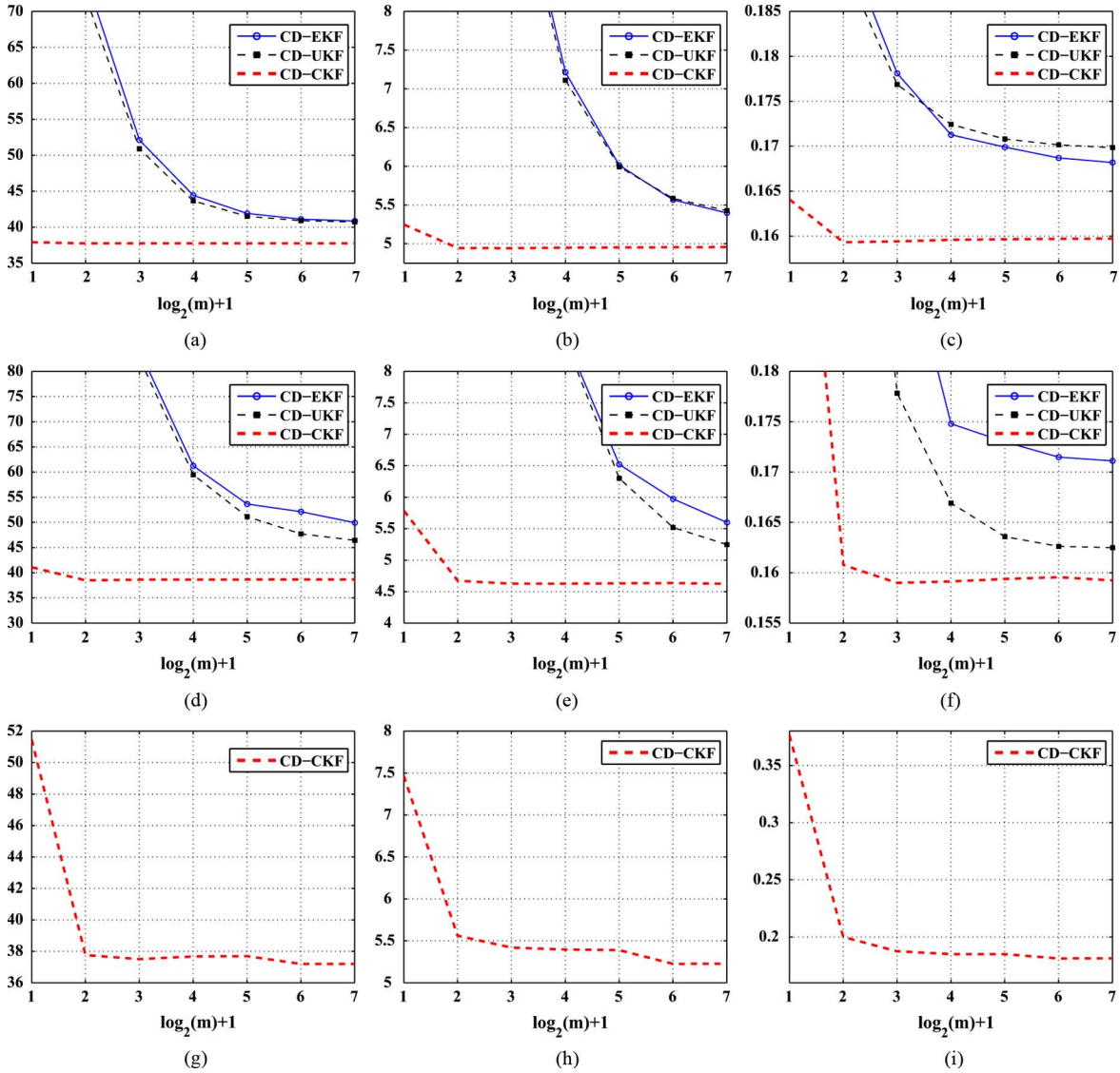


Fig. 4. Accumulative RMSE $\forall$s. $(\log_2(m)+1)$, where m is the number of time-update iterations/sampling interval, for a fixed sampling interval $T = 6$ s and varying turn rates. (a)–(c) $\omega = 3^\circ/\text{s}$; (d)–(f) $\omega = 4.5^\circ/\text{s}$; (g)–(i) $\omega = 6^\circ/\text{s}$. (a) RMSE in Position. (b) RMSE in Velocity. (c) RMSE in Turn Rate. (d) RMSE in Position. (e) RMSE in Velocity. (f) RMSE in Turn Rate. (g) RMSE in Position. (h) RMSE in Velocity. (i) RMSE in Turn Rate.

measurements weighted by their association probabilities. The continuous-discrete version of the JPDA may use the time-update of the CD-CKF in order to propagate state statistics in time; to fuse measurements with the predicted density, it may utilize a number of steps of the measurement update of the CD-CKF, e.g., to compute the predicted measurements and the filter gain, which is the function of the innovations covariances and cross covariances.

