AAS 09–305

GENERALIZED COVARIANCE ANALYSIS OF ADDITIVE DIVIDED–DIFFERENCE SIGMA–POINT FILTERS

J. Russell Carpenter*, **Sun Hur–Diaz**†, **and F. Landis Markley**‡

The divided-difference sigma-point filter is a sequential estimator that replaces first-order truncations of Taylor series approximations with second-order numerical differencing equations to approximate nonlinear dynamics and measurement models. If the process and measurement noise enter the system additively, several simplifications are possible, including a substantial reduction in the number of sigma-points. As a consequence of the additive noise assumption, a generalized covariance analysis approach that partitions the contributions to the total error of *a priori*, process, and measurement noise may be applied to the additive divided-difference sigma-point filter. The Cholesky decompositions of the true and formal initial covariances provide true and formal *a priori* Cholesky factors, and true and formal measurement and process Cholesky partitions are initialized to zero. Two sets of sigma points, truth and formal, are spawned and propagated from the joint set of all three partitions for each. Divided differences are separately extracted from each partition, and factorized to derive time updates for each partition separately, as well as merged to form propagated states. This process is repeated for the measurement update, with a filter gain similarly derived from a joint set of all three partitions. The states ignored by the filter are not updated by this gain. The entire algorithm is formulated using only Cholesky factors. As an example, a simulated highly elliptical orbit is estimated from nonlinear Global Positioning System measurements. In this example, there is a significant nonlinearity at perigee.

INTRODUCTION

Derivative-free state estimation techniques have received increasing attention in recent years. A particular class of such estimators make use of the columns of the factors of the estimators' error covariance matrices, which are scaled to form vectors that have become generally known as "sigma points." The so-called "Unscented Kalman Filter"[1–3] is a particular example of a sigma-point filter. A more general form is the divided-difference sigma-point filter, which is a sequential estimator that replaces first-order truncations of Taylor series approximations with second-order numerical differencing equations to approximate nonlinear dynamics and measurement models.[4,5] If the process and measurement noise enter the system additively, Lee and Alfriend recently showed

*Aerospace Engineer, Navigation and Mission Design Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

†Director of Guidance, Navigation and Control, Emergent Space Technologies, Inc., Greenbelt, MD 20770.

‡Aerospace Engineer, Attitude Control Systems Engineering Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771.

that several simplifications are possible, including a substantial reduction in the number of sigma-points.[6] As a consequence of the additive noise assumption, a generalized covariance analysis approach that partitions the contributions to the total error of *a priori* error, process noise, and measurement noise[7] may be applied to the additive divided-difference sigma-point filter, which this paper shall abbreviate as ADDSPF hereafter. Similar to, but more general than the work of Lisano,[8,9] the generalized approach allows for a subset of the state elements to be solved for, while the statistical effect of the other is considered, within the sigma-point filtering context.

This paper begins with a review of the sigma-point filtering technique, based on the approach of Nørgaard, et al.,[4,5] highlighting the simplifications pointed out by Lee and Alfriend[6] when the measurement and process noise enter additively. The subsequent section applies the generalized covariance analysis approach of Carpenter and Markley[7] to the ADDSPF. A penultimate section applies the method to a simplified orbit determination problem involving a highly elliptical two-body orbit with biased position measurements, and the paper concludes with a summary.

BACKGROUND

This section highlights some broad aspects of sigma-point filtering, then briefly reviews how the ADDSPF works. It concludes with some brief comments comparing the ADDSPF to other sequential filters.

The Sigma-Point Filter

In its most general form, the sigma-point filter performs sequential estimation of the n-dimensional state, x, whose nonlinear dynamics over the time interval $[t_k, t_{k+1}]$ are given by

$$\mathsf{x}_{k+1} = \mathsf{f}(\mathsf{x}_k, \mathsf{w}_k) \tag{1}$$

The process noise input, w, consists of independent increments whose first two moments are $\mathrm{E}[\mathsf{w}_k] = 0$ and $\mathrm{E}[\mathsf{w}_k \mathsf{w}_\ell] = \mathsf{Q}_k \delta_{k\ell}$, where $\delta_{k\ell}$ is the Kronecker delta. Although the second moment may be a function of the time index, this estimator assumes that all of the samples of w arise from the same type of distribution, and this work further assumes that this distribution is Gaussian, so that higher-order moments may be neglected.

The filter sequentially processes an ordered set of measurements, $\mathfrak{Y}_k = [\mathsf{y}_0, \mathsf{y}_1, ..., \mathsf{y}_k]$ of the form

$$\mathsf{y}_k = \mathsf{h}(\mathsf{x}_k, \mathsf{v}_k) \tag{2}$$

where the measurement noise input, v, consists of independent and identically distributed (again, in this work, Gaussian) increments whose first two moments are $\mathrm{E}[\mathsf{v}_k] = 0$ and $\mathrm{E}[\mathsf{v}_k \mathsf{v}_\ell] = \mathsf{R}_k \delta_{k\ell}$. By contrast, the ADDSPF utilizes models where the noise sources enter additively:

$$\mathsf{x}_{k+1} = \mathsf{f}(\mathsf{x}_k) + \mathsf{g}(\mathsf{x}_k)\mathsf{w}_k \tag{3}$$

and

$$\mathsf{y}_k = \mathsf{h}(\mathsf{x}_k) + \mathsf{v}_k. \tag{4}$$

All sigma-point filters utilize a linear measurement update equation of the form

$$\hat{\mathsf{x}}_k^+ = \hat{\mathsf{x}}_k^- + \mathsf{K}_k\left(\mathsf{y}_k - \hat{\mathsf{y}}_k^-\right) \tag{5}$$

where the accented variables in Eq. 5 denote conditional expectations, as in the Kalman filter:

$$\hat{\mathsf{x}}_k^+ = \mathrm{E}\left[\mathsf{x}_k|\mathfrak{Y}_k\right] \tag{6}$$

$$\hat{\mathsf{x}}_k^- = \mathrm{E}\left[\mathsf{x}_k|\mathfrak{Y}_{k-1}\right] \tag{7}$$

$$\hat{\mathsf{y}}_k^- = \mathrm{E}\left[\mathsf{y}_k|\mathfrak{Y}_{k-1}\right] \tag{8}$$

