Linear Covariance Analysis and Epoch State Estimators

F. Landis Markley* **and J. Russell Carpenter**†

ABSTRACT

This paper extends in two directions the results of prior work on generalized linear covariance analysis of both batch least-squares and sequential estimators. The first is an improved treatment of process noise in the batch, or epoch state, estimator with an epoch time that may be later than some or all of the measurements in the batch. The second is to account for process noise in specifying the gains in the epoch state estimator. We establish the conditions under which the latter estimator is equivalent to the Kalman filter.

INTRODUCTION

Reference 1 gave an overview of a comprehensive approach to linear covariance analysis, closely following the approach of Refs. 2 and 3, and including references to earlier studies of this subject. The models presented include both batch least-squares methods, which originated with Gauss,[4] and sequential estimators, which were pioneered by Kalman.[5] Both batch least-squares estimation and Kalman filtering methods produce the covariance matrix encapsulating both the expected errors of the estimated parameters and the correlations among them. The number of parameters describing a dynamical system can be quite large, and it is often impractical to solve for all of them, especially in real-time applications. "Consider covariance analysis," the subject of Refs. 1–3, treats estimators that solve for only a subset of the parameters, but considers the effect on the covariance of the parameters that are not solved for. Our approach differs from the Schmidt Kalman filter,[6] which takes the consider parameters into account in computing the update gains, but updates neither their values nor their covariances. Our consider parameters, in contrast, are completely ignored by the estimation algorithm. Their effect on the accuracy of the estimates obtained is considered in the covariance analysis, though, accounting for their name. This treatment is very similar to that described in Bierman[7] and in Tapley, et al.[8] The method described here follows Refs. 1 – 3 and differs from many other treatments in explicitly partitioning the errors into subspaces containing only the influence

*Attitude Control Systems Engineering Branch, NASA Goddard Space Flight Center, Code 591, Greenbelt, MD 20771
†Navigation and Mission Design Branch, NASA Goddard Space Flight Center, Code 595, Greenbelt, MD 20771.

of the measurement noise, process noise, and *a priori* solve-for and consider covariances. McReynolds[9] has also implemented such a partitioning in both batch and sequential filter/smoother error budget analysis. Explicitly exhibiting the *a priori* component of the covariance is useful for observability analysis because *a priori* information about states that are observable is "forgotten" as more data are processed.[10]

The treatment of the batch estimation in Refs. 1–3 assumed that the estimator did not account for process noise in computing the update gains, as is customary in batch least-squares estimation. The estimators considered in this paper, which we call epoch-state estimators to distinguish them from the usual batch estimators, take process noise into consideration. We show that the usual batch estimator is recovered when process noise is ignored by the estimator. Process noise may or may not be included in the estimation process, but it is always accounted for in the linear covariance analysis, as in Refs. 1 – 3, because it can contribute to estimation errors whether the estimator knows about it or not.

We define the epoch time to be the time at which the *a priori* information is specified and at which the state update is performed. The analysis of the batch estimator in Refs. 2 and 3 implicitly assumed that the epoch time is prior to all of the measurements, which is the usual case, and which the term *a priori* implies. In this case the *a priori* errors at the update time are uncorrelated with the process noise during the measurement span. It is sometimes convenient to allow the epoch time to be later than some or all of the measurements, though, and Ref. 1 considered this case. It continued to assume zero correlations between the *a priori* errors and the process noise, however. It may be more reasonable to assume that the *a priori* errors include the process noise accumulated between the beginning of the observation span and the epoch time, if the latter is later than some or all of the measurements. This modifies the manner in which process noise appears in the covariance analysis of the epoch-state estimator, and is the principal subject of this paper.

To make this paper self-contained and to establish notation, the first several sections of this paper contain a review of the material in Ref. 1. We have, however, attempted to eliminate unnecessary duplication with the earlier paper. We first consider the linear state dynamics models, and the partitioning of state errors into those arising from *a priori* errors, from measurement errors, and from process noise. The dynamics does not distinguish between solve-for and consider parameters, so this distinction is not made until we turn to the processing of measurement information. The measurement update equations involve the projection matrix onto the solve-for subspace, because the update does not affect the consider parameters. We then compute the contributions from *a priori* errors, measurement errors, and process noise to the covariance matrix of the full state. The measurement errors are uncorrelated with both the *a priori* errors and the process noise,

so the contribution of the measurement errors to the covariance can be defined unambiguously. However, we now account for correlations between the *a priori* errors and the process noise, and we include these correlations in the process noise part of the covariance.