IX. Conclusion

In this paper, we have expanded the CKF to extend its applicability to continuous-discrete systems and named the resulting algorithm the "continuous-discrete cubature Kalman filter (CD-CKF)." For its time-update, we have employed two different numerical integration tools:

- Discretization of the continuous-time process equation modeled by a stochastic differential equation based on the *Itô-Taylor expansion of order 1.5*; this tool transforms the process equation into a familiar stochastic difference equation.
- Computing the second-order statistics of various conditional densities using the *third-degree cubature rule*; this second tool accurately computes Gaussian-weighted integrals whose integrands are well-behaved nonlinear functions.

For improved numerical stability, we also developed a square-root version of the CD-CKF.

To demonstrate the computing power of the square-root CD-CKF in terms of reliability and accuracy, we tested it in tackling a seven-dimensional radar tracking problem, where an aircraft executes a coordinated turn. The experiment also involved comparing performance of the CD-CKF with the CD-EKF and CD-UKF with all three approximate Bayesian filters having comparable computational complexities. The experiment was repeated under two different conditions: varying turn rate for fixed measurement sampling time interval

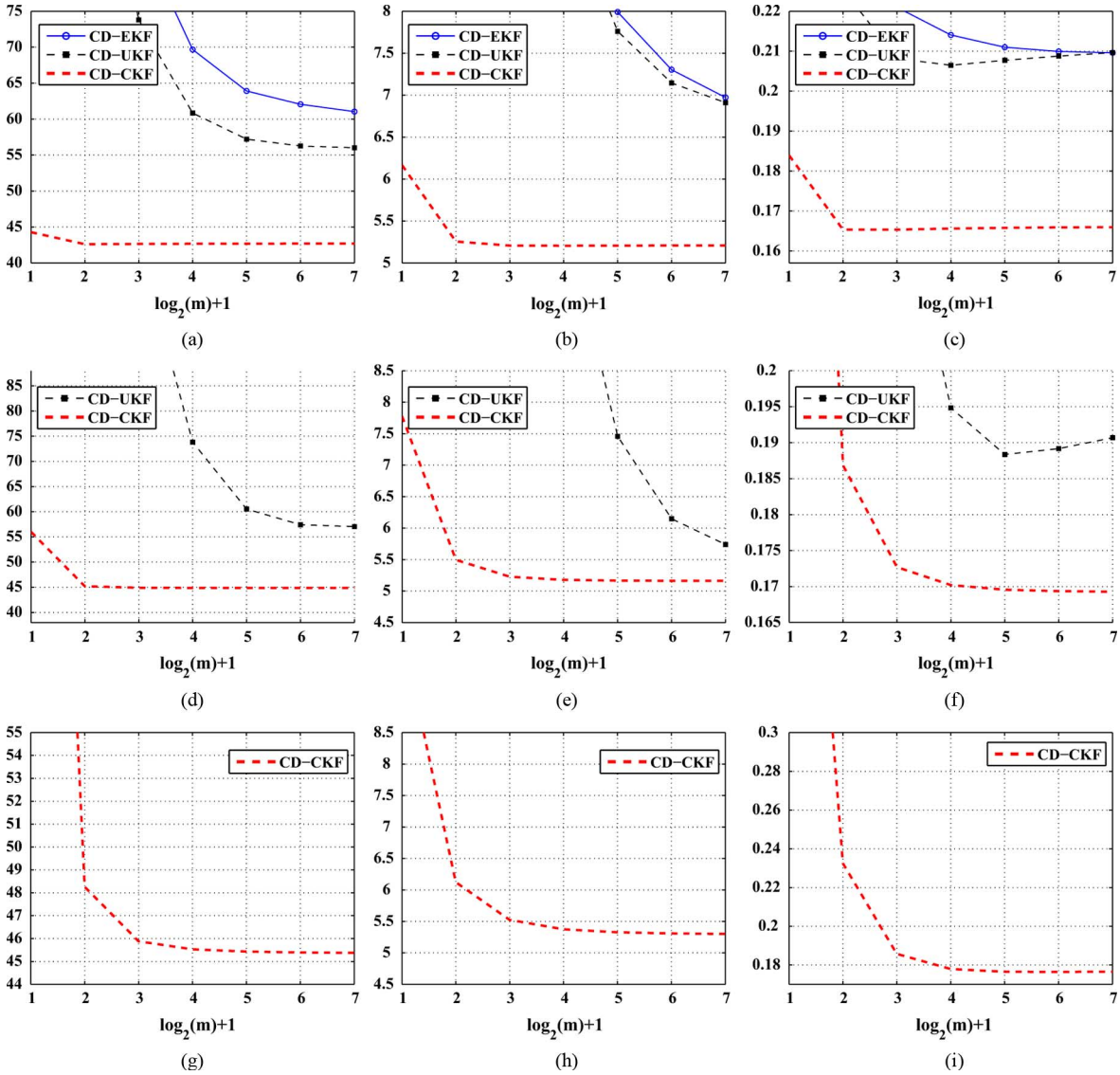


Fig. 5. Accumulative RMSE $\forall$s. $(\log_2(m)+1)$, where m is the number of time-update iterations/sampling interval, for a fixed sampling interval $T = 8$ s and varying turn rates. (a)–(c) $\omega = 3^\circ/\text{s}$; (d)–(f) $\omega = 4.5^\circ/\text{s}$; (g)–(i) $\omega = 6^\circ/\text{s}$. (a) RMSE in Position. (b) RMSE in Velocity. (c) RMSE in Turn Rate. (d) RMSE in Position. (e) RMSE in Velocity. (f) RMSE in Turn Rate. (g) RMSE in Position. (h) RMSE in Velocity. (i) RMSE in Turn Rate.

