A Multipurpose Consider Covariance Analysis for Square-Root Information Smoothers

Joanna C. Hinks* and Mark L. Psiaki†

Cornell University, Ithaca, NY, 14853-7501

A new form of consider covariance analysis for studying incorrectly-modeled square-root information smoothers is presented and demonstrated. The value of this technique is its ability to compute the true estimation error covariance when the smoother has an incorrect dynamics model, an incorrect measurement model, or an incorrect statistical model. The new smoother analysis casts systems with a wide range of possible modeling errors into a special form and exploits square-root techniques to provide both generality and compactness. A consider covariance analysis can improve smoother design or characterize an existing smoother's true accuracy. Areas of application include incorrect initial state covariance, colored or correlated noise statistics, unestimated disturbance states, and erroneous system matrices. Several simple examples are developed to illustrate the application of the consider smoother analysis to models with a variety of error types, and these examples are independently validated by Monte Carlo simulation.

I. Introduction

DYNAMIC estimation methods compute two main outputs: Estimates of a system state vector and some metric of uncertainty associated with those estimates. This general statement holds for Kalman filters, square-root information filters (SRIFs), unscented filters, and particle filters, among other filter varieties. Most commonly, estimation error covariance is the chosen uncertainty metric. Filters typically form estimates and estimation error covariances recursively as measurements arrive, and thus they only use measurements up to and including the sample time at which the estimate applies. Conversely, smoothers employ measurements from both past and future times to compute better estimates and reduce estimation error covariance by "smoothing out" the uncertainty in the system. In particular, fixed-interval smoothers form estimates over an interval based on all the measurements obtained during that interval. When a filter or smoother is based on an incorrect system model, both the estimates and the estimation error covariance may be affected. The estimates may be degraded in some way, and the estimator will report an estimation error covariance that does not correspond to the true level of error.

One approach to the problem of estimators with incorrect models is consider covariance analysis. The name "consider covariance analysis" stems from the analysis' ability to "consider" the effects of various kinds of modeling errors on filter and smoother behavior. The analysis does this by computing the "true" estimation error covariance of the mismodeled estimator. Note that the resulting estimation error covariance is *true* relative to a particular mismodeled estimator and a particular assumed "truth" model. It is not the true covariance in an absolute sense unless the "truth" model perfectly describes the real system. Many of this paper's references give the name "covariance analysis" to any analysis that investigates covariances, whether or not the filter or smoother is based on an incorrect system model. Others use the name "sensitivity analysis", because comparison of the optimal and actual estimation error covariances provides a metric of the sensitivity of the system to modeling error. This paper uses the phrase "consider covariance analysis" for all such analyses.

Filter and smoother model errors may result from intentional simplifications, or they may be unavoidable due to poor understanding of the system dynamics or measurements. When simplifications have been made,

---

*Postdoctoral Assistant, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York. Email: jchinks@buffalo.edu, Member, AIAA. (Formerly Graduate Student, Sibley School of Mechanical & Aerospace Engineering, Cornell University)

†Professor, Sibley School of Mechanical & Aerospace Engineering, 206 Upson Hall. Associate Fellow, AIAA.

such as model order reduction or decoupling of dynamics, consider covariance analysis is valuable as a design tool. For each candidate simplified estimator, a consider analysis can indicate its expected estimation accuracy. In other situations where the true system characteristics are not well-known, even the most detailed filter or smoother is likely to deviate from the true system significantly. A simpler estimator design, while suboptimal, may be more robust and less sensitive to the potential errors. One can explore these trade-offs by an iterative consider analysis process, evaluating multiple combinations of proposed filters and hypothetical "truth" models. Consider covariance analysis may even be helpful in determining whether to implement a filter only or a filter/smoother for a given system. In the presence of model mismatch, smoothing may not improve estimation accuracy as much as it otherwise would. Alternatively, depending on the system and type of error, smoothing may eliminate most of the error effects and restore much of the performance of the optimal system.

As defined here, consider covariance analysis can be used to study the effects of many different classes of model errors. One of the most commonly addressed error classes is that of unmodeled constant biases in the dynamics or measurements.[1–3] Another function of consider analysis is to investigate systems with incorrect *a priori* state estimate covariance, incorrect process noise, or incorrect measurement noise.[4–7] This includes the situation where the noise is colored rather than white. One may also study systems that have unestimated dynamic states.[8–11] Finally, some analyses examine the situation where some or all of the state-space system matrices used by the estimator have errors.[9–16] Some methods for consider covariance analysis are narrow algorithms that focus on a single class of filter errors; others are general approaches that can analyze all of these filter error types simultaneously.

Another paper by the authors, Ref. 17, presents a new square-root information form of consider covariance analysis. This method treats all of the discussed error classes within a single framework. The present paper extends the consider covariance analysis of Ref. 17 to analyze a discrete, fixed-interval square-root information smoother (SRIS). In keeping with common smoother practice, this paper's consider smoother analysis operates on a Rauch-Tung-Striebel (RTS) square-root information smoother.[18] That is, the analysis requires that there is first a forward filtering pass using SRIF techniques, followed by a backward smoothing pass also using square-root techniques. The consider smoother analysis uses the same architecture: a forward-pass consider analysis of the SRIF, followed by a backward-pass consider analysis of the SRIS. The forward filtering pass is analyzed via the algorithms of Ref. 17. The present paper develops the backward-pass consider covariance analysis using data generated during the forward pass, in the spirit of the RTS method.

Several of the previous consider covariance analyses have investigated the smoother problem. Reference 9 analyzes a continuous-time RTS smoother and allows general model errors, but it does not present an equivalent discrete-time analysis. Both 16 and 19 operate on different non-standard two-filter smoothers rather than RTS-style smoothers. To the authors' knowledge, the only general consider analysis of a discrete-time RTS smoother is given in Refs. 14, 15. The derivation, which is performed in the covariance domain, is given only in Ref. 14, and the provided algorithms in both sources are very complicated with many intermediate calculations. While neither of Refs. 14, 15 explicitly allows for unestimated state vector elements or mutually-correlated process and measurement noise, the provided algorithms could be modified by an experienced analyst to accommodate these cases.

All of the above consider smoother algorithms operate in the covariance domain. In the information domain, Refs. 3, 4 analyze RTS square-root information smoothers, but only for systems with unestimated random biases. These sources do not address smoothers with other types of model error, such as unestimated dynamic disturbances, incorrect noise covariances, or erroneous system matrices.

This paper makes two main contributions. First, the new consider covariance analysis of Ref. 17 is extended to analyze fixed-interval smoothers with incorrect system models. The resulting smoother analysis is the only generalized consider covariance analysis for discrete RTS square-root information smoothers. The analysis derivation, which is analogous to the derivation of the SRIS equations, is believed to be more understandable than that of Ref. 14, and the final form is simpler and more compact. Second, this paper provides several simple, concrete numerical examples beyond those of Ref. 17. These examples demonstrate how the smoother analysis procedure can be applied to common modeling error situations. Some of these types of errors, such as non-independent process and measurement noise, cannot be handled by the algorithms of Refs. 14, 15 without significant pre-processing/model augmentation. Monte Carlo techniques validate the consider analyses of these examples.

Implementation of this paper's algorithms requires two building blocks. The first building block is the consider filter analysis of Ref. 17, which is used for the forward-filtering-pass stage of the present paper's

American Institute of Aeronautics and Astronautics

analysis. The second building block is the basic method of square-root information smoothing. Readers unfamiliar with SRIS techniques are encouraged to refer to Refs. 4, 20 for more background on the subject.

The remainder of this paper is organized as follows: Section II summarizes some of the results of Ref. 17 that are needed for this paper's analysis. In Section III, the algorithms for the consider covariance analysis of a smoother backward pass are developed and explained. Section IV discusses several interesting example problems and validates their consider analyses using Monte Carlo simulations. Conclusions are drawn in Section V.

II.   Background: Consider System Model & Forward-Pass Filter Analysis

A number of key ideas and results from Ref. 17 are necessary for the development of this paper's smoothing algorithms. These concepts are briefly reviewed in this section. Section A describes a special "consider form" of the system equations. Select equations from the forward-pass filter analysis are presented in Section B.

A.   Consider Form

The consider analysis developed in this paper and in Ref. 17 begins by casting each problem into a special, pre-defined consider form. The algorithms of Section III operate directly on matrices and equations in this special form without regard to how the problem statement was originally framed. As a result, a wide variety of model errors can be studied without modifying the core algorithms.

The philosophy of the consider form is simple. For each set of filter/smoother modeling errors, one can write both the estimator's assumed model equations and the "truth" model equations. All but one consider form equation is just a version of the corresponding "truth" equation, rewritten as the filter/ smoother equation with modified noise terms or an additive perturbation. The equation pairs are all listed for convenience in Table 1, which shows the parallel structure of the two model formulations. The filter/

Table 1.   Consider system model summary.

Equation	Filter/Smoother Version		Consider Version
State dynamics	$\boldsymbol{x}_{k+1} = \Phi_{fk}\boldsymbol{x}_k + \Gamma_{fk}\boldsymbol{w}_k$	$\Rightarrow$	$\boldsymbol{x}_{k+1} = \Phi_{fk}\boldsymbol{x}_k + \Gamma_{fk}\boldsymbol{w}_k + \Gamma_{xck}\boldsymbol{x_{c}}_k + \boldsymbol{b}_{xk}$
Measurement	$\boldsymbol{y}_k = H_{fk}\boldsymbol{x}_k + \boldsymbol{\nu}_k$	$\Rightarrow$	$\boldsymbol{y}_k = H_{fk}\boldsymbol{x}_k + H_{ck}\boldsymbol{x_{c}}_k + \boldsymbol{b}_{yk}$
Consider dynamics			$\boldsymbol{x_{c}}_{k+1} = \Phi_{ck}\boldsymbol{x_{c}}_k + \Gamma_{cck}\boldsymbol{w_{c}}_k$
Process noise information	$\mathcal{R}_{fwwk}\boldsymbol{w}_k = -\boldsymbol{\nu}_{wk}$	$\Rightarrow$	$\mathcal{R}_{fwwk}\boldsymbol{w}_k = -S_{wck}\boldsymbol{x_{c}}_k - \boldsymbol{b}_{wk}$
Initial state information	$\bar{\mathcal{R}}_{fxx0}\boldsymbol{x}_0 = \bar{\boldsymbol{z}}_0 - \bar{\boldsymbol{\nu}}_{x0}$	$\Rightarrow$	$\bar{\mathcal{R}}_{fxx0}\boldsymbol{x}_0 = \bar{\boldsymbol{z}}_0 - \bar{S}_{xc0}\boldsymbol{x_{c}}_0 - \bar{\boldsymbol{b}}_{c0}$

smoother equations in the center column of Table 1 are those of a standard square-root information filter/ smoother (SRIF/S). In the dynamics equation, the state vector $\boldsymbol{x}_k$ evolves as dictated by the smoother's state transition matrix Φ_{fk} and process noise influence matrix Γ_{fk}. The dynamics are driven by the stochastic process noise vector $\boldsymbol{w}_k$. Likewise, the smoother's assumed measurement sensitivity matrix H_{fk} is standard. Here and throughout this paper, the subscript "f" designates the versions of the given quantities assumed by the filter/smoother, and symbols without this subscript represent the corresponding "truth" quantities. This notation is chosen to be consistent with that of Ref. 17. Without loss of generality, the smoother's measurement equation is assumed to have been normalized so that the filter models $\boldsymbol{\nu}_k$ as a zero-mean, identity-covariance, white, Gaussian random vector, which is uncorrelated with the process noise. The smoother's assumed statistics for the process noise $\boldsymbol{w}_k$ are captured in the third equation of the center column, the process noise information equation. Its matrix $\mathcal{R}_{fwwk}$ is the inverse-square-root of the smoother's assumed process noise covariance matrix Q_{fk}. Because the process noise is also assumed to be Gaussian and white with zero mean, the process noise error term $\boldsymbol{\nu}_{wk}$ has zero mean and identity covariance. Finally, the fourth equation in the center column of Table 1 parameterizes the filter/smoother's initial estimation uncertainty. The vector $\bar{\boldsymbol{z}}_0$ is the *a priori* information state at sample $k = 0$, and $\bar{\mathcal{R}}_{fxx0}$ is the *a priori* inverse-square-root of the assumed initial estimation error covariance $\overline{P}_{fxx0}$. By construction, the state

American Institute of Aeronautics and Astronautics

information error $\bar{\boldsymbol{\nu}}_{x0}$ is assumed to be zero-mean, identity-covariance, and Gaussian.

