The Journal of the Astronautical Sciences, Vol. 56, No. 3, July–September 2008, pp. 401–440

A High Order Method for Estimation of Dynamic Systems[1,2]

Manoranjan Majji,[3] John L. Junkins,[4] and James D. Turner[5]

Abstract

An analytical approach is presented for developing an estimation framework (called the J*th* Moment Extended Kalman Filter (JMEKF)). This forms an important component of a class of architectures under investigation to study the interplay of major issues in nonlinear estimation such as model nonlinearity, measurement sparsity and initial condition uncertainty in the presence of low process noise. Utilizing an automated nonlinear expansion of the model about the current best estimated trajectory, a J*th* order approximate solution for the departure motion dynamics about a nominal trajectory is derived in the form of state transition tensors. This solution is utilized in evaluating the evolution of statistics of the departure motion as a function of the statistics of initial conditions. The statistics thus obtained are used in the determination of a state estimate assuming a Kalman update structure. Central to the state transition tensor integration about a nominal trajectory, is the high order sensitivity calculations of the nonlinear models (dynamics and measurement), being automated by OCEA (Object Oriented Coordinate Embedding Method), a computational tool generating the required various order partials of the system differential equations without user intervention. Working in tandem with an OCEA automation of the derivation of the state transition tensor differential equations is a vector matrix representation structure of tensors of arbitrary rank, facilitating faster and more accurate computations. High order moment update equations are derived to incorporate the statistical effects of the innovations process more rigorously, improving the effectiveness of the estimation scheme. Numerical simulations on an orbit estimation example investigate the gain obtained in using the proposed methodology in situations where the classical extended Kalman filter's domain of convergence is smaller. The orbit estimation example presented examines a situation that requires us to determine the position and velocity state of the orbiter from range, azimuth and elevation measurements being made available sparsely.

[1]Dedicated to Professor Byron D. Tapley, celebrating 50 years of his academic and research contributions.
[2]Earlier versions of the paper were presented at the 2008 AAS/AIAA Space Flight Mechanics Meeting, Galveston, TX, January, 27–31, and at the Symposium Honoring Byron D. Tapley's 50 years of contributions of Aerospace Education, Research, and Service, February 1, 2008.
[3]Graduate Research Assistant, Aerospace Engineering Department, 3141 TAMU, Student Member, AAS.
[4]Regents Professor, Distinguished Professor of Aerospace Engineering, Holder of the Royce E. Wisenbaker Chair in Engineering, Aerospace Engineering Department, 3141 TAMU, Fellow, AAS.
[5]Research Professor, Director of Operations, Consortium for Autonomous Space Systems (CASS), Aerospace Engineering Department, 3141, TAMU, Senior Member, AAS.

Introduction

Estimation theory deals with the problem of estimating the state of a dynamical system from sensor measurements, usually corrupted by noise. The famous paper by Kalman [1], which solved the estimation problem for a linear dynamical system with measurements corrupted with Gaussian noise, was quickly applied to nonlinear problems via a local linearization of the dynamic system and measurement equations arriving at what is today known as the Extended Kalman Filter (EKF). One of the most important problems in nonlinear estimation involves the determination of position and velocity estimates of orbiting spacecraft. Teaching researchers the science of Extended Kalman filtering and most importantly, developing adaptive methods to account for the unmodeled dynamics in the propagation and update phases of the filter constitute the hallmark achievements of the Tapley era. Methods like the Myers–Tapley technique [2, 3] for estimation of process noise parameters, considered to be a gold standard to this day, are a testimony to the fundamental contributions of Professor Tapley and his co-workers [4–7] that lead to form the strong foundations of statistical estimation theory. Detailed derivations of these and other foundational algorithms associated with statistical estimation theory have now become an integral part of university curriculum and are documented very widely [7–10].

Recently, due to advancement of computational capabilities, more stringent estimator design requirements are being placed on engineers. This led to a path for engineers to investigate possible generalizations of extended filtering strategies to accommodate the assumptions taken by the classical Kalmanesque framework. Work of Julier et al. [11], motivated by Athans [12], developed the unscented Kalman filter that takes into account the second-order terms in the Extended Kalman filter can be considered a testimony of the resurgence of interest in high-order methods. Several other methods quickly followed, including the so called particle methods [13]. Although particle filters are an important approach, in most physical systems, the continuity of solutions and associated properties of the models makes analytic methods of estimation remain extremely attractive. A recent investigation by Park and Scheeres [14, 15] exploits this feature of mathematical models and proposes a semi-analytic framework, based on the same. It was shown that in certain problems of orbit determination, analytic approaches are often desirable since the nominal motion is known *a priori.* The fact that the models governing the motion of celestial bodies are highly precise also contributes to the gain in accuracy in this problem. We take the approach of Park and Scheeres [14] and incorporate the state transition tensor approach to propagate the evolution of the statistical moments. Efficient tensor representations are proposed to integrate the tensor differential equations, while the Object Oriented Coordinate Embedding Algorithm (OCEA) provides the required sensitivities to drive the process. To incorporate the changes in the moments due to Kalman update, we introduce the high-order update equations.

In most engineering applications, the estimator design is fraught with the interaction of three very important factors (challenges). First is the nonlinearity of the dynamic system in question and the measurement model. Dynamic system nonlinearity forces non-Gaussian error propagation through the dynamical system and causes deviation of the case from the Extended Kalman framework. Measurement system nonlinearity plays an important part in the decision process to enable the determination of the "best" state estimate. The second factor influencing the importance of high-order terms in the state estimation is the sparsity of measurements. Long propagation phases enable the nonlinearities in the dynamic system to "matter" and the linearized covariance propagation might under-predict otherwise. The third factor is the initial condition uncertainty and levels of process and measurement noise. While large uncertainty in initial conditions amplifies the effect of high-order terms, the increase in process noise statistics tends to wash-out the model errors and, if an appropriately large estimate of the second moments of the noise is used, to inflate the covariance to retain a near-Gaussian nature. Larger measurement noise on the other hand emphasizes the believability of the model and this in-turn calls for a more accurate account of the nonlinearity of the system model whenever possible. The study of these three important issues in a pristine environment is facilitated by retaining high-order terms in the Extended Kalman Filter and producing an analytic filtering framework out of the same. This paper presents an alternative path to the development of such an analytic filtering architecture using the state transition tensors to propagate and update the statistical moments governing the estimation error process. These moments we subsequently used in a Kalman Gain calculation for state estimation. A unified framework, which we call the J*th* Moment Extended Kalman Filter (JMEKF) is realized to facilitate a systematic study of the said interactions. The realization of such an architecture among other things, also provides an answer to the long standing debate as to which order is necessary at what desired level of accuracy for a given problem. This classical debate dates back to Athans et al., [16] and the references there-in and is extremely relevant today in the light of the novel approximate high-order filters being developed (Julier et al. [11], Norgaard et al., [17] and Park and Scheeres [14] for example).

The organization of the paper is as follows. The derivation of equations governing the dynamics of departure motion using the concepts of state transition tensors is presented in the first section. Following this discussion, we present a scalar example detailing the important ideas of the estimation architecture (JMEKF). The subsequent section discusses the difficulties involved in extending the process in a general vector setting. OCEA is proposed as a tool to compute the necessary sensitivities accurately and efficiently. A vector matrix representation of tensors is proposed to handle the tensor operations proposed in the paper. The state transition tensor evolution equations are then expressed in this vector matrix representation. Solutions of the tensors using this representation for some classical problems are detailed in the appendix along with a discussion on the speed-up achieved in those problems. We use these solutions to compute the evolution of statistical moments of the departure motion dynamics. Incorporation of these evolved moments in a Kalman decision framework yields an estimation paradigm which is the main topic of subsequent discussion. High-order tensor updates are developed and incorporated into the algorithm at this stage. An orbit determination problem example is presented where the performance of the classical extended Kalman filter is enhanced by including the high-order terms in accordance with the proposed JMEKF paradigm. To improve the flow of the paper, discussions on the vector matrix representation of the tensors, notational development for this representation and the associated computational advantages are detailed wherever appropriate. A brief exploration of solutions of differential equations in the presence of forcing functions is done in Appendix II, to clarify (to some extent) the separation of the process noise statistics.

High Order State Transition Tensors

In this section, we derive the higher-order state transition tensors and discuss their utility in the uncertainty propagation problem. We remark that the methodology presented herein is not novel and was in use as early as the 1970s. Athans et al. [12] developed a second-order analytical propagation methodology to implement a second-order Kalman filter. A similar technique has been developed by Wiberg [18] that is restricted to third statistical moment propagation. The methodology also forms an important motivation for the modern unscented filtering technique (see Julier and Uhlmann [11]). The method is also suggested by Gelb [10], as a technique for development of nonlinear filters for specific problems. In addition to the notations and developments of Griffith [19] and Park [14], we denote vectors with boldfaced letters in the small alphabet. Matrices are denoted by italicized capital letters and tensors are denoted similarly but with additional subscripts and a superscript $(\cdot)^{(r)}_{i_1 i_2 \ldots i_r}$ within parenthesis denoting the index of the component and the rank of the tensor. A note on what we mean by order and rank is due here for some important developments of the paper. By order, we mean the order of the sensitivity of the object in question. Note that this is in general different from the rank of the associated tensor. For our purposes rank indicates the number of independent dimensions of a multi-dimensional or an N-way array (refer to [20] for more on multi-dimensional array operations). For instance if $\mathbf{f} \in \mathbb{R}^n$ was a vector function, its Hessian is a tensor of rank three $f^*_{i_1, j_1 j_2}$ representable for all practical purposes of computation, by a three dimensional array with an assumed convention, but it is a second-order term. Throughout the document, the Einstein summation convention is used along with the now standard free and dummy index definitions (Chapter 2 in Lawden [21]).

Consider the nonlinear system

$$\dot{x}_i(t) = f_i(t, x(t)), \quad x_i(t_0) = x_i^0 \tag{1}$$

Let the solution to such a system exist, and be given by $x_i(t) = x_i(t, t^0, \mathbf{x}^0)$, with the property that $\mathbf{x}(t_0) = \mathbf{x}(t_0, t_0, \mathbf{x}^0) = \mathbf{x}^0$. The neighborhood of an initial condition of interest, say $\mathbf{x}^0$, denoted by $\mathbf{x}^0 + \delta\mathbf{x}^0$ is required to be studied. Hence the solutions of the nonlinear system with this continuous neighborhood of initial conditions are assumed to exist and are given by, $\mathbf{x}(t, t^0, \mathbf{x}^0 + \delta\mathbf{x}^0)$. Time evolution of the neighborhood of initial conditions is then given by

$$\delta x_i(t) := x_i(t, t^0, \mathbf{x}^0 + \delta\mathbf{x}^0) - x_i(t, t^0, \mathbf{x}^0) \tag{2}$$

which satisfy the differential equations

$$\begin{aligned} \delta\dot{x}_i(t) &:= \dot{x}_i(t, t^0, \mathbf{x}^0 + \delta\mathbf{x}^0) - \dot{x}_i(t, t^0, \mathbf{x}^0) \\ \delta\dot{x}_i(t) &:= f_i(\mathbf{x}(t, t^0, \mathbf{x}^0 + \delta\mathbf{x}^0), t) - f_i(\mathbf{x}(t, t^0, \mathbf{x}^0), t) \end{aligned} \tag{3}$$

This can be expanded as an m^{th} order Taylor Series

$$\delta\dot{x}_i(t) = \sum_{p=1}^{m} \frac{1}{p!} f^*_{i,k_1 k_2 \ldots k_p} \delta x_{k_1 \ldots} \delta x_{k_p} \tag{4}$$

where the state trajectories, $\delta x_i(t)$, are mapped from their initial conditions by the state transition tensors

$$\delta x_i(t) = \sum_{p=1}^{m} \frac{1}{p!} \Phi_{i,k_1 \ldots k_p} \delta x^0_{k_1 \ldots} \delta x^0_{k_p} \tag{5}$$

with $f^*_{i,k_1\ldots k_p} := \frac{\partial^p f_i}{\partial x_{k_1}\ldots\partial x_{k_p}}\Big|_{x=x^*}$ denoting the higher-order sensitivity evaluated along a nominal trajectory, $\mathbf{x}^*(t)$ and $\Phi_{i,k_1\ldots k_p} := \frac{\partial^p x_i}{\partial x^0_{k_1}\ldots\partial x^0_{k_p}}\Big|_{x=x^*}$ being the state transition matrix, that calculates the change in the solution of the nonlinear dynamical system with changes to initial conditions and whose time evolution is given by the differential equations

$$\dot{\Phi}_{i,j_1} = f^*_{i,k_1}\Phi_{k_1,j_1} \tag{6}$$

$$\dot{\Phi}_{i,j_1 j_2} = f^*_{i,k_1}\Phi_{k_1,j_1 j_2} + f^*_{i,k_1k_2}\Phi_{k_1,j_1}\Phi_{k_2,j_2} \tag{7}$$

$$\begin{aligned}\dot{\Phi}_{i,j_1j_2j_3} &= f^*_{i,k_1}\Phi_{k_1j_1j_2j_3} + f^*_{i,k_1k_2}(\Phi_{k_1j_1j_2}\Phi_{k_2,j_3} + \Phi_{k_1,j_1j_3}\Phi_{k_2,j_2} + \Phi_{k_1,j_2j_3}\Phi_{k_1,j_1}) \\ &\quad + f^*_{i,k_1k_2k_3}\Phi_{k_1,j_1}\Phi_{k_2,j_2}\Phi_{k_3,j_3}\end{aligned} \tag{8}$$

$$\begin{aligned}\dot{\Phi}_{i,j_1j_2j_3j_4} &= f^*_{1,k1}\Phi_{k_1,j_1j_2j_3j_4} \\ &\quad + f^*_{i,k_1k_2}\begin{pmatrix}\Phi_{k_1,j_1j_2j_3}\Phi_{k_2,j_4} + \Phi_{k_1,j_1j_2j_4}\Phi_{k_2,j_3} + \Phi_{k_2,j_2j_3j_4}\Phi_{k_1,j_1} \\ +\Phi_{k_1,j_1j_3j_4}\Phi_{k_2,j_2}\end{pmatrix} \\ &\quad + f^*_{i,k_1k_2}(\Phi_{k_1,j_1j_2}\Phi_{k_2,j_3j_4} + \Phi_{k_1,j_1j_3}\Phi_{k_2,j_2j_4} + \Phi_{k_2,j_2j_3}\Phi_{k_1,j_1j_4}) \\ &\quad + f^*_{i,k_1k_2k_3}\begin{pmatrix}\Phi_{k_1,j_1j_2}\Phi_{k_2,j_3}\Phi_{k_3,j_4} + \Phi_{k_1,j_1j_3}\Phi_{k_2,j_2}\Phi_{k_3,j_4} \\ +\Phi_{k_2,j_2j_3}\Phi_{k_1,j_1}\Phi_{k_3,j_4}\end{pmatrix} \\ &\quad + f^*_{i,k_1k_2k_3}\begin{pmatrix}\Phi_{k_1,j_1j_4}\Phi_{k_2,j_3}\Phi_{k_3,j_3} + \Phi_{k_1,j_1}\Phi_{k_2,j_2j_4}\Phi_{k_3,j_3} \\ +\Phi_{k_2,j_2}\Phi_{k_1,j_1}\Phi_{k_3,j_3j_4}\end{pmatrix} \\ &\quad + f^*_{i,k1\,k2\,k3\,k4}\Phi_{k_1,j_1}\Phi_{k_2,j_2}\Phi_{k_3,j_3}\Phi_{k_4,j_4}\end{aligned} \tag{9}$$

with initial conditions

$$\Phi_{i,j}(t_0,t_0) = \delta_{i,j}, \quad \Phi_{i,j_1j_2\cdot\cdot j_p}(t_0,t_0) = 0, \forall p > 1$$

These equations can be derived by comparison of coefficients and by applying to the basic definition of the state transition tensors and taking the time derivatives. Notice that these differential equations can be solved sequentially, since they are in cascade form. However, the most important issue is to establish a methodology to derive the needed partial derivative coefficients (e.g., $f^*_{i,k_1k_2k_3k_4}$), and to systematically solve these differential equations. In subsequent sections we propose a method to represent these equations systematically and use the representation to numerically integrate them.

However, the solution presented here is unforced. In most problems of interest for filtering applications, where nonlinearity has a say, the Gaussian random process that forces the dynamic system (process noise) is assumed to have a low power. High levels of process noise tend to "balloon" the statistics and drown the nonlinear effects in the propagation phase of the filter. The following section proposes a heuristic way of justifying the dominance of initial condition uncertainty (unforced motion and thus the state transition tensor solution) in the presence of low levels of random forcing. This consequently lays foundation for the filtering architecture that is developed subsequently.

Estimation of Nonlinear Dynamical Systems (Scalar Problem): JMEKF Approach

Let us consider a scalar nonlinear dynamical system being measured by a nonlinear measurement model to derive clearly the essential components of the proposed estimation framework. Generalization to the multivariable case with dynamic models in an N-dimensional state space with m measured outputs is conceptually equivalent with exception of a few difficulties in practical application. This extension to multivariate models will be systematically developed in subsequent sections.

The said nonlinear system and measurement model are assumed to be given by

$$\begin{aligned} \dot{x} &= f(t, x(t)) + Gw(t) \\ \tilde{y} &= h(x) + v \end{aligned} \tag{10}$$

The nonlinear filtering problem being presented in the current discussion is best suited for continuous discrete estimation problems, where there is a continuous time phase involving the propagation of the statistics of estimation error and a discrete phase, where with the arrival of new measurements, the propagated statistics are used to make a decision about the state estimate. The measurement equations are most frequently nonlinear and additive discrete Gaussian random variables are assumed to corrupt the measurements. The characteristics of the additive white noise, which is assumed to be uncorrelated with itself and other random variables governing the process noise, are given by $E(v_k) = 0, E(v_k v_j^{\mathrm{T}}) = R_k \delta_{kj}$. The utilization of the state transition tensor approach (presented in the previous section and originally due to Griffith [19] and Park [14, 15]) to develop a set of tools that enable the study of interacting effects of nonlinearity, sparsity of measurements in the presence of low process noise, is the topic of current discussion.