and varying measurement sampling time interval for fixed turn rate. The conclusions to be drawn from the experiments are summarized as follows: Among the three approximate Bayesian filters considered herein, the CD-CKF is the most accurate and reliable, followed by the CD-UKF and then the CD-EKF. The CD-CKF becomes the method of choice for the air-traffic-control problem considered in this paper as the degree of nonlinearity increases or the measurements are available in substantially long time intervals. This improved performance of the CD-CKF compared to the CD-UKF is attributed to the basic differences between these two filters as highlighted in Section VII.

Note that this paper considers different algorithms as the number of steps, m, varies; however, given m, the sampling-time is equally divided to yield a fixed step-size time update. In contrast, for improved accuracy and fast convergence, the authors of [41] implement the CD-UKF that uses an adaptive step-size Runge-Kutta method as a tool for numerical integration. It will be an interesting future research topic to undertake a detailed comparative study on the feasibility of various adaptive step-size numerical tools for the development of an approximate continuous-discrete filter.

Appendix A
Continuous-Discrete Cubature Filter

m-step Time-Update: Initialize j to be zero. That is, $\hat{\mathbf{x}}^0_{k|k} = \hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}^0_{k|k} = \mathbf{P}_{k|k}$.

1) Factorize

$$\mathbf{P}^j_{k|k} = (\mathbf{S}^j_{k|k})(\mathbf{S}^j_{k|k})^T.$$

2) Evaluate the cubature points $(i = 1, 2, \ldots, 2n)$

$$X^j_{i,k|k} = \mathbf{S}^j_{k|k}\xi_i + \hat{\mathbf{x}}^j_{k|k}.$$

where $\{\xi_i\}$ refers to the set of cubature points as defined in Section V-A.

TABLE II
DIVERGENCE TABLE FOR A FIXED $T = 4$ s: THREE ARGUMENTS WITHIN EACH PARENTHESIS GOING FROM LEFT TO RIGHT DENOTE THE NUMBER OF DIVERGENCES OUT OF 100 EXPERIMENTS FOR THE CD-EKF, CD-UKF, AND CD-CKF, RESPECTIVELY

$T = 4$ s	$m = 1$	$m = 2$	$m = 4$	$m = 8$	$m = 16$	$m = 32$	$m = 64$
$\omega = 4.5$ °/s	(15,3,0)	(10,3,0)	(9,2,0)	(9,1,0)	(9,0,0)	(8,0,0)	(8,0,0)
$\omega = 6$ °/s	(100,100,21)	(100,90,29)	(97,89,21)	(96,87,27)	(96,88,21)	(93,84,28)	(93,85,27)

TABLE III
DIVERGENCE TABLE FOR A FIXED $T = 6$ s: THREE ARGUMENTS WITHIN EACH PARENTHESIS GOING FROM LEFT TO RIGHT DENOTE THE NUMBER OF DIVERGENCES OUT OF 100 EXPERIMENTS FOR THE CD-EKF, CD-UKF, AND CD-CKF, RESPECTIVELY