The consider dynamics equation, the third equation in the right-hand column, is unique to the new consider system form; it has no filter/smoother counterpart in the center column of Table 1. It describes the behavior of the "consider state" $\boldsymbol{x}_{ck}$ using its own state transition matrix Φ_{ck}, process noise influence matrix Γ_{cck}, and consider process noise $\boldsymbol{w}_{ck}$. The consider vector $\boldsymbol{x}_{ck}$ captures in one quantity all the sources of randomness and uncertainty contained within the "truth" system. This typically includes the "truth" process and measurement noise, as well as the "truth" initial state uncertainty at sample $k = 0$. Also included in the consider vector are stochastic quantities that are not modeled by the filter/smoother, such as unmodeled zero-mean stochastic biases and unmodeled dynamic states that influence the "truth" system dynamics. It is possible to model quantities such as biases or white noise sequences using the consider dynamics equation by including rows of zeros in Γ_{cck} or in Φ_{ck}. The consider state transition matrix Φ_{ck} need not be invertible or even square, and this flexibility allows for rows of zeros or a changing number of consider state elements. Without loss of generality, the consider process noise $\boldsymbol{w}_{ck}$ is constrained to be zero-mean, identity-covariance, white, and Gaussian. The consider analysis algorithms require, again without loss of generality, that the initial consider vector $\boldsymbol{x}_{c0}$ be zero-mean, identity-covariance, Gaussian, and uncorrelated with the $\boldsymbol{w}_{ck}$ noise sequence.

Each of the remaining four consider equations in the right-hand column of Table 1 differs from the corresponding filter/smoother version by its last two terms. These terms have the same structure in all equations: a zero-mean stochastic part plus a deterministic, possibly-time-varying bias part. In this context, the term "bias" is used in its statistical sense, as a known value that perturbs the mean from zero. It is not necessarily a dynamical bias in the sense of being a constant independent of k. To make matters even more confusing, elements of $\boldsymbol{x}_{ck}$ sometimes represent constant biases that are sampled from zero-mean distributions, as will be apparent in one of the examples.

Each zero-mean stochastic component in the consider model is written as a coefficient matrix multiplying the consider state vector $\boldsymbol{x}_{ck}$. As previously stated, $\boldsymbol{x}_{ck}$ contains all the sources of randomness and uncertainty in the "truth" system. The coefficient matrices Γ_{xck}, H_{ck}, S_{wck}, and $\bar{S}_{xc0}$ have two roles. They select those portions of $\boldsymbol{x}_{ck}$ relating to a particular equation's noise terms, and they weight the stochastic perturbation appropriately so that the "truth" noise or error enters the equation in place of the assumed uncertainty. The deterministic biases $\boldsymbol{b}_{xk}$, $\boldsymbol{b}_{yk}$, $\boldsymbol{b}_{wk}$, and $\bar{\boldsymbol{b}}_{c0}$ do not depend on a particular noise or state estimate history, and can be precomputed. Among other roles, they allow the user to consider non-zero-mean noise.

An intuitive understanding of the consider form and the consider state vector is most easily gained by studying the examples in Section IV and Ref. 17. These examples demonstrate how specific filter/smoother error classes can be represented using the equations of Table 1. In particular, one of the examples in Ref. 17 shows how to handle the case where the "truth" system matrices Φ_k, Γ_k, and H_k differ from those assumed by the filter/smoother. This analysis is possible even though the consider form equations explicitly contain only the assumed system matrices Φ_{fk}, Γ_{fk}, and H_{fk}.

B.   Needed Outputs from Forward-Pass Consider Filter Analysis

In a traditional RTS smoother, select quantities computed during the forward filter pass must be stored for later use in the backward smoother pass. The consider smoother analysis likewise requires some matrices and vectors from the forward consider filter analysis to complete its algorithm. Specifically, two sets of quantities are required: There is a set of quantities related to the consider state vector $\boldsymbol{x}_{ck}$ and the way it propagates through the consider analysis, and a set that is the consider-analysis analog of the quantities stored for a traditional SRIF/S. The matrices Φ_{fk}, Γ_{fk}, and Γ_{xck} and the vector bias $\boldsymbol{b}_{xk}$ from the consider-form state dynamics equation of Table 1 are also needed.

During the forward filter pass, the consider analysis must store matrices related to the consider state vector $\boldsymbol{x}_{ck}$. Within the filter analysis, $\boldsymbol{x}_{ck}$ is replaced by a new vector $\begin{bmatrix} \boldsymbol{\alpha}_k^{\mathrm{T}} & \boldsymbol{\beta}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$, which satisfies the relationship:

$$\boldsymbol{x}_{ck} = \begin{bmatrix} L_{\alpha k} & L_{\beta k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \end{bmatrix} = L_{\alpha k} \boldsymbol{\alpha}_k + L_{\beta k} \boldsymbol{\beta}_k \tag{1}$$

The coefficient matrices $L_{\alpha k}$ and $L_{\beta k}$ are computed recursively at each filter sample k, and are required as inputs for this paper's smoother analysis algorithms. The vectors $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ are never explicitly computed. Rather, they are mathematical constructs with certain useful properties. Equation (1) gives the consider

American Institute of Aeronautics and Astronautics

vector $\boldsymbol{x}_{\boldsymbol{c}k}$ as a linear combination of the two vectors $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$, but the relationship cannot be inverted. The quantities $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ do not depend on $\boldsymbol{x}_{\boldsymbol{c}k}$ alone, but rather on the entire consider state history up to sample k: $\boldsymbol{x}_{\boldsymbol{c}0}, \boldsymbol{x}_{\boldsymbol{c}1}, \ldots, \boldsymbol{x}_{\boldsymbol{c}k}$. The vector $\boldsymbol{\alpha}_k$ can be thought of as the part of the consider state history that has an effect on the estimate of the state $\boldsymbol{x}_k$ at sample k. Conversely, the vector $\boldsymbol{\beta}_k$ is the part of the consider state history that has no effect on the estimate of $\boldsymbol{x}_k$ at sample k, while having the potential to affect the filter's estimates at later sample times.

An advantage of this representation is compactness. By construction, the vector $\boldsymbol{\alpha}_k$ has at most n_x elements, the number of elements in the state vector $\boldsymbol{x}_k$. Similarly, the combined vector $\begin{bmatrix} \boldsymbol{\alpha}_k^{\mathrm{T}} & \boldsymbol{\beta}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ has at most $n_x + n_{xck}$ elements, where n_{xck} is the number of elements in $\boldsymbol{x}_{\boldsymbol{c}k}$. Thus, $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ parameterize the error effects of the consider state history without requiring an increasing number of storage elements at every new sample.

One other significant feature makes $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ the preferred consider analysis variables: The composite vector formed from these quantities and the consider process noise $\boldsymbol{w}_{\boldsymbol{c}k}$ has zero mean and identity covariance.

$$\begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{\boldsymbol{c}k} \end{bmatrix} \sim \mathcal{N}(0, \mathrm{I}) \tag{2}$$

The composite vector in Eq. (2) is propagated according to:

$$\begin{bmatrix} \boldsymbol{\alpha}_{k+1} \\ \boldsymbol{\beta}_{k+1} \\ \boldsymbol{\gamma}_{k+1} \end{bmatrix} = C_{k+1} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{\boldsymbol{c}k} \end{bmatrix} \tag{3}$$

where the matrix C_{k+1} is orthonormal, and is used extensively in the smoother algorithms. The vector $\boldsymbol{\gamma}_{k+1}$ is not needed in further calculations. This propagation equation is derived in the filter analysis from the *a posteriori* state information equation and from the dynamics of the consider vector $\boldsymbol{x}_{\boldsymbol{c}k}$.

The second set of stored filter analysis quantities is analogous to the set of quantities stored by a standard SRIF/S. Fixed-interval RTS square-root information smoothing algorithms run backwards from $k = N$, the final sample, to $k = 0$. This backward recursion relies on the post-dynamic-propagation version of the process noise information equation. This paper's consider smoother derivations require the consider version of this propagated process noise information equation at every sample k. It takes the form:

$$\bar{\mathcal{R}}_{wwk}\boldsymbol{w}_k + \bar{\mathcal{R}}_{wxk+1}\boldsymbol{x}_{k+1} = \bar{\boldsymbol{z}}_{wk} - \begin{bmatrix} \bar{S}_{w\alpha k} & \bar{S}_{w\beta k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \end{bmatrix} - \bar{\boldsymbol{b}}_{wk} \tag{4}$$

This equation contains information about the process noise $\boldsymbol{w}_k$ that can be inferred from the estimate of the state vector $\boldsymbol{x}_{k+1}$. In the consider smoother analysis, the matrix coefficients $\bar{S}_{w\alpha k}$ and $\bar{S}_{w\beta k}$ and the bias vector $\bar{\boldsymbol{b}}_{wk}$ also convey information. They are used to smooth the calculations of the consider error effects in the smoothed state information equations.

The consider smoother analysis requires one additional equation from the forward-pass consider filter analysis: the terminal-sample consider form of the *a posteriori* state information equation. It is described in the next section in the context of the backward smoothing pass initialization.

III.   Consider Smoother Analysis Algorithms

The consider smoother analysis is based on a modification of a standard SRIF/S. This type of modification already has been portrayed in the second, fourth, and fifth rows of Table 1. In the center column, the noise terms $\boldsymbol{\nu}_k$, $\boldsymbol{\nu}_{wk}$, and $\bar{\boldsymbol{\nu}}_{x0}$ are each simple, zero-mean, identity-covariance Gaussian random vectors. Each of these simple noise terms is replaced in the right-hand column by its consider-analysis counterpart, which can have non-zero mean and non-identity covariance. This generalization to arbitrary noise means and covariances flows through the entire consider smoother analysis.