Using the heuristic arguments made in the context of the derivation of the forced response of the state transition tensor solution in the appendix, we neglect the higher-order effects of the forcing function and develop the solution for the unforced departure motion. This expression for the scalar case is given by

$$\delta x(t) = \phi_{(1)} \delta x_k^{+} + \frac{1}{2!} \phi_{(2)} \delta x_k^{+2} + \frac{1}{3!} \phi_{(3)} \delta x_k^{+3} + \ldots \tag{11}$$

where $\phi_{(1)}$ are state transition tensors satisfying the differential equations

$$\begin{aligned} \dot{\phi}_{(1)} &= f_x^* \phi_{(1)} \\ \dot{\phi}_{(2)} &= f_x^* \phi_{(2)} + f_{xx}^* \phi_{(1)}^2 \\ \dot{\phi}_{(3)} &= f_x^* \phi_{(3)} + 3 f_{xx}^* \phi_{(2)} \phi_{(1)} + f_{xxx}^* \phi_{(1)}^3 \\ &\ldots \end{aligned} \tag{12}$$

to be integrated with initial conditions, $\phi_{(1)}(t_k) = 1$, $\phi_{(j)}(t_k) = 0, \forall j > 1$ (the scalar versions of (6)–(9)), during the propagation phase $\forall t \in [t_k, t_{k+1}]$, where the sensitivity of the dynamical system, $f^*_{xxx\ldots(p \text{ times})} := \frac{\partial^p f_i}{\partial x^p}\big|_{x=\hat{x}_k^-}$ and the measurement model $h^*_{xxx\ldots(p \text{ times})} := \frac{\partial^p h_i}{\partial x^p}\big|_{x=\hat{x}_k^-}$ is evaluated along the trajectory propagated from the state estimate updated at the instant of the arrival of the last set of measurements, t_k, through the exact differential equations of motion (in the classical Extended Kalman Filtering framework). We note in passing that the truncation of the Taylor series expansion of the nonlinear equations of motion and the measurement equations leads us to incur errors in the corresponding state transition tensor calculations. A detailed audit of the errors incurred is a topic of investigation by the authors and a corresponding analysis will be reported separately. Using the expression for the approximate departure motion dynamics (equation (11)), the evolution of statistical moments from their values at the last update instant can be determined as

$$\begin{aligned}
P^{(2)}(t) &= E(\delta x^2(t)) \\
&= \phi_{(1)}^2 P^{(2)+}(t_k) + \phi_{(1)}\phi_{(2)}P^{(3)+}(t_k) \\
&\quad + \left[\frac{1}{2!2!}\phi_{(2)}\phi_{(2)} + \frac{2}{3!}\phi_{(3)}\phi_{(1)}\right]P^{(4)+}(t_k) + HOT + \int_{t_k}^{t}\phi_{(1)}^2(t_k,\tau)g^2 Q(\tau)\mathrm{d}\tau \\
P^{(3)}(t) &= E(\delta x^3(t)) \qquad (13)\\
&= \phi_{(1)}^3 P^{(3)+}(t_k) + \frac{3}{2!}\phi_{(2)}\phi_{(1)}^2 P^{(4)+}(t_k) + HOT \\
P^{(4)}(t) &= E(\delta x^4(t)) \\
&= \phi_{(1)}^4 P^{(4)+}(t_k) + HOT
\end{aligned}$$

where the terms $P^{(j)+}(t_k) = E[(\delta x_k^+)^j]$ refer to the values of the corresponding updated jth statistical moment at time t_k and $P^{(j)+}(t) = E[(\delta x_k^-)^j]$ refer to the value of the corresponding jth statistical moment at time t propagated from a certain initial time and a corresponding initial moment value.

Having set up a method to compute the evolution of the statistical moments, we proceed to the second important phase of the estimation framework. Keeping in perspective the Kalman way of thinking, let us consider a state update which incorporates the innovations process to correct the prediction by the use of a linear gain given by

$$\hat{x}^+ = \hat{x}^- + K\nu \qquad (14)$$

with $\nu = \tilde{y}(t_{k+1}) - \hat{y}(t_{k+1}) = h(x) - h(\hat{x}^-) + v$, where the Kalman gain is calculated as

$$K = P^{x\nu}(P^{\nu\nu})^{-1} \qquad (15)$$

The time subscript $(\cdot)_k$ is dropped in the developments and the update phase is assumed to be taking place at a general time instant t_k when a new set of measurements are made available to the estimator. State estimation error before and after the update is defined as $\delta x^\pm := x - \hat{x}^\pm$. We point out at this stage that this linear update structure (derived from the implicit assumption that the closed loop estimation error dynamics is a Gaussian random process, depending only on the first two moments), being an important artifact of the Kalman era in estimation is not necessary, although it has served human-kind very well thus far. One can indeed define the correction to be a higher-order function of the innovations process and choose the gains to minimize appropriate high-order statistics. However, for the initial developments, we follow the same path, as treaded by our fellow nonlinear filter architectures proposed by Julier and Uhlmann [11] and Park and Scheeres [14], and postpone such exercises to future communications. However, it is important to realize that this architecture has the flexibility to accommodate such ideas with no loss of generality. The updated state is always denoted by the superscript $(\,)^+$ and wherever convenient, we choose to omit the $(\,)^-$ superscript for clarity. The terms $P^{x\nu}$ and $P^{\nu\nu}$ are the statistics of the so-called innovations process with itself and with the state estimation error, given by

$$\begin{aligned}
P^{x\nu} &= E[\delta x^- \nu] \\
&= h_x^* E[(\delta x^-)^2] + \frac{1}{2}h_{xx}^* E[(\delta x^-)^3] + \frac{1}{3!}h_{xxx}^* E[(\delta x^-)^4] + \ldots \qquad (16)
\end{aligned}$$

$$
\begin{aligned}
P^{\nu\nu} &= E[\nu\nu] \\
&= h_x^* h_x^* E[(\delta x^-)^2] + \frac{2}{2!} h_x^* h_{xx}^* E[(\delta x^-)^3] \\
&\quad + \left(\frac{2}{3!} h_{xxx}^* h_x^* + \frac{1}{2!2!} h_{xx}^{*2}\right) E[(\delta x^-)^4] + \ldots + R
\end{aligned}
\tag{17}
$$

Once the state is updated using the measurements in equation (14), we calculate the corresponding changes in the statistics of departure motion (e.g., in equations (11) and (13)), in preparation for the next propagation phase. These calculations are given by the high-order tensor update equations that are defined below. In the unscented Kalman filter framework, this calculation is inbuilt as the new sigma points are calculated about the updated state estimate. This is indeed an essential step in the filter design, because, without the incorporation of the new information arriving via the innovations process, the high-order moments keep growing in undesirable directions which, over the progress of time, become physically meaningless and the filter asymptotically approaches the Extended Kalman Filter. To the authors' best knowledge, this is the first instance of such a high-order moment update in a filter and is one of the main contributions of this paper. In the scalar case, all the statistics collapse to a scalar value. But in the vector case, due to the presence of the interaction terms, these update equations have to be handled efficiently, as they are high dimensioned tensors. Such extensions are presented in the next section.

The bias update (first moment update) is given by

$$
\begin{aligned}
E[\delta x^+] &= E[\delta x^- - K\nu] \\
&= [1 - Kh_x^*]E(\delta x^-) - \frac{1}{2} K h_{xx}^* E[(\delta x^-)^2] - \frac{1}{3!} K h_{xxx}^* E[(\delta x^-)^3] \cdots
\end{aligned}
\tag{18}
$$

The second moment update is given by

$$
P^{(2)+}(t_k) = P^{(2)}(t_k) - KP^{\nu\nu}K^{\mathrm{T}} \tag{19}
$$

In a manner similar to the classical covariance update equation, the third moment update equation follows as

$$
P^{(3)+}(t_k) = P^{(3)}(t_k) + 3K^2P^{x\nu\nu} - 3KP^{xx\nu} - K^3P^{\nu\nu\nu} \tag{20}
$$

where the high-order tensors on the right hand side are given by

$$
\begin{aligned}
P^{x\nu\nu} &= E[\delta x \nu^2] \\
&= h_x^{*2} P^{(3)}(t_k) + \frac{2}{2!} h_{xx}^* h_x^* P^{(4)}(t_k) + \ldots + E(\delta x v^2)
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
P^{xx\nu} &= E[\delta x^2 \nu] \\
&= h_x^* P^{(3)}(t_k) + \frac{1}{2!} h_{xx}^* P^{(4)}(t_k) + \frac{1}{3!} h_{xxx}^* P^{(5)}(t_k) + \ldots
\end{aligned}
\tag{22}
$$

$$
\begin{aligned}
P^{\nu\nu\nu} &= E[\nu^3] \\
&= h_x^{*3} P^{(3)}(t_k) + \frac{3}{2!} h_{xx}^* h_x^{*2} P^{(4)}(t_k) + \ldots + E(\nu^3)
\end{aligned}
\tag{23}
$$

The fourth moment update equation consequently becomes

$$P^{(4)+}(t_k) = P^{(4)}(t_k) - 4K\,P^{xxx\nu} + 6K^2P^{xx\nu\nu} - 4K^3P^{\nu\nu\nu x} + K^4P^{\nu\nu\nu\nu} \tag{24}$$

The correlations being defined by

$$\begin{aligned} P^{xxx\nu} &= E[\delta x^3\nu] \\ &= h_x^*P^{(4)}(t_k) + \frac{1}{2!}h_{xx}^*P^{(5)}(t_k) + \frac{1}{3!}h_{xxx}^*P^{(4)}(t_k) + \ldots \end{aligned} \tag{25}$$

$$\begin{aligned} P^{xx\nu\nu} &= E[\delta x^2\nu^2] \\ &= h_x^{*2}P^{(4)}(t_k) + \frac{2}{2!}h_{xx}^*h_x^*P^{(5)}(t_k) + \ldots + E(\delta x^2\nu^2) \end{aligned} \tag{26}$$

$$\begin{aligned} P^{x\nu\nu\nu} &= E[\delta x\nu^3] \\ &= h_x^{*3}P^{(4)}(t_k) + \frac{3}{2!}h_{xx}^*h_x^{*2}P^{(5)}(t_k) + \ldots \end{aligned} \tag{27}$$

$$\begin{aligned} P^{\nu\nu\nu\nu} &= E[\nu^4] \\ &= h_x^{*4}P^{(4)}(t_k) + \frac{4}{2!}h_{xx}^*h_x^{*3}P^{(5)}(t_k) + \ldots + E(\nu^4) \end{aligned} \tag{28}$$

and so on, in general. Having updated as indicated above, the process of integrating the state transition tensors is repeated for the next set of measurements and the subsequent decision of the best state estimate. This scalar example summarizes the central idea of the paper. The inclusion of the high-order terms along with the high-order statistic updates within the classical Extended Kalman Filtering framework is the main result of this paper. Subsequent sections extend this formally to the general vector problem of engineering importance. However, the central idea remains unchanged even in the vector case. Note that in the developments, we include a term to monitor the bias of the estimator. This is important in problems where the governing equations are nonlinear (in the measurement model and the plant dynamics). As shown in the propagation part of an orbit problem in the appendix, for some nonlinear problems the probability density function (pdf) governing the estimation error of the problem is inherently biased. An estimator inherited from a linear-Gaussian mindset tries to be unbiased by incorporating corrections. However, such a correction is not required. Consider an imaginary situation with a banana shaped pdf about the true trajectory. Since the expected value of the pdf in such a situation lies outside, a bias correction tends to throw the estimate farther away from the truth, which not only increases the estimation error but is also completely non-physical. Hence we propose the tracking of this bias term as a health monitor for the estimator and also to calculate the high-order correlations required in the statistical moment calculations. Let us now look at a vector-matrix representation method of tensors and express the tensor differential equations using the same.

Implementation Using OCEA and Array Manipulation Techniques

We notice that there are two important difficulties in extending the scalar estimator design proposed in the previous section to the general vector models. First is the calculation of the high-order partials required to integrate the differential equations and the algebraic transformations involving the moments. The Object-oriented Coordinate Embedding Algorithm (OCEA) [22], with its extensive use of operator overloaded embedded functions, provides a greatly simplified modeling development environment for the analysis, providing just these high order partials.

The higher-order partial derivatives (of arbitrary order) associated with the dynamical equations of motion are made available within this framework, by a simple declaration of symbolic variables and functions. The general goals of the programming environment are to facilitate the development of a generalized optimization solution by using advances in object oriented programming and automatic differentiation.

The second important obstacle is the management of multidimensional arrays and tensor operations required for computing the state transition tensors and the statistical moment tensor calculations at every propagation and update step. This is achieved by a special representation of tensors that is proposed in this paper. Similar developments exist in applied mathematics and associated toolboxes are in development and active usage (one of the popular tensor toolboxes comes out of Sandia National Labs, refer to [20]). To develop the core expressions required for our operations, we first setup a convention to handle tensor data structures. This notation is partly based on the way MATLAB® chooses to represent and manipulate multidimensional arrays and hence enables us to write pseudo-code at important places for a clear presentation. We then define some matrix and vector representations of these tensors and use them to represent the state transition tensor differential equations in this framework. Advantages and disadvantages of such a representation along with some application problems are presented in Appendix A.

Consider a tensor of arbitrary rank (a matrix can be thought of as a second-order tensor), say, $\zeta_{i_1 i_2 i_3 i_4}$. It is convenient to represent this tensor in a vector form for numerical operations. Now, let $\mathrm{Vec}(\cdot)$ be the operation, that transforms a tensor of arbitrary order into a vector (stacking operation). We will follow the convention of stacking the tensor components (a multidimensional array) in columns. For the given tensor $\zeta_{i_1 i_2 i_3 i_4}$, this results in a vector of dimension $N^4 \times 1$, defined according to

$$\psi_\zeta^{(4)}(i_1 + N(i_2 - 1) + N^2(i_3 - 1) + N^3(i_4 - 1), 1) := \mathrm{Vec}(\zeta_{i_1 i_2 i_3 i_4}) \tag{29}$$

where the $[i_1 + N(i_2 - 1) + N^2(i_3 - 1) + N^3(i_4 - 1)]^{th}$ element of the vector $\psi_\zeta^{(4)}$ is $\zeta_{i_1 i_2 i_3 i_4}$. We sometimes use additional superscripts on top of the tensors to additionally qualify whether the tensor is updated or propagated (*a posteriori* statistical moment $(\cdot)^{(r)+}$ or an *a priori* moment $(\cdot)^{(r)t}_{i_1 \ldots i_r}$, $(\cdot)^{(r)-}_{i_1 \ldots i_r}$, or simply $(\cdot)^{(r)}_{i_1 \ldots i_r}$).

Also, let the matrix representation of a high order tensor be given by Θ. This means in this case that

$$\Theta_\zeta^{(4)}(i_1, i_2 + N(i_3 - 1) + N^2(i_4 - 1)) := \mathrm{Mat}(\zeta_{i_1 i_2 i_3 i_4}) \tag{30}$$

The matrix representation of a general tensor of rank k, ζ, thus defines a matrix $\Theta_{(k)\zeta} \in \mathbb{R}^{N \times N^{k-1}}$ having the components identical to the tensor arranged appropriately. This is demonstrated in the pictorial example in Fig. 1.

With these definitions, the tensor differential equations can be vectorized. Notice that this process is a generalization of vectorizing matrix equations, using Kronecker products (see Graham [23]). Methodologies of this type have been utilized for faster computations historically by Rauhala [24] (and several references therein), and recently by Van Loan et al. [25]. Faster computations using these architectures happen to be the result of sparse matrices that are a consequence of this representation. A speed comparison on 2D oscillator problems is present in the appendix. However, when the dimensionality of the problem increases, the matrices become too large even for efficient manipulations using sparse data structures. Notice, however, that there are a variety of ways of representation other than the one proposed by us. So in such a situation switching to a more appropriate representation is advised. By all means, what we propose is the beginning of a whole new paradigm and not the ultimate

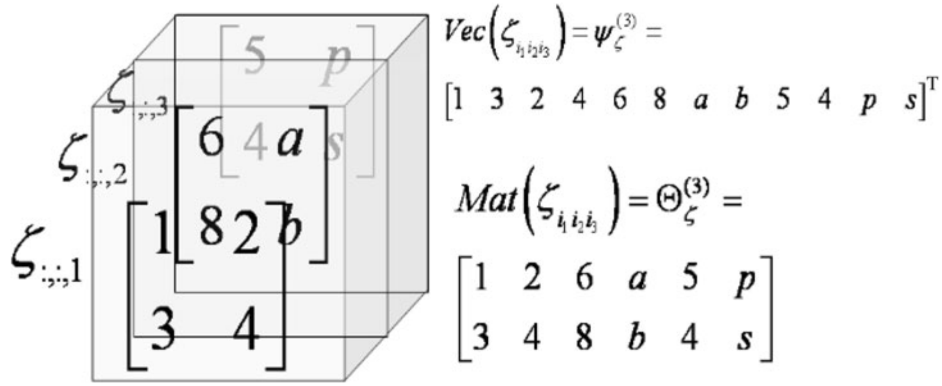

FIG. 1. Pictorial Demonstration of the Vector-Matrix Representation of Tensors.

panacea of all problems. Proposing an exciting new mixture of ideas that enable practical computations on an important problem is the central objective of this paper.