The gain matrix, K, is based on conditional covariances, as in the Kalman filter:

$$\mathsf{K}_k = \mathsf{P}_{xy_k}^-\left(\mathsf{P}_{yy_k}^-\right)^{-1} \tag{9}$$

$$\mathsf{P}_k^- = \mathsf{P}_{xx_k}^- = \mathrm{E}\left[\left(\mathsf{x}_k - \hat{\mathsf{x}}_k^-\right)\left(\mathsf{x}_k - \hat{\mathsf{x}}_k^-\right)^\mathsf{T}|\mathfrak{Y}_{k-1}\right] \tag{10}$$

$$\mathsf{P}_{xy_k}^- = \mathrm{E}\left[\left(\mathsf{x}_k - \hat{\mathsf{x}}_k^-\right)\left(\mathsf{y}_k - \hat{\mathsf{y}}_k^-\right)^\mathsf{T}|\mathfrak{Y}_{k-1}\right] \tag{11}$$

$$\mathsf{P}_{yy_k}^- = \mathrm{E}\left[\left(\mathsf{y}_k - \hat{\mathsf{y}}_k^-\right)\left(\mathsf{y}_k - \hat{\mathsf{y}}_k^-\right)^\mathsf{T}|\mathfrak{Y}_{k-1}\right] \tag{12}$$

$$\tag{13}$$

and the covariance associated with state estimate $\hat{\mathsf{x}}_k^+$ is

$$\mathsf{P}_k^+ = \mathsf{P}_{xx_k}^+ = \mathrm{E}\left[\left(\mathsf{x}_k - \hat{\mathsf{x}}_k^+\right)\left(\mathsf{x}_k - \hat{\mathsf{x}}_k^+\right)^\mathsf{T}|\mathfrak{Y}_k\right] \tag{14}$$

Hereafter, equations will suppress the time index if it is the same for all variables in the equation.

Estimators such as the Kalman filter estimate these conditional expectations by approximating the nonlinear functions f and h with first-order Taylor series truncations, e.g.:

$$\mathsf{f}(\mathsf{x}) \approx \mathsf{f}(\hat{\mathsf{x}}^-) + \mathsf{f}'(\mathsf{x})(\mathsf{x} - \hat{\mathsf{x}}^-) \tag{15}$$

where f' is an exact gradient. By contrast, the divided difference filter uses a second-order truncation along with numerical differencing formulas for the derivatives:

$$\mathsf{f}(\mathsf{x}) \approx \mathsf{f}(\hat{\mathsf{x}}^-) + \tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(1)}\mathsf{f}(\hat{\mathsf{x}}^-) + \tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(2)}\mathsf{f}(\hat{\mathsf{x}}^-) \tag{16}$$

where the divided difference operators, $\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(i)}\mathsf{f}(\hat{\mathsf{x}}^-)$, approximate the coefficients of the multidimensional Taylor series expansion using Stirling interpolations. These interpolators difference perturbations of $\mathsf{f}(\hat{\mathsf{x}}^-)$ across an interval, h, over a spanning basis set. Whether they are first-order, such as the unscented filter, or second order, sigma-point filters choose the interval so as to better approximate the moments required for the gain calculation, and choose as the spanning basis a set of sigma points, which are derived from $\hat{\mathsf{x}}^-$ and the columns of the Cholesky factors of P^- as follows.

The ADDSPF

Let $\hat{\mathcal{X}}$ denote the array whose columns are a particular ordering of the sigma points derived from $\hat{\mathsf{x}}$ and its corresponding covariance, P. Then

$$\hat{\mathcal{X}} = \left[\hat{\mathsf{x}},\, \hat{\mathsf{x}} + h\sqrt[c]{\mathsf{P}}_{(:,1)},\, \hat{\mathsf{x}} + h\sqrt[c]{\mathsf{P}}_{(:,2)}, \ldots, \hat{\mathsf{x}} - h\sqrt[c]{\mathsf{P}}_{(:,1)}, \hat{\mathsf{x}} - h\sqrt[c]{\mathsf{P}}_{(:,2)}, \ldots\right] \tag{17}$$

where the subscript $(:,i)$ denotes column i of the corresponding array, and $\mathsf{P} = \sqrt[c]{\mathsf{P}}\sqrt[c]{\mathsf{P}}^{\mathsf{T}}$ denotes a Cholesky factorization. In the sequel, the shorthand notation $\hat{\mathsf{x}} \pm h\sqrt[c]{\mathsf{P}}$ will denote the array on the right-hand side of the equation above. Then, for the ADDSPF, Ref. 6 shows that as each new measurement becomes available, an array of sigma points generated from the prior update should be propagated to the new measurement time:

$$\hat{\mathcal{X}}_k^- = \mathsf{f}(\hat{\mathcal{X}}_{k-1}^+) \tag{18}$$

These propagated sigma points are then merged to form the state estimate just prior to incorporating the new measurement as follows:

$$\hat{\mathsf{x}}^- = \mu_h(\hat{\mathcal{X}}^-) = \frac{h^2 - n}{h^2}\hat{\mathcal{X}}_{(:,1)}^- + \frac{1}{2h^2}\sum_{i=2}^{2n+1}\hat{\mathcal{X}}_{(:,i)}^- \tag{19}$$

To form an associated covariance, the following divided-differences are next computed:

$$\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(1)}\mathsf{f}(\hat{\mathsf{x}}^-)_{(:,i)} = \frac{1}{2h}\left[\hat{\mathcal{X}}_{(:,i+1)}^- - \hat{\mathcal{X}}_{(:,i+1+n)}^-\right] \tag{20}$$

$$\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(2)}\mathsf{f}(\hat{\mathsf{x}}^-)_{(:,i)} = \frac{\sqrt{h^2-1}}{2h^2}\left[\hat{\mathcal{X}}_{(:,i+1)}^- + \hat{\mathcal{X}}_{(:,i+1+n)}^- - 2\hat{\mathcal{X}}_{(:,1)}^-\right] \tag{21}$$

Ref. 6 shows that the covariance may then be computed from

$$\mathsf{P}^- = \left[\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(1)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(2)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \sqrt[c]{\mathsf{Q}_d}\right]\left[\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(1)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(2)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \sqrt[c]{\mathsf{Q}_d}\right]^{\mathsf{T}} \tag{22}$$