The consider analysis derivations for the smoother proceed in a manner that resembles mathematical induction. That is, one begins by defining a particular form for the smoother state information equation at generic sample $k + 1$ and by showing that this form holds for some specific value of $k + 1$. Next, one shows how to derive the smoother state information equation at preceding sample k by processing the equation at sample $k + 1$. This backward transition step also employs the consider-form model equations and various stored matrix and vector quantities from the forward-pass filter analysis.

A. Initial Definitions and Setup

At each sample $k+1$, the consider smoother analysis is assumed to have a state information equation of the generic form:

$$\mathcal{R}^*_{xxk+1}\boldsymbol{x}_{k+1} = \boldsymbol{z}^*_{k+1} - \begin{bmatrix} S^*_{x\alpha k+1} & S^*_{x\beta k+1} & S^*_{x\psi k+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{k+1} \\ \boldsymbol{\beta}_{k+1} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} - \boldsymbol{b}^*_{ck+1} \tag{5}$$

Furthermore, the error effects vector, $\begin{bmatrix} \boldsymbol{\alpha}^{\mathrm{T}}_{k+1} & \boldsymbol{\beta}^{\mathrm{T}}_{k+1} & \boldsymbol{\psi}^{\mathrm{T}}_{k+1} \end{bmatrix}^{\mathrm{T}}$, is assumed to have zero mean and identity covariance at sample $k+1$. The task of the consider smoother analysis derivation is to show how one can obtain a sample-k version of Eq. (5). This sample-k version must have an error effects vector $\begin{bmatrix} \boldsymbol{\alpha}^{\mathrm{T}}_{k} & \boldsymbol{\beta}^{\mathrm{T}}_{k} & \boldsymbol{\psi}^{\mathrm{T}}_{k} \end{bmatrix}^{\mathrm{T}}$ with zero mean and identity covariance.

The form of Eq. (5) is non-obvious and merits some discussion. The consider version of the filter's *a posteriori* state information equation for sample $k+1$, as given in Ref. 17, takes the form:

$$\widehat{\mathcal{R}}_{xxk+1}\boldsymbol{x}_{k+1} = \hat{\boldsymbol{z}}_{k+1} - S_{x\alpha k+1}\boldsymbol{\alpha}_{k+1} - \hat{\boldsymbol{b}}_{ck+1} \tag{6}$$

This state information equation resembles the consider form of the initial state information equation in Table 1, except that the stochastic part of its error has been written in terms of $\boldsymbol{\alpha}_{k+1}$ rather than $\boldsymbol{x}_{ck+1}$. Recall that the vector $\boldsymbol{\alpha}_{k+1}$ captures the estimation error effects of the entire consider state history up to and including sample $k+1$, i.e., the effects of $\boldsymbol{x}_{c0}, \boldsymbol{x}_{c1}, \ldots, \boldsymbol{x}_{ck+1}$. Thus, Eq. (6) is adequate for the forward filter stage of the consider analysis. The principles of smoothing, however, dictate that the smoothed estimate of $\boldsymbol{x}_{k+1}$ may be degraded by the uncertainty contained in the consider state history over the entire interval $\boldsymbol{x}_{c0}, \boldsymbol{x}_{c1}, \ldots, \boldsymbol{x}_{cN}$. This includes the effects up to sample $k+1$ represented by $\boldsymbol{\beta}_{k+1}$. It also includes the uncertainty introduced by the consider process noise vector at later samples, i.e., $\boldsymbol{w}_{ck+1}, \boldsymbol{w}_{ck+2}, \ldots, \boldsymbol{w}_{cN-1}$. By construction, the new vector $\boldsymbol{\psi}_{k+1}$ parameterizes the uncertainty contained in this consider process noise sequence at samples *after* k.

Before propagating Eq. (5) from $k+1$ back to sample k, it is necessary to initialize the induction by showing that the equation is satisfied for some specific $k+1$. With suitable definitions, the $k+1 = N$ version of Eq. (6) provides the initialization for Eq. (5). Thus, the consider filter analysis results at the terminal sample serve to initialize the consider smoother analysis. This choice for the initialization of the backward-smoother consider pass should not be surprising. In a traditional RTS SRIS, the backward-smoother pass starts from the final *a posteriori* state information equation that is produced by the filter. At the final sample $k+1 = N$, the filter has processed all of the measurements and the filter's estimate is identical to the smoothed estimate.

To show the equivalence of consider smoother analysis Eq. (5) and consider filter analysis Eq. (6) at sample $k+1 = N$, one must first equate the square-root information matrices, information state vectors, and deterministic bias vectors at this final sample time. One must also equate the coefficient matrix for the vector $\boldsymbol{\alpha}_N$:

$$\mathcal{R}^*_{xxN} = \widehat{\mathcal{R}}_{xxN}, \quad \boldsymbol{z}^*_N = \hat{\boldsymbol{z}}_N, \quad \boldsymbol{b}^*_{cN} = \hat{\boldsymbol{b}}_{cN}, \quad S^*_{x\alpha N} = S_{x\alpha N} \tag{7}$$

where the notation "$( )^*$" designates a smoother-analysis matrix or vector. The resulting version of Eq. (5) still has terms containing $\boldsymbol{\beta}_N$ and $\boldsymbol{\psi}_N$ that have no counterpart in filter-analysis Eq. (6). To complete the initialization, one sets the coefficient matrix $S^*_{x\beta N}$ to zero. The new vector $\boldsymbol{\psi}_N$ is defined as an empty vector for this terminal sample, and its coefficient matrix $S^*_{x\psi N}$ is defined as an empty matrix:

$$S^*_{x\beta N} = 0, \quad , \boldsymbol{\psi}_N = [\,], \quad S^*_{x\psi N} = [\,] \tag{8}$$

The definition of $\boldsymbol{\psi}_N$ as an empty vector also satisfies the identity-covariance requirement for $k+1 = N$. The composite vector has the correct initial distribution because $\begin{bmatrix} \boldsymbol{\alpha}^{\mathrm{T}}_N & \boldsymbol{\beta}^{\mathrm{T}}_N & \boldsymbol{\psi}^{\mathrm{T}}_N \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \boldsymbol{\alpha}^{\mathrm{T}}_N & \boldsymbol{\beta}^{\mathrm{T}}_N \end{bmatrix}^{\mathrm{T}}$, and $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ were previously defined by the filter analysis to be uncorrelated with identity covariances for all samples k.

B. Backwards Propagation

The backwards propagation of Eq. (5) from sample $k+1$ to sample k smoothes the estimates by incorporating the process noise information of Eq. (4). Both are repeated here for convenience:

$$\mathcal{R}^*_{xxk+1}\boldsymbol{x}_{k+1} = \boldsymbol{z}^*_{k+1} - \begin{bmatrix} S^*_{x\alpha k+1} & S^*_{x\beta k+1} & S^*_{x\psi k+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_{k+1} \\ \boldsymbol{\beta}_{k+1} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} - \boldsymbol{b}^*_{ck+1} \tag{9a}$$

$$\bar{\mathcal{R}}_{wwk}\boldsymbol{w}_k + \bar{\mathcal{R}}_{wxk+1}\boldsymbol{x}_{k+1} = \bar{\boldsymbol{z}}_{wk} - \begin{bmatrix} \bar{S}_{w\alpha k} & \bar{S}_{w\beta k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \end{bmatrix} - \bar{\boldsymbol{b}}_{wk} \tag{9b}$$

All vectors with index $k+1$ must be eliminated in these two equations in favor of their sample-k versions. In particular, $\boldsymbol{\alpha}_{k+1}$, $\boldsymbol{\beta}_{k+1}$, $\boldsymbol{\psi}_{k+1}$, and $\boldsymbol{x}_{k+1}$ must be replaced.

The zero-mean, identity covariance $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ have been propagated between samples by means of Eq. (3). This equation also involves the consider process noise vector $\boldsymbol{w}_{ck}$. In greater detail with individual blocks of the orthonormal propagation matrix C_{k+1} shown, the equation is:

$$\begin{bmatrix} \boldsymbol{\alpha}_{k+1} \\ \boldsymbol{\beta}_{k+1} \\ \boldsymbol{\gamma}_{k+1} \end{bmatrix} = C_{k+1} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \end{bmatrix} = \begin{bmatrix} C_{\alpha\alpha k+1} & C_{\alpha\beta k+1} & C_{\alpha wck+1} \\ C_{\beta\alpha k+1} & C_{\beta\beta k+1} & C_{\beta wck+1} \\ C_{\gamma\alpha k+1} & C_{\gamma\beta k+1} & C_{\gamma wck+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \end{bmatrix} \tag{10}$$

In order to eliminate $\boldsymbol{\alpha}_{k+1}$ and $\boldsymbol{\beta}_{k+1}$ from Eq. (9a), only the first two rows of Eq. (10) are required:

$$\begin{bmatrix} \boldsymbol{\alpha}_{k+1} \\ \boldsymbol{\beta}_{k+1} \end{bmatrix} = \begin{bmatrix} C_{\alpha\alpha k+1} & C_{\alpha\beta k+1} & C_{\alpha wck+1} \\ C_{\beta\alpha k+1} & C_{\beta\beta k+1} & C_{\beta wck+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \end{bmatrix} \tag{11}$$

Substitution of Eq. (11) into Eq. (9a) involves the multiplication of two large block matrices and the resulting block matrix is complicated. For the sake of compactness, it helps to first define an intermediate expression:

$$\begin{bmatrix} \tilde{S}^*_{x\alpha k} & \tilde{S}^*_{x\beta k} & \tilde{S}^*_{xwck} \end{bmatrix} \equiv \begin{bmatrix} S^*_{x\alpha k+1} & S^*_{x\beta k+1} \end{bmatrix} \begin{bmatrix} C_{\alpha\alpha k+1} & C_{\alpha\beta k+1} & C_{\alpha wck+1} \\ C_{\beta\alpha k+1} & C_{\beta\beta k+1} & C_{\beta wck+1} \end{bmatrix} \tag{12}$$

With this notation, Eq. (9a) can be written without reference to $\boldsymbol{\alpha}_{k+1}$ or $\boldsymbol{\beta}_{k+1}$:

$$\mathcal{R}^*_{xxk+1}\boldsymbol{x}_{k+1} = \boldsymbol{z}^*_{k+1} - \begin{bmatrix} \tilde{S}^*_{x\alpha k} & \tilde{S}^*_{x\beta k} & \tilde{S}^*_{xwck} & S^*_{x\psi k+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} - \boldsymbol{b}^*_{ck+1} \tag{13}$$