In tandem with the vector matrix representations defined above, let us look at a set of generalized transpose operators useful in tensor manipulation. We note in passing that this set in general is much more complex than that of the standard matrix transposition operation. The MATLAB® command pair "permute" and "ipermute" allows us to do these generalized transposition operations very easily. Let us demonstrate the utility of these permute functions for the state transition tensor differential equation calculations. Considering equation (8), we realize, using the vector matrix representation above, the terms $f^*_{i,k_1k_2}\Phi_{k_1,j_1j_2}\Phi_{k_2,j_3}$ and $f^*_{i,k_1k_2}\Phi_{k_2,j_2j_3}\Phi_{k_1,j_1}$ can be vectorized as

$$\mathrm{Vec}\,(f^*_{i,k_1k_2}\Phi_{k_1,j_1j_2}\Phi_{k_2,j_3}) = [\Theta_\Phi^{(2)\mathrm{T}} \otimes \Theta_\Phi^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \tag{31}$$

$$\mathrm{Vec}\,(f^*_{i,k_1k_2}\Phi_{k_2,j_2j_3}\Phi_{k_1,j_1}) = [\Theta_\Phi^{(3)\mathrm{T}} \otimes \Theta_\Phi^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \tag{32}$$

However, the vector representation of $f^*_{i,k_1k_2}\Phi_{k_1,j_1j_3}\Phi_{k_2,j_2}$ becomes non-trivial and we would need to resort to column splitting or row splitting of the matrices involved. To offset this representational difficulty, we introduce the "permute" operator. (It turns out to be more useful than just for representation purposes in most situations.)

$$\mathrm{permute}(\zeta_{i_1i_2i_3i_4},\ [1 \quad 2 \quad 4 \quad 3]) = \zeta_{i_1i_2i_4i_3} \tag{33}$$

Using the permute operator, we can immediately write

$$\begin{aligned} \mathrm{Vec}(f^*_{i,k_1k_2}\Phi_{k_1,j_1j_3}\Phi_{k_2j_2}) &= \mathrm{Vec}\,(\mathrm{permute}(f^*_{i,k_1k_2}\Phi_{k_1,j_1j_2}\Phi_{k_2,j_3}, [1 \quad 2 \quad 4 \quad 3])) \\ \mathrm{Vec}(f^*_{i,k_1k_2}\Phi_{k_1,j_1j_3}\Phi_{k_2j_2}) &= \mathrm{permute}([\Theta_\Phi^{(2)\mathrm{T}} \otimes \Theta_\Phi^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)}, [1 \quad 2 \quad 4 \quad 3]) \\ &= ([\Theta_\Phi^{(2)\mathrm{T}} \otimes \Theta_\Phi^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 4\ 3]} \end{aligned} \tag{34}$$

The notational developments of this section thus enable us to write the state transition tensor differential equations (written in (6) through (9)) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_\Phi^{(2)} = [I_N \otimes \Theta_f^{(2)}]\psi_\Phi^{(2)} \tag{35}$$

where $\Theta_f^{(2)}$ is the familiar Jacobian matrix. Similarly, high-order tensor rates are given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_\Phi^{(3)} = [I_N \otimes I_N \otimes \Theta_f^{(2)}]\psi_\Phi^{(3)} + [\Theta_\Phi^{(2)\mathrm{T}} \otimes \Theta_\Phi^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \tag{36}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi_{\Phi}^{(4)} = [I_N \otimes I_N \otimes I_N \otimes \Theta_f^{(2)}]\psi_{\Phi}^{(4)} + \begin{pmatrix} [\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N] \\ + [\Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N] \end{pmatrix}\psi_f^{(3)}$$
$$([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 4\ 3]} + [\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)T} \otimes I_N]\psi_f^{(4)} \tag{37}$$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\psi_{\Phi}^{(5)} = {} & [I_N \otimes I_N \otimes I_N \otimes \Theta_f^{(2)}]\psi_{\Phi}^{(5)} + [\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(4)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(4)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 3\ 5\ 4]} + [\Theta_{\Phi}^{(4)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(4)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 4\ 5\ 3]} + [\Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)} \\
& + ([\Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 4\ 3\ 5]} \\
& + ([\Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(3)})^{\pi[1\ 2\ 5\ 3\ 4]} \\
& + [\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(4)} + [\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(4)} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(4)})^{\pi[1\ 2\ 4\ 3\ 5]} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes I_N]\psi_f^{(4)})^{\pi[1\ 2\ 5\ 3\ 4]} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(4)})^{\pi[1\ 2\ 3\ 5\ 4]} \\
& + ([\Theta_{\Phi}^{(3)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(4)} \\
& + ([\Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes \Theta_{\Phi}^{(2)\mathrm{T}} \otimes I_N]\psi_f^{(5)}
\end{aligned} \tag{38}$$

It is important to realize that only a few terms in the above equations are independent and need to be calculated. The tensors that are permutations of each other are obtained trivially by appropriately accessing the elements specified uniquely due to the argument list (in the case of MATLAB the permute command does this automatically for a multidimensional array). No arithmetic operations are required whatsoever. At this stage we remind ourselves of the fact that the superscript $(\cdot)^{(3)}$ in $(\cdot)_{\Phi}^{(3)}$ should not be confused to be indicative of the "order" of the state transition tensor involved. It is the rank of the tensor. In this case the tensor in question, $(\Phi_{i_1 j_1 j_2})$, is to have three indices and hence is to have a rank of three. However, it is indeed a second-order state transition tensor.

Nonlinear Filtering (General Case): Propagation Phase

The scalar example can be repeated in the general vector case. To facilitate the appropriate book keeping for the vector problem, we resort to the indicial notation of tensor calculus [21]. We now write the vector versions of the moment propagation equations given in equation (13) as

$$\begin{aligned}
P_{i_1,i_2}^{(2)}(t) := E(\delta x_{i_1}\delta x_{i_2}) = {} & \Phi_{i_1,j_1}\Phi_{i_2,j_2}P_{j_1 j_2}^{(2)+} + \frac{1}{2!}(\Phi_{i_1,j_1 j_2}\Phi_{i_2,j_3} + \Phi_{i_1,j_1}\Phi_{i_2,j_2 j_3})P_{j_1 j_2 j_3}^{(3)+} \\
& + \left[\frac{1}{3!}(\Phi_{i_1,j_1}\Phi_{i_2,j_2 j_3 j_4} + \Phi_{i_1,j_1 j_2 j_3}\Phi_{i_2,j_4}) + \frac{1}{2!2!}(\Phi_{i_1,j_1 j_2}\Phi_{i_2,j_3 j_4})\right]P_{j_1 j_2 j_3 j_4}^{(4)+} \\
& + \ldots + \int_{t_0}^{t}\Phi_{i_1,k_1}(t,\tau)G_{k_1 j_1}(\tau)Q_{j_1 j_2}G_{k_2 j_2}(\tau)\Phi_{i_2,k_2}(t,\tau)\mathrm{d}\tau
\end{aligned} \tag{39}$$

Evolution of the skewness tensor is given by

$$P_{i_1 i_2 i_3}^{(3)t} := \Phi_{i_1,j_1}\Phi_{i_2,j_2}\Phi_{i_3,j_3}P_{j_1 j_2 j_3}^{(3)+} + \frac{1}{2!}\begin{bmatrix} \Phi_{i_1,j_1 j_2} & \Phi_{i_2,j_3} & \Phi_{i_3,j_4} \\ +\Phi_{i_1,j_1} & \Phi_{i_2,j_2 j_3} & \Phi_{i_3,j_4} \\ +\Phi_{i_1,j_1} & \Phi_{i_2,j_2} & \Phi_{i_3,j_3 j_4} \end{bmatrix}P_{j_1 j_2 j_3 j_4}^{(4)+} + \ldots \tag{40}$$

and similarly, the evolution of kurtosis and the 5th moment are given by

$$P^{(4)t}_{i_1i_2i_3i_4} := \Phi_{i_1,j_1}\Phi_{i_2,j_2}\Phi_{i_3,j_3}\Phi_{i_4,j_4}P^{(4)+}_{j_1j_2j_3j_4} + \frac{1}{2!}\begin{pmatrix}\Phi_{i_1,j_1j_2}\Phi_{i_2,j_3}\Phi_{i_3,j_4}\Phi_{i_4,j_5} + \Phi_{i_1,j_1}\Phi_{i_2,j_2j_3}\Phi_{i_3,j_4}\Phi_{i_4,j_5} + \\ \Phi_{i_1,j_1}\Phi_{i_2,j_2}\Phi_{i_3,j_3j_4}\Phi_{i_4,j_5} + \Phi_{i_1,j_1}\Phi_{i_2,j_2}\Phi_{i_3,j_3}\Phi_{i_4,j_4j_5}\end{pmatrix}P^{(5)+}_{j_1j_2j_3j_4j_5} + \ldots$$

$$P^{(5)t}_{i_1i_2i_3i_4i_5} := \Phi_{i_1,j_1}\Phi_{i_2,j_2}\Phi_{i_3,j_3}\Phi_{i_4,j_4}\Phi_{i_5,j_5}P^{(5)+}_{j_1j_2j_3j_4j_5} + \ldots \tag{41}$$

These equations are conveniently expressed in the vector–matrix representation convention set up in the previous sections for tensors. The utility of such a setup is the computational savings we incur due to the matrix operations in addition to the notational simplicity. Using the said vector–matrix representation of tensors, the moment propagation equations are given by

$$\begin{aligned}\psi_P^{(2)t} &= [\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}]\psi_P^{(2)+} + \frac{1}{2!}\begin{bmatrix}\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \\ +\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)}\end{bmatrix}\psi_P^{(3)+} \\ &\quad + \left\{\begin{array}{l}\dfrac{1}{3!}\begin{bmatrix}\Theta_\Phi^{(4)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(4)}\end{bmatrix} \\ + \dfrac{1}{2!2!}[\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(3)}]\end{array}\right\}\psi_P^{(4)+} + \left\{\begin{array}{l}\dfrac{1}{4!}\begin{bmatrix}\Theta_\Phi^{(5)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(5)}\end{bmatrix} \\ + \dfrac{1}{3!2!}\begin{bmatrix}\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(4)} \\ +\Theta_\Phi^{(4)} \otimes \Theta_\Phi^{(3)}\end{bmatrix}\end{array}\right\}\psi_P^{(5)+} \\ &\quad + \ldots + \psi_F^{(2)t}\end{aligned} \tag{42}$$

$$\begin{aligned}\psi_P^{(3)t} &= [\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}]\psi_P^{(3)+} + \frac{1}{2!}\begin{bmatrix}\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)}\end{bmatrix}\psi_P^{(4)+} \\ &\quad + \left\{\begin{array}{l}\dfrac{1}{3!}\begin{bmatrix}\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(4)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(4)} + \Theta_\Phi^{(4)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}\end{bmatrix} \\ + \dfrac{1}{2!2!}\begin{bmatrix}\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} + \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(3)}\end{bmatrix}\end{array}\right\}\psi_P^{(5)+} + \ldots\end{aligned} \tag{43}$$

$$\begin{aligned}\psi_P^{(4)t} &= [\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}]\psi_P^{(4)+} + \frac{1}{2!}\begin{bmatrix}\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \\ +\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(3)} \\ +\Theta_\Phi^{(3)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}\end{bmatrix}\psi_P^{(5)+} \\ &\quad + \ldots\end{aligned} \tag{44}$$

$$\psi_P^{(5)t} = [\Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)} \otimes \Theta_\Phi^{(2)}]\psi_P^{(5)+} + \ldots \tag{45}$$

where

$$\psi_F^{(2)t} := \mathrm{Vec}\left(\int_{t_k}^{t}\Phi_{i_1,k_1}(t,\tau)G_{k_1,j_1}Q_{j_1,j_2}G_{k_2,j_2}\Phi_{i_2,k_2}(t,\tau)\mathrm{d}\tau\right)$$

Nonlinear Filtering: The Update Step

We now develop the vector counterparts of the update equations ((18) through (28)) detailed by the scalar problem previously. It is pointed out again that the central equation updating the state estimate and incorporating the measurement sets is set within the Kalman paradigm, following closely the developments due to Julier and Uhlman [11] and more recently Park and Scheeres [14].

The nonlinear dynamics of the plant model are given by the vector differential equations

$$\dot{x}_i = f_i(\mathbf{x}, t) + G_{ij_1} w_{j_1} \tag{46}$$

with a known distribution of initial conditions, $E(\mathbf{x}(t_0)) = \mu_0$, $E(\delta\mathbf{x}_{j_1}(t_0)\ldots\delta x_{i_n}(t_0)) = P_0^{i_1\ldots i_n}$ and being forced by a zero-mean random process given by the intensity $E(\mathbf{w}(t)\mathbf{w}^{\mathrm{T}}(\tau)) = Q\delta(t-\tau)$. For filter development, similar to the scalar situation, consider the two time instances, t_k, t_{k+1}, between which the propagation takes place. The state estimate is related to the unknown "truth" by $\mathbf{x}(t) := \hat{\mathbf{x}}(t) + \delta\mathbf{x}$ with the estimator dynamics being given by

$$\dot{\hat{x}}_i = f_i(\hat{\mathbf{x}}, t) \tag{47}$$

operating with the initial conditions as the best estimates at time step t_k, $\hat{x}_i(t_k) = \hat{x}_i^+$. With the arrival of new measurement sets and the corresponding innovations process, given by $\nu_i = \tilde{y}_i - h_i(\hat{\mathbf{x}}^-) = h_i(\mathbf{x}) - h_i(\hat{\mathbf{x}}^-) + v_i$, the state update equation is given by

$$\begin{aligned} \hat{x}_i^+ &= \hat{x}_i^- + K_{ij}\nu_j \\ P_{ij}^{(2)+}(t_k) &= P_{ij}^{(2)-}(t_k) - K_{ij_1} P_{j_1 j_2}^{\nu\nu} K_{j_2 j} \end{aligned} \tag{48}$$

where the Kalman gain is computed using the equation

$$K = P^{x\nu}(P^{\nu\nu})^{-1} \tag{49}$$

with the corresponding statistical correlation matrices being given by

$$\begin{aligned} (P_k^{x\nu})_{i,j} &= E[\delta x_i(\tilde{y}_j - \hat{y}_j)] \\ (P_k^{\nu\nu})_{i,j} &= E[(\tilde{y}_i - \hat{y}_i)(\tilde{y}_j - \hat{y}_j)] \end{aligned} \tag{50}$$

To compute these correlation matrices, the measurement equations are expanded in a Taylor series about the propagated state estimate, given by

$$\begin{aligned} \nu_i(t_k) &= y_i - \hat{y}_i^- + v_i = h_i(\mathbf{x}_k, t_k) - h_i(\hat{\mathbf{x}}_k^-, t_k) + v_i \\ &\approx h_{i,j_1}\delta x_{j_1} + \frac{1}{2} h_{i,j_1 j_2}\delta x_{j_1}\delta x_{j_2} + \ldots + v_i \end{aligned} \tag{51}$$

This helps us express the correlations as functions of the statistics of the state estimate and the measurement sensitivities evaluated about the propagated state as

$$\begin{aligned} (P_k^{x\nu})_{i,j} &= E\left[\delta x_i\left(h_{j,h}\delta x_{j_1} + \frac{1}{2}h_{j,j_1 j_2}\delta x_{j_1}\delta x_{j_2} + \ldots + v_j\right)\right] \\ &= h_{j,j_1}P_{ij_1}^{(2)t} + \frac{1}{2}h_{j,j_1 j_2}P_{ij_1 j_2}^{(3)t} + \frac{1}{3!}h_{j,j_1 j_2 j_3}P_{ij_1 j_2 j_3}^{(4)t} + \ldots + \frac{1}{N!}h_{j,j_1 j_2\ldots j_N}P_{i,j_1\ldots j_N}^{(N+1)t} \end{aligned} \tag{52}$$

$$\begin{aligned} (P_k^{\nu\nu})_{i,j} &= E\left[\left(h_{i,i_1}\delta x_{i_1} + \frac{1}{2}h_{i,i_1 i_2}\delta x_{i_1}\delta x_{i_2} + \ldots + v_i\right)(h_{j,j_1}\delta x_{j_1} + \ldots + v_j)\right] \\ &= h_{i,i_1}P_{i_1 j_1}^{(2)t}h_{j,j_1} + \ldots + \frac{1}{p!q!}h_{i,i_1 i_2\ldots i_p}P_{i_1\ldots i_p j_1\ldots j_q}^{(p+q)t}h_{j,j_1\ldots j_q} + \ldots + R_{i,j} \end{aligned} \tag{53}$$

Using the vector–matrix representation of tensors setup in the paper, the above equations take the form

$$\psi^{(2)t}_{P_k^{x\nu-}} := \mathrm{Vec}\,(P_k^{x\nu-}) = [\Theta_h^{(2)} \otimes I_N]\psi_P^{(2)t} + \frac{1}{2!}[\Theta_h^{(3)} \otimes I_N]\psi_P^{(3)t} + \frac{1}{3!}(\Theta_h^{(4)} \otimes I_N)\psi_P^{(4)t} + \ldots \tag{54}$$

$$\begin{aligned}\psi^{(2)t}_{P^{\nu\nu}} :&= \mathrm{Vec}\,(P^{\nu\nu}) = [\Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_P^{(2)t} + \frac{1}{2!}\begin{bmatrix}\Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \\ +\Theta_h^{(2)t} \otimes \Theta_h^{(3)t}\end{bmatrix}\psi_P^{(3)t} + \\ &\left(\frac{1}{3!}\begin{bmatrix}\Theta_h^{(4)t} \otimes \Theta_h^{(2)t} \\ +\Theta_h^{(2)t} \otimes \Theta_h^{(4)t}\end{bmatrix} + \frac{1}{2!2!}[\Theta_h^{(3)t} \otimes \Theta_h^{(3)t}]\right)\psi_P^{(4)t} \\ &\left(\frac{1}{2!3!}\begin{bmatrix}\Theta_h^{(3)} \otimes \Theta_h^{(4)} \\ +\Theta_h^{(4)} \otimes \Theta_h^{(3)}\end{bmatrix} + \frac{1}{4!}\begin{bmatrix}\Theta_h^{(5)} \otimes \Theta_h^{(2)} \\ +\Theta_h^{(2)} \otimes \Theta_h^{(5)}\end{bmatrix}\right)\psi_P^{(5)t} + \ldots + \mathrm{Vec}(R)\end{aligned} \tag{55}$$

where the matrix representation $\Theta_h^{(k)t} := \begin{bmatrix} h_{1,1\ldots(k-1)terms\ldots1} & \ldots & h_{1,n\ldots n} \\ h_{2,1\ldots(k-1)terms\ldots1} & & h_{2,n\ldots n} \\ \vdots & \ddots & \vdots \\ h_{m,1\ldots(k-1)terms\ldots1} & \ldots & h_{m,n\ldots n}\end{bmatrix} \in \mathbb{R}^{m \times n^{k-1}}$ of the high-order partial derivatives of the measurement equations has been used. Note that the notation allows non-homogeneity of the associated index sets (first index varies from $1\ldots m$, while the others vary from $1\ldots n$). The compatibility of the dimensions and the validity of the expressions can be verified proving the generality of the underlying structure being presented herein.

Using this representation, we further develop the vector generalizations of the equations involving high-order statistical moment updates *(a posteriori* error moments–scalar case equations (18–28)). These definitions are dependent upon high-order correlations between the innovations process and the estimation error statistics.