$T = 6$ s	$m = 2$	$m = 4$	$m = 8$	$m = 16$	$m = 32$	$m = 64$
$\omega = 3$ °/s	(16,9,0)	(12,3,0)	(12,1,0)	(12,1,0)	(11,1,0)	(11,1,0)
$\omega = 4.5$ °/s	(99,48,4)	(90,33,4)	(88,33,4)	(88,32,4)	(88,31,4)	(79,30,0)
$\omega = 6$ °/s	(100,100,29)	(100,100,28)	(100,100,28)	(100,100,27)	(100,100,25)	(100,100,27)

3) Evaluate the propagated cubature point set ($i = 1, 2, \ldots, 2n$)

$$X^{*(j+1)}_{i,k|k} = \mathbf{f}_d(X^j_{i,k|k}, kT + j\delta)$$

4) Estimate the predicted state

$$\hat{\mathbf{x}}^{j+1}_{k|k} = \frac{1}{2n}\sum_{i=1}^{2n} X^{*(j+1)}_{i,k|k}.$$

5) Estimate the predicted error covariance matrix

$$\begin{aligned}\mathbf{P}^{j+1}_{k|k} = {} & \left(\mathsf{X}^{*(j+1)}_{k|k}\right)\left(\mathsf{X}^{*(j+1)}_{k|k}\right)^T + \frac{\delta^2}{2}\Big[\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, \right. \\ & \left. kT + j\delta\right)\mathbf{Q}^T + \mathbf{Q}\left(\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, kT + j\delta\right)\right)^T\Big] \\ & + \frac{\delta^3}{3}\left(\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, kT + j\delta\right)\right) \\ & \times \left(\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, kT + j\delta\right)\right)^T + \delta\mathbf{Q}\end{aligned}$$

where the weighted-centered matrix

$$\mathsf{X}^{*(j+1)}_{k|k} = \frac{1}{\sqrt{2n}}\left[X^{*(j+1)}_{1,k|k} - \hat{\mathbf{x}}^{j+1}_{k|k} \quad X^{*(j+1)}_{2,k|k} - \hat{\mathbf{x}}^{j+1}_{k|k} \quad \cdots \quad X^{*(j+1)}_{2n,k|k} - \hat{\mathbf{x}}^{j+1}_{k|k}\right].$$

6) Increase j by one and repeat the steps 1)–5) until j reaches m (that is, until time t_{k+1}).

Measurement-Update

1) Factorize

$$\mathbf{P}_{k+1|k} = \mathbf{S}_{k+1|k}\mathbf{S}^T_{k+1|k}.$$

2) Evaluate the cubature points ($i = 1, 2, \ldots, 2n$)

$$X_{i,k+1|k} = \mathbf{S}_{k+1|k}\xi_i + \hat{\mathbf{x}}_{k+1|k}.$$

3) Evaluate the propagated cubature points ($i = 1, 2, \ldots, 2n$)

$$Z_{i,k+1|k} = \mathbf{h}(X_{i,k+1|k}, k+1).$$

4) Estimate the predicted measurement

$$\hat{\mathbf{z}}_{k+1|k} = \frac{1}{2n}\sum_{i=1}^{2n} Z_{i,k+1|k}.$$

5) Estimate the innovations covariance matrix

$$\mathbf{P}_{zz,k+1|k} = \mathsf{Z}_{k+1|k}\mathsf{Z}^T_{k+1|k} + \mathbf{R}_{k+1},$$

where the weighted-centered matrix

$$\mathsf{Z}_{k+1|k} = \frac{1}{\sqrt{2n}}\left[Z_{1,k+1|k} - \hat{\mathbf{z}}_{k+1|k} \quad Z_{2,k+1|k} - \hat{\mathbf{z}}_{k+1|k} \quad \cdots \quad Z_{2n,k+1|k} - \hat{\mathbf{z}}_{k+1|k}\right].$$

6) Estimate the cross-covariance matrix

$$\mathbf{P}_{xz,k+1|k} = \mathsf{X}_{k+1|k}\mathsf{Z}^T_{k+1|k},$$

where the weighted-centered matrix

$$\mathsf{X}_{k+1|k} = \frac{1}{\sqrt{2n}}\left[X_{1,k+1|k} - \hat{\mathbf{x}}_{k+1|k} \quad X_{2,k+1|k} - \hat{\mathbf{x}}_{k+1|k} \quad \cdots \quad X_{2n,k+1|k} - \hat{\mathbf{x}}_{k+1|k}\right].$$

7) Estimate the continuous-discrete cubature gain

$$\mathbf{W}_{k+1} = \mathbf{P}_{xz,k+1|k}\mathbf{P}^{-1}_{zz,k+1|k}.$$

8) Estimate the updated state

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{W}_{k+1}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}).$$

9) Estimate the corresponding error covariance matrix

$$\mathbf{P}_{k+1|k+1} = \mathbf{P}_{k+1|k} - \mathbf{W}_{k+1}\mathbf{P}_{zz,k+1|k}\mathbf{W}^T_{k+1}.$$

APPENDIX B
SQUARE-ROOT CD-CKF

Below, we summarize the square-root CD-CKF writing only those steps when they differ from the CD-CKF.