We next discuss the selection of the update gains to minimize the covariance of the estimation errors at the epoch time, taking into account a model of the process noise in the gain specification. We show that our results recover the usual batch least-squares filter when the process noise is zero. If process noise is not neglected, the epoch-state estimators fall into three classes: 1) when no measurements follow the epoch time, the estimator can be operated sequentially, but the covariance analysis of Refs. 1–3 must be modified; 2) when no measurements precede the epoch time, the covariance analysis of Refs. 1–3 is valid, but the estimator cannot be operated sequentially; 3) when some measurements precede and some follow the epoch time, sequential operation of the estimator is precluded, and the covariance analysis of Refs. 1–3 must be modified.

The sequential formulation of this epoch-state estimator has been applied to an analysis of satellite close approaches.[11] For this application, two filters are run concurrently, one constrained by the condition that the distance of closest approach is greater than some effective hard body radius, and the other constrained so that the closest approach distance is less than the hard body radius. The constraint is implemented using the technique of Zanetti, et al.[12] An epoch state filter with epoch time in the future is required because the time of closest approach, where the constrained update is enforced, is in the future of all the measurements.

We close the paper by showing that the sequential epoch-state estimator is equivalent to the Kalman filter.[5] and by presenting a summary and our conclusions.

MODELING APPROACH

General Dynamics Model

The state vector $\mathbf{x}$ is an n-dimensional vector of parameters that completely characterizes a system, including both time-dependent and constant parameters. By assumption, the state vector evolves according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{w}(t) \tag{1}$$

where the dynamic noise $\mathbf{w}(t)$ is a Gaussian white noise process with mean and covariance given by

$$\mathrm{E}\left[\mathbf{w}(t)\right] = \mathbf{0}_n \quad \text{and} \quad \mathrm{E}\left[\mathbf{w}(t)\mathbf{w}(\tau)^{\mathbf{T}}\right] = Q(t)\delta(t-\tau), \tag{2}$$

with $\mathrm{E}[...]$ denoting the expectation operator and $\mathbf{0}_n$ denoting an n-dimensional vector of zeros. In the covariance equation, Q is the $n \times n$ dynamic noise spectral density matrix and $\delta(t-\tau)$ denotes the Dirac delta, or unit impulse, function.

The true value of the state vector is never exactly known, but can only be estimated. The state estimate vector, $\hat{\mathbf{x}}$, evolves according to

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{x}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), t). \tag{3}$$

The state error vector, given by

$$\mathbf{e}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t), \tag{4}$$

is assumed to always remain small. Then first-order linear error analysis techniques, as commonly found in an extended Kalman filter, give

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}(t) = A(t)\mathbf{e}(t) + \mathbf{w}(t), \tag{5}$$

where

$$A(t) = \left.\frac{\partial \mathbf{f}(\mathbf{x}(t), t)}{\partial \mathbf{x}(t)}\right|_{\hat{\mathbf{x}}(t)}. \tag{6}$$

Formal integration of equation (5) gives

$$\mathbf{e}(t) = \Phi(t, t')\mathbf{e}(t') + \mathbf{w}_d(t, t') \tag{7}$$

where $\mathbf{e}(t')$ the error vector at time t', the state transition matrix, $\Phi(t, t')$, is the solution of

$$\frac{\mathrm{d}}{\mathrm{d}t}\Phi(t, t') = A(t)\Phi(t, t') \tag{8}$$

with the initial condition

$$\Phi(t', t') = I_n = \text{the } n \times n \text{ identity matrix}, \tag{9}$$

and the random excitation vector, $\mathbf{w}_d(t, t')$, is given by the formal integral

$$\mathbf{w}_d(t, t') = \int_{t'}^{t} \Phi(t, \tau)\mathbf{w}(\tau)\mathrm{d}\tau. \tag{10}$$

We will never need to evaluate this mathematically ill-defined integral, but we use it to formally compute

covariances, which are the well-defined quantities of interest.