Third-order correlations of the estimation error with the innovations process are given by the tensor equations

$$\begin{aligned} P^{\nu xx}_{i_1i_2i_3} &= h_{i_1,j_1}P^{(3)t}_{j_1i_2i_3} + \frac{1}{2!}h_{i_1,j_1j_2}P^{(4)t}_{j_1j_2i_2i_3} + \frac{1}{3!}h_{i_1,j_1j_2j_3}P^{(5)t}_{j_1j_2j_3i_2i_3} + \ldots \\ P^{\nu\nu x}_{i_1i_2i_3} &= h_{i_1,j_1}h_{i_2,j_2}P^{(3)t}_{j_1j_2j_3} + \frac{1}{2!}(h_{i_1,j_1j_2}h_{i_2,j_3} + h_{i_1,j_1}h_{i_2,j_2j_3})P^{(4)t}_{j_1j_2j_3i_3} + \\ &\left(\frac{1}{3!}[h_{i_1,j_1j_2j_3}h_{i_2,j_4} + h_{i_1,j_1}h_{i_2,j_2j_3j_4}] + \frac{1}{2!2!}[h_{i_1,j_1j_2}h_{i_2,j_3j_4}]\right)P^{(5)t}_{j_1j_2j_3j_4i_3} + \ldots \\ P^{\nu\nu\nu}_{i_1i_2i_3} &= h_{i_1,j_1}h_{i_2,j_2}h_{i_3,j_3}P^{(3)t}_{j_1j_2j_3} + \frac{1}{2!}\begin{pmatrix}h_{i_1,j_1j_2}h_{i_2,j_3}h_{i_3,j_4} + h_{i_1,j_1}h_{i_2,j_2j_3}h_{i_3,j_4} \\ +h_{i_1,j_1}h_{i_2,j_2}h_{i_3,j_3j_4}\end{pmatrix}P^{(4)t}_{j_1j_2j_3j_4} + \ldots\end{aligned} \tag{56}$$

Expressing the same in the vector matrix representation for tensors, we get the linear algebraic equations

$$\begin{aligned}\psi_{p^{\nu xx(t)}} &= [I \otimes I \otimes \Theta_h^{(2)t}]\psi_{p^{(3)t}} + \frac{1}{2!}[I \otimes I \otimes \Theta_h^{(3)t}]\psi_{p^{(4)t}} \\ &+ \frac{1}{3!}[I \otimes I \otimes \Theta_h^{(4)t}]\psi_{P^{(5)t}} + \ldots\end{aligned}$$

$$\psi_{P^{\nu\nu x(t)}} = [I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_{P^{(3)t}} + \frac{1}{2!}\begin{pmatrix}[I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[I \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}]\end{pmatrix}\psi_{P^{(4)t}} + \left(\frac{1}{3!}\begin{pmatrix}[I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(4)t}] \\ +[I \otimes \Theta_h^{(4)t} \otimes \Theta_h^{(2)t}]\end{pmatrix} + \frac{1}{2!2!}[I \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(3)t}]\right)\psi_{P^{(5)t}} + \ldots \tag{57}$$

$$\psi_{P^{\nu\nu\nu(t)}} = [\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_{P^{(3)t}} + \frac{1}{2!}\begin{pmatrix}[\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[\Theta_h^{(2)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}] \\ +[\Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\end{pmatrix}\psi_{P^{(4)t}} + \left(\frac{1}{3!}\begin{pmatrix}[\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(4)t}] \\ +[\Theta_h^{(2)t} \otimes \Theta_h^{(4)t} \otimes \Theta_h^{(2)t}] \\ +[\Theta_h^{(4)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\end{pmatrix} + \frac{1}{2!2!}\begin{pmatrix}[\Theta_h^{(2)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(3)t}] \\ +[\Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[\Theta_h^{(3)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}]\end{pmatrix}\right)\psi_{P^{(5)t}} + \ldots$$

These correlation terms are used in updating the third-order moments. This third-order moment update equation, derived from first principles just as in the scalar case, is given by

$$\begin{aligned} P_{i_1 i_2 i_3}^{(3)+}(t_k) &= E(\delta x_{i_1}^{+} \delta x_{i_2}^{+} \delta x_{i_3}^{+}] \\ &= P_{i_1 i_2 i_3}^{(3)+} - K_{i_1 j_1} P_{j_1 i_2 i_3}^{\nu xx} - K_{i_2 j_2} P_{i_1 j_2 j_3}^{x\nu x} - K_{i_3 j_3} P_{i_1 i_2 j_3}^{xx\nu} \\ &\quad + K_{i_1 j_1} K_{i_2 j_2} P_{j_1 j_2 i_3}^{\nu\nu x} + K_{i_1 j_1} K_{i_3 j_3} P_{j_1 i_2 j_3}^{\nu x\nu} + K_{i_2 j_2} K_{i_3 j_3} P_{i_1 j_2 j_3}^{x\nu\nu} \\ &\quad - K_{i_1 j_1} K_{i_2 j_2} K_{i_3 j_3} P_{j_1 j_2 j_3}^{\nu\nu\nu} \end{aligned} \tag{58}$$

When expressed in the vector matrix representation, these equations are written as

$$\begin{aligned} \psi_{P^{(3)+}} = \psi_{P^{(3)t}} &- [I \otimes I \otimes K]\psi_{P^{\nu xx}} - [I \otimes K \otimes I]\psi_{P^{x\nu x}} \\ &- [K \otimes I \otimes I]\psi_{P^{xx\nu}} + [I \otimes K \otimes K]\psi_{P^{\nu\nu x}} + [K \otimes I \otimes K]\psi_{P^{\nu x\nu}} \\ &+ [K \otimes K \otimes I]\psi_{P^{x\nu\nu}} + [K \otimes K \otimes K]\psi_{P^{\nu\nu\nu}} \end{aligned} \tag{59}$$

Substituting correlation equations derived previously (57) into the above set (59) and using the identity involving Kronecker products (64), the third-order moment update equations can be directly given as

$$\begin{aligned} \psi_{P^{(3)+}} = \psi_{P^{(3)t}} &- \psi_{Z_1^{(3)}} + \psi_{Z_2^{(3)}} \\ &- [K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\psi_{P^{(3)t}} \\ &- \frac{1}{2!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\end{pmatrix}\psi_{P^{(4)t}} \\ &- \begin{pmatrix} -\frac{1}{3!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(4)t}] \\ +[K\Theta_h^{(2)t} \otimes K\Theta_h^{(4)t} \otimes K\Theta_h^{(2)t}] \\ +[K\Theta_h^{(4)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\end{pmatrix} + \\ \frac{1}{2!2!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}]\end{pmatrix}\end{pmatrix}\psi_{P^{(5)t}} + \ldots \end{aligned} \tag{60}$$

where $\psi_{Z_1^{(3)}}$, $\psi_{Z_2^{(3)}}$, are given by

$$
\begin{aligned}
\psi_{Z_1^{(3)}}: = & [I \otimes I \otimes K\Theta_h^{(2)t}]\psi_{P(3)t} + \frac{1}{2!}[I \otimes I \otimes K\Theta_h^{(3)t}]\psi_{P(4)t} \\
& + \frac{1}{3!}[I \otimes I \otimes K\Theta_h^{(4)t}]\psi_{P(5)t} \\
& + [I \otimes K\Theta_h^{(2)t} \otimes I]\psi_{P(3)t} + \frac{1}{2!}[I \otimes K\Theta_h^{(3)t} \otimes I]\psi_{P(4)t} \\
& + \frac{1}{3!}[I \otimes K\Theta_h^{(4)t} \otimes I]\psi_{P(5)t} \\
& + [K\Theta_h^{(2)t} \otimes I \otimes I]\psi_{P(3)t} + \frac{1}{2!}[K\Theta_h^{(3)t} \otimes I \otimes I]\psi_{P(4)t} \\
& + \frac{1}{3!}[K\Theta_h^{(4)t} \otimes I \otimes I]\psi_{P(5)t} + \dots
\end{aligned}
\tag{61}
$$

$$
\begin{aligned}
\psi_{Z_2^{(3)}}: = & [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\psi_{p(3)t} + \frac{1}{2!}\begin{pmatrix} [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}] \end{pmatrix}\psi_{p(4)t} \\
& + \begin{bmatrix} \frac{1}{3!}\begin{pmatrix} [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(4)t}) \\ +[I \otimes K\Theta_h^{(4)t} \otimes K\Theta_h^{(2)t}] \end{pmatrix} \\ +\frac{1}{2!2!}[I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(3)t}] \end{bmatrix}\psi_{P(5)t} + \dots \\
& + [K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(2)t}]\psi_{P(3)t} + \frac{1}{2!}\begin{pmatrix} [K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(3)t} \otimes I \otimes K\Theta_h^{(2)t}] \end{pmatrix}\psi_{P(4)t} \\
& + \begin{bmatrix} \frac{1}{3!}\begin{pmatrix} [K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(4)t}) \\ +[K\Theta_h^{(4)t} \otimes I \otimes K\Theta_h^{(2)t}] \end{pmatrix} \\ +\frac{1}{2!2!}[K\Theta_h^{(3)t} \otimes I \otimes K\Theta_h^{(3)t}] \end{bmatrix}\psi_{P(5)t} + \dots \\
& + [K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes I]\psi_{P(3)t} + \frac{1}{2!}\begin{pmatrix} [K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes I] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes I] \end{pmatrix}\psi_{P(4)t} \\
& + \begin{bmatrix} \frac{1}{3!}\begin{pmatrix} [K\Theta_h^{(2)t} \otimes K\Theta_h^{(4)t} \otimes I) \\ +[K\Theta_h^{(4)t} \otimes K\Theta_h^{(2)t} \otimes I] \end{pmatrix} \\ +\frac{1}{2!2!}[I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(3)t}] \end{bmatrix}\psi_{P(5)t} + \dots
\end{aligned}
\tag{62}
$$

$$
(A \otimes B)(C \otimes D) = AC \otimes BD \tag{63}
$$

Following an analogous methodology, the fourth order correlation tensors are given by

$$
\psi_{P^{\nu xx(t)}} = [I \otimes I \otimes I \otimes \Theta_h^{(2)t}]\psi_{P(4)t} + \frac{1}{2!}[I \otimes I \otimes I \otimes \Theta_h^{(3)t}]\psi_{P(5)t} + \dots
$$

$$
\begin{aligned}
\psi_{P^{\nu\nu xx(t)}} = & [I \otimes I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_{p(4)t} + \frac{1}{2!}\begin{pmatrix} [I \otimes I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[I \otimes I \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}] \end{pmatrix}\psi_{P(5)t} \\
& + \dots + \psi_{P(2)t} \otimes \psi_R
\end{aligned}
$$

$$\psi_{P\nu\nu\nu x(t)} = [I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_{P^{(4)t}} + \frac{1}{2!}\begin{pmatrix} [I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[I \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}] \\ +[I \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}] \end{pmatrix}\psi_{P^{(5)t}} + \ldots \tag{64}$$

$$\begin{aligned}\psi_{P\nu\nu\nu\nu(t)} = {} & [\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}]\psi_{P^{(4)t}} \\ & + \frac{1}{2!}\begin{pmatrix} [\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t}] \\ +[\Theta_h^{(2)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t}] \\ +[\Theta_h^{(2)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}] \\ +[\Theta_h^{(3)t} \otimes \Theta_h^{(3)t} \otimes \Theta_h^{(2)t} \otimes \Theta_h^{(2)t}] \end{pmatrix}\psi_{P^{(5)t}} + \ldots \\ & + \psi_R \otimes \psi_R + (\psi_R \otimes \psi_R)^{\pi[1\ 3\ 2\ 4]} + (\psi_R \otimes \psi_R)^{\pi[1\ 4\ 2\ 3]}\end{aligned}$$

The update equation for the fourth moment follows as (in the vector–matrix representation setup earlier)

$$\begin{aligned}\psi_{P^{(4)+}} = {} & \psi_{P^{(4)t}} - [I \otimes I \otimes I \otimes K]\psi_{P^{\nu xxx}} - [I \otimes I \otimes K \otimes I]\psi_{P x\nu xx} \\ & - [I \otimes K \otimes I \otimes I]\psi_{P^{xx\nu x}} - [K \otimes I \otimes I \otimes I]\psi_{P^{xxx\nu}} \\ & + [I \otimes I \otimes K \otimes K]\psi_{P^{\nu\nu xx}} + [I \otimes K \otimes I \otimes K]\psi_{P^{\nu x\nu x}} + [K \otimes I \otimes I \otimes K]\psi_{P^{\nu xx\nu}} \\ & + [I \otimes K \otimes K \otimes I]\psi_{P^{x\nu\nu x}} + [K \otimes I \otimes K \otimes I]\psi_{P^{x\nu x\nu}} + [K \otimes K \otimes I \otimes I]\psi_{P^{xx\nu\nu}} \\ & - [I \otimes K \otimes K \otimes K]\psi_{P^{\nu\nu\nu x}} - [K \otimes K \otimes I \otimes K]\psi_{P^{\nu x\nu\nu}} \\ & - [K \otimes I \otimes K \otimes K]\psi_{P^{\nu\nu x\nu}} - [K \otimes K \otimes K \otimes I]\psi_{P^{x\nu\nu\nu}} \\ & + [K \otimes K \otimes K \otimes K]\psi_{P^{\nu\nu\nu\nu}}\end{aligned} \tag{65}$$

Substituting the vectorized tensor moment equations from (64) and using the identity (63), we get the fourth-order update equation to take the form

$$\begin{aligned}\psi_{P^{(4)+}} = {} & \psi_{P^{(4)t}} - \psi_{Z_1^{(4)}} + \psi_{Z_2^{(4)}} \\ & - [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\psi_{P^{(4)t}} \\ & - \frac{1}{2!}\begin{pmatrix} [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}] \\ +[I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}] \end{pmatrix}\psi_{P^{(5)t}} \\ & - \text{permutation terms (3 in number)} \\ & + [K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\psi_{P^{(4)t}} \\ & + \frac{1}{2!}\begin{pmatrix} [K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}] \\ +[K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}] \end{pmatrix}\psi_{P^{(5)t}} + \ldots \\ & + [K \otimes K \otimes K \otimes K](\psi_R \otimes \psi_R + (\psi_R \otimes \psi_R)^{\pi[1\ 3\ 2\ 4]} + (\psi_R \otimes \psi_R)^{\pi[1\ 4\ 2\ 3]})\end{aligned} \tag{66}$$

with the expressions $\psi_{Z_1^{(4)}}$, $\psi_{Z_2^{(4)}}$ being defined as

$$\begin{aligned}\psi_{Z_1^{(4)}} := {} & [I \otimes I \otimes I \otimes K\Theta_h^{(2)t}]\psi_{P^{(4)t}} + \frac{1}{2!}[I \otimes I \otimes I \otimes K\Theta_h^{(3)t}]\psi_{P^{(5)t}} \\ & + [I \otimes I \otimes K\Theta_h^{(2)t} \otimes I]\psi_{P^{(4)t}} + \frac{1}{2!}[I \otimes I \otimes K\Theta_h^{(3)t} \otimes I]\psi_{P^{(5)t}}\end{aligned}$$

$$+ [I \otimes K\Theta_h^{(2)t} \otimes I \otimes I]\psi_{P(4)t} + \frac{1}{2!}[I \otimes K\Theta_h^{(3)t} \otimes I \otimes I]\psi_{P(5)t}$$
$$+ [K\Theta_h^{(2)t} \otimes I \otimes I \otimes I]\psi_{P(4)t} + \frac{1}{2!}[K\Theta_h^{(3)t} \otimes I \otimes I \otimes I]\psi_{P(5)t} \tag{67}$$

and

$$\begin{aligned}
\psi_{Z_2^{(4)}} :=& [I \otimes I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t}]\psi_{P(4)t} \\
&+ \frac{1}{2!}\begin{pmatrix}[I \otimes I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t}] \\ +[I \otimes I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t}]\end{pmatrix}\psi_{P(5)t} \\
&+ [I \otimes K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(2)t}]\psi_{P(4)t} + \frac{1}{2!}\begin{pmatrix}[I \otimes K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(3)t}] \\ +[I \otimes K\Theta_h^{(3)t} \otimes I \otimes K\Theta_h^{(2)t}]\end{pmatrix}\psi_{P(5)t} \\
&+ [K\Theta_h^{(2)t} \otimes I \otimes I \otimes K\Theta_h^{(2)t}]\psi_{P(4)t} + \frac{1}{2!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes I \otimes I \otimes K\Theta_h^{(3)t}] \\ +[K\Theta_h^{(3)t} \otimes I \otimes I \otimes K\Theta_h^{(2)t}]\end{pmatrix}\psi_{P(5)t} \\
&+ [I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes I]\psi_{P(4)t} + \frac{1}{2!}\begin{pmatrix}[I \otimes K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes I] \\ +[I \otimes K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes I]\end{pmatrix}\psi_{P(5)t} \\
&+ [K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(2)t} \otimes I]\psi_{P(4)t} + \frac{1}{2!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes I \otimes K\Theta_h^{(3)t} \otimes I] \\ +[K\Theta_h^{(3)t} \otimes I \otimes K\Theta_h^{(2)t} \otimes I]\end{pmatrix}\psi_{P(5)t} \\
&+ [K\Theta_h^{(2)t} \otimes K\Theta_h^{(2)t} \otimes I \otimes I]\psi_{P(4)t} + \frac{1}{2!}\begin{pmatrix}[K\Theta_h^{(2)t} \otimes K\Theta_h^{(3)t} \otimes I \otimes I] \\ +[K\Theta_h^{(3)t} \otimes K\Theta_h^{(2)t} \otimes I \otimes I]\end{pmatrix}\psi_{P(5)t} \\
&+ [I \otimes I \otimes K \otimes K](\psi_{P(2)t} \otimes \psi_{R)} + [I \otimes K \otimes I \otimes K] + (\psi_{P(2)t} \otimes \psi_R)^{\pi[1\ 3\ 2\ 4]} \\
&+ [K \otimes I \otimes I \otimes K](\psi_{P(2)t} \otimes \psi_R)^{\pi[2\ 3\ 4\ 1]} \\
&+ [K \otimes K \otimes I \otimes I](\psi_R \otimes \psi_{P(2)t}) + [K \otimes I \otimes K \otimes I] + (\psi_R \otimes \psi_{P(2)t})^{\pi[1\ 3\ 2\ 4]} \\
&+ [I \otimes K \otimes K \otimes I](\psi_R \otimes \psi_{P(2)t})^{\pi[2\ 3\ 4\ 1]}
\end{aligned} \tag{68}$$

Similarly, the fifth-order update equation can also be written if the appropriate state transition tensors are available after integration through the nominal path. At this stage of the development, it is clear that there is a certain structure to the equations governing the departure motion of the trajectory from the given nominal path (in this case, the best estimate, with the arrival of new measurements). In general, we would need to track an infinite number of such moments to describe the evolution of the uncertainty, which in turn helps us to get a better estimate for a given problem. However, as a practical matter, it is impossible to track infinite moments. In the current development, we follow along the lines of Athans [12] and Gelb [10] in defining the order. Athans develops a filter, which retains the second-order terms in the Taylor series expansion of the measurement and dynamics models of the system, leading to consideration of high-order terms up to second order in covariance calculations. Skewness and higher moments are assumed to be small in that development and the filter thus derived is appropriately called the second-order filter. We, in addition to maintaining the high-order terms (in the measurement model and the dynamics model of the system), perform high-order moment propagation and updates up to fourth order (although explicit expressions are obtained up to fourth order, the structures presented here generalize up to Nth order for a finite N). The question of how many terms to be retained is an important one and we postpone the issue to a future correspondence, where schemes to legitimize truncation will be presented as the central topic of discussion. The central aim of this paper is to report the existence of simple data structures and methods to perform high-order moment calculations in a computationally and analytically elegant and tractable format. The next section reports the application of such architecture to a problem of orbit estimation.