TABLE IV
DIVERGENCE TABLE FOR A FIXED $T = 8$ s: THREE ARGUMENTS WITHIN EACH PARENTHESIS GOING FROM LEFT TO RIGHT DENOTE THE NUMBER OF DIVERGENCES OUT OF 100 EXPERIMENTS FOR THE CD-EKF, CD-UKF, AND CD-CKF, RESPECTIVELY

$T = 8$s	$m = 2$	$m = 4$	$m = 8$	$m = 16$	$m = 32$	$m = 64$
$\omega = 3\,^\circ$/s	(43,9,0)	(27,8,0)	(24,4,0)	(23,3,0)	(24,3,0)	(25,1,0)
$\omega = 4.5\,^\circ$/s	(100,78,4)	(100,63,4)	(100,53,4)	(100,52,4)	(100,51,4)	(100,49,0)
$\omega = 6\,^\circ$/s	(100,100,59)	(100,100,58)	(100,100,58)	(100,100,55)	(100,100,57)	(100,100,57)

TABLE V

Time Update for the Case of $m = 1$			
Steps	$+/-$	$\times/\div$	$\sqrt{}$
1. $X_{i,k\|k} = \mathbf{S}_{k\|k}\xi_i + \hat{\mathbf{x}}_{k\|k}$ $(i = 1, \dots 2n)$	$n^3 + n^2$	$n^3 + n^2$	$-$
2. $\hat{\mathbf{x}}_{k+1\|k} = \frac{1}{2n}\sum_{i=1}^{2n} X^*_{i,k+1\|k}$	$2n^2 - n$	n	$-$
3. $\mathsf{X}^*_{k+1\|k} = \frac{1}{\sqrt{2n}}[X^*_{1,k+1\|k} - \dots - \hat{\mathbf{x}}_{k+1\|k}]$	$2n^2$	$2n^2$	$-$
4. $\mathbf{S}_{k+1\|k} = \mathbf{Tria}([\mathsf{X}^*_{k+1\|k} \quad \dots])$	$3n^3 - \frac{1}{2}n^2 - \frac{1}{2}n$	$3n^3 + 3n^2$	n
Total cost	$8n^3 + \frac{21}{2}n^2 + \frac{1}{2}n$		

Measurement Update			
Steps	$+/-$	$\times/\div$	$\sqrt{}$
1. $X_{i,k\|k} = \mathbf{S}_{k\|k}\xi_i + \hat{\mathbf{x}}_{k\|k}$ $(i = 1, \dots 2n)$	$n^3 + n^2$	$n^3 + n^2$	$-$
2. $\hat{\mathbf{z}}_{k+1\|k} = \frac{1}{2n}\sum_{i=1}^{2n} Z_{i,k+1\|k}$	$(2n-1)d$	$-$	$-$
3. $\mathsf{X}_{k+1\|k} = \frac{1}{\sqrt{2n}}[X_{1,k+1\|k} - \dots - \hat{\mathbf{x}}_{k+1\|k}]$	$2n^2$	$2n^2$	$-$
$\mathsf{Z}_{k+1\|k} = \frac{1}{\sqrt{2n}}[Z_{1,k+1\|k} - \dots - \hat{\mathbf{z}}_{k+1\|k}]$	$2nd$	$2nd$	$-$
4. $\mathbf{Tria}\begin{pmatrix} \mathsf{Z}_{k+1\|k} & \mathbf{S}_{R,k+1} \\ \mathsf{X}_{k+1\|k} & \mathbf{O} \end{pmatrix}$	$2n^3 + (5d - \frac{1}{2})n^2$ $+(4d^2 - d - \frac{1}{2})n$ $+d^3 - \frac{1}{2}d^2 - \frac{1}{2}d$	$2n^3 + (5d + 2)n^2$ $+(4d^2 + 3d)n$ $+d^3 + d^2$	$n + d$
5. $\mathbf{W}_{k+1} = \mathbf{T}_{21}/\mathbf{T}_{11}$	$\frac{1}{2}d(d-1)n + \frac{1}{6}d^3$ $-\frac{1}{2}d^2 + \frac{1}{3}d$	$\frac{1}{2}d(d+1)n + \frac{1}{6}d^3$ $+\frac{1}{2}d^2 - \frac{2}{3}d$	d
6. $\hat{\mathbf{x}}_{k+1\|k+1} = \hat{\mathbf{x}}_{k+1\|k} + \mathbf{W}_{k+1}\dots$	$d(n+1)$	nd	$-$
Total cost	$6n^3 + (10d + \frac{15}{2})n^2 + (9d^2 + 10d + \frac{1}{2})n$ $+\frac{7}{3}d^3 + \frac{1}{2}d^2 + \frac{7}{6}d$		