One advantage of sigma-point filters is that the full covariance need not be maintained, but rather only its Cholesky factor. Although the factors in square brackets in Eq. 22 are not Cholesky factors, since each is a full $n \times 3n$ matrix, one may extract an $n \times n$ triangular factor from it using the so-called "thin" version[10] of the QR decomposition*, or alternatively using a Householder factorization.[11] Thus,

$$\mathsf{M}\begin{bmatrix}\sqrt[c]{\mathsf{P}^-}^{\mathsf{T}} \\ \mathsf{O}_{2n\times n}\end{bmatrix} = \left[\tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(1)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}_{\Delta\mathsf{x}}^{(2)}\mathsf{f}(\hat{\mathsf{x}}^-),\, \sqrt[c]{\mathsf{Q}_d}\right]^{\mathsf{T}} \tag{23}$$

where M is a full $3n \times 3n$ orthonormal matrix, and $\mathsf{O}_{2n\times n}$ is a $2n \times n$ matrix of zeros.

*For *Matlab* users, this may be accomplished in several ways, e.g. by passing the transpose of this matrix to the `qr` function, then keeping the first n non-zero rows from the second output, and transposing this result.

For the measurement update, a new array of sigma points must be generated from $\hat{\mathsf{x}}^-$ and P^-; this array is denoted $\hat{\mathcal{X}}^*$. These sigma points are used to generate a set of sigma points representing the measurement:

$$\hat{\mathcal{Y}}^- = \mathsf{h}(\hat{\mathcal{X}}^*) \tag{24}$$

In similar fashion to the time update, the sigma points of the measurement are then merged to form the estimated measurement:

$$\hat{\mathsf{y}}^- = \mu_h(\hat{\mathcal{Y}}^-) \tag{25}$$

the corresponding divided-differences are computed as:

$$\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)_{(:,i)} = \frac{1}{2h}\left[\hat{\mathcal{Y}}^-_{(:,i+1)} - \hat{\mathcal{Y}}^-_{(:,i+1+n)}\right] \tag{26}$$

$$\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)_{(:,i)} = \frac{\sqrt{h^2-1}}{2h^2}\left[\hat{\mathcal{Y}}^-_{(:,i+1)} + \hat{\mathcal{Y}}^-_{(:,i+1+n)} - 2\hat{\mathcal{Y}}^-_{(:,1)}\right] \tag{27}$$

and the covariances required for the gain calculation may then be computed from

$$\mathsf{P}^-_{yy} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\ \sqrt[c]{\mathsf{R}}\right]\left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\ \sqrt[c]{\mathsf{R}}\right]^\mathsf{T} \tag{28}$$

$$\mathsf{P}^-_{xy} = \sqrt[c]{\mathsf{P}^-}\left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)\right]^\mathsf{T} \tag{29}$$

Note that the second-order divided difference for the measurement function, $\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)$, does *not* appear in the cross-covariance update. As with the time update, through the use of the thin QR factorization, only triangular factors need be maintained for P^-_{yy} [†] :

$$\mathsf{M}_{yy}\begin{bmatrix}\sqrt[c]{\mathsf{P}^-_{yy}}^\mathsf{T} \\ \mathsf{O}_{2n\times n}\end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\, \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\ \sqrt[c]{\mathsf{R}}\right]^\mathsf{T} \tag{30}$$

Now, all of the terms required for the state update (Eqs. 5 and 9), are available. Ref. 6 shows that the corresponding Cholesky factor of the covariance is extracted from

$$\mathsf{M}^+\begin{bmatrix}\sqrt[c]{\mathsf{P}^+}^\mathsf{T} \\ \mathsf{O}_{2n\times n}\end{bmatrix} = \left[\sqrt[c]{\mathsf{P}^-} - \mathsf{K}\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\, \mathsf{K}\left[\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-),\ \sqrt[c]{\mathsf{R}}\right]\right] \tag{31}$$

The ADDSPF vs. Other Sequential Filters

To conclude this section, some observations concerning the ADDSPF in comparison to other filters are offered. These observations concern the number of sigma points, the order of approximation, and the existence and method of choice of free parameters in the algorithms.

[†] Although it might seem that the full matrix P^-_{yy} is required for the gain computation of Eq. 9, Ref. 5 points out that, rather than inverting the product of the factors to compute the gain, the gain may be solved from forward and back substitution directly using the Cholesky factor.

Although in many problems of practical interest the noise enters the system additively, if this is not the case, then either the original divided difference filter or the unscented filter may provide superior results to the ADDSPF, at the cost of requiring more sigma points. In both of the former algorithms, the nonlinear functions must be perturbed not only over a basis spanning the state space, but also over the discrete process noise and measurement noise spaces. Thus, rather than $2n + 1$ sigma points, the more general algorithms require $2n_a + 1$, where $n_a = n + n_w + n_v$, and n_w and n_v are the dimensions of the discrete process noise and measurement noise inputs.

The Kalman filter is an exact algorithm for linear stochastic systems driven by Gaussian noise, and nothing is to be gained from the use the sigma-point filters for such purely linear systems. First-order sigma-point filters such as the UKF retain the Kalman filter's first-order truncation, but avoid the need for the designer to supply explicit gradients. The divided difference filter is comparable to a derivative-free version of the modified second-order Gaussian filter[12] in that, for symmetric distributions, it retains some terms as high as order four.

Unlike the Kalman filter, for which all of the parameters in principle can be associated with properties of the underlying stochastic system, all of the sigma point filters involve at least one free parameter. In the unscented filter, the weights for combining the sigma points involve three parameters whose physical interpretation is perhaps less clear than with the single parameter in the divided difference algorithms, where the free parameter h is clearly associated with the size of the perturbation in the numerical differencing formulae. Ref. 5 shows that h should be bounded below by $h > 1$, and that for symmetric distributions, $\sqrt{h}$ should be equal to the kurtosis, which for a Gaussian distribution is three[‡].