Application: Orbit Estimation

The J*th* Moment Extended Kalman Filter (JMEKF) architecture presented in the paper is applied to the problem of estimation of position and velocity of a particle in orbit from range, azimuth and elevation measurements. True measurements are generated by considering the solution of the two body problem of the corresponding true orbit parameters.

The equations of motion are given by

$$\ddot{\mathbf{r}} = -\frac{\mu}{\|\mathbf{r}\|^3}\mathbf{r} \tag{69}$$

Range, azimuth and elevation measurements as taken as

$$\begin{aligned} \|\rho\| &= \sqrt{\rho_u^2 + \rho_e^2 + \rho_n^2} \\ az &= \tan^{-1}\left(\frac{\rho_e}{\rho_n}\right) \\ el &= \tan^{-1}\left(\frac{\rho_u}{\sqrt{\rho_n^2 + \rho_e^2}}\right) \end{aligned} \tag{70}$$

A standard up, north, east coordinate system was assumed at the site location for simulated measurements and the computation of the measurement sensitivities. Measurements were assumed to be taken from College Station, Texas ($\phi_{gd} = 30.588°$, $\theta_0 = -96.3638°$). The coordinate definitions and the measurement system is sketched in the schematic in Fig. 2. The orbit considered has an eccentricity $e = 0.8$, with a perigee altitude $r_p = 300$ km. Inclination of the orbit plane is assumed at $i = 5°$, with the longitude of ascending node at $\Omega = 45°$ and the argument of perigee $\omega = 30°$.

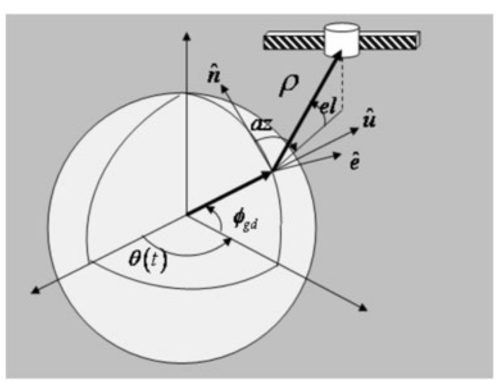

FIG. 2. Observations Setup and Coordinate Definitions for Orbit Estimation Problem.

Two representative situations were considered for demonstration of the effectiveness of the filter. In most circumstances, it is reasonable to expect the level of accuracy of the measurements in state estimates. For instance, if the sensor resolution is on the order of a few kilometers, it is unreasonable to expect the filtered results to have accuracy of order of a few centimeters. In the same spirit, we consider two sets of measurements, one of them having an order of magnitude greater resolution than the other and compare the resulting filtered estimates. Range is fixed to have a resolution of $\sigma_{\text{range}} = 0.025$ km. While the first set of measurements has azimuth and elevation with standard deviations $\sigma_{\text{azm}} = \sigma_{\text{ele}} = 0.015°$, the second set of measurements is assumed to have a lower noise level with $\sigma_{\text{azm}} = \sigma_{\text{ele}} = 0.0015°$. This order of magnitude difference is anticipated in the estimation errors of filtered states.

To start the filter, it is assumed that a geometry-based calculation gives initial state estimates and the corresponding covariance estimates. Gaussian Least Square Differential Corrections (GLSDC) is an important and highly efficient tool to give such initial estimates. Alternate methods based on Kalman filtering do exist which smooth the solution over the short span of initial time period. Refer to Chapters 4 and 7 of [8] for a detailed discussion and example applications of these classical methods of initial orbit determination. The levels of initial position and velocity error are set at least an order of magnitude worse when compared to the "most common" levels of convergence of the GLSDC. In this particular situation, $\mathbf{x}(t_0) = \mathbf{x_{true}} + [1 \quad -1.1 \quad 1.4]^{\mathrm{T}}\,\text{km}$ and $\mathbf{v}(t_0) = \mathbf{v_{true}} + [0.1 \quad 0.2 \quad 0.1]^{\mathrm{T}}\,\text{km/s}$ were used, which are approximately an order of magnitude worse than the 1σ of a GLSDC converged ball of position initial conditions (Example 4.3 of [8]) and almost two orders of magnitude worse than the 1σ of a GLSDC converged ball of velocity initial conditions. The levels of initial uncertainty are set to be $P_{\delta x \delta x}(t_0) = 25 \times \mathrm{I}_3\,\text{km}^2$, $P_{\delta v \delta v}(t_0) = 10^{-2}\mathrm{I}_3(\text{km/sec})^2$. As a consequence we can study the nonlinear effects of the propagation phase of the filter operating in the absence of process noise, which was noted to spoil the true nonlinear effects of the process. Also, we have to consider the fact that only a fraction ($1/3^{\text{rd}}$) of measurements was made available for estimation.

The measurements are made so sparse as to make the classical Extended Kalman Filter, running with the same simulation parameters operate in a fragile manner, in the presence of low process noise (Case 1 has $Q_{\text{LowRes}} = 9.407 \times 10^{-18}\text{N}^2/\text{sec}$ and Case 2 has $Q_{\text{HighRes}} = 9.407 \times 10^{-18}\text{N}^2/\text{sec}$). This is to amplify the nonlinear effects of the problem and further study the filter operation as it makes use of the nonlinear effects in the propagation phase affecting the decision process after considerable evolution. Considering these issues the update rate of the filter was fixed to be about $\Delta t \approx 2400\,\text{sec}\,(40\,\text{min})$ for both situations. Due to the use of non-dimensional canonical length units (RE–Earth Radii) and time units $\left(\text{TU} = \sqrt{\frac{\mu_\oplus}{\text{RE}^3}}\right.$, see Appendix A of [26]), some of the process noise values have been stated using more than a few digits in precision. These are summarized in Table 1.

Synthetic measurements were generated using simulation parameters detailed in Table 1 for both the cases. The JMEKF and classical Extended Kalman Filter schemes were subsequently used to produce state estimates based on the measurement sets thus generated. The process was repeated 100 times and the filtered estimates were collected with corresponding filters operating on a different set of measurements each time. Such a Monte Carlo simulation (with sample size = 100

TABLE 1. Summary of the Simulation Parameters for JMEKF vs. EKF Comparison

Simulation Parameter	Case 1 (Fine Pointing)	Case 2 (Rough Pointing)
Filter Initial Condition	$x(t_0) = \mathbf{x}_{true} + [1 \quad -1.1 \quad 1.4]^{\mathrm{T}}$ km $\mathbf{v}(t_0) = \mathbf{v}_{true} + [0.1 \quad -0.2 \quad 0.1]^{\mathrm{T}}$ km/sec	$x(t_0) = \mathbf{x}_{true} + [1 \quad -1.1 \quad 1.4]^{\mathrm{T}}$ km $\mathbf{v}(t_0) = \mathbf{v}_{true} + [0.1 \quad -0.2 \quad 0.1]^{\mathrm{T}}$ km/sec
Initial Covariance	$P_{\delta x \delta x}(t_0) = 25 \times I_3$ km^2, $P_{\delta v \delta v}(t_0) = 10^{-1} I_3$ km^2/sec^2	$P_{\delta x \delta x}(t_0) = 25 \times I_3$ km^2, $P_{\delta v \delta v}(t_0) = 10^{-1} I_3$ km^2/sec^2
Measurement Noise Covariances	$\sigma_{\mathrm{range}} = 25$ m $\sigma_{\mathrm{az}} = 0.0015^0$ $\sigma_{\mathrm{ele}} = 0.0015^0$	$\sigma_{\mathrm{range}} = 25$ m $\sigma_{\mathrm{az}} = 0.015^0$ $\sigma_{\mathrm{ele}} = 0.015^0$
Process Noise Covariances	$Q_{\mathrm{HighRes}} = 9.407 \times 10^{-18}\, \mathrm{I}_3\, \mathrm{N}^2/\mathrm{sec}$	$Q_{\mathrm{LowRes}} = 9.407 \times 10^{-18}\, \mathrm{I}_3\, \mathrm{N}^2/\mathrm{sec}$

in our case) is expected to bring out the most likely behavior of the filter in operation. Conclusions based on the statistics of the Monte Carlo results are known to be representative of the general trend of the filter operation, in most situations. Third-order state transition tensor description of the departure motion (expansion of Taylor series of the dynamic system) was used in the filter architecture, along with a second-order expansion of the measurement equation. A computational artifact of using updated covariance in the subsequent moment update equations (skewness and kurtosis) was found to improve the performance of the filter and hence was employed. It was found to be equivalent to regrouping neglected terms in the tensor update equations. This technique was also found to enhance the radius of convergence of Taylor series in several example problems.

The average position and velocity errors incurred by the JMEKF along with corresponding covariance bounds (3σ) calculated by the framework are plotted in Figs. 3 and 9 for measurement sets generated using parameters listed as Case 1. The errors incurred by EKF with the same simulation parameters are plotted in Figs. 5 and 6. We note in passing that the EKF estimation errors being plotted here-in are for reference purposes to demonstrate the operating conditions of the JMEKF. In practice, the EKF is "tuned" and this tuning process yields considerably different simulation parameters (especially process noise and initial condition covariance values). State estimation errors and the corresponding covariance bounds for measurement sets corresponding to Case 2 of JMEKF are plotted in Figs. 7 and 8. EKF results for the corresponding Case 2 are plotted in Figs. 9 and 10. A comparison of errors incurred by both filters is shown in a log scale in Fig. 11 (position) and Fig. 12 (velocity) for Case 1 and in Fig. 13 (position) and Fig. 14 (velocity) for Case 2. Note that the dotted values are updated instances while the solid line indicates the propagated values of the best estimates through the corresponding true nonlinear models for computation of residuals and corresponding methods of covariance propagation.

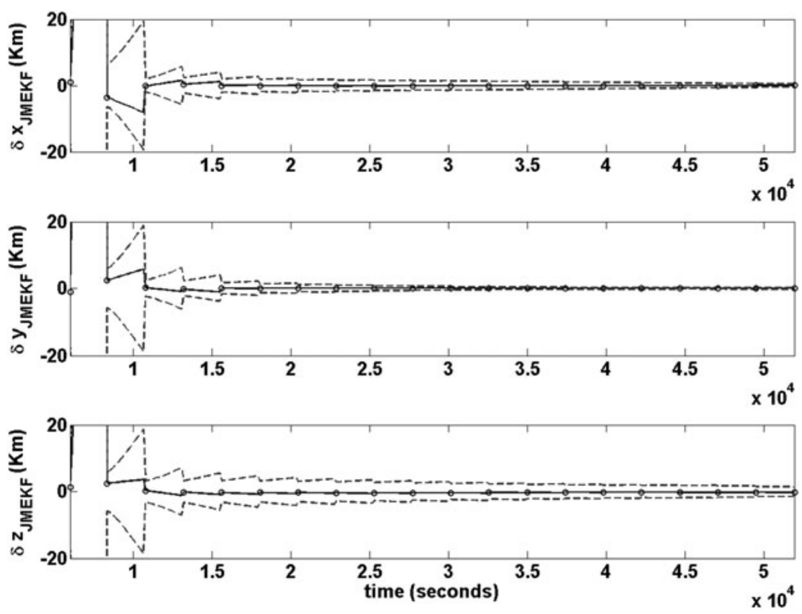

FIG. 3. Position Estimation Error: JMEKF (Case 1).

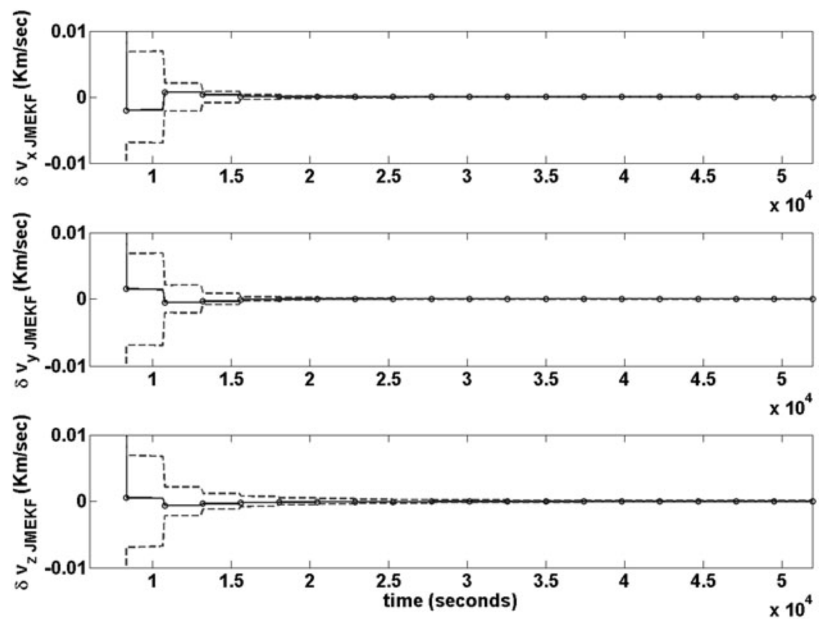

FIG. 4. Velocity Estimation Errors: JMEKF (Case 1).

The Monte Carlo simulations gave rise to a wealth of additional information. State estimates from the JMEKF and the EKF were used in the nonlinear measurement model to calculate the *a posteriori* range, azimuth and elevation estimates $\hat{\rho}|_{\hat{x}^+} := \rho^+$, $\widehat{az}|_{\hat{x}^+} := az^+$, $\widehat{el}|_{\hat{x}^+} := el^+$. The corresponding output errors were determined and the sample mean was calculated (using the sample mean

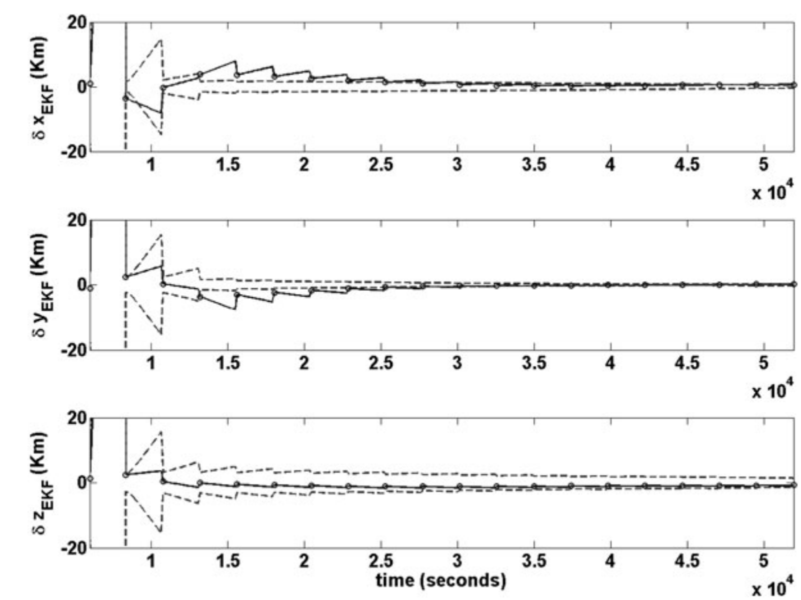

FIG. 5. Position Estimation Errors: EKF (Case 1).

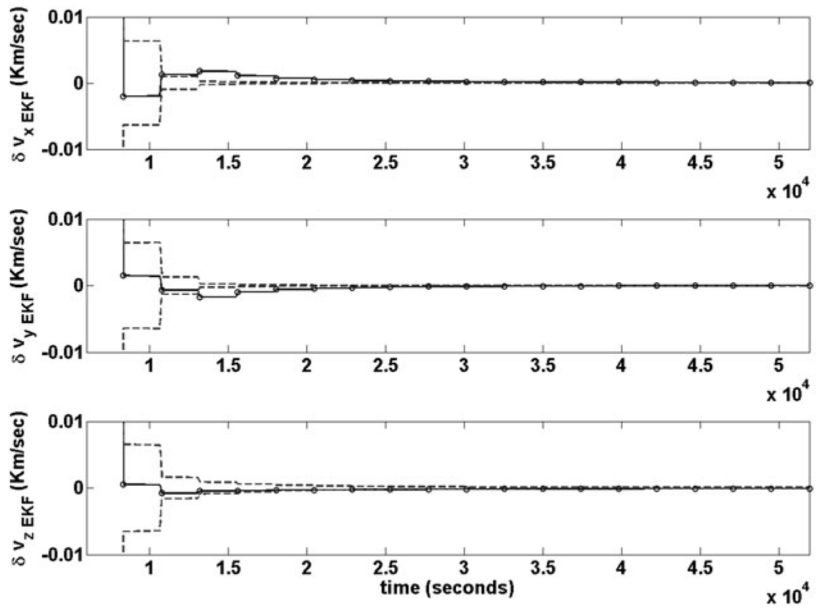

FIG. 6. Velocity Estimation Errors: EKF (Case 1).

statistic $\mu_x = \frac{1}{N}\sum_{i=1}^{N} X_i$). An estimation framework is expected to perform well if the reconstructed measurements are of the same order of magnitude of the standard deviations of the original measurement noise level (additive) most of the time (on an "average"). The comparison of the achieved average errors incurred by the estimated

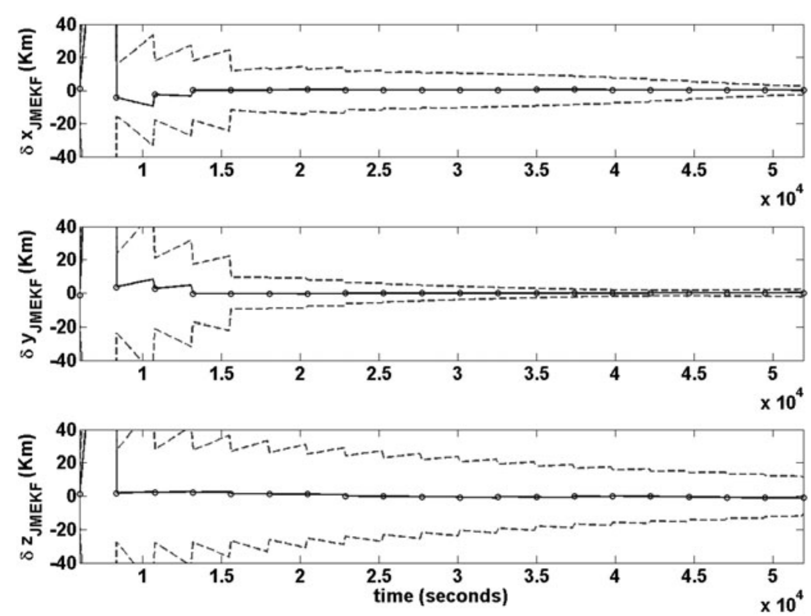

FIG. 7. Position Estimation Error: JMEKF (Case 2).