Next, the state $\boldsymbol{x}_{k+1}$ can be written in terms of quantities referenced to sample k by means of the consider version of the state dynamics equation in Table 1. In order to make the substitution, the consider state $\boldsymbol{x}_{ck}$ must first be replaced in this dynamics equation by $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ using the relationship defined in Eq. (1). The resulting state dynamics equation is:

$$\boldsymbol{x}_{k+1} = \Phi_{fk}\boldsymbol{x}_k + \Gamma_{fk}\boldsymbol{w}_k + \Gamma_{xck} \begin{bmatrix} L_{\alpha k} & L_{\beta k} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \end{bmatrix} + \boldsymbol{b}_{xk} \tag{14}$$

After substituting Eq. (14) into Eqs. (13) and (9b), rearranging terms, and stacking the results, one obtains:

$$\begin{bmatrix} (\bar{\mathcal{R}}_{wwk} + \bar{\mathcal{R}}_{wxk+1}\Gamma_{fk}) & \bar{\mathcal{R}}_{wxk+1}\Phi_{fk} \\ \mathcal{R}^*_{xxk+1}\Gamma_{fk} & \mathcal{R}^*_{xxk+1}\Phi_{fk} \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{x}_k \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{z}}_{wk} \\ \boldsymbol{z}^*_{k+1} \end{bmatrix}$$

$$- \begin{bmatrix} (\bar{S}_{w\alpha k} + \bar{\mathcal{R}}_{wxk+1}\Gamma_{xck}L_{\alpha k}) & (\bar{S}_{w\beta k} + \bar{\mathcal{R}}_{wxk+1}\Gamma_{xck}L_{\beta k}) & 0 & 0 \\ (\tilde{S}^*_{x\alpha k} + \mathcal{R}^*_{xxk+1}\Gamma_{xck}L_{\alpha k}) & (\tilde{S}^*_{x\beta k} + \mathcal{R}^*_{xxk+1}\Gamma_{xck}L_{\beta k}) & \tilde{S}^*_{xwck} & S^*_{x\psi k+1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix}$$

$$- \begin{bmatrix} \bar{\boldsymbol{b}}_{wk} + \bar{\mathcal{R}}_{wxk+1}\boldsymbol{b}_{xk} \\ \boldsymbol{b}^*_{ck+1} + \mathcal{R}^*_{xxk+1}\boldsymbol{b}_{xk} \end{bmatrix} \tag{15}$$

In Eq. (15), the left-hand side and the first term on the right-hand side are exactly the terms formed by a standard RTS SRIS prior to completing a backwards propagation step. The remaining terms take the place of a zero-mean, identity-covariance information error vector. They describe the combined effects of all of the filter/smoother model errors on the smoother's estimation error.

To complete the backwards propagation, one next uses standard orthonormal/upper-triangular (QR) factorization to compute an orthonormal matrix T_k^* such that the following relationship is satisfied:

$$T_k^* \begin{bmatrix} (\bar{\mathcal{R}}_{wwk} + \bar{\mathcal{R}}_{wxk+1}\Gamma_{fk}) & \bar{\mathcal{R}}_{wxk+1}\Phi_{fk} \\ \mathcal{R}_{xxk+1}^*\Gamma_{fk} & \mathcal{R}_{xxk+1}^*\Phi_{fk} \end{bmatrix} = \begin{bmatrix} \mathcal{R}_{wwk}^* & \mathcal{R}_{wxk}^* \\ 0 & \mathcal{R}_{xxk}^* \end{bmatrix} \tag{16}$$

In Eq. (16), the block matrix on the left-hand side is the input to the factorization, and T_k^* is the transpose of the orthonormal matrix output. The upper-triangular matrices $\mathcal{R}_{wwk}^*$ and $\mathcal{R}_{xxk}^*$ and the general matrix $\mathcal{R}_{wxk}^*$ are additional QR-factorization outputs. Every term in Eq. (15) is multiplied by the transformation T_k^*, and the resulting matrices and vectors are named to yield:

$$\begin{bmatrix} \mathcal{R}_{wwk}^* & \mathcal{R}_{wxk}^* \\ 0 & \mathcal{R}_{xxk}^* \end{bmatrix}\begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{x}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{z}_{wk}^* \\ \boldsymbol{z}_k^* \end{bmatrix} - \begin{bmatrix} S_{w\alpha k}^* & S_{w\beta k}^* & S_{wwck}^* & S_{w\psi k+1}^* \\ S_{x\alpha k}^* & S_{x\beta k}^* & \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} - \begin{bmatrix} \boldsymbol{b}_{wk}^* \\ \boldsymbol{b}_{ck}^* \end{bmatrix} \tag{17}$$

In this equation, the transformed matrix and vector terms appear in exactly the same order as the corresponding original terms from Eq. (15).

Neglecting the top row of Eq. (17), which is no longer needed, the bottom row is:

$$\mathcal{R}_{xxk}^*\boldsymbol{x}_k = \boldsymbol{z}_k^* - \begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix}\begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} - \boldsymbol{b}_{ck}^* \tag{18}$$

This information equation is almost in the desired form, which is identical to Eq. (9a) except that it is written in terms of quantities at sample k rather than sample $k + 1$. The current form differs from the desired equation by the presence of the two vectors $\boldsymbol{w}_{ck}$ and $\boldsymbol{\psi}_{k+1}$ rather than the single vector $\boldsymbol{\psi}_k$. To complete the recursion step, it is necessary to define a $\boldsymbol{\psi}_k$ and to compute a matrix $S_{x\psi k}^*$ such that the following relationship holds:

$$S_{x\psi k}^*\boldsymbol{\psi}_k = \begin{bmatrix} \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix}\begin{bmatrix} \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} \tag{19}$$

There is more than one possible $\left(\boldsymbol{\psi}_k, S_{x\psi k}^*\right)$ pair that satisfies Eq. (19). For any sensible choice, however, the vector $\boldsymbol{\psi}_k$ should not have more elements than necessary. To capture the estimation error effects of the consider process noise vector $\boldsymbol{w}_{cj}$ at all samples $j \geq k$, $\boldsymbol{\psi}_k$ requires at most n_x elements. Thus there are two cases to handle.

In the first and simplest case, the composite vector $\begin{bmatrix} \boldsymbol{w}_{ck}^{\mathrm{T}} & \boldsymbol{\psi}_{k+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ already has dimension less than or equal to the maximum dimension n_x. This is more likely to occur near the beginning of a smoothing pass since $\boldsymbol{\psi}_N$ is initialized as an empty vector. In this situation, the recursion can be completed by defining the composite vector and computing its matrix coefficient according to:

$$\boldsymbol{\psi}_k = \begin{bmatrix} \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix}, \quad S_{x\psi k}^* = \begin{bmatrix} \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix} \tag{20}$$

It is further possible to verify that the resulting composite error effects vector $\begin{bmatrix} \boldsymbol{\alpha}_k^{\mathrm{T}} & \boldsymbol{\beta}_k^{\mathrm{T}} & \boldsymbol{\psi}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ has the desired identity covariance. The new $\boldsymbol{\psi}_k$ has identity covariance because $\boldsymbol{\psi}_{k+1}$ and $\boldsymbol{w}_{ck}$ individually have identity covariance, and because $\boldsymbol{\psi}_{k+1}$ depends only on the history of the consider process noise at future samples and so is uncorrelated with $\boldsymbol{w}_{ck}$. As $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$ depend on the history of the consider state vector $\boldsymbol{x}_{ck}$ up to and including sample k, they are uncorrelated with this $\boldsymbol{\psi}_k$ which depends only on consider process noise at samples greater than or equal to k.

American Institute of Aeronautics and Astronautics

The second case is slightly more complicated. If $\begin{bmatrix} \boldsymbol{w}_{ck}^{\mathrm{T}} & \boldsymbol{\psi}_{k+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ has dimension greater than n_x, then the error effects contained in the composite vector can be compressed into a vector of dimension n_x by means of lower-triangular/orthonormal (LQ) factorization. This factorization procedure is somewhat less common than QR factorization, but is closely related. It is described in more detail in the Appendix of Ref. 17. In this situation, the LQ factorization computes an orthonormal transformation C_k^* that satisfies:

$$\begin{bmatrix} S_{x\psi k}^* & 0 \end{bmatrix} C_k^* = \begin{bmatrix} \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix} \tag{21}$$

The block sensitivity matrix on the right-hand side is the input to the LQ factorization. In addition to the orthonormal C_k^*, the block lower-triangular matrix on the left-hand side is an output. The matrix $S_{x\psi k}^*$ is thus square and lower-triangular, with dimension n_x. When the composite vector $\begin{bmatrix} \boldsymbol{w}_{ck}^{\mathrm{T}} & \boldsymbol{\psi}_{k+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ is transformed by C_k^*, the result is defined as:

$$\begin{bmatrix} \boldsymbol{\psi}_k \\ \boldsymbol{\zeta}_k \end{bmatrix} = C_k^* \begin{bmatrix} \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} \tag{22}$$

The resulting $\boldsymbol{\psi}_k$ has the desired dimension n_x. The transformation given by Eqs. (21) and (22) can be shown to satisfy Eq. (19):

$$\begin{bmatrix} \bar{S}_{xwck}^* & \bar{S}_{x\psi k+1}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} = \begin{bmatrix} S_{x\psi k}^* & 0 \end{bmatrix} C_k^* \begin{bmatrix} \boldsymbol{w}_{ck} \\ \boldsymbol{\psi}_{k+1} \end{bmatrix} = \begin{bmatrix} S_{x\psi k}^* & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\psi}_k \\ \boldsymbol{\zeta}_k \end{bmatrix} = S_{x\psi k}^* \boldsymbol{\psi}_k \tag{23}$$

Furthermore, the orthonormality of the transformation C_k^* preserves the identity covariance of $\boldsymbol{\psi}_k$, and it remains uncorrelated with $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k$.

In either of the two cases represented by Eqs. (20)-(22), smoother state information Eq. 18 reduces to

$$\mathcal{R}_{xxk}^* \boldsymbol{x}_k = \boldsymbol{z}_k^* - \begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\psi}_k \end{bmatrix} - \boldsymbol{b}_{ck}^* \tag{24}$$

which is just where the consider smoother analysis started, except indexed to k instead of $k+1$. The RTS process is thus complete, and the next iteration of the backward recursion may begin.