m-step Time-Update

1) Skip the factorization step 1) because the square-root of the error covariance matrix $\mathbf{S}^j_{k|k}$ is available. Compute steps 2)-4).
2) Estimate the square-root factor of the predicted error covariance matrix

$$\mathbf{S}^{j+1}_{k|k} = \mathbf{Tria}\left(\left[\mathsf{X}^{*(j+1)}_{k|k} \quad \sqrt{\delta}\left(\sqrt{\mathbf{Q}} + \tfrac{\delta}{2}\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, kT + j\delta\right)\right) \quad \sqrt{\tfrac{\delta^3}{12}}\mathbb{L}\mathbf{f}\left(\hat{\mathbf{x}}^j_{k|k}, kT + j\delta\right)\right]\right).$$

3) Increase j by one and repeat the above steps 1)-2) until j reaches m (that is, until time t_{k+1}).

Measurement-Update

1) Evaluate the cubature points $(i = 1, 2, \dots 2n)$

$$X_{i,k+1|k} = \mathbf{S}_{k+1|k}\xi_i + \hat{\mathbf{x}}_{k+1|k}.$$

2) Evaluate the propagated cubature points $(i = 1, 2, \dots 2n)$

$$Z_{i,k+1|k} = \mathbf{h}(X_{i,k+1|k}, k+1).$$

3) Estimate the predicted measurement

$$\hat{\mathbf{z}}_{k+1|k} = \frac{1}{2n}\sum_{i=1}^{2n} Z_{i,k+1|k}.$$

4) Compute the matrices $\mathbf{T}_{11}$, $\mathbf{T}_{21}$, and $\mathbf{T}_{22}$ using the triangularization algorithm

$$\begin{pmatrix} \mathbf{T}_{11} & \mathbf{O} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{pmatrix} = \mathbf{Tria}\begin{pmatrix} \mathsf{Z}_{k+1|k} & \mathbf{S}_{R,k+1} \\ \mathsf{X}_{k+1|k} & \mathbf{O} \end{pmatrix}$$

where the weighted-centered matrices present on the RHS of the above equation are given in Steps 5) and 6) of the measurement update of the CD-CKF algorithm.
5) Estimate the continuous-discrete cubature gain

$$\mathbf{W}_{k+1} = \mathbf{T}_{21}/\mathbf{T}_{11}.$$

6) Estimate the updated state

$$\hat{\mathbf{x}}_{k+1|k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{W}_{k+1}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1|k}).$$

7) The square-root factor of the corresponding error covariance matrix is given by

$$\mathbf{S}_{k+1|k+1} = \mathbf{T}_{22}.$$

APPENDIX C
COMPUTATIONAL COST OF THE SQUARE-ROOT CD-CKF IN FLOPS

See Table V.

ACKNOWLEDGMENT

The authors wish to express their gratitude to the associate editor and the anonymous reviewers for their constructive comments, which shaped the paper into its final form.