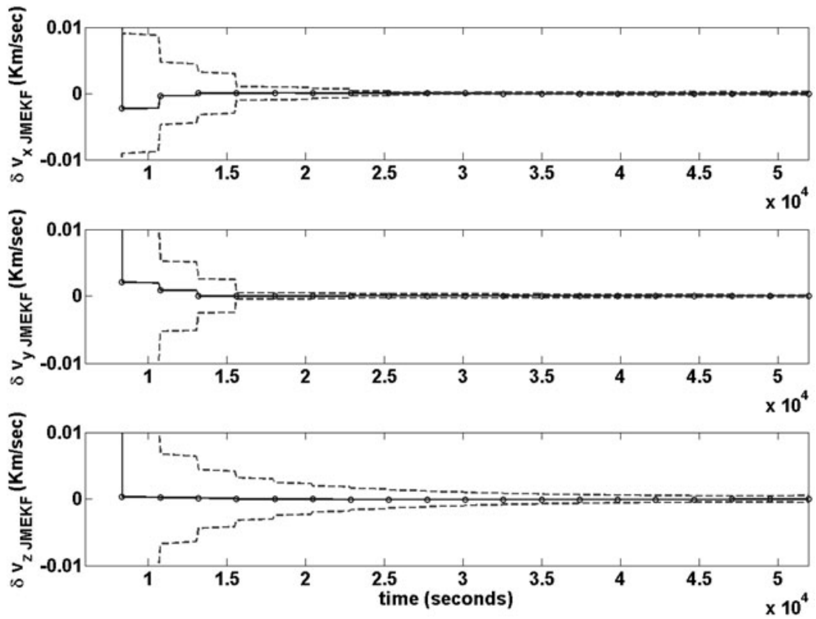

FIG. 8. Velocity Estimation Error: JMEKF (Case 2).

measurements reconstructed from JMEKF and EKF estimates are now presented. Case 1 error comparisons for range azimuth and elevation errors are plotted in Fig. 15, while a similar plot for Case 2 measurement sets is shown in Fig. 16. It can be noted from the average residuals in both cases that some systematic nature can be

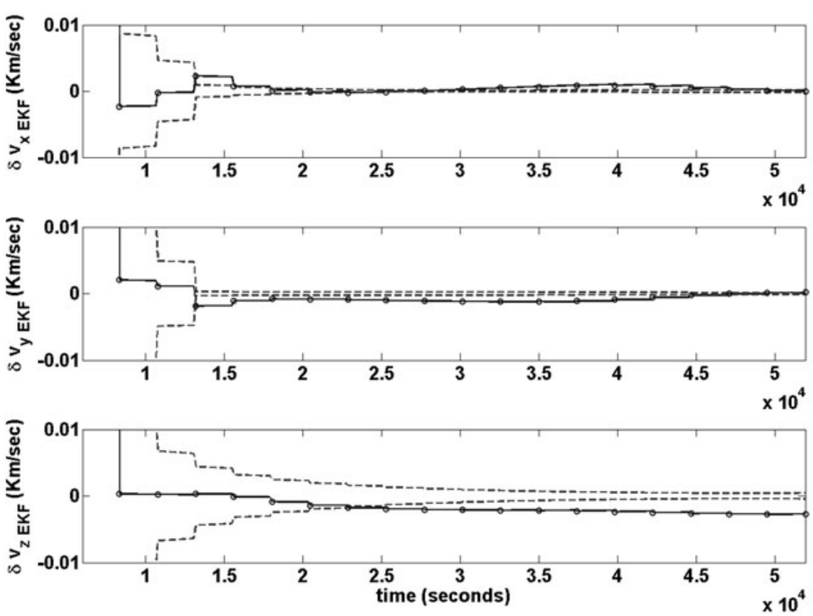

FIG. 9. Position Estimation Error: EKF (Case 2).

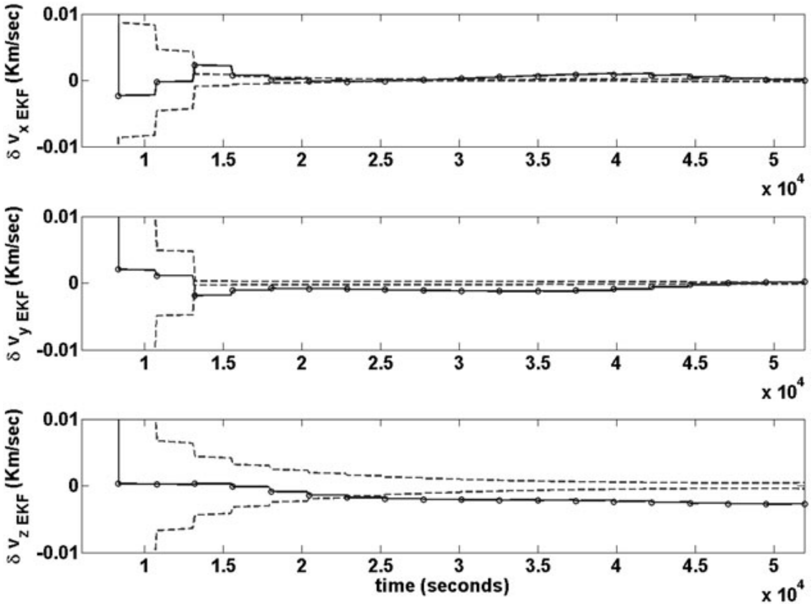

FIG. 10. Velocity Estimation Error: EKF (Case 2).

observed in the steady state performance of the Extended Kalman Filter. An alternative means of investigating the efficacy of the filtering architecture is to employ one of the consistency calculations used in linear estimation. Notice that since the consistency checks examine the residual error (whiteness of it or the lack there of),

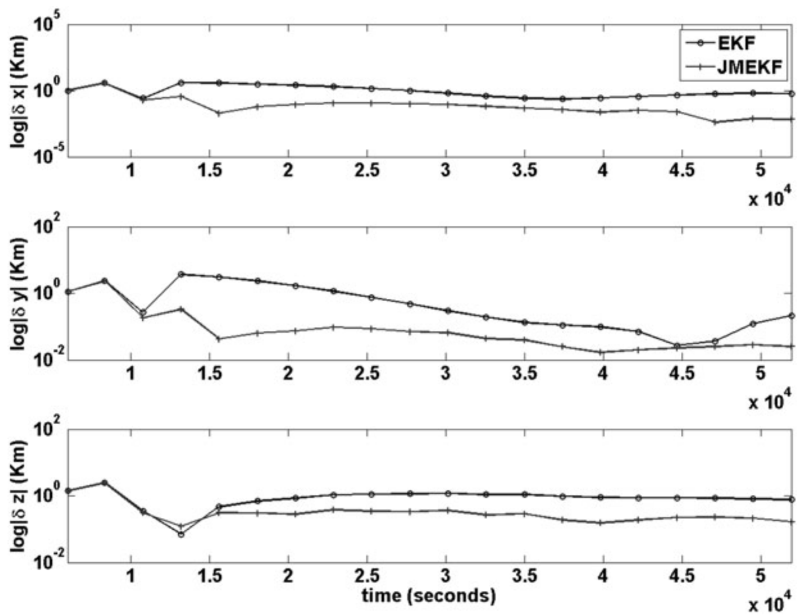

FIG. 11. Position Error Comparison JMEKF Vs. EKF in Log Scale (Case 1).

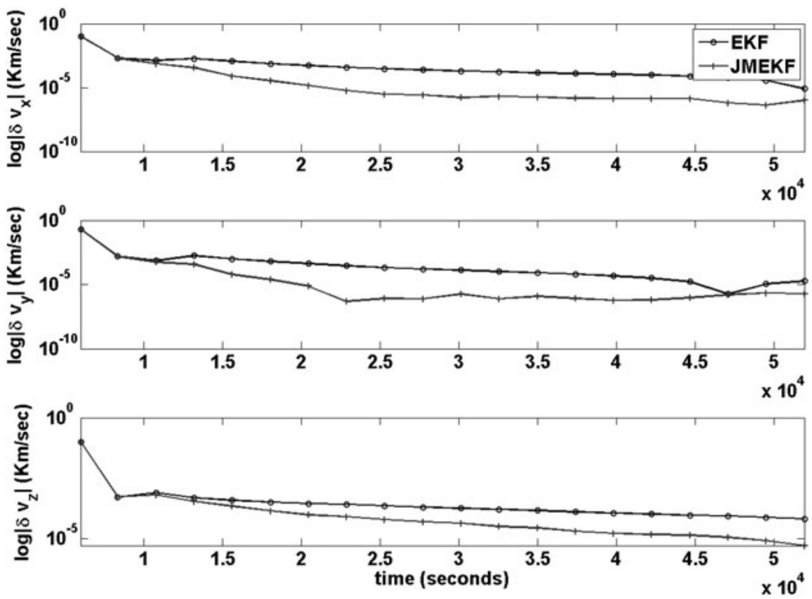

FIG. 12. Velocity Error Comparison JMEKF Vs. EKF in Log Scale (Case 1).

the tests are equally relevant and valid in the nonlinear case. We employ the Normalized Error Square (NME) test for the state and measurement residual processes independently. A simplified form of the test is employed where each component of the residual error vector (state/measurement) is treated as a one-dimensional random process and normalized by the corresponding variance. If the process is consistent,

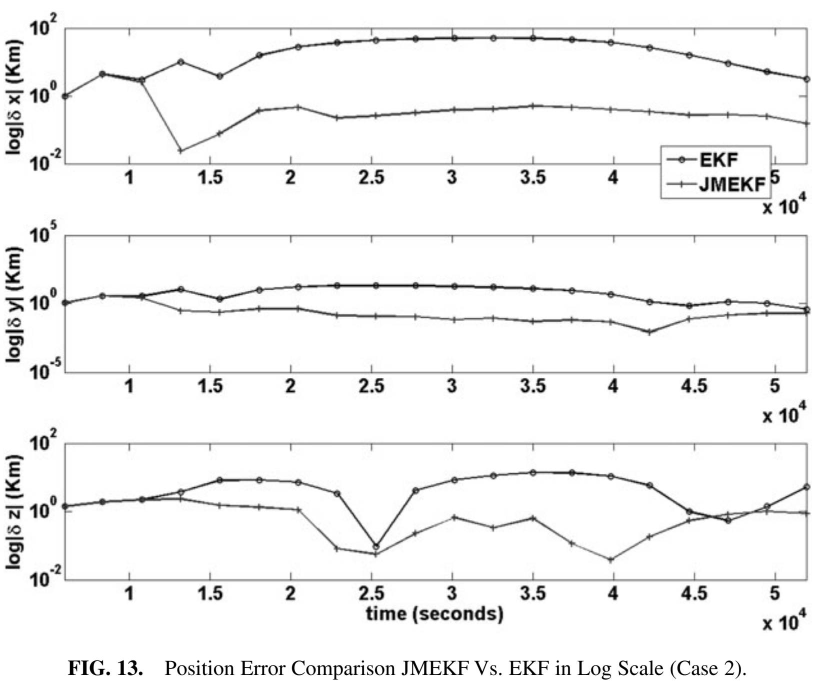

FIG. 13. Position Error Comparison JMEKF Vs. EKF in Log Scale (Case 2).

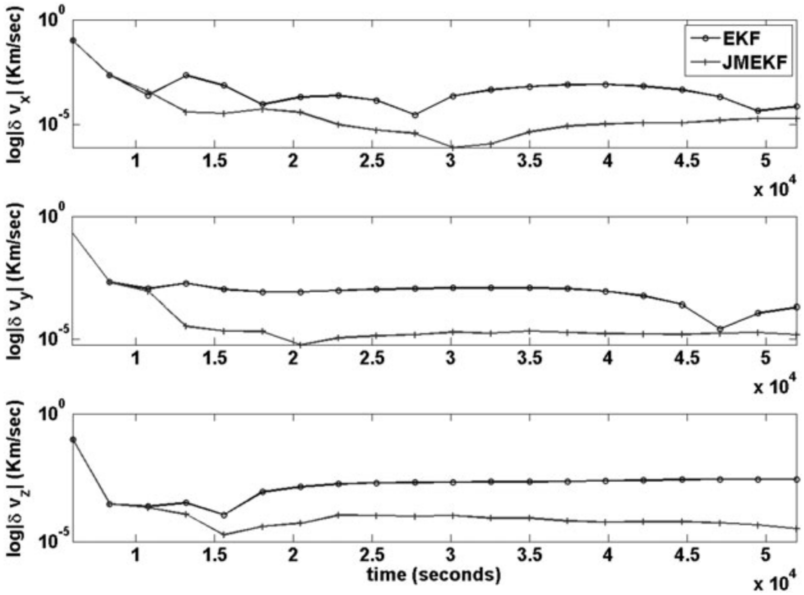

FIG. 14. Velocity Error Comparison JMEKF Vs. EKF in Log Scale (Case 2).

the expected value (calculated using the sample mean metric from all the Monte Carlo simulation results) should be close to one (dimension of the random process). For more on the test, see Section 5.7.3 of [8] and [27]. The correlation terms are not considered for simplicity. Normalized Error Square comparisons for position and velocity error processes are plotted in Fig. 17 (position) and Fig. 18 (velocity) for Case 1 simulated data. NES comparisons for Case 2 are plotted in Figs. 19 and 20

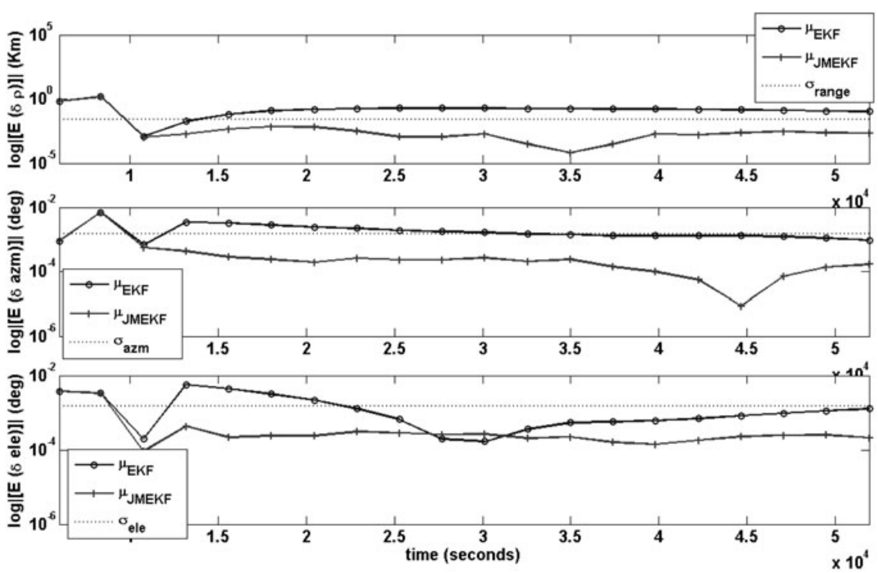

FIG. 15. Mean Measurement Error Residual (Range, Azimuth and Elevation) from Monte-Carlo Simulations (Case 1).

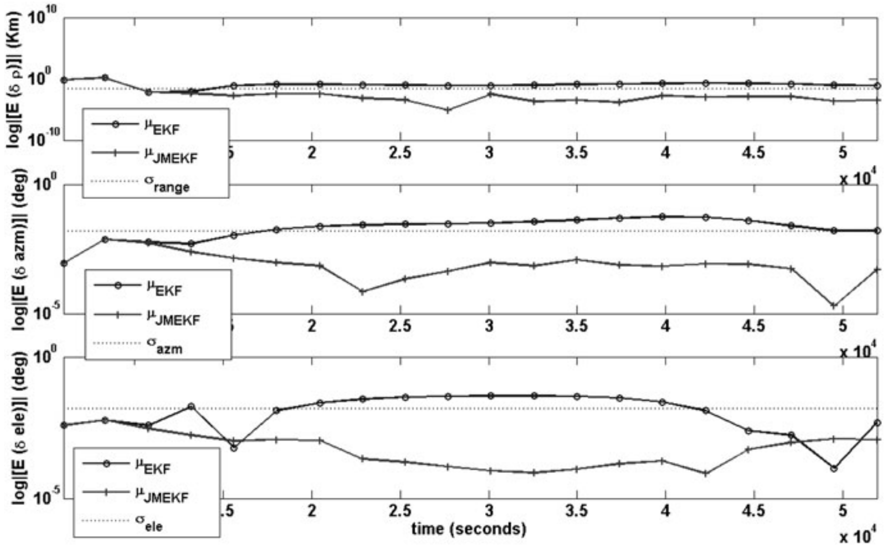

FIG. 16. Mean Measurement Error Residual (Range, Azimuth and Elevation) from Monte-Carlo Simulations (Case 2)

for the cases of position and velocity state errors respectively. It can be noted that the state estimation errors incurred by the JMEKF framework are considerably more "white" than the classical EKF counterparts. This demonstrates that the influence of nonlinearity is accounted in a much better fashion by the JMEKF framework. Similar comparisons have been studied for the Normalized Error Squares of the measurement residual errors. In this case however, for normalization, the standard deviation of the discrete measurement noise (level used to generate synthetic measurements) was used instead of the *a posterori* error auto covariance of the

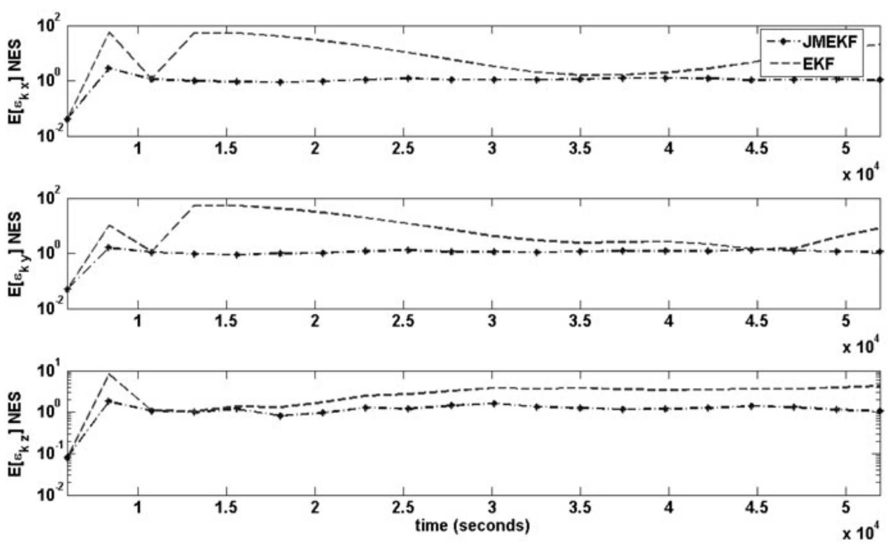

FIG. 17. NES Consistency Test : Position Errors (Case 1).