C.   Smoothed Estimation Error Covariances

In a typical square-root information smoother, a smoothed state information equation is computed at each step of the backward pass. It takes the form:

$$\mathcal{R}_{xxk}^* \boldsymbol{x}_k = \boldsymbol{z}_k^* - \boldsymbol{\nu}_{xk}^* \tag{25}$$

where $\mathcal{R}_{xxk}^*$ is the inverse-square-root of the smoothed estimation error covariance, $\boldsymbol{z}_k^*$ is the smoothed information vector, and the error vector $\boldsymbol{\nu}_{xk}^*$ is assumed to have zero mean and identity covariance. The smoothed estimate can be computed from Eq. (25) as:

$$\boldsymbol{x}_k^* = \mathcal{R}_{xxk}^{*-1} \boldsymbol{z}_k^* \tag{26}$$

and the smoothed estimation error is

$$\boldsymbol{x}_k^* - \boldsymbol{x}_k = \mathcal{R}_{xxk}^{*-1} \boldsymbol{\nu}_{xk}^* \tag{27}$$

The smoothed estimation error covariance is therefore:

$$P_{fxxk}^* = \mathrm{E}\left[\left(\mathcal{R}_{xxk}^{*-1} \boldsymbol{\nu}_{xk}^*\right) \left(\mathcal{R}_{xxk}^{*-1} \boldsymbol{\nu}_{xk}^*\right)^{\mathrm{T}} \right] = \mathcal{R}_{xxk}^{*-1} \mathrm{E}\left[\boldsymbol{\nu}_{xk}^* \boldsymbol{\nu}_{xk}^{*\mathrm{T}} \right] \mathcal{R}_{xxk}^{*-\mathrm{T}} = \mathcal{R}_{xxk}^{*-1} \mathcal{R}_{xxk}^{*-\mathrm{T}} \tag{28}$$

The consider analysis version of this calculation substitutes its own complicated error term in place of the simplistic $\boldsymbol{\nu}_{xk}^*$. It starts with the consider version of the smoothed information equation.

$$\mathcal{R}_{xxk}^* \boldsymbol{x}_k = \boldsymbol{z}_k^* - \boldsymbol{\eta}_k^* \tag{29}$$

where the error vector $\boldsymbol{\eta}_k^*$ replaces $\boldsymbol{\nu}_{xk}^*$ of the standard smoother and is defined as

$$\boldsymbol{\eta}_k^* \equiv \begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\psi}_k \end{bmatrix} + \boldsymbol{b}_{ck}^* \tag{30}$$

Equation (29) is just a more compact way of writing Eq. (24) from the end of the preceding subsection. Substitution of $\boldsymbol{\eta}_k^*$ and its statistics into the smoother covariance calculations of Eq. (28) yields:

$$\begin{aligned}
P_{xxk}^* &= \mathrm{E}\left[\left(\mathcal{R}_{xxk}^{*-1} \boldsymbol{\eta}_k^* \right) \left(\mathcal{R}_{xxk}^{*-1} \boldsymbol{\eta}_k^* \right)^{\mathrm{T}} \right] \\
&= \mathcal{R}_{xxk}^{*-1} \mathrm{E}\left[\left(\begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\psi}_k \end{bmatrix} + \boldsymbol{b}_{ck}^* \right) \right. \\
&\qquad \left. \times \left(\begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\psi}_k \end{bmatrix} + \boldsymbol{b}_{ck}^* \right)^{\mathrm{T}} \right] \mathcal{R}_{xxk}^{*-\mathrm{T}} \\
&= \mathcal{R}_{xxk}^{*-1} \begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \mathrm{E}\left\{ \begin{bmatrix} \boldsymbol{\alpha}_k \\ \boldsymbol{\beta}_k \\ \boldsymbol{\psi}_k \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha}_k^{\mathrm{T}} & \boldsymbol{\beta}_k^{\mathrm{T}} & \boldsymbol{\psi}_k^{\mathrm{T}} \end{bmatrix} \right\} \begin{bmatrix} S_{x\alpha k}^{*\mathrm{T}} \\ S_{x\beta k}^{*\mathrm{T}} \\ S_{x\psi k}^{*\mathrm{T}} \end{bmatrix} \mathcal{R}_{xxk}^{*-\mathrm{T}} \\
&\qquad + \mathcal{R}_{xxk}^{*-1} \boldsymbol{b}_{ck}^* \boldsymbol{b}_{ck}^{*\mathrm{T}} \mathcal{R}_{xxk}^{*-\mathrm{T}} \\
&= \mathcal{R}_{xxk}^{*-1} \begin{bmatrix} S_{x\alpha k}^* & S_{x\beta k}^* & S_{x\psi k}^* \end{bmatrix} \begin{bmatrix} S_{x\alpha k}^{*\mathrm{T}} \\ S_{x\beta k}^{*\mathrm{T}} \\ S_{x\psi k}^{*\mathrm{T}} \end{bmatrix} \mathcal{R}_{xxk}^{*-\mathrm{T}} + \mathcal{R}_{xxk}^{*-1} \boldsymbol{b}_{ck}^* \boldsymbol{b}_{ck}^{*\mathrm{T}} \mathcal{R}_{xxk}^{*-\mathrm{T}} \tag{31}
\end{aligned}$$

The resulting quantity, P_{xxk}^*, is the true matrix mean square error (MSE) for the smoother. It is equal to the true smoothed estimation error covariance plus a rank-one term relating to the deterministic bias $\boldsymbol{b}_{ck}^*$. The true covariance is thus the left-hand term of the last line of Eq. (31). Although the use of "P_{xx}" to represent matrix MSE rather than covariance is non-standard, it emphasizes the role of P_{xxk}^* in the consider analysis. The matrix MSE P_{xxk}^*, in contrast to the estimation error covariance, contains the full mismodeling effect from both random and deterministic errors. Consequently, it is the most appropriate quantity for comparison with the smoother-assumed covariance P_{fxxk}^* of Eq. 28.

IV.   Examples

In order to demonstrate the consider smoother analysis, several concrete examples have been developed. These examples, while simple, illustrate several common classes of filter/smoother modeling errors. They complement Ref. 17 by addressing some error varieties not investigated by that paper's examples. The examples in this section further clarify how one can pose consider analyses in the consider model form of Section A.

For each example, the techniques of Section A have been employed to write the "truth" system equations in the defined consider form. This procedure is often the most challenging part of the consider analysis. Once in the correct form, the equations can be manipulated to perform a forward consider filter analysis and a backward consider smoother analysis, as per the algorithms of Ref. 17 and Section III of the present paper. The results of each consider smoother analysis have been independently validated by Monte Carlo analysis.

The examples to be addressed are as follows: The first example investigates the effects of a biased initial state estimate and incorrect initial covariance. The next example is a case with mutually correlated process and measurement noise, both with incorrect covariances. A third example examines a system that has dynamics perturbed by unmodeled colored process noise with a sinusoidal influence matrix. This case also has an unmodeled random bias in its measurements. All of these examples share the same incorrect filter/ smoother for easy comparison, but the nature of its errors varies.

American Institute of Aeronautics and Astronautics

A. Filter/Smoother Assumed Model Description

The filter/smoother for all of the examples has assumed dynamics and measurement models given by

$$\begin{bmatrix} r_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_k \tag{32a}$$

$$y_k = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \nu_k \tag{32b}$$

This system can be thought of as describing some one-dimensional motion, with position state r_k and velocity state v_k. The sample interval, Δt, is assumed to be 0.5 s. The scalar y_k is a measurement. The filter/smoother's assumed system matrices Φ_{fk}, Γ_{fk}, and H_{fk} can be directly extracted from these equations.

Scalar process noise w_k and measurement noise ν_k are assumed to be zero-mean, Gaussian, and white, with joint covariance given by:

$$\mathrm{E}\left\{ \begin{bmatrix} w_k \\ \nu_k \end{bmatrix} \begin{bmatrix} w_k & \nu_k \end{bmatrix} \right\} = \begin{bmatrix} Q_{fk} & 0 \\ 0 & R_{fk} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{33}$$

where Q_{fk} and R_{fk} are the nominal process and measurement noise covariances, respectively. In other words, Eq. (33) specifies that the noise processes are assumed to be uncorrelated, and both are assumed to have unit variance. The filter/smoother's assumed square-root information matrices for process noise and measurement noise are the inverse square roots of the corresponding assumed covariances matrices. In the general case, such matrix square roots can be computed by a standard method such as Cholesky factorization. For this scalar case, however, they are just $\mathcal{R}_{fwwk} = 1/\sqrt{1} = 1$ and $\mathcal{R}_{f\nu\nu k} = 1/\sqrt{1} = 1$. Note that the assumed measurement model form in the center column of Table 1 requires, without loss of generality, that $\mathcal{R}_{f\nu\nu k}$ be the identity matrix in order to be consistent with standard SRIF/S practice.

The initial state estimate and initial state error covariance are assumed to be:

$$\bar{x}_0 = \begin{bmatrix} \bar{r}_0 \\ \bar{v}_0 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad \overline{P}_{fxx0} = \begin{bmatrix} \sigma_{r0}^2 & \sigma_{rv0} \\ \sigma_{rv0} & \sigma_{v0}^2 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \tag{34}$$

The initial estimation error is further assumed to be uncorrelated with process or measurement noise at any sample time. From Eq. (34), the assumed initial state square-root information matrix $\mathcal{R}_{fxx0}$ and the initial information state $\bar{z}_0$ can be computed. They are:

$$\bar{\mathcal{R}}_{fxx0} = \begin{bmatrix} 1/\sqrt{10} & 0 \\ 0 & 1/\sqrt{5} \end{bmatrix}, \quad \bar{z}_0 = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{5} \end{bmatrix} \tag{35}$$

The filter/smoother operates on measurements available once per sample interval Δt for a total of 50 s (100 discrete-time samples). Note that all of the filter/smoother equations conform to the assumed forms in Table 1 of Section A.

B. Example: Incorrect Initialization

The first example has error only in its initialization. Its dynamics and measurement models are assumed to be correctly given by the filter/smoother's models of Eqs. (32a) and (32b). Likewise, the Gaussian process and measurement noise are assumed to be correctly modeled as zero-mean, white, and uncorrelated, with unit variance. Process and measurement noise are uncorrelated with the initial estimation error.

Two kinds of error enter the initialization: An incorrect initial covariance and a deterministic estimation bias. The filter/smoother assumes that the initial estimation error has zero mean and covariance $\overline{P}_{fxx0}$. This example's "truth" system instead has an initial estimate that relates to the "truth" initial state according to:

$$\bar{x}_0 = x_0 + \tilde{x}_0 + x_{b0} \tag{36}$$

where $\tilde{x}_0$ is the zero-mean random component of the error in the initial estimate, and x_{b0} is a deterministic non-zero bias error. The covariance of the random error $\tilde{x}_0$ and the value of the bias x_{b0} are given by:

$$\overline{P}_{xx0} = \mathrm{E}\big[\tilde{x}_0 \tilde{x}_0^{\mathrm{T}}\big] = \begin{bmatrix} 16 & 0 \\ 0 & 9 \end{bmatrix}, \quad x_{b0} = \begin{bmatrix} -20 \\ 30 \end{bmatrix} \tag{37}$$

American Institute of Aeronautics and Astronautics

Note that the covariance of $\tilde{\boldsymbol{x}}_0$ is the "truth" initial state error covariance $\overline{P}_{xx0}$, which in this case is not equal to the filter/smoother's initial state error covariance $\overline{P}_{fxx0}$. The corresponding "truth" initial square-root information matrix for the state is

$$\bar{\mathcal{R}}_{xx0} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/3 \end{bmatrix} \tag{38}$$

To write this example's system in consider form, one must first choose an appropriate consider state vector $\boldsymbol{x_{ck}}$. It must contain all the random uncertainty that enters via the initial estimate, the process noise, and the measurement noise. One such definition is:

$$\boldsymbol{x_{ck}} \equiv \begin{cases} \begin{bmatrix} \bar{\mathcal{R}}_{xx0}\tilde{\boldsymbol{x}}_0 \\ \mathcal{R}_{ww0}w_0 \\ \mathcal{R}_{\nu\nu0}\nu_0 \end{bmatrix} & k = 0 \\ \begin{bmatrix} w_k \\ \nu_k \end{bmatrix} & k > 0 \end{cases} \tag{39}$$

Note how only the stochastic part of the initial estimation error enters this definition. Premultiplication by the various square-root information matrices at sample $k = 0$ is required to ensure that $\boldsymbol{x_{c0}}$ has identity covariance.