REFERENCES

[1] I. Arasaratnam and S. Haykin, "Cubature Kalman filters," *IEEE Trans. Autom. Control*, vol. 54, pp. 1254–1269, Jun. 2009.
[2] I. Arasaratnam and S. Haykin, "Square-root quadrature Kalman filtering," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2589–2593, Jun. 2008.
[3] I. Arasaratnam, "Cubature Kalman filtering: Theory and applications," Ph.D., Dep. Elect. Comput. Eng., McMaster Univ., Ontario, Canada, 2009.
[4] K. J. Aström, *Introduction to Stochastic Control Theory*. New York: Academic, 1970.
[5] M. Athans, R. P. Wishner, and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noise measurements," *IEEE Trans. Autom. Control*, vol. 13, pp. 504–514, 1968.
[6] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*. Boston, MA: Birkhauser, 2003.
[7] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*. San Diego, CA: Academic, 1988.
[8] Y. B. Shalom, X. R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. New York: Wiley, 2001.
[9] R. Beard, J. Kennedy, J. Gunther, J. Lawton, and W. Stirling, "Nonlinear projection filter based on Galerkin approximation," *J. Guid., Contr. Dyn.*, vol. 22, no. 2, Mar. 1992.
[10] V. E. Beneš, "Exact finite-dimensional filters with certain diffusion nonlinear drift," *Stochastics*, vol. 5, pp. 65–92, 1981.
[11] C. Berzuini, N. G. Best, W. Gilks, and C. Larizza, "Dynamic conditional independent models and Markov chain Monte Carlo methods," *J. Amer. Stat. Assoc.*, vol. 92, pp. 1403–1412, 1997.
[12] S. Blackman, *Multiple Target Tracking With Radar Applications*. Norwood, MA: Artech House, 1986.
[13] A. Budhiraja, L. Chen, and C. Lee, "A survey of numerical methods for nonlinear filtering problems," *Physica D*, vol. 230, pp. 27–36, 2007.
[14] F. Campillo, F. Cérou, and F. L. Gland, Particle and Cell Approximations for Nonlinear Filtering, Int. Pub. INRIA Tech. Rep. 2567.
[15] R. Cools, "Constructing cubature formulas: The science behind the art," *Acta Numerica 6*, pp. 1–54, 1997.
[16] F. Daum, "Exact finite-dimensional nonlinear filters," *IEEE Trans. Autom. Control*, vol. 31, pp. 616–622, 1986.
[17] F. E. Daum, "Nonlinear filters: beyond the Kalman filter," *IEEE Aerosp. Elect. Syst. Mag.*, Aug. 2005.
[18] F. E. Daum and M. Krichman, "Meshfree adjoint methods for nonlinear filtering," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, Mar. 2006.
[19] A. T. Fuller, "Analysis of nonlinear stochastic systems by means of the Fokker-Planck equation," in *Proc. Int. J. Contr. 9*, 1969, pp. 603–655.
[20] M. Grewal and A. Andrews, *Kalman Filtering: Theory and Practice Using Matlab*, 2nd ed. New York: Wiley, 2001.
[21] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global Positioning Systems, Inertial Navigation and Integration*. New York: Wiley, 2001.
[22] Y. C. Ho and R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. Autom. Control*, vol. AC-9, pp. 333–339, Oct. 1964.
[23] J. D. Hoffman, *Numerical Methods for Engineers and Scientists*. New York: McGraw-Hill, 1992.
[24] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.
[25] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
[26] S. J. Julier, J. K. Ulhmann, and H. F. Durrant-Whyte, "A new method for nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, pp. 472–482, Mar. 2000.
[27] P. Kaminski, A. Bryson, and S. Schmidt, "Discrete sqaure root filtering: A survey of current techniques," *IEEE Trans. Autom. Control*, vol. AC-16, Dec. 1971.
[28] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Eng.*, vol. 83, pp. 95–108, Mar. 1961.
[29] K. Kastella, "Finite difference methods for nonlinear filtering and automatic target recoginition," in *Multitarget/Multisensor Tracking*, Y. B. Shalom and W. D. Blair, Eds. Norwood, MA: Artech House, 2000, vol. III.
[30] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*. Berlin, Germany: Springer, 1999.
[31] H. J. Kushner and A. S. Budhiraja, "A nonlinear filtering algorithm based on an approximation of the conditional distribution," *IEEE Trans. Autom. Control*, vol. 45, no. 3, Mar. 2000.
[32] H. J. Kushner, "Dynamical equations for optimal nonlinear filter," *J. Differen. Equat.*, vol. 3, no. 2, pp. 179–190, 1967.
[33] X. R. Li and V. P. Jilkov, "A survey of manuevering target tracking: Part II: Ballistic target models," *Proc. SPIE*, vol. 4473, pp. 559–581, 2001.
[34] R. Mahler, *Statistical Multisource Multitarget Information Fusion*. Norwood, MA: Artech House, 2007.
[35] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*. New York: Springer, 2003.
[36] P. A. Raviart, "An analysis of particle methods," in *Numerical Methods in Fluid Dynamics, Lecture Notes Math*, F. Brezzi, Ed. New York: Springer-Verlag, 1985, vol. 1127, pp. 243–324.
[37] H. Risken, *The Fokker-Planck Equation: Methods of Solutions and Applications*, 2nd ed. New York: Springer-Verlag, 1989.
[38] S. Särkkä, "On unscented Kalman filtering for state estimation of continuos-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 52, no. 9, pp. 1631–1641, Sep. 2007.
[39] H. Singer, Continuous-Discrete Unscented Kalman Filtering, Diskussionsbeiträge FachbereichWirtschaftswissenschaft 384, FernUniversität in Hagen [Online]. Available: http:www.fernuni-hagen.de/FB-WIWI/forschung/beitraege/pdf/db384.pdf
[40] A. H. Stroud, *Approximate Calculation of Multiple Integrals*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
[41] B. O. S. Teixeira, M. A. Santillo, R. S. Erwin, and D. S. Bernstein, "Spacecraft tracking using sampled-data Kalman filters," *IEEE Contr. Syst. Mag.*, vol. 28, no. 4, pp. 78–94, Aug. 2008.
[42] B. N. Vo and W. K. Ma, "The Gaussian-mixture probability hypothesis density filter," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.
[43] B. T. Vo, B. N. Vo, and A. Cantoni, "Bayesian filtering with random finite set observations," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1313–1326, Apr. 2008.