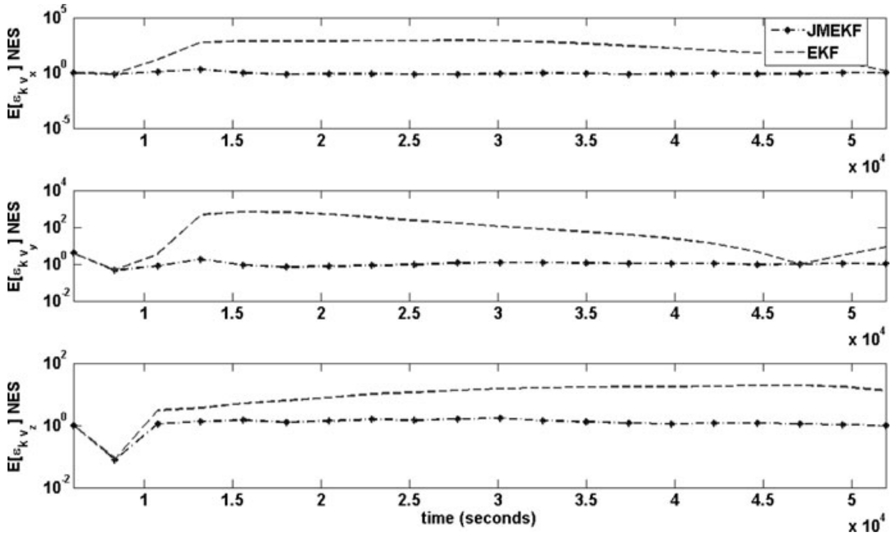

FIG. 18. NES Consistency Test : Velocity Errors (Case 1).

innovations process. This yields a more stringent comparison as such normalization includes several terms originally neglected by the Taylor series expansion of auto covariance of the innovations process. Case 1 results are plotted in Fig. 21 and Case 2 comparisons are plotted in Fig. 22.

This definitively demonstrates the fact that the JMEKF method presented in the paper incorporates the effects of some high-order terms ignored by the classical Extended Kalman Filter. As mentioned earlier, these high-order term effects are more pronounced in the problems involving larger propagation phases and it was found that for high update rates both filters perform equivalently, since in such situations,

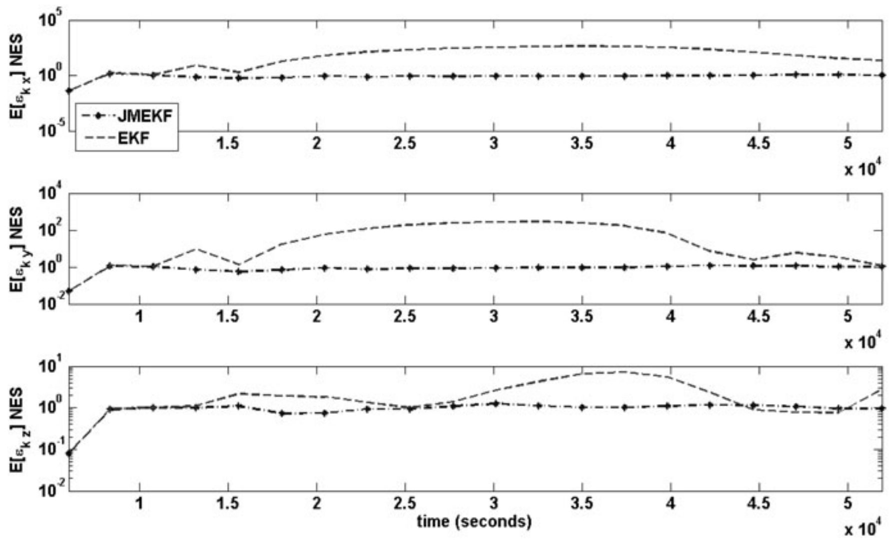

FIG. 19. NES Consistency Test : Position Errors (Case 2).

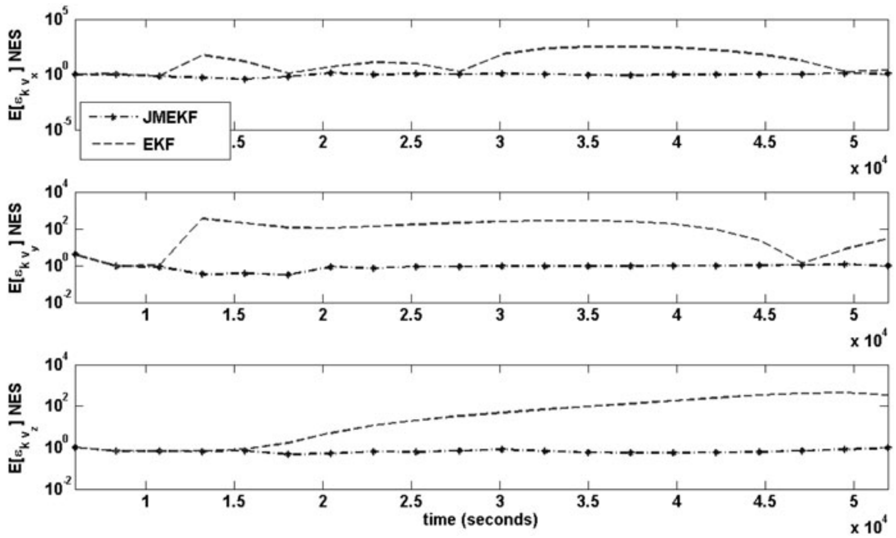

FIG. 20. NES Consistency Test : Velocity Errors (Case 2).

depending on the levels of process noise and initial uncertainty under consideration, first-order terms sufficiently describe the covariance propagation and update. The fact that the residual errors (state and measurement) are “whiter” provides the required numerical evidence to the purported claim that the nonlinear effects are being incorporated into the propagation and update phase of the filter with the new algorithm.

Extra calculations were performed systematically leading to better results and a more robust filter architecture involving lesser tuning efforts. A pertinent question to ask at this stage is about the associated cost overhead incurred to perform the calculations outlined. In some problems the answer to this question is certainly

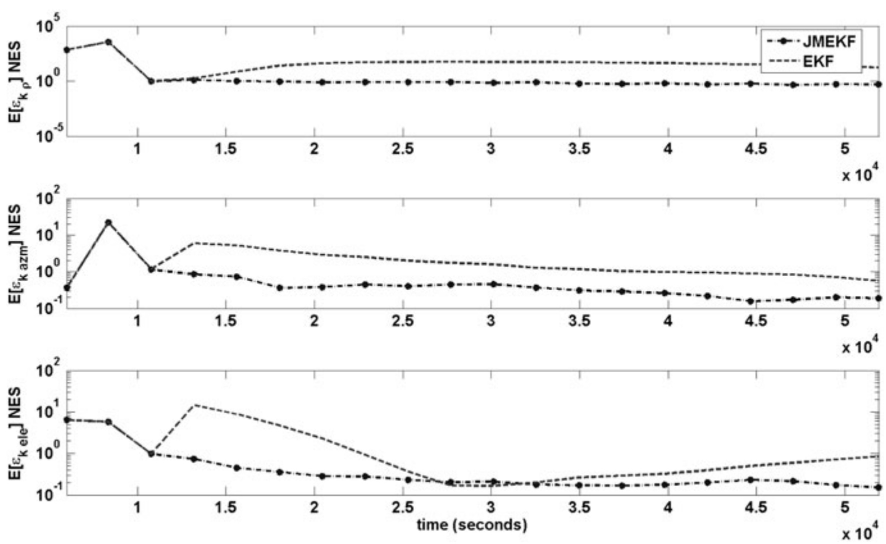

FIG. 21. NES Consistency Test : Measurement Residuals (Case 1).

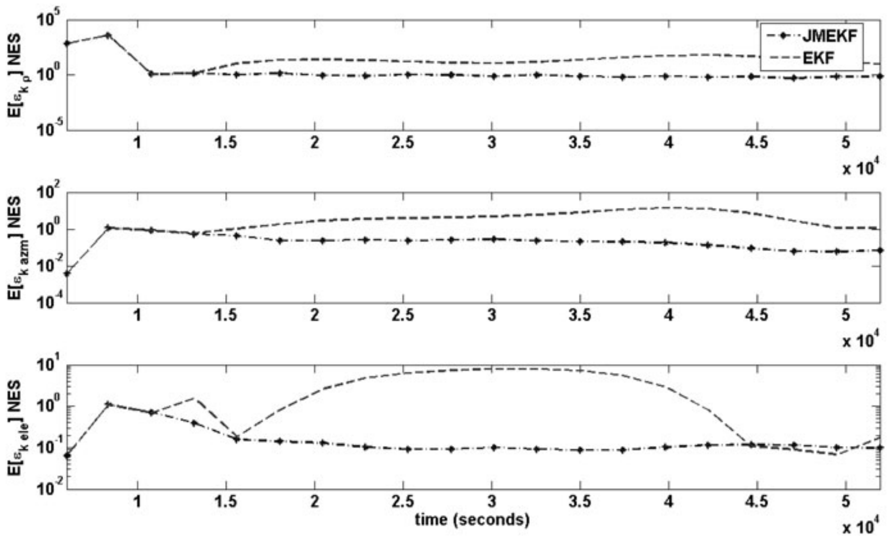

FIG. 22. NES Consistency Test : Measurement Residuals (Case 2).

affirmative and the authors are investigating alternative computational structures to make the associated computations more efficient [28]. However, the authors have been able to find an important set of nonlinear problems (namely the two body problem) where the state transition partials of high-order can be computed analytically [29] in closed-form. In such situations, obviously, the calculations can be performed instantly and the method presented herein has important applications (even in the presence of orbit perturbations) since the perturbation tensors are used to compute statistics for relatively short periods of time, while the best model (inclusive of all known perturbation sub-models) is used to propagate the nominal estimated motion of the particle in the presence of a central force field.

Conclusion

This paper provides an answer to the following question:

How can the classical Extended Kalman Filter be modified to accommodate the Jth order statistical moments (not merely the second-order moments), including recursive updates to process new measurements and associated statistical moments consistently?

Without introducing new approaches to handle the multidimensional algebra and calculus, the answer to this question becomes computationally cumbersome above second-order. However, for moderately dimensioned systems, the formulations introduced in this paper render the computations tractable. We illustrate the methodology for an orbit estimation problem which was structured to challenge the classical Extended Kalman Filter through sparse measurements and large errors. The JMEKF formulation was shown to provide significantly enhanced robustness to large nonlinearities and sparsity of measurements.

Acknowledgment

The authors are pleased to acknowledge the historical inspiration of Byron D. Tapley and discussions with John L. Crassidis.

References

[1] KALMAN, R. E. "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME Journal of Basic Engineering*, 1960, Vol. 82, No. D, pp. 35–45.

[2] MYERS, K. A. and TAPLEY, B. D. "Dynamical Model Compensation for Near-Earth Satellite Orbit Determination," *AIAA Journal*, 1975, Vol. 13, No. 3, pp. 343–349.

[3] MYERS, K. A. and TAPLEY, B. D. "Adaptive Sequential Estimation with Unknown Noise Statistics," *IEEE Transactions on Automatic Control*, 1976, Vol. 21, No. 4, pp. 520–523.

[4] TAPLEY, B. D. and BORN, G. H. "Sequential Estimation of the State and the Observation-Error Covariance Matrix," *AIAA Journal*, 1971, Vol. 9, No. 2, pp. 212–217.

[5] TAPLEY, B. D. and INGRAM, D. S. "Orbit Determination in the Presence of Unmodeled Accelerations," *IEEE Transactions on Automatic Control*, 1973, Vol. 18, No. 4, pp. 364–373.

[6] TAPLEY, B. D. and SCHUTZ, B. E. "Estimation of Unmodeled Forces on a Lunar Satellite," *Celestial Mechanics and Dynamical Astronomy*, 1975, Vol. 12, No. 4, pp. 409–424.

[7] TAPLEY, B. D., SCHUTZ, B. E., and BORN, G. H. "Statistical Orbit Determination," Elsevier Science and Technology Books, April, 2004, Burlington, MA, Elsevier Academic Press.

[8] CRASSIDIS, J. L. and JUNKINS, J. L. *Optimal Estimation of Dynamic Systems*, Applied Mathematics and Nonlinear Science Series, 2004, Boca Raton, FL, Chapman and Hall/CRC Press.

[9] JAZWINSKI, A. H. *Stochastic Processes and Filtering Theory*, 2007, Mineola, NY, Dover Publications.

[10] GELB, A. ed. *Applied Optimal Estimation*, 1974, The M.I.T. Press, Reading, MA.

[11] JULIER, S. and UHLMANN, J. "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE*, 2002, Vol. 92, No. 3, pp. 401–422.

[12] ATHANS, M., WISHNER, R., and BERTOLINI, A. "Suboptimal State Estimation for Continuous-Time Nonlinear Systems from Discrete Noisy Measurements," *IEEE Transactions of Automatic Control*, 1968, Vol. 13, No. 5, pp. 504–514.

[13] GUSTAFSSON, F., GUNNARSSON, F., BERGMAN, N., JANSSON, U. F. J., KARLSON, R., and NORDLUND, P. J. "Particle Filters for Positioning, Navigation and Tracking," *IEEE Transactions on Signal Processing*, 2002, Vol. 50, No. 2, pp. 425–437.

[14] PARK, R. S. and SCHEERES, D. J. "Nonlinear Semi-Analytic Methods for Trajectory Estimation," *Journal of Guidance, Control and Dynamics,* 2007, Vol. 30, No. 6, pp. 1668–1676.

[15] PARK, R. S. and SCHEERES, D. J. "Nonlinear Mapping of Gaussian Statistics: Theory and Applications to Spacecraft Trajectory Design," *Journal of Guidance, Control and Dynamics*, 2006, Vol. 29, No. 6, pp. 1367–1375.

[16] WISHNER, R. P., TABACZYNSKI, J. A., and ATHANS, M. "A Comparison of Three Nonlinear Filters," *Automatica*, 1969, Vol. 5, pp. 487–496.

[17] NORGAARD, M., POULSEN, N.K., and RAVN, O. "New Developments in State Estimation of Nonlinear Systems," *Automatica*, 2000, Vol. 36, pp. 1627–1638.

[18] WIBERG, D. M. "The MIMO Wiberg Estimator," *Proceedings of the IEEE 28th Conference on Decision and Control,* 1989, Tampa, FL.

[19] GRIFFITH, D. T., TURNER, J. D., and JUNKINS, J. L. "An Embedded Function Tool for Modeling and Simulating Estimation Problems in Aerospace Engineering," *Proceedings of the 2004 AAS-AIAA Space Flight Mechanics Meeting*, 2004, Advances in Astronautical Science, American Astronautical Society, Maui, Hawaii.

[20] BADER, B. W. and KOLDA, T. G. "Algorithm 862: MATLAB Tensor Classes for Fast Algorithm Prototyping," *ACM Transactions on Mathematical Software*, 2006, Vol. 32, No. 4, pp. 635–653.

[21] LAWDEN, D. F. "Introduction to Tensor Calculus, Relativity and Cosmology," 3rd ed., 2002, Mineola, NY, Dover Publications.

[22] TURNER, J. D. "Automatic Generation of High-Order Partial Derivative Models," *AIAA Journal*, 2003, Vol. 41, No. 8.

[23] GRAHAM, A. "Kronecker Products and Matrix Calculus with Applications," Ellis Horwood Series: Mathematics and its Applications, ed. G.M. Bell. 1981, New York, NY, John Wiley & Sons.

[24] RAUHALA, U. A. "Array Algebra Expansion of Matrix and Tensor Calculus: Part 1," *SIAM Journal of Matrix Analysis and Applications*, 2002, Vol. 24, No. 2, pp. 490–508.

[25] GOLUB, G. H. and VANLOAN, C. F. *Matrix Computations,* Johns Hopkins Studies in Mathematical Sciences, 1996, Baltimore, MD, The Johns Hopkins University Press.

[26] BATE, R. R., MUELLER, D. D., and WHITE, J. E. *Fundamentals of Astrodynamics,* Dover Books on Astronomy, 1971, New York, Dover Publications.

[27] BAR-SHALOM, Y., LI, X. R., and KIRUBARAJAN, T. *Estimation with Applications to Tracking and Navigation,* 2001, New York, NY, John Wiley and Sons.

[28] TURNER, J. D., MAJJI, M., and JUNKINS, J. L. "High-Order State and Parameter Transition Tensor Calculations," *Proceedings of the 2008 AIAA Guidance Navigation and Control Conference and Exhibit,* 2008, American Institute of Aeronautics and Astronautics, Honolulu, HI.

[29] JUNKINS, J. L., MAJJI, M., and TURNER, J.D. "High Order Keplerian State Transition Tensors," *Proceedings of the F. Landis Markley Astronautics Symposium,* 2008, Cambridge, MD, Univelt Publishing Company, American Astronautical Society.

[30] JUNKINS, J. L., AKELLA, M. R., and ALFRIEND, K. T. "Non-Gaussian Error Propagation in Orbital Mechanics," *The Journal of the Astronautical Sciences*, 1996, Vol. 44, No. 4, pp. 541–563.

[31] JUNKINS, J. L. "Adventures on the Interface of Dynamics and Control," *Journal of Guidance, Control and Dynamics,* 1997, Vol. 20, No. 6, pp. 1058–1071.

[32] LAKSMIKANTHAM, V. and DEO, S. G. *Method of Variation of Parameters for Dynamic Systems,* Series in Mathematical Analysis and Applications, ed. R. P. Agarwal, and O'Regan, D. Vol. 1. 1998, Boca Raton, FL, Gordon and Breach Science Publishers/CRC Press. 328.