COVARIANCE ANALYSIS

This section applies the generalized covariance analysis method described in Ref. 7 to the ADDSPF. The method of Ref. 7 depends on two types of projections. One is the projection of the total covariance into partitions arising solely from *a priori* errors, errors due to measurement noise, and errors due to process noise. The other is a projection from the full state vector of all parameters that affect the estimation, to a "solve-for" state vector containing a linear mapping of some subset of the full state. Application of the generalized covariance analysis method to the additive divided-difference sigma-point filter involves a fairly straightforward substitution of these terms into the algorithms specified in Refs. 4 and 6.

Following Ref. 7, define the n_s-dimensional solve-for state, s, and n_c-dimensional "consider" state, c, as general linear mappings of the full state:

$$\mathsf{s} = \mathsf{S}\mathsf{x} \text{ and } \mathsf{c} = \mathsf{C}\mathsf{x} \tag{32}$$

with $n = n_s + n_c$. The filter's models for the time and measurement update are

$$\mathsf{s}_{k+1} = \mathsf{f}_s(\mathsf{s}_k) + \mathsf{g}_s(\mathsf{s}_k)\mathsf{w}_k \tag{33}$$

‡Some authors subtract three from the definition of kurtosis, so that Gaussian distributions have zero kurtosis.

and

$$\mathsf{z}_k = \mathsf{h}_s(\mathsf{s}_k) + \mathsf{v}_k. \tag{34}$$

The filter maintains an $n_s \times n_s$ formal covariance, $\hat{\mathsf{P}}$, which is based on the filter's assumed values for the *a priori* state error covariance, $\hat{\mathsf{P}}_o$, process noise covariance, $\hat{\mathsf{Q}}$, and measurement noise covariance, $\hat{\mathsf{R}}$. The formal covariance is used by the filter to update the solve-for state using

$$\hat{\mathsf{s}}^+ = \hat{\mathsf{s}}^- + \hat{\mathsf{K}}\left(\mathsf{y} - \hat{\mathsf{z}}^-\right) \tag{35}$$

where $\hat{\mathsf{s}}^-$ and $\hat{\mathsf{z}}^-$ are formed by merging the sigma points derived from Eqs. 33 and 34, analogous to the procedures of Eqs. 18-19, and Eqs. 24-25, as will be shown below. The full state is updated per Eq. 5, except with a gain that only applies update information to the solve-for projection:

$$\hat{\mathsf{x}}^+ = \hat{\mathsf{x}}^- + \tilde{\mathsf{S}}\hat{\mathsf{K}}\left(\mathsf{y} - \hat{\mathsf{y}}^-\right) \tag{36}$$

where $\tilde{\mathsf{S}}$ is formed from the first n_s columns of the inverse of $[\mathsf{S},\, \mathsf{C}]^{\mathsf{T}}$. Thus, the consider states are not updated by the gain. The difference between the projection of the true covariance into the solve-for space and the formal covariance,

$$\Delta\mathsf{P} = \mathsf{S}\mathsf{P}\mathsf{S}^{\mathsf{T}} - \hat{\mathsf{P}}, \tag{37}$$

is a measure of the error committed by the filter by ignoring the consider states. Since the noise inputs enter the system additively in the ADDSPF, the *a priori* errors, errors due to measurement noise, and errors due to process noise are independent, then as Ref. 7 shows, the covariances may be written as the sum the covariances due to each of these error sources[§]:

$$\mathsf{P} = \mathsf{P}_a + \mathsf{P}_v + \mathsf{P}_w \tag{40}$$

$$\hat{\mathsf{P}} = \hat{\mathsf{P}}_a + \hat{\mathsf{P}}_v + \hat{\mathsf{P}}_w \tag{41}$$

These covariance projections may be Cholesky-factored and the results collected so as to write the full covariance in terms of the Cholesky factors of the projections:

$$\mathsf{P} = \left[\sqrt[c]{\mathsf{P}}_a,\ \sqrt[c]{\mathsf{P}}_v,\ \sqrt[c]{\mathsf{P}}_w\right]\left[\sqrt[c]{\mathsf{P}}_a,\ \sqrt[c]{\mathsf{P}}_v,\ \sqrt[c]{\mathsf{P}}_w\right]^{\mathsf{T}} \tag{42}$$

$$\hat{\mathsf{P}} = \left[\sqrt[c]{\hat{\mathsf{P}}}_a,\ \sqrt[c]{\hat{\mathsf{P}}}_v,\ \sqrt[c]{\hat{\mathsf{P}}}_w\right]\left[\sqrt[c]{\hat{\mathsf{P}}}_a,\ \sqrt[c]{\hat{\mathsf{P}}}_v,\ \sqrt[c]{\hat{\mathsf{P}}}_w\right]^{\mathsf{T}} \tag{43}$$

[§]To study the sensitivity to individual channels, or groups, of the the *a priori* errors, errors due to measurement noise, and errors due to process noise, one may choose to use even more partitions, e.g.:

$$\mathsf{P} = \sum_i \mathsf{P}_{ai} + \sum_j \mathsf{P}_{vj} + \sum_k \mathsf{P}_{wk} \tag{38}$$

$$\hat{\mathsf{P}} = \sum_i \hat{\mathsf{P}}_{ai} + \sum_j \hat{\mathsf{P}}_{vj} + \sum_k \hat{\mathsf{P}}_{wk} \tag{39}$$

To perform a general covariance analysis of the ADDSPF, one merely applies the ADDSPF algorithm to the covariance projections individually, as follows. First, sigma points may be generated from the columns of the factors of all the covariance projections:

$$\hat{\mathcal{S}} = \hat{\mathsf{s}} \pm h \left[\sqrt[c]{\hat{\mathsf{P}}}_a,\ \sqrt[c]{\hat{\mathsf{P}}}_v,\ \sqrt[c]{\hat{\mathsf{P}}}_w \right] \tag{44}$$

$$\hat{\mathcal{X}} = \hat{\mathsf{x}} \pm h \left[\sqrt[c]{\mathsf{P}}_a,\ \sqrt[c]{\mathsf{P}}_v,\ \sqrt[c]{\mathsf{P}}_w \right] \tag{45}$$