Ienkaran Arasaratnam received the B.Sc. degree (first-class honors) from the Department of Electronics and Telecommunication Engineering, University of Moratuwa, Sri Lanka, in 2003, and the M.A.Sc. and Ph.D. degrees from the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, in 2006 and 2009, respectively. For his Ph.D. dissertation titled "Cubature Kalman Filtering: Theory and Applications," he formulated a new approximate nonlinear Bayesian filter called the "Cubature Kalman Filter."

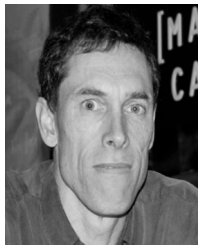
Currently, he is a Postdoctoral Researcher with the Ford Motor Company of Canada. His main research interests include signal processing, control and machine learning with applications to target tracking and fault diagnosis of automotive engines.

Dr. Arasaratnam received the Mahapola Merit Scholarship during his undergraduate studies. He is the recipient of the Outstanding Thesis Research Award at M.A.Sc. Level in 2006, the Ontario Graduate Scholarship in Science and Technology in 2008, and NSERC's Industrial Research and Development Fellowship in 2010.

Simon Haykin (SM'70–F'82–LF'01) received the B.Sc. (first-class honors), Ph.D., and D.Sc. degrees, all in electrical engineering, from the University of Birmingham, Birmingham, U.K.

Currently, he is the University Professor at McMaster University, Hamilton, ON, Canada. He is a pioneer in adaptive signal-processing with emphasis on applications in radar and communications, an area of research which has occupied much of his professional life. In the mid-1980s, he shifted the thrust of his research effort in the direction of Neural Computation, which was reemerging at that time. All along, he had the vision of revisiting the fields of radar and communications from a brand new perspective. That vision became a reality in the early years of this century with the publication of two seminal journal papers: "Cognitive Radio: Brain-Empowered Wireless Communications," appearing in the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONSin February 2005 and "Cognitive Radar: A Way of the Future," appearing in the *IEEE Signal Processing Magazine* in February 2006. Cognitive radio and cognitive radar are two important parts of a much wider and multidisciplinary subject: cognitive dynamic systems, the research of which has become his passion.

Prof. Haykin is a Fellow of the Royal Society of Canada. He is the recipient of the Henry Booker Medal of 2002, the Honorary Degree of Doctor of Technical Sciences from ETH Zentrum, Zurich, Switzerland, in 1999, and many other medals and prizes.



Thomas R. Hurd received the B.Sc. degree in mathematical physics from Queen's University, Canada, and the D.Phil. degree in mathematics from Oxford University, U.K.

He is a professor of mathematics with McMaster University, Hamilton, ON, Canada. He has been an active researcher in the field of mathematical finance for more than 12 years. In an earlier phase of his academic career, he concentrated on mathematical physics, where he attempted to solve problems at the foundations of quantum field theory, inventing rigorous mathematical methods for constructing models proposed by physicists. Finance applications have proved to be more popular with graduate students, and he is now heavily involved with running M- Phimac, the Masters in Mathematical Finance program at McMaster University, and supervising a number of M.Sc. and Ph.D. degree students in Phimac, the McMaster financial mathematics laboratory.

Prof. Hurd has also been involved in international aspects of his research field. In collaboration with Matheus Grasselli, he is an Organizer of the current Fields Institute 6 Month Thematic Program on Quantitative Finance: Foundations and Applications. He is the lead organizer, together with Sebastian Jaimungal, of the 2010 World Congress of the Bachelier Finance Society held in Toronto, Canada, June 22–26, 2010. He has also been the chair of the Organizing Committee for the Fields Institute Quantitative Finance Seminar Series for several years .