Appendix A: Tensor Representations: Applications, Advantages and Disadvantages

Application Case Study 1: Propagation of Initial Condition Balls through Duffing Oscillator Dynamics

Arbitrarily small balls of initial conditions were propagated through the dynamics of several representative problems to investigate the efficacy of the solution reconstructed by the state transition tensor approach. Keeping in mind the fact that the state transition tensor approach is being utilized to propagate the statistics of the neighboring flow of solutions in phase space, these "open-loop" type propagations demonstrate up to what approximate sized balls are transmitted accurately for how long. This is very important qualitative information regarding the characteristics of the solution of the problem. The information thus obtained is useful in determining the levels of initial condition uncertainty and propagation times of filters of a given order. If the state transition tensor solution of a given order cannot transmit a given sized ball of initial conditions at a desired level of accuracy, then the statistical moment propagation accuracy will be of the same order, and hence the process noise levels could be set appropriately to facilitate a better Kalman decision process. The computational advantage realized in using the vector matrix representation of the tensors for the propagation demonstrations is shown in Table 2 for representative application problems.

We now consider the classical Duffing oscillator example and investigate the propagation of a ball of initial conditions through its dynamics. The dynamics of the oscillator are given by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - \epsilon x_1^3 \end{aligned} \tag{71}$$

where the parameter (which is a gross measure of the nonlinearity of the oscillator) is set to unity. The propagation of a ball of radius 0.25 is plotted in Fig. 23.

Application Case Study 2: Third Order State Transition Tensor Solution Propagation of Initial Condition Balls through Two Body Orbital Dynamics

This problem has a more physical connotation. This can be thought of as a cloud of particles dispersed about the nominal trajectory (debris) and being tracked by the solution (similar to the discussions of [30] and [31]). The orbital parameters of the associated orbit are the same as the true orbit associated with the numerical demonstration (Orbit Estimation example). Two dimensional projection of the evolution of

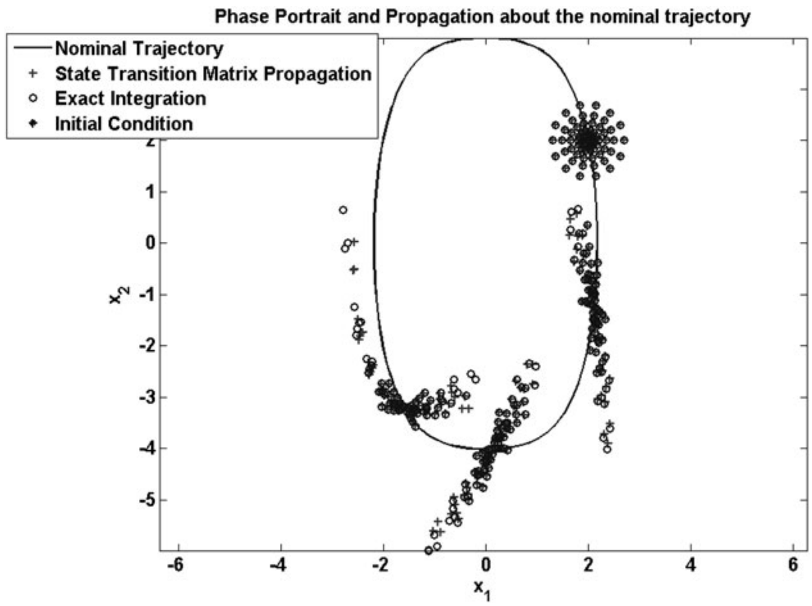

FIG. 23. Fourth order State Transition Tensor Solution: Propagation Through Duffing Oscillator Dynamics.

spherical ball of initial conditions (760 km of initial position error, no velocity error was assumed) is shown in Fig. 24. This plot also compares the state transition tensor solution with that of the exact integration. A three-dimensional evolution of the classic "ball evolving in to a banana schematic of [31]" is also plotted in Fig. 25 (760 km errors of initial position and 0.1 km/sec ball of initial velocity errors propagated until approximately 10% relative error incurred from the true values).

Appendix B: Forced Departure Motion

In this appendix, we derive, using variation of parameters, the effect of forcing functions on the solution of unforced motion for a general system of nonlinear differential equations whose solutions depend continuously on initial conditions. This facilitates the examination of the heuristic assumption we make in the filter derivation for the special case of low process noise.

Consider the nonlinear system

$$\dot{x}_i(t) = f_i(t, \mathbf{x}(t)) + G_{i,j_1} w_{j_1}(t) \tag{72}$$

with uncertainty in the initial conditions whose statistics are given by $E(\delta x_0) = \mu_0$(mean) and $E(\delta x_{i_1}(t_0)...\delta x_{i_n}(t_0)) = P^{(n)}_{i_1...i_n}(t_0)$ (higher moments). The process noise is characterized by the statistics, $E(w_i(t) = \mu^*_i(t)$ and $E(w_{i_1}(t_1)...w_{i_n}(t_n)) = P^{(n)}_{wi_1...i_n}(t_1, \ldots, t_n)$. Although the developments below could be carried out to completion using the above arbitrarily distributed random variables, we consider a special class of random forcing functions that play an important role in most dynamical systems of practical importance. We consider the so-called Gaussian white noise process, uncorrelated in time to force our dynamical system. In addition to its practical importance, the Gaussian white noise process also has

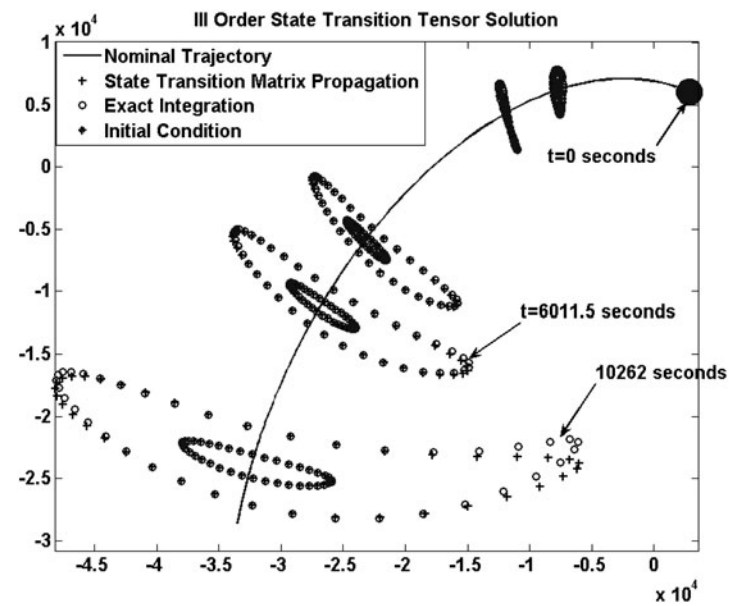

FIG. 24. Third Order State Transition Tensor Solution: XY Plane Projection of Propagation of Initial Conditions Through Two-Body Dynamics.

several attractive properties that make it possible to develop elegant expressions and several simplifications in the theoretical developments. Simplified developments introduce the transparency to enable a better investigation, discussion and remedy the three interacting factors of nonlinear estimation proposed at the outset. Some properties of Gaussian uncorrelated white noise are

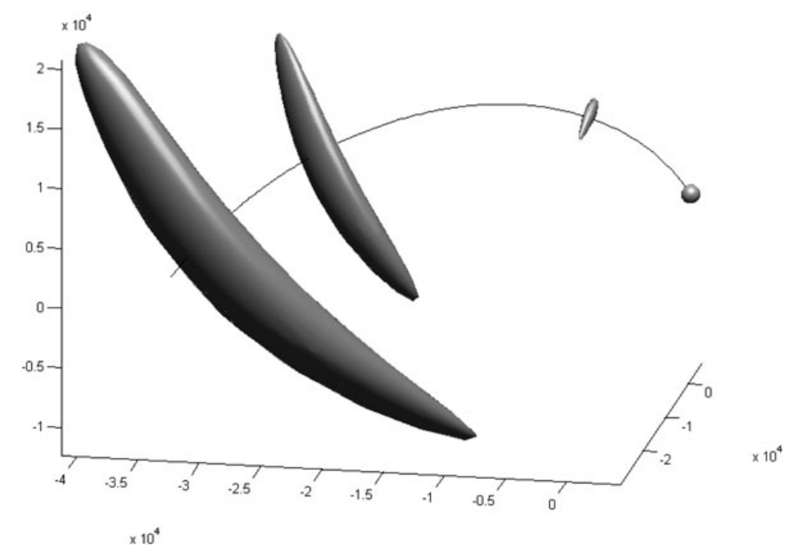

FIG. 25. Third Order State Transition Tensor Solution: Propagation Through Two-Body Dynamics (3D Ball).

TABLE 2. Table Comparing Computational Performance

	Case Study 1: Duffing Oscillator problem	Case Study 2: van der pol Oscillator problem
Computational Time (classical indexing and looping)	282.0940 seconds	148.9540 seconds
Computation Time (array representation)	4.0780 seconds **(X 70 speed up)**	4.25 seconds **(X 35 speed up)**

$$
\begin{aligned}
&E(w_i(t)) = \mu_i^w(t) \\
&E((w_i(t) - \mu_i^w(t))(w_j(\tau) - \mu_j^w(t))) = Q_{i,j}\delta(t - \tau)
\end{aligned} \tag{73}
$$

Since the process is Gaussian, all other higher moments can be derived from the mean and covariance specified above. The state therefore becomes a random variable and the evolution of the statistics associated with this state vector with time is the problem of interest in this paper. This evolution (typically we need to be able to track infinite moments) of moments of the state vector random variable, is tantamount to the evolution of the probability density function (pdf) of the associated continuous random process. It is of central importance in problems of prediction and filtering, where the accuracy of the solution heavily depends on the uncertainty propagation methodology. Most of the challenges in the classical Kalman Filter theory arise due to approximations made in propagating the moments of the pdf, and in using these to process measurements recursively and update state estimates. The central goal of the paper is to investigate the importance of the higher-order terms. Since it is known that higher process noise "washes-out" the high-order effects, we retain the *power of this process noise to a small level*, and consider it to be a zero mean process.

Considering the estimator dynamics

$$
\dot{\hat{x}}_i(t) = f_i(t, \hat{\mathbf{x}}(t)), \quad \hat{x}_i(t_k) = \hat{x}_i^+ \tag{74}
$$

defining the state estimation error, $\delta\mathbf{x}' = \mathbf{x} - \hat{\mathbf{x}}$ associated error dynamics can be written as

$$
\begin{aligned}
&\delta\dot{\mathbf{x}}'(t) := \dot{\mathbf{x}}(t, t_k, \hat{\mathbf{x}}(t_k) + \delta\mathbf{x}(t_k)) - \dot{\hat{\mathbf{x}}}(t, t_k, \hat{\mathbf{x}}(t_k)) \\
&\delta\dot{\mathbf{x}}'(t) := \mathbf{f}(\mathbf{x}(t, t_k, \hat{\mathbf{x}}(t_k) + \delta\mathbf{x}(t_k)), t) - \mathbf{f}(\hat{\mathbf{x}}(t, t_k, \hat{\mathbf{x}}(t_k))) + G\mathbf{w}(t)
\end{aligned} \tag{75}
$$

The unforced motion of the estimation error dynamics above can therefore be given in terms of high order state transition tensors operating on a corresponding ball of initial conditions. Such a map, which was developed in detail in the previous sections (following the developments of Griffith et al. [19], and Park and Scheeres [15]) is written as

$$
\delta x_i(t) = \sum_{p=1}^{m} \frac{1}{p!} \Phi_{i,k_1 \ldots k_p} \delta x_{k_1}^0 \ldots \delta x_{k_p}^0 \tag{76}
$$

Using variation of parameters (for a different derivation please refer to the text book by Lakshmikantham [32]), the forced motion (due to the stochastic forcing component) is given by

$$\delta\mathbf{x}'(t) = \delta\mathbf{x}(t) + \boldsymbol{\xi}(t) \tag{77}$$

where we dictate the solution $\xi_i(t) := \sum_{p=1}^{m} \frac{1}{p!} \Phi_{i,k_1\ldots k_p}\kappa_{k_1}(t)\ldots\kappa_{k_p}(t)$, consistent with Lagrange's variation of parameters framework. Substituting back in to the solution, we see that

$$\begin{aligned}\dot{\xi}_i &= \sum_{p=1}^{m} \frac{1}{p!} \dot{\Phi}_{i,k_1\ldots k_p}\kappa_{k_1}(t)\ldots\kappa_{k_p}(t) + \sum_{p=1}^{m} \frac{1}{p!} \Phi_{i,k_1\ldots k_p} \frac{\mathrm{d}}{\mathrm{d}t}(\kappa_{k_1}(t)\ldots\kappa_{k_p}(t)) \\ &= \sum_{p=1}^{m} \frac{1}{p!} f^*_{i,k_1k_2\ldots k_p}\kappa_{k_1}(t)\ldots\kappa_{k_p}(t) + G_{i,j_1}w_{j_1}(t)\end{aligned} \tag{78}$$

Where $f^*_{i,k_1\ldots k_p} := \frac{\partial^p f_i}{\partial x_{k_1}\ldots\partial x_{k_p}}\Big|_{x=x^*}$ is the high-order sensitivity of the equations of motion, evaluated along the nominal trajectory (estimator dynamics starting from the *a posteriori* state estimate $\hat{\mathbf{x}}^+(t_k)$) and $\Phi_{i,k_1\ldots k_p} := \frac{\partial^p xi}{\partial x^0_{k_1}\ldots\partial x^0_{k_p}}\Big|_{x=x^*}$ is the high-order state transition tensor along the same trajectory. The governing equations and methods to compute them are discussed in the main body of this paper (see also Park and Scheeres [14]). The solution for the forced motion is therefore obtained by solving for $\kappa(t)$ from

$$\sum_{p=1}^{m} \frac{1}{p!} \Phi_{i,k_1\ldots k_p} \frac{\mathrm{d}}{\mathrm{d}t}(\kappa_{k_1}(t)\ldots k_{k_p}(t)) = G_{i,j_1}w_{j_1}(t) \tag{79}$$

with initial conditions, $\delta\mathbf{x}'(t_0) = \delta\mathbf{x}^0$. This would entail us defining the inverse state transition tensor map and solve this highly coupled nonlinear system of equations. Under the assumption of *small process noise*, (which is the situation of interest for the present investigation) this can be further simplified by neglecting the higher state transition tensors. This is legitimate in the initial propagation phase since the high-order tensors start from zero initial conditions and are really forced by the model error. The process noise at this stage is treated as an artificial tuning knob to compensate the filter performance for lacking the infinite moments that are needed to capture a general nonlinear uncertainty propagation problem.

$$\Phi_{i,j_1}\kappa(t) = G_{i,j_1}w_{j_1}(t) \tag{80}$$

In a similar manner, neglecting the high order state transition tensor participation in $\xi(t)$, we write the solution of the forced departure motion as

$$\delta\mathbf{x}'(t) = \delta\mathbf{x}(t) + \int_{t_0}^{t} \Phi^{-1}(\tau, t_0)G\mathbf{w}(\tau)\mathrm{d}\tau \tag{81}$$

Using the above relation, let us now consider the evolution of covariance matrix of the forced departure motion by differentiating $E[\delta\mathbf{x}'\delta\mathbf{x}'^{\mathrm{T}}]$. Now, using the fact that (to the degree of approximation made above), $\delta\dot{\mathbf{x}}'(t) = \delta\dot{\mathbf{x}}(t) + \Phi^{-1}(t, t_0)G\mathbf{w}(t)$.

$$\frac{\mathrm{d}}{\mathrm{d}t}P^{(2)'}_{i_1i_2} = \frac{\mathrm{d}}{\mathrm{d}t}E[\delta x_{i_1}'(t)\delta x_{i_2}'(t)] = E[\delta\dot{x}_{i_1}'(t)\delta x_{i_2}'(t)] + E[\delta x_{i_1}'(t)\delta\dot{x}_{i_2}'(t)] \tag{82}$$

Considering the expression

$$
\begin{aligned}
E[\delta\dot{x}_{i_1}'(t)\delta x_{i_2}'(t)] &= E[\delta\dot{x}_{i_1}'(t)\delta x_{i_2}'(t)] + E[\Phi^{-1}_{i_1,j_1}(t,t_0)G_{j_1,j_2}w_{j_2}(t)\delta x_{j_2}'(t)] \\
&= E(\delta\dot{x}_{i_1}(t)\delta x_{i_2}(t)] \\
&\quad + E\left[\Phi^{-1}_{i_1,j_1}(t,t_0)G_{j_1,j_2}w_{j_2}(t)\left(\int_{t_0}^{t}\Phi^{-1}(\tau,t_0)G\mathbf{w}(\tau)\mathrm{d}\tau\right)_{i_2}\right] \\
&= E[\delta\dot{x}_{i_1}(t)\delta x_{i_2}(t)] + \Phi^{-1}_{i_1,j_1}(t,t_0)G_{j_1,j_2}Q_{j_2 j_3}(t)G_{j_1,j_2}\Phi^{-1}_{j_1,j_2}(t,t_0) \\
&\approx E[\delta\dot{x}_{i_1}(t)\delta x_{i_2}(t)] + \frac{1}{2}G_{j_1,j_2}Q^{*}_{j_2 j_3}(t)G_{j_1,j_2}
\end{aligned}
\tag{83}
$$

The assumption that the process noise $w(t)$ is uncorrelated with all moments of the state variable (and therefore $\delta x(t)$) is made and the additional process noise covariance $Q^{*}(t)$ is introduced, the statistics of which, upon careful tuning would "cover-up" the errors made in truncation of the series (which are presented shortly) and other approximations (like the assumptions in forcing term calculations from the variation of parameters calculation above). Hence the covariance propagation is given by the differential equation

$$
\frac{\mathrm{d}}{\mathrm{d}t}P^{(2)'} = \frac{\mathrm{d}}{\mathrm{d}t}P^{(2)} + GQ^{*}G^{\mathrm{T}} = \frac{\mathrm{d}}{\mathrm{d}t}E[\delta\mathbf{x}(t)\delta\mathbf{x}^{\mathrm{T}}(t)] + GQ^{*}G^{\mathrm{T}} \tag{84}
$$

To determine the differential equations governing the evolution of statistical moments of the departure motion dynamics (pure initial condition response), $\delta x(t)$, we use the state transition tensor approach presented early on in the paper. Thus, for the problems of engineering interest, it is entirely legitimate to investigate the propagation of initial condition uncertainty and "add" the process noise to the propagation in the form of a convolution integral, analogous to the case of a linear dynamical system (similar to the discussion in Chapter 5 of Crassidis and Junkins [8]).