The "truth" system matrices Φ_k, Γ_k, and H_k are identical to those of the filter, as are the "truth" noise square-root information matrices $\mathcal{R}_{wwk}$ and $\mathcal{R}_{\nu\nu k}$. There is no dynamics bias or measurement bias, i.e., $\boldsymbol{b}_{xk} = 0$ and $\boldsymbol{b}_{yk} = 0$. Because there are no unmodeled disturbances to the system dynamics, the matrix Γ_{xck} is an appropriately-sized matrix of zeros for all samples k.

Next, one can define the dynamics of the consider state $\boldsymbol{x_{ck}}$. The first component at $k = 0$, $\bar{\mathcal{R}}_{xx0}\tilde{\boldsymbol{x}}_0$, exists only at this sample and thus has no true dynamics. Likewise, the noise components w_k and ν_k are both white, and therefore they do not depend on noise at previous samples. Consequently, the consider state transition matrix Φ_{ck} is also an appropriately-sized matrix of zeros for all samples k. In the special case of $k = 0$, the matrix Φ_{c0} is rectangular, with fewer rows than columns in order to omit the initial uncertainty component at the next sample. The consider process noise influence matrix Γ_{cck} gives w_k and ν_k their proper statistics. Specifically,

$$\Gamma_{cck} = \begin{bmatrix} \mathcal{R}_{wwk+1}^{-1} & 0 \\ 0 & \mathcal{R}_{\nu\nu k+1}^{-1} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \forall\, k \tag{40}$$

The process noise information equation models the way the "truth" process noise statistics enter the system. Because the process noise is unbiased, $\boldsymbol{b}_{wk}$ is zero. As per the consider-form process noise information equation of Table 1, the matrix S_{wck} is given by

$$S_{wck} = \begin{cases} \begin{bmatrix} \boldsymbol{0} & -\mathcal{R}_{fww0}\mathcal{R}_{ww0}^{-1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 0 \end{bmatrix} & k = 0 \\ \begin{bmatrix} -\mathcal{R}_{fwwk} & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \end{bmatrix} & k > 0 \end{cases} \tag{41}$$

which picks out the properly weighted process noise component of $\boldsymbol{x_{ck}}$. Note that the $\boldsymbol{0}$ on the top line of Eq. (41) indicates a matrix of zeros with appropriate dimensions in all of this paper's examples. The previous sections did not use this convention because no dimensional ambiguities arose for 0 matrices. In a similar manner to S_{wck}, H_{ck} extracts the properly weighted measurement noise component of $\boldsymbol{x_{ck}}$:

$$H_{ck} = \begin{cases} \begin{bmatrix} \boldsymbol{0} & 0 & \mathcal{R}_{\nu\nu0}^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} & k = 0 \\ \begin{bmatrix} 0 & 1 \end{bmatrix} & k > 0 \end{cases} \tag{42}$$

The remaining and most difficult part of the consider form setup for this example relates to the initial state information equation, where the only considered model errors enter. It is necessary to place the "truth" initial information equation into its consider form, repeated here for convenience:

$$\bar{\mathcal{R}}_{fxx0}\boldsymbol{x}_0 = \bar{\boldsymbol{z}}_0 - \bar{S}_{xc0}\boldsymbol{x_{c0}} - \bar{\boldsymbol{b}}_{c0} \tag{43}$$

In order to determine the correct $\bar{S}_{xc0}$ and $\bar{\boldsymbol{b}}_{c0}$, one multiplies Eq. (36) by the filter/smoother's presumed initial square root information matrix $\bar{\mathcal{R}}_{fxx0}$, and one rearranges the result into something similar to Eq. (43).

$$\bar{\mathcal{R}}_{fxx0}\boldsymbol{x}_0 = \bar{\mathcal{R}}_{fxx0}\bar{\boldsymbol{x}}_0 - \bar{\mathcal{R}}_{fxx0}\tilde{\boldsymbol{x}}_0 - \bar{\mathcal{R}}_{fxx0}\boldsymbol{x}_{b0} \tag{44}$$

The first term on the right-hand side of Eq. (44) is by definition equal to the filter/smoother's initial *a priori* information state $\bar{z}_0$. The second term is zero-mean and stochastic, so by comparison of Eqs. (43) and (44), $\bar{S}_{xc0}$ must be chosen such that $\bar{S}_{xc0}\boldsymbol{x}_{c0} = \mathcal{R}_{fxx0}\tilde{\boldsymbol{x}}_0$. Consistent with the first line of Eq. (39), this is accomplished by the definition:

$$\bar{S}_{xc0} = \begin{bmatrix} (\mathcal{R}_{fxx0}\mathcal{R}_{xx0}^{-1}) & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{10}} & 0 & 0 & 0 \\ 0 & \frac{3}{\sqrt{5}} & 0 & 0 \end{bmatrix} \tag{45}$$

Likewise, the final terms in Eqs. (43) and (44) are both deterministic initial information biases. Setting them equal yields:

$$\bar{\boldsymbol{b}}_{c0} = \mathcal{R}_{fxx0}\boldsymbol{x}_{b0} = \begin{bmatrix} -\frac{20}{\sqrt{10}} \\ \frac{30}{\sqrt{5}} \end{bmatrix} \tag{46}$$

At this point, all the consider form quantities have been defined, and the filter/smoother consider analysis algorithms of Ref. 17 and the present paper's Section III can be implemented.

In addition to implementation of the analytical consider algorithms, numerical Monte Carlo simulations have been used to independently verify the computed smoothed matrix mean square errors. This technique forms a large number of simulated estimation errors by comparing "truth" state vectors to estimated states produced by the mismodeled filter/smoother. Each simulated estimation error history is a sample of a random process, and sample means and matrix MSEs can be computed. As the number of trials increases, the estimated matrix MSE or covariance approaches the true value. For each of this paper's examples 5000 trials have been used.

To implement the Monte Carlo simulations, a "truth" model is created for the given example based on the defined "truth" system matrices and "truth" noise covariances. The "truth" model takes as inputs the initial state and histories of the measurement and process noise, which are constrained to have the "truth" statistical properties as specified in the example. It outputs a "truth" history of the state vector and a series of noisy measurements. The noisy measurements from the "truth" model are next used as the inputs to the filter/smoother described in Section A, and it outputs histories of state estimates. Estimation error histories are computed as the difference of these estimates and the states from the "truth" model.

Note that the "truth" model implementation does not use the special consider form defined in Section A. This independence from the consider model form is an important feature of the Monte Carlo tests. Close correspondence between the Monte Carlo and consider analysis results is unlikely unless the consider analysis algorithms and the consider model form both have been implemented correctly.

Figure 1 plots representative results from the Monte Carlo and consider analyses for the incorrect initialization example. It displays three versions of the smoothed root mean square (RMS) position error, which is the square root of the diagonal element of the matrix MSE corresponding to position. The dashed line is the
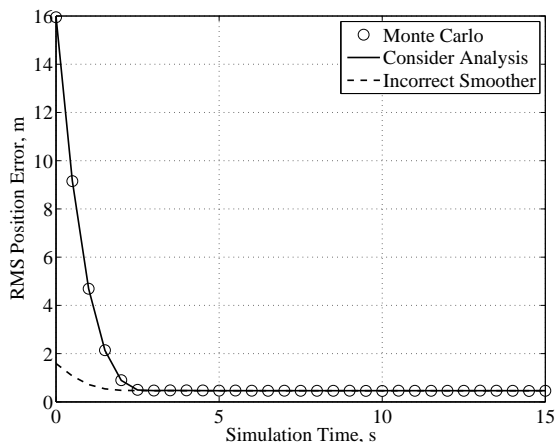

Figure 1. RMS smoother position errors for incorrect initialization example.

smoothed RMS error reported by the mismodeled filter/smoother, and the solid line is the true RMS error for the smoother as computed by the consider analysis. The circles are the smoothed RMS errors computed

American Institute of Aeronautics and Astronautics

from the Monte Carlo simulations. Note that only the first 15 s of the 50 s simulation time are shown in Fig. 1. Where the consider and nominal smoother errors diverge, the Monte Carlo results follow the consider analysis closely. The biased initial estimate and incorrect initial covariance only affect estimation uncertainty for a relatively short time. After this initial transient behavior, the true filter/smoother accuracy returns to the same level that it would have had with no model errors. The other filtering and smoothing results for this example are qualitatively similar to those of Fig. 1.

In addition to true matrix MSE, the distributions of estimation error from the Monte Carlo simulations have been used to compute 99% confidence bounds on the results. Such confidence bounds can indicate whether a sufficiently large number of Monte Carlo trials has been used, so as to achieve a desired level of accuracy. For the present example, the Monte Carlo and consider analysis results generally agree to well within the confidence bounds. The confidence bounds are not shown in Fig. 1 in order to reduce clutter.

C.   Example: Correlated Noise

The second numerical example is that of correlated process and measurement noise. In this example, the "truth" system's dynamics and measurements are still described by the models of Eqs. (32a) and (32b). Also, the filter/smoother's initial estimate and initial state error covariance are accurate, and the initial estimation error is uncorrelated with process or measurement noise. The process and measurement noise, however, both have much larger covariances than the assumed values, and they are correlated. Specifically,

$$\mathrm{E}\left\{\begin{bmatrix} w_k \\ \nu_k \end{bmatrix}\begin{bmatrix} w_k & \nu_k \end{bmatrix}\right\} = \begin{bmatrix} Q_k & P_{w\nu k} \\ P_{w\nu k} & R_k \end{bmatrix} = \begin{bmatrix} 10 & -3 \\ -3 & 8 \end{bmatrix} \tag{47}$$

This case can be handled by defining the composite noise vector $\boldsymbol{\mu}_k = \begin{bmatrix} w_k & \nu_k \end{bmatrix}^{\mathrm{T}}$. With this definition, the composite noise covariance matrix $P_{\mu\mu k}$ is just the rightmost matrix of Eq. (47). Its inverse square root, $\mathcal{R}_{\mu\mu k}$, is found by Cholesky factorization of $P_{\mu\mu k}^{-1}$ to be approximately:

$$\mathcal{R}_{\mu\mu k} = \begin{bmatrix} 0.3357 & 0.1259 \\ 0 & 0.3536 \end{bmatrix} \tag{48}$$

While one could compute the individual square-root information matrices for process and measurement noise, $\mathcal{R}_{wwk}$ and $\mathcal{R}_{\nu\nu k}$, they are less meaningful in this example than the joint square-root information matrix $\mathcal{R}_{\mu\mu k}$.