Note that the partitioning of the covariance information has required that there are now $3n_s + 1$ n_s-dimensional sigma point vectors associated with Eq. 44, and $3n + 1$ additional n-dimensional sigma point vectors associated with Eq. 45. As before, as each measurement becomes available, pass the sets of sigma points through the nonlinear time update functions,

$$\hat{\mathcal{S}}_k^- = \mathsf{f}_s(\hat{\mathcal{S}}_{k-1}^+) \tag{46}$$

$$\hat{\mathcal{X}}_k^- = \mathsf{f}(\hat{\mathcal{X}}_{k-1}^+) \tag{47}$$

and compute the first- and second-order divided differences of the resulting propagated sigma points, $\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)$, $\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)$, $\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)$ and $\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)$. Because of the manner in which the sigma point arrays have been ordered, each of the divided difference matrices contains an ordered array of the divided differences corresponding to the various covariance partitions:

$$\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-) = \left[\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_a,\ \tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_v,\ \tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_w \right] \tag{48}$$

$$\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-) = \left[\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_a,\ \tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_v,\ \tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_w \right] \tag{49}$$

i.e., the first n_s columns of $\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)$ contain the divided differences associated with the sigma points arising from the *a priori* projection, the next n_s columns contain the divided differences associated with sigma points arising from the measurement noise projection, and the final n_s columns contain the divided differences associated with the sigma points arising from the process noise projection, *et sim.* for $\tilde{\mathsf{D}}^{(i)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)$. Thus, the time updates for the covariance projections may be extracted as thin QR factors as follows:

$$\hat{\mathsf{M}}_a^- \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_a^-}^{\mathsf{T}} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_a,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_a \right]^{\mathsf{T}} \tag{50}$$

$$\hat{\mathsf{M}}_v^- \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_v^-}^{\mathsf{T}} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_v,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_v \right]^{\mathsf{T}} \tag{51}$$

$$\hat{\mathsf{M}}_w^- \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_w^-}^{\mathsf{T}} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_w,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{f}_s(\hat{\mathsf{s}}^-)_w,\ \sqrt[c]{\hat{\mathsf{Q}}_d} \right]^{\mathsf{T}} \tag{52}$$

and

$$\mathsf{M}_a^- \begin{bmatrix} \sqrt[c]{\mathsf{P}_a^-}^{\mathsf{T}} \\ \mathsf{O}_{2n\times n} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_a,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_a\right]^{\mathsf{T}} \tag{53}$$

$$\mathsf{M}_v^- \begin{bmatrix} \sqrt[c]{\mathsf{P}_v^-}^{\mathsf{T}} \\ \mathsf{O}_{2n\times n} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_v,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_v\right]^{\mathsf{T}} \tag{54}$$

$$\mathsf{M}_w^- \begin{bmatrix} \sqrt[c]{\mathsf{P}_w^-}^{\mathsf{T}} \\ \mathsf{O}_{2n\times n} \end{bmatrix} = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_w,\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{f}(\hat{\mathsf{x}}^-)_w,\ \sqrt[c]{\mathsf{Q}_d}\right]^{\mathsf{T}} \tag{55}$$

Time updates for the states may be computed by merging the sigma points:

$$\hat{\mathsf{s}}^- = \mu_h(\hat{\mathcal{S}}^-) \tag{56}$$

$$\hat{\mathsf{x}}^- = \mu_h(\hat{\mathcal{X}}^-) \tag{57}$$

For the measurement update, new arrays of sigma points are generated and used to generate sets of sigma points representing the measurement:

$$\hat{\mathcal{Z}}^- = \mathsf{h}_s(\hat{\mathcal{S}}^*) \tag{58}$$

$$\hat{\mathcal{Y}}^- = \mathsf{h}(\hat{\mathcal{X}}^*) \tag{59}$$

In similar fashion to the time update, the sigma points of the measurement are then merged to form the estimated measurement,

$$\hat{\mathsf{z}}^- = \mu_h(\hat{\mathcal{Z}}^-) \tag{60}$$

$$\hat{\mathsf{y}}^- = \mu_h(\hat{\mathcal{Y}}^-) \tag{61}$$

and the corresponding divided-differences, $\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-)$, $\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-)$, $\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)$, and $\tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{x}}\mathsf{h}(\hat{\mathsf{x}}^-)$ are computed. To compute the filter's gain, only the covariances associated with the solve-for projection are required:

$$\hat{\mathsf{P}}_{zz}^- = \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-),\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-),\ \sqrt[c]{\hat{\mathsf{R}}}\right]\ \left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-),\ \tilde{\mathsf{D}}^{(2)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-),\ \sqrt[c]{\hat{\mathsf{R}}}\right]^{\mathsf{T}} \tag{62}$$

$$\hat{\mathsf{P}}_{sz}^- = \left[\sqrt[c]{\hat{\mathsf{P}}_a^-},\ \sqrt[c]{\hat{\mathsf{P}}_v^-},\ \sqrt[c]{\hat{\mathsf{P}}_w^-}\right]\left[\tilde{\mathsf{D}}^{(1)}_{\Delta\mathsf{s}}\mathsf{h}_s(\hat{\mathsf{s}}^-)\right]^{\mathsf{T}} \tag{63}$$

Note that all of the columns of the divided difference arrays are used. The filter's gain is then

$$\hat{\mathsf{K}} = \hat{\mathsf{P}}_{sz}^-(\hat{\mathsf{P}}_{zz}^-)^{-1}. \tag{64}$$

Finally, the measurement updates for the covariance projections may be extracted as thin QR factors as follows:

$$\hat{\mathsf{M}}_a^+ \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_a^+}^\mathsf{T} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\sqrt[c]{\hat{\mathsf{P}}_a^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(1)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_a, \; \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(2)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_a \right]^\mathsf{T} \tag{65}$$

$$\hat{\mathsf{M}}_v^+ \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_v^+}^\mathsf{T} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\sqrt[c]{\hat{\mathsf{P}}_v^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(1)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_v, \; \mathsf{K} \left[\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(2)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_v, \; \sqrt[c]{\hat{\mathsf{R}}} \right] \right]^\mathsf{T} \tag{66}$$

$$\hat{\mathsf{M}}_w^+ \begin{bmatrix} \sqrt[c]{\hat{\mathsf{P}}_w^+}^\mathsf{T} \\ \mathsf{O}_{2n_s \times n_s} \end{bmatrix} = \left[\sqrt[c]{\hat{\mathsf{P}}_w^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(1)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_w, \; \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{s}}^{(2)} \mathsf{h}_s(\hat{\mathsf{s}}^-)_w \right]^\mathsf{T} \tag{67}$$