The state vector error at time t can be written as the sum of three contributions: the error $\mathbf{e}_a(t)$, due to an *a priori* error at an epoch time t_*, the error $\mathbf{e}_v(t)$ due to noise in the measurements used for the estimation, and the error $\mathbf{e}_w(t)$ due to dynamic noise.

$$\mathbf{e}(t) = \mathbf{e}_a(t) + \mathbf{e}_v(t) + \mathbf{e}_w(t). \tag{11}$$

Equation (7) can be used to propagate the components of the state error vector from the epoch time t_* to some other time t:

$$\mathbf{e}_a(t) = \Phi(t, t_*)\mathbf{e}_a(t_*), \tag{12a}$$

$$\mathbf{e}_v(t) = \Phi(t, t_*)\mathbf{e}_v(t_*), \tag{12b}$$

$$\mathbf{e}_w(t) = \Phi(t, t_*)\mathbf{e}_w(t_*) + \mathbf{w_d}(t, t_*). \tag{12c}$$

General Estimation Model

It is usually not necessary to produce estimates of all of the state parameters, but only of a subset of n_s solve-for parameters. The remaining $n_c = n - n_s$ parameters are called consider parameters because they are not estimated but contain uncertainties that must be considered in the error analysis. The solve-for parameter vector $\mathbf{s}(t)$ and the consider parameter vector $\mathbf{c}(t)$ are given by

$$\mathbf{s}(t) = S(t)\mathbf{x}(t) \quad \text{and} \quad \mathbf{c}(t) = C(t)\mathbf{x}(t) \tag{13}$$

where the $n_s \times n$ matrix $S(t)$ and the $n_c \times n$ matrix $C(t)$ are such that the matrix

$$M = \begin{bmatrix} S \\ C \end{bmatrix} \tag{14}$$

is non-singular*. The inverse of M is partitioned into an $n \times n_s$ matrix $\tilde{S}$ and an $n \times n_c$ matrix $\tilde{C}$:

$$M^{-1} = \left[\tilde{S},\, \tilde{C}\right]. \tag{15}$$

*Here and subsequently we suppress the time arguments where there is no ambiguity so as to simplify the notation.

The properties of the matrix inverse then lead immediately to the identities

$$S\tilde{S} = I_{n_s}, \quad C\tilde{C} = I_{n_c}, \quad S\tilde{C} = 0_{n_s \times n_c}, \quad C\tilde{S} = 0_{n_c \times n_s}, \tag{16}$$

and

$$\tilde{S}S + \tilde{C}C = I_n. \tag{17}$$

In the usual case that the elements of the solve-for and consider vectors are merely selected and possibly permuted components of the state vector, the matrix M is an orthogonal permutation matrix. In this case, and in any case for which M is orthogonal, the matrices $\tilde{S}$ and $\tilde{C}$ are just the transposes of S and C, respectively, which makes inversion of M unnecessary and simplifies many of the following equations.

It follows from equations (13) and (17) that

$$\mathbf{x}(t) = \tilde{S}(t)\mathbf{s}(t) + \tilde{C}(t)\mathbf{c}(t). \tag{18}$$

We also have an estimated value $\hat{\mathbf{s}}(t)$ of the solve-for vector and an assumed value $\hat{\mathbf{c}}(t)$ of consider vector. Both are propagated by the dynamics, but only the solve-for vector is updated by the estimator. Relations similar to equation (13) give the estimated solve-for vector and the assumed consider vector the in terms of the estimate of the full state $\hat{\mathbf{x}}(t)$. Thus, the error vector can also be expressed in terms of its solve-for and consider components as

$$\mathbf{e}_s(t) = \mathbf{s}(t) - \hat{\mathbf{s}}(t) = S(t)\mathbf{e}(t) \tag{19a}$$

$$\mathbf{e}_c(t) = \mathbf{c}(t) - \hat{\mathbf{c}}(t) = C(t)\mathbf{e}(t) \tag{19b}$$

and

$$\mathbf{e}(t) = \tilde{S}(t)\mathbf{e}_s(t) + \tilde{C}(t)\mathbf{e}_c(t). \tag{20}$$

The subscripts s and c are independent of the subscripts a, v, and w in the previous section. Both the solve-for errors and the consider errors have contributions arising from the *a priori* estimate, the measurement noise, and the process noise. We will avoid double subscripts, however.