To describe this system in consider form, the consider state vector $\boldsymbol{x}_{\boldsymbol{c}k}$ must contain the "true", correlated process and measurement noise. It must also capture the initial estimation error at sample $k = 0$. A suitable definition is:

$$\boldsymbol{x}_{\boldsymbol{c}k} \equiv \begin{cases} \begin{bmatrix} \bar{\boldsymbol{\nu}}_{x0} \\ \mathcal{R}_{\mu\mu 0}\boldsymbol{\mu}_0 \end{bmatrix} & k = 0 \\[2ex] \begin{bmatrix} \boldsymbol{\mu}_k \end{bmatrix} & k > 0 \end{cases} \tag{49}$$

where $\bar{\boldsymbol{\nu}}_{x0}$ is the initial information state error vector. It appears in the filter/smoother version of the initial information equation in Table 1, and in this scenario is unchanged in the "truth" system. The vector $\boldsymbol{\mu}_k$ is the previously defined composite noise vector containing w_k and ν_k. At sample $k = 0$, premultiplication of $\boldsymbol{\mu}_0$ by $\mathcal{R}_{\mu\mu 0}$ causes the consider state vector $\boldsymbol{x}_{\boldsymbol{c}0}$ to have identity covariance, as required by the consider model form.

As in the previous example, the "truth" system matrices Φ_k, Γ_k, and H_k are the same as those used by the filter/smoother. Also, there is no dynamics disturbance so Γ_{xck} is an appropriately-dimensioned matrix of zeros. There are no non-zero deterministic biases for this example, so $\boldsymbol{b}_{xk}$, $\boldsymbol{b}_{yk}$, $\boldsymbol{b}_{wk}$, and $\bar{\boldsymbol{b}}_{c0}$ are all zero-valued.

Based on the definition of the consider state in Eq. (49), the consider state dynamics equation can be constructed. The noise contained in $\boldsymbol{\mu}_k$ is white, and the initial information error $\bar{\boldsymbol{\nu}}_{x0}$ is only present at sample $k = 0$, so Φ_{ck} is a matrix of zeros for all samples k. At $k = 0$, Φ_{c0} is rectangular with more columns than rows in order to transition to the new version of the consider state vector with fewer components. The matrix Γ_{cck} models the correlated statistics of w_k and ν_k. Thus for all samples k it takes the form:

$$\Gamma_{cck} = \mathcal{R}_{\mu\mu k+1}^{-1} = \begin{bmatrix} 2.9791 & -1.0607 \\ 0 & 2.8284 \end{bmatrix} \tag{50}$$

The matrix S_{wck} models how the "true", correlated process noise enters the filter/smoother. It extracts the first component of the composite noise vector $\boldsymbol{\mu}_k$ from $\boldsymbol{x_{ck}}$ and properly weights the result. In terms of the defined quantities, it is:

$$S_{wck} = \begin{cases} -\mathcal{R}_{fww0}\begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} \mathbf{0} & \mathcal{R}_{\mu\mu0}^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & -2.9791 & 1.0607 \end{bmatrix} & k = 0 \\ -\mathcal{R}_{fwwk}\begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \end{bmatrix} & k > 0 \end{cases}$$ (51)

In the same manner, H_{ck} selects the "true", correlated measurement noise by extracting the second component of $\boldsymbol{\mu}_k$ from $\boldsymbol{x_c}k$. The matrix that accomplishes this is:

$$H_{ck} = \begin{cases} \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \mathbf{0} & \mathcal{R}_{\mu\mu0}^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 2.8284 \end{bmatrix} & k = 0 \\ \begin{bmatrix} 0 & 1 \end{bmatrix} & k > 0 \end{cases}$$ (52)

Finally, the initial estimation error covariance assumed by the filter/smoother is correct. This is modeled by $\bar{S}_{xc0}$, which selects the unweighted $\bar{\boldsymbol{\nu}}_{x0}$ that forms the first component of $\boldsymbol{x_{c0}}$:

$$\bar{S}_{xc0} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$ (53)

After applying the consider filter/smoother algorithms to this consider-form system, Monte Carlo simulations have been used to demonstrate the correctness of the calculated true smoother covariances. Figures 2 and 3 display the filtered (*a posteriori*) and smoothed velocity error standard deviations, respectively. As in
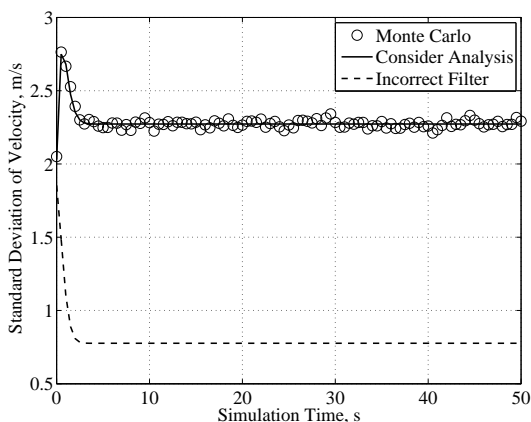

Figure 2. Standard deviations of filtered *a posteriori* velocity error for correlated noise example.
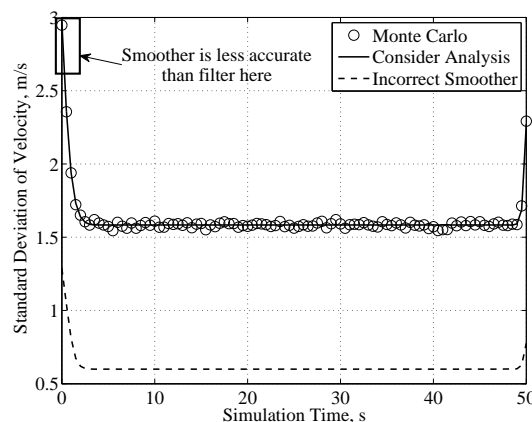

Figure 3. Standard deviations of smoothed velocity error for correlated noise example.

Fig. 1, the solid lines are the analytical results from the consider analysis and the circles are the numerical results from the Monte Carlo simulations. The dashed lines show the incorrect standard deviations reported by the original filter/smoother.

This example's results demonstrate close agreement between the analytical and numerical techniques. They further illustrate an important and counterintuitive feature of the consider analysis of a smoother. A traditional linear filter/smoother has the property that its smoothed covariance is always less than its filtered *a posteriori* covariance in a positive definite sense: $P_{fxxk}^* \leq P_{fxxk}$. One can understand this property in terms of the information used to form each estimate; information from future measurements always reduces uncertainty rather than increasing it. This principle no longer holds for the consider analysis covariances: $P_{xxk}^* \lesseqqgtr P_{xxk}$. The smoothed velocity standard deviations of Fig. 3 are higher than the *a posteriori* standard deviations of Fig. 2 near the beginning of the simulation. Although not shown, the same situation occurs for this example's position standard deviations. Not only does the smoother incorporate information from future measurements into its estimates, but it also incorporates model errors that apply at future samples. For this example, the additional model error effects added to the early estimates are greater than the uncertainty eliminated by additional measurements. In other words, a smoother with the wrong kind of modeling error may degrade rather than enhance estimation quality.

American Institute of Aeronautics and Astronautics

D.   Example: Unestimated Disturbances

As with the first two examples, the filter/smoother for the third example is the one described in Section A. In this example, the deterministic biases $\boldsymbol{b}_{xk}$, $\boldsymbol{b}_{yk}$, $\boldsymbol{b}_{wk}$, and $\bar{\boldsymbol{b}}_{c0}$ are all zero. There are also no errors associated with initialization or noise statistics. That is, $\overline{P}_{xx0} = \overline{P}_{fxx0}$, $Q_k = Q_{fk}$, and $R_k = R_{fk}$. The noise elements and initial estimation error are uncorrelated, and both the process noise and the measurement noise are white and Gaussian.

Although the "truth" system matrices are identical to those assumed by the filter/smoother, the "truth" dynamics and measurement models are affected by additional unmodeled disturbances. The "truth" dynamics equation is disturbed by unmodeled, colored process noise that has a sinusoidally time-varying influence on the velocity state. The measurements contain a constant random bias. Neither of these effects is estimated by the filter/smoother. The "truth" dynamics and measurement models are given by:

$$\begin{bmatrix} r_{k+1} \\ v_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_k + \begin{bmatrix} 0 \\ \sin\left(\frac{2\pi k \Delta t}{T_{per}}\right) \end{bmatrix} a_k \tag{54a}$$

$$y_k = \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} r_k \\ v_k \end{bmatrix} + \nu_k + b_k \tag{54b}$$

where the period of the sinusoidal disturbance influence is $T_{per} = 11.15$ s and where the colored disturbance a_k is a first-order Markov process. The bias $b_k = b_0$ is assumed to be drawn from a zero-mean random distribution with standard deviation $\sigma_b = 2/3$, and the initial noise state a_0 is drawn from a zero-mean random distribution with standard deviation $\sigma_a = 4$. The dynamics of the joint disturbance vector $\begin{bmatrix} a_k & b_k \end{bmatrix}^{\mathrm{T}}$ are

$$\begin{bmatrix} a_{k+1} \\ b_{k+1} \end{bmatrix} = \begin{bmatrix} e^{-\frac{\Delta t}{\tau_a}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_k \\ b_k \end{bmatrix} + \begin{bmatrix} \gamma_a \\ 0 \end{bmatrix} w_{ak} \tag{55}$$

with Markov time constant $\tau_a = 16.725$ s and parameter $\gamma_a = \sigma_a \sqrt{1 - e^{-2\Delta t/\tau_a}}$ (approximately $\gamma_a = 0.9636$). The process noise w_{ak} that drives the Markov process has zero mean and unit variance. Note, the steady-state standard deviation of a_k equals the initial standard deviation σ_a by construction.