and

$$\mathsf{M}_a^+ \begin{bmatrix} \sqrt[c]{\mathsf{P}_a^+}^\mathsf{T} \\ \mathsf{O}_{2n \times n} \end{bmatrix} = \left[\sqrt[c]{\mathsf{P}_a^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(1)} \mathsf{h}(\hat{\mathsf{x}}^-)_a, \; \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(2)} \mathsf{h}(\hat{\mathsf{x}}^-)_a \right]^\mathsf{T} \tag{68}$$

$$\mathsf{M}_v^+ \begin{bmatrix} \sqrt[c]{\mathsf{P}_v^+}^\mathsf{T} \\ \mathsf{O}_{2n \times n} \end{bmatrix} = \left[\sqrt[c]{\mathsf{P}_v^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(1)} \mathsf{h}(\hat{\mathsf{x}}^-)_v, \; \mathsf{K} \left[\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(2)} \mathsf{h}(\hat{\mathsf{x}}^-)_v, \; \sqrt[c]{\mathsf{R}} \right] \right]^\mathsf{T} \tag{69}$$

$$\mathsf{M}_w^+ \begin{bmatrix} \sqrt[c]{\mathsf{P}_w^+}^\mathsf{T} \\ \mathsf{O}_{2n \times n} \end{bmatrix} = \left[\sqrt[c]{\mathsf{P}_w^-} - \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(1)} \mathsf{h}(\hat{\mathsf{x}}^-)_w, \; \mathsf{K}\tilde{\mathsf{D}}_{\Delta \mathsf{x}}^{(2)} \mathsf{h}(\hat{\mathsf{x}}^-)_w \right]^\mathsf{T} \tag{70}$$

EXAMPLE

The example that this section describes demonstrates the approach to analysis of an ADDSPF that this paper proposes. This example studies GPS orbit determination for spacecraft in a highly elliptical, 1.2×12 Earth radii orbit, similar to the Phase 1 orbit for the Magnetospheric Multi-Scale mission, which Figure 1 shows. Table 1 lists the force modeling assumptions the example uses, and Table 2 lists its GPS measurement modeling assumptions[¶]. Figure 2 shows the GPS signal availability that results from the assumptions of Tables 1 and 2. As Figure 2 shows, for most of the orbit, at least one GPS signal is available, but only near the perigees does the number of signals exceed the minimum of four required for a deterministic fix.

Figure 3 shows results from covariance analysis of the ADDSPF, and Figure 4 shows results from a 10-case Monte Carlo analysis. For this case, the filter and truth models were the same, and the GPS receiver clock bias was not modeled. Estimation starts at apogee with spherical *a priori* uncertainty of 1 km in position and 10 m/sec in velocity. The filter processes all available GPS signals. Figure 3 shows that the influence of *a priori* uncertainty decays approximately 3 hours after apogee, well before the time when four or more GPS signals are visible. Once the *a priori* uncertainty decays, measurement noise contributes most of the total variance in the estimate, until after the first perigee. Thereafter, process noise becomes a somewhat larger contributer to total variance

[¶]One should *not* assume that the data listed in Tables 1 and 2 correspond to the MMS mission.

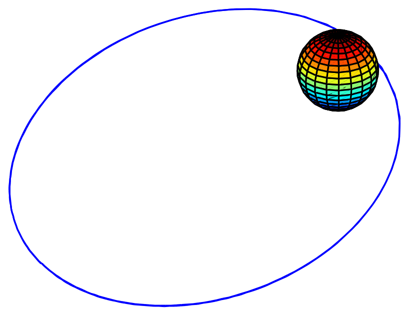

Figure 1. Highly Elliptical Orbit Used for Example.

Table 1. Force Modeling

Parameter	Value
Mass [kg]	200
Area [m^2]	1.12
Gravity Model	JGM–3
Gravity Degree and Order	8×8
Sun and Moon Point Mass Gravity	enabled
Drag	disabled
Solar Radiation Pressure	enabled
Solar Radiation Pressure Coefficient	1.4
Random Walk Acceleration Process Noise Intensity	$10\ \frac{\text{micro-gee}}{\sqrt{\text{sec}}}$

Table 2. GPS Measurement Modeling

Parameter	Value
Frequency	L_1
Number of Antennas	2
Antenna Pointing	Zenith and Nadir
Antenna Pattern	Hemispherical: 4 dB peak gain, $157\,\mathrm{deg}$ half beamwidth
Receiver antenna mask angle	$180\,\mathrm{deg}$
Atmosphere mask altitude	50 km
System noise temperature	300 K
Atmospheric attenuation	0 dB
GPS SV Transmit Power	Typical Block II-A
GPS SV Transmit Power Offset	0 dB
Transmitter Antenna Mask Angle	$180\,\mathrm{deg}$
Receiver Noise Figure	-3 dB
Receiver A/D Conversion Loss	-1.5 dB
System Loss	0 dB
Low-Noise Amplifier Gain	40 dB
Cable Loss	-2 dB
Receiver Acquisition Threshold	32 dB-Hz
Receiver Tracking Threshold	27 dB-Hz
Receiver Dynamic Tracking Range Limit	15 dB
Pseudo-range noise (1σ)	1 m

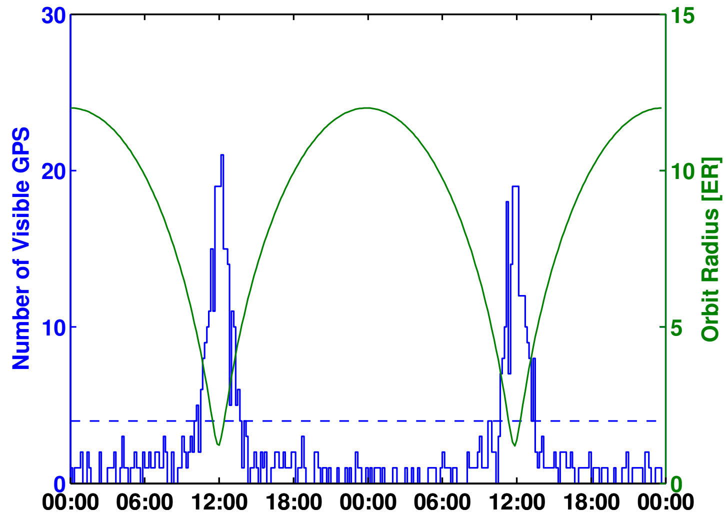

Figure 2 GPS signal availability over two orbits with 24-hour period; perigees occur near noon, and apogees near midnight.