To write this system in consider form, the consider state vector must include not only the correctly-modeled initial uncertainty, process noise, and measurement noise, but also the unestimated dynamically varying a_k and the random bias b_k. A suitable definition is

$$\boldsymbol{x}_{ck} \equiv \begin{cases} \begin{bmatrix} \bar{\boldsymbol{\nu}}_{x0} \\ a_0/\sigma_a \\ b_0/\sigma_b \\ \mathcal{R}_{ww0} w_0 \\ \mathcal{R}_{\nu\nu 0} \nu_0 \end{bmatrix} & k = 0 \\ \begin{bmatrix} a_k \\ b_k \\ w_k \\ \nu_k \end{bmatrix} & k > 0 \end{cases} \tag{56}$$

As before, the consider vector at sample $k = 0$ contains an additional component for initial estimate uncertainty, and its elements are defined such that it has identity covariance. With this definition of $\boldsymbol{x}_{ck}$, one can write the matrices for the consider dynamics equation for $k > 0$:

$$\Phi_{ck} = \begin{bmatrix} e^{-\frac{\Delta t}{\tau_a}} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0.9705 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{57a}$$

$$\Gamma_{cck} = \begin{bmatrix} \gamma_a & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \mathcal{R}_{wwk+1}^{-1} & 0 \\ 0 & 0 & \mathcal{R}_{\nu\nu k+1}^{-1} \end{bmatrix} = \begin{bmatrix} 0.9636 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{57b}$$

Note how the consider dynamics matrices capture both the dynamic behavior of a_k and b_k and the statistical behavior of w_k and ν_k. At $k = 0$, the Γ_{cck} matrix is unchanged, but Φ_{ck} becomes

$$\Phi_{c0} = \begin{bmatrix} \mathbf{0} & \sigma_a e^{-\frac{\Delta t}{\tau_a}} & 0 & 0 & 0 \\ \mathbf{0} & 0 & \sigma_b & 0 & 0 \\ \mathbf{0} & 0 & 0 & 0 & 0 \\ \mathbf{0} & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 3.8822 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{58}$$

in order to transition from the original consider vector at $k = 0$ to its new form for $k > 0$.

In contrast to the previous two examples, Γ_{xck} is a non-zero matrix. It specifies how the unmodeled disturbance affects the filter's state. It is given by:

$$\Gamma_{xck} = \begin{cases} \begin{bmatrix} \mathbf{0} & 0 & 0 & 0 & 0 \\ \mathbf{0} & \sigma_a \sin\left(\frac{2\pi k \Delta t}{T_{per}}\right) & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} & k = 0 \\[1em] \begin{bmatrix} 0 & 0 & 0 & 0 \\ \sin\left(\frac{2\pi k \Delta t}{T_{per}}\right) & 0 & 0 & 0 \end{bmatrix} & k > 0 \end{cases} \tag{59}$$

In the measurement equation, H_{ck} must select the component of $\boldsymbol{x_{ck}}$ corresponding to measurement noise ν_k. For this example, it must additionally select the component corresponding to the random measurement bias b_k. Note that this analysis considers the effects of a zero-mean random measurement bias rather than a non-zero deterministic one. In other words, it predicts the true covariance when a bias with a given statistical distribution is present, rather than the matrix MSE resulting from a specific deterministic bias. The matrix H_{ck} that extracts the measurement noise and bias is given by:

$$H_{ck} = \begin{cases} \begin{bmatrix} \mathbf{0} & 0 & \sigma_b & 0 & \mathcal{R}_{\nu\nu0}^{-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{2}{3} & 0 & 1 \end{bmatrix} & k = 0 \\[1em] \begin{bmatrix} 0 & 1 & 0 & 1 \end{bmatrix} & k > 0 \end{cases} \tag{60}$$

The matrix S_{wck} selects the components of $\boldsymbol{x_{ck}}$ related to process noise and indicates that "truth" and modeled process noise are identical for this scenario:

$$S_{wck} = \begin{cases} \begin{bmatrix} \mathbf{0} & 0 & 0 & -\mathcal{R}_{fww0}\mathcal{R}_{ww0}^{-1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 & 0 \end{bmatrix} & k = 0 \\[1em] \begin{bmatrix} 0 & 0 & -\mathcal{R}_{fwwk} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 0 \end{bmatrix} & k > 0 \end{cases} \tag{61}$$

Finally, the matrix $\bar{S}_{xc0}$ specifies that the error term in the initial state information equation is statistically correct.

$$\bar{S}_{xc0} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{62}$$

At this point, the system has been written in consider form, and the consider algorithms and Monte Carlo analysis can proceed in standard fashion.

Figures 4 and 5 show the filtered and smoothed velocity estimation error standard deviations. They are labeled in the same manner as the corresponding results in Figs. 2 and 3. In Fig. 4, it is clear that the unestimated sinusoidal disturbance in the velocity state corresponds to an unpredicted, roughly sinusoidal variation of the velocity error standard deviation. The oscillation frequency in this plot is approximately twice that of the disturbance, however. Both peaks and valleys of the sinusoidal disturbance result in high uncertainty, and the uncertainty drops to near its nominal value when the disturbance passes through zero.

Similar to the preceding correlated noise example, the smoothed standard deviations of Fig. 5 are mostly lower than the filtered *a posteriori* standard deviations of Fig. 4, but not at the beginning of the interval. It is significant that the sinusoidal variation in standard deviation is greatly reduced by the smoother. For this example, the consider analyses show that the filter is very sensitive to the modeling error while the smoother is not: Note in Fig. 5 how the consider analysis standard deviation is nearly equal to that of the incorrect smoother except at the beginning of the interval. This reduced smoother sensitivity stands in sharp contrast with the filter results of Fig. 4.
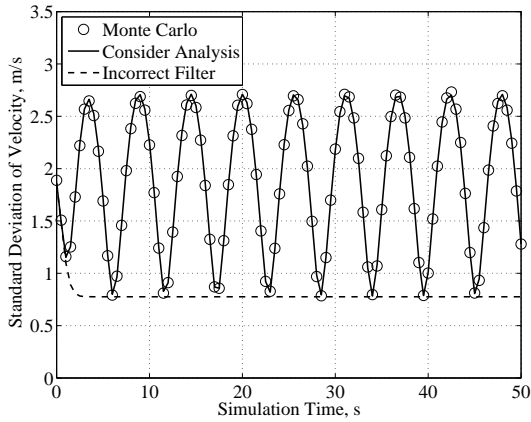
American Institute of Aeronautics and Astronautics

Figure 4. Standard deviations of filtered *a posteriori* velocity error for unestimated disturbances example.
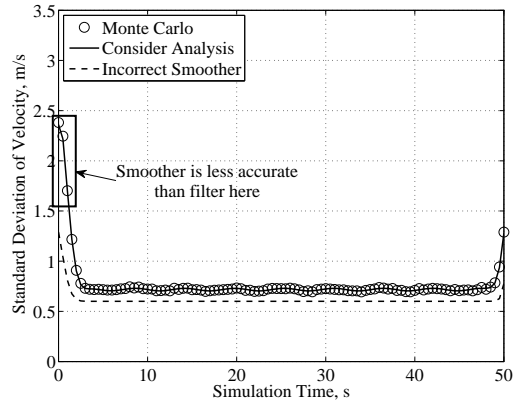

Figure 5. Standard deviations of smoothed velocity error for unestimated disturbances example.

V. Conclusions

This paper has developed a new form of consider covariance analysis that can be applied to Rauch-Tung-Striebel discrete square-root information smoothers. Smoother analysis is accomplished by performing a backward analysis pass after having performed a forward-pass consider filter analysis. The analysis computes the smoother's true estimation error covariance or matrix mean square error, whichever is most relevant. It generalizes to a wide variety of possible model errors. This generalization capability is enabled by the definition of a new standard system form. Once the system modeling errors have been written in this standard form, the analysis algorithms can be applied in a straightforward manner. A special feature of the consider smoother analysis is its ability to "consider" the effects of both past and future model errors on the smoothed estimates.

Several concrete examples illustrate the power of the new method while clarifying its implementation. They include a system model with a biased initial estimate and an incorrect initial covariance, a model with cross-correlated process and measurement noise, and a model with dynamically-varying unestimated disturbance states. The ability to model three distinct varieties of errors in the same framework is thus demonstrated. Monte Carlo simulations provide independent verification of the smoother consider analysis equations. Comparisons of the filter and smoother results for two of the examples show that the addition of a smoothing pass sometimes mitigates the effects of filter/smoother modeling errors, but can exacerbate model error effects in other instances.

References

[1]Tapley, B. D., Schutz, B. E., and Born, G. H., *Statistical Orbit Determination*, Academic Press, 2004.

[2]Montenbruck, O. and Gill, E., *Satellite Orbits: Models, Methods, and Applications*, Springer-Verlag, 2000.

[3]Bierman, G. J., "The Treatment of Bias in the Square-Root Information Filter/Smoother," *Journal of Optimization Theory and Applications*, Vol. 16, No. 1, 1975, pp. 165–178.

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

[5]Curkendall, D. W., *Problems in Estimation Theory with Applications to Orbit Determination*, Ph.D. thesis, UCLA School of Engineering and Applied Science, Sept. 1972.

[6]Nishimura, T., "On the a priori Information in Sequential Estimation Problems," *IEEE Transactions on Automatic Control*, Vol. 11, No. 2, 1966, pp. 197–204.

[7]Heffes, H., "The Effect of Erroneous Models on the Kalman Filter Response," *IEEE Transactions on Automatic Control*, Vol. 11, No. 3, July 1966, pp. 541–543.

[8]Huddle, J. and Wismer, D., "Degradation of Linear Filter Performance Due to Modeling Error," *IEEE Transactions on Automatic Control*, Vol. 13, No. 4, 1968, pp. 421–423.

[9]Gelb, A., editor, *Applied Optimal Estimation*, The Analytical Sciences Corporation, The M.I.T. Press, 1974.

[10]Lewis, F. L., *Optimal Estimation*, John Wiley & Sons, 1986.

[11]Bar-Shalom, Y., Li, X. R., and Kirubarajan, T., *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, Inc., 2001.

American Institute of Aeronautics and Astronautics

[12]Fagin, S. L., "Recursive Linear Regression Theory, Optimal Filter Theory and Error Analysis of Optimal Systems," *IEEE international convention record*, Vol. 12, 1964, pp. 216–245.

[13]Griffin, R. E. and Sage, A. P., "Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms," *IEEE Transactions on Automatic Control*, Vol. 13, No. 4, Aug. 1968, pp. 320–329.

[14]Griffin, R. E. and Sage, A. P., "Sensitivity Analysis of Discrete Filtering and Smoothing Algorithms," *AIAA Journal*, Vol. 7, No. 10, Oct. 1969, pp. 1890–1897.

[15]Sage, A. P. and Melsa, J. L., *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, 1971.

[16]Morf, M., Verriest, E., Dobbins, J., and Kailath, T., "Square-Root Algorithms for Model Sensitivity Analysis," *Proc. of the 1977 Conf. on Information Sciences and Systems*, Vol. 11, Mar. 30-Apr. 1, 1977, pp. 64–69.

[17]Hinks, J. C. and Psiaki, M. L., "A Multipurpose Consider Covariance Analysis for Square-Root Information Filters," *AIAA Guidance, Navigation, and Control Conference*, 13-16 Aug. 2012.

[18]Rauch, H. E., Tung, F., and Striebel, C. T., "Maximum Likelihood Estimates of Linear Dynamic Systems," *AIAA Journal*, Vol. 3, No. 8, Aug. 1965, pp. 1445–1450.

[19]Wall, J. E., Willsky, A. S., and Sandell, N. R., "On the Fixed-Interval Smoothing Problem," *Stochastics*, Vol. 5, No. 1-2, 1981, pp. 1–41.

[20]Bierman, G. J., "Sequential Square Root Filtering and Smoothing of Discrete Linear Systems," *Automatica*, Vol. 10, 1974, pp. 147–158.