near subsequent apogees, while measurement noise dominates the velocity variance at subsequent perigees. Observations such as these are enabled by the analysis approach this paper presents. In Figure 4, which shows the 10-case Monte Carlo results, red bands show the 95% confidence intervals on the empirical ensemble standard deviations. These confidence intervals significantly overlap the formal errors derived from the covariance analysis, which the figure shows as cyan bands, just visible at the margins of the red bands. Thus, Figure 4 confirms that the covariance analysis results are consistent with the Monte Carlo statistics, to the degree of confidence available from the limited number of Monte Carlo cases.

SUMMARY AND CONCLUSIONS

This paper has presented a generalized covariance analysis approach for additive divided difference sigma-point filters that partitions the contributions to the total error of *a priori*, process, and measurement noise. As an example, a simulated highly elliptical orbit was estimated from Global Positioning System measurements. Enabled by the unique approach that this paper developed, analysis of the example showed the relative significance of *a priori* uncertainty, measurement noise, and process noise to the total variance in the estimated states. The covariance analysis was confirmed by a 10-case Monte Carlo study.

ACKNOWLEDGMENTS

The authors wish to thank the reviewers of this paper.

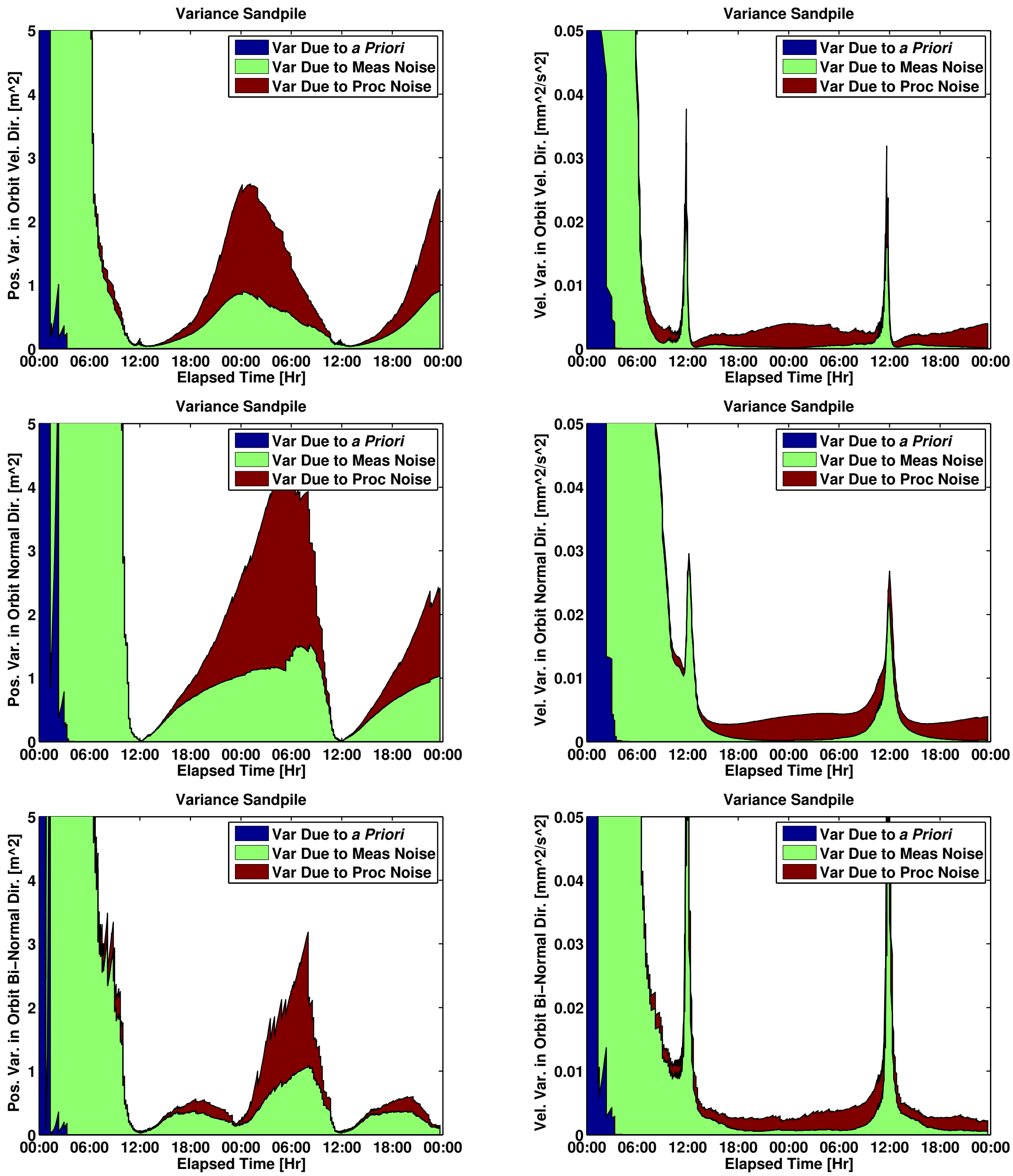

Figure 3 Variance sandpiles for highly elliptical (1.2 by 12 Earth radii) orbit with 24-hour period; perigees occur near noon, and apogees near midnight.

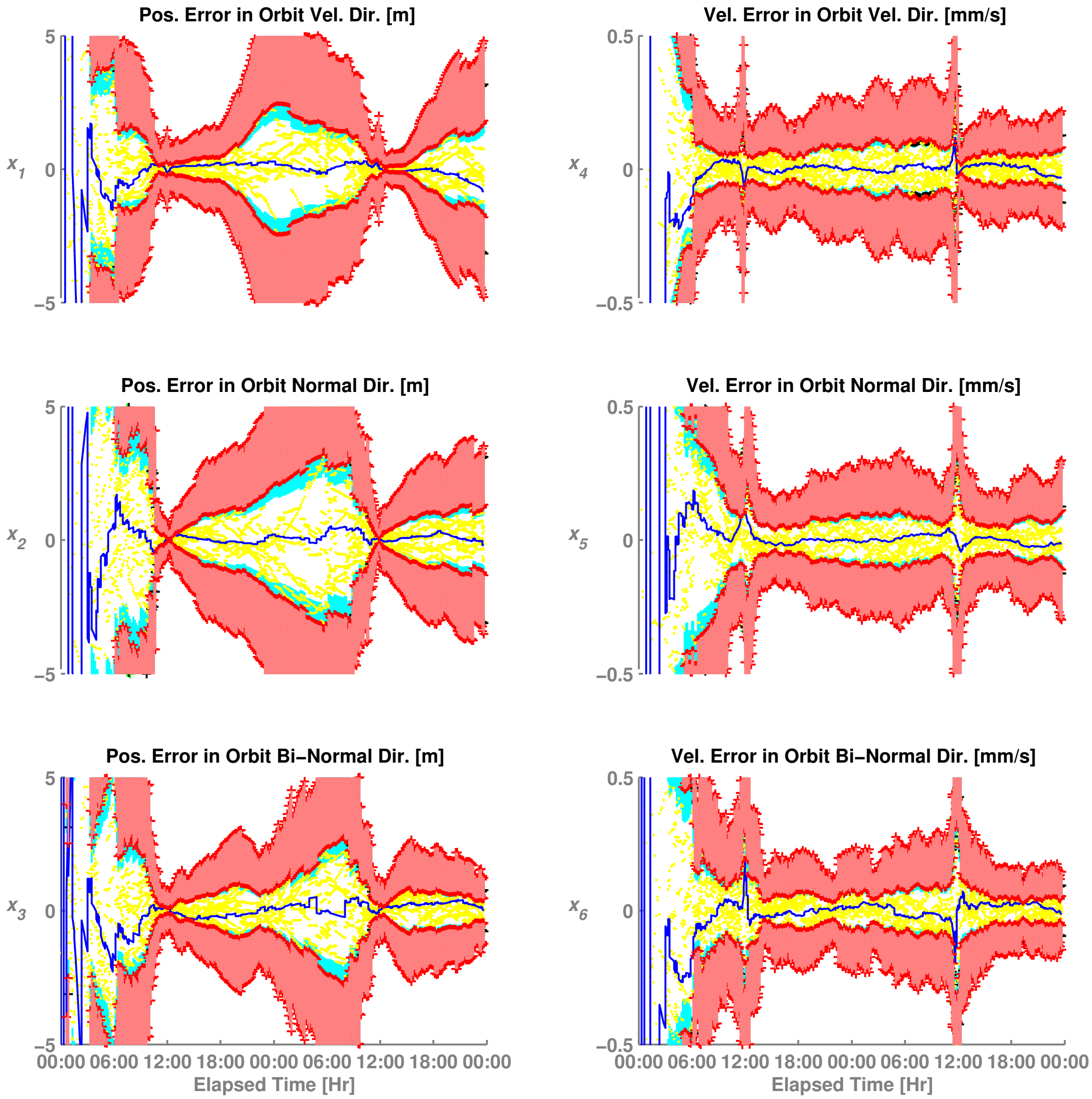

Figure 4 Formal and empirical estimation errors from 10-case Monte Carlo. Plot legend is as follows: yellow traces are the individual case time series; the blue trace is the empirical ensemble mean; red bands are 95% confidence intervals on the empirical ensemble standard deviations; and cyan bands bound $\pm 1\sigma$ to $\pm 3\sigma$ formal errors.

REFERENCES

[1] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, Vol. 45, No. 3, March 2000, pp. 477–482.

[2] T. Lefebvre, H. Bruyninckx, and J. D. Schutter, "Comment on 'A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators'," *IEEE Transactions on Automatic Control*, Vol. 47, No. 8, August 2002, pp. 1406–1408.

[3] S. Julier and J. Uhlmann, "Authors' Reply," *IEEE Transactions on Automatic Control*, Vol. 47, No. 8, August 2002, pp. 1408–1409.

[4] M. Nørgaard, N. K. Poulsen, and O. Ravn, "New Developments in State Estimation for Nonlinear Estimation Problems," *Automatica*, Vol. 36, No. 11, November 2000, pp. 1627–1638.

[5] M. Nørgaard, N. K. Poulsen, and O. Ravn, "Advances in Derivative-Free State Estimation for Nonlinear Systems," Tech. Rep. IMM–REP–1998–15, Technical University of Denmark, 2800 Lyngby, Denmark, April 7, 2000.

[6] D.-J. Lee and K. T. Alfriend, "Additive Divided Difference Filtering for Attitude Estimation Using Modified Rodrigues Parameters," *Proceedings of the F. Landis Markley Astronautics Symposium*, edited by J. L. Crassidis et al., Vol. 132 (CD–ROM Supplement) of *Advances in the Astronautical Sciences*, American Astronautical Society, Univelt, 2008.

[7] J. R. Carpenter and F. L. Markley, "Generalized Linear Covariance Analysis," *Proceedings of the F. Landis Markley Astronautics Symposium*, edited by J. L. Crassidis et al., Vol. 132 (CD–ROM Supplement) of *Advances in the Astronautical Sciences*, American Astronautical Society, Univelt, 2008.

[8] M. E. Lisano, "Nonlinear Consider Covariance Analysis Using a Sigma-Point Filter Formulation," *Guidance and Control 2006*, edited by S. D. Jolly and R. D. Culp, Vol. 125 of *Advances in the Astronautical Sciences*, American Astronautical Society, Univelt, San Diego, CA, 2006, pp. 129–144.

[9] M. E. Lisano, "Comparing Consider-Covariance Analysis With Sigma-Point Consider Filter and Linear-Theory Consider Filter Formulations," *Proceedings of the 20th International Symposium on Space Flight Dynamics (CD–ROM)*, No. NASA/CP–2007–214158, National Aeronautics and Space Administration, 2007.

[10] G. H. Golub and C. F. V. Loan, *Matrix Computations*, JHU Press, 3rd ed., 1996.

[11] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Academic Press, New York, 1977.

[12] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.