An estimator produces its state estimates based on an *a priori* estimate at an epoch time t_* and information obtained from measurements obtained at discrete times t_i. Let $\mathbf{y}_i$ be an m_i-dimensional vector of

measurement values obtained at times t_i. Measurements are related to the state vector by the following measurement model:

$$\mathbf{y}_i = \mathbf{h}_i(\mathbf{x}_i, t_i) + \mathbf{v}_i, \tag{21}$$

where $\mathbf{v}_i$ is a Gaussian white noise process with mean and covariance given by

$$\mathrm{E}\left[\mathbf{v}_i\right] = \mathbf{0}, \quad \mathrm{E}\left[\mathbf{v}_i\mathbf{v}_i^{\mathbf{T}}\right] = R_i, \quad \text{and} \quad \mathrm{E}\left[\mathbf{v}_i\mathbf{v}_j^{\mathbf{T}}\right] = 0, \quad \text{for } i \neq j. \tag{22}$$

The functions $\mathbf{h}_i$ are assumed to be known functions of imprecisely known arguments. Therefore, it is possible to compute predicted measurements from

$$\hat{\mathbf{y}}_i^- = \mathbf{h}_i(\hat{\mathbf{x}}_i^-, t_i) \tag{23}$$

where $\hat{\mathbf{x}}_i^-$ is the state estimate before incorporating the measurement information. Conversely, $\hat{\mathbf{x}}_i^+$ denotes the state estimate at time t_i including the measurement information.[†] A Taylor series expansion gives the difference between the observed and computed measurements, or "o-minus-c," to first order in $\mathbf{e}$, as

$$\mathbf{r}_i^- = \mathbf{y}_i - \hat{\mathbf{y}}_i^- = H_i\mathbf{e}_i^- + \mathbf{v}_i, \tag{24}$$

where the measurement sensitivity matrix H_i is given by

$$H_i = \left.\frac{\partial \mathbf{h}_i(\mathbf{x}_i, t_i)}{\partial \mathbf{x}_i}\right|_{\hat{\mathbf{x}}_i^-}. \tag{25}$$

This linearization is familiar from the extended Kalman filter. The o-minus-c before the update, $\mathbf{r}_i^-$, is commonly referred to as the "innovation," and the o-minus-c after the update as the "residual," though this terminology is by no means universal.

[†]We use this notation in the sequel to denote *a priori* and *a posteriori* values of various quantities.

Epoch State Estimator

An epoch-state estimator produces an estimate by performing a single update at the epoch time t_* based on a single batch measurement $\mathbf{y}$ that combines the N measurements $\mathbf{y}_i$ made at various times. Thus,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix}, \quad \hat{\mathbf{y}} = \begin{bmatrix} \hat{\mathbf{y}}_1 \\ \vdots \\ \hat{\mathbf{y}}_N \end{bmatrix}, \quad \mathbf{r} = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} \mathbf{r}_1 \\ \vdots \\ \mathbf{r}_N \end{bmatrix} \tag{26}$$

The update is given by:

$$\hat{\mathbf{s}}_*^+ = \hat{\mathbf{s}}_*^- + K\mathbf{r} = \hat{\mathbf{s}}_*^- + \sum_{i=1}^{N} K_i \mathbf{r}_i, \tag{27}$$

where $\hat{\mathbf{s}}_*^-$ is the *a priori* estimate of $\mathbf{s}(t_*)$ and K is a gain matrix consisting of a "row" of gain matrices for each measurement, K_i. Substituting equations (24) and (7), with $t = t_i$ and $t' = t_*$, into equation (26) gives

$$\mathbf{r} = \tilde{H}\mathbf{e}_*^- + \mathbf{u}_d + \mathbf{v} \tag{28}$$

where

$$\tilde{H} = \begin{bmatrix} H_{1,*} \\ \vdots \\ H_{N,*} \end{bmatrix} = \begin{bmatrix} H_1 \Phi(t_1, t_*) \\ \vdots \\ H_N \Phi(t_N, t_*) \end{bmatrix}, \quad \mathbf{u}_d = \begin{bmatrix} \mathbf{u}_{d1} \\ \vdots \\ \mathbf{u}_{dN} \end{bmatrix} = \begin{bmatrix} H_1 \mathbf{w}_d(t_1, t_*) \\ \vdots \\ H_N \mathbf{w}_d(t_N, t_*) \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_N \end{bmatrix}. \tag{29}$$

The state estimates must be propagated from the epoch time t_* to the measurement times t_i in order to compute the measurement sensitivity matrices H_i by equation (25), and these are propagated back to the epoch time to obtain the matrices $H_{i,*}$.

Since the estimation process does not alter the consider parameters, the estimation error after an update is related to its value before the update by

$$\mathbf{e}_*^+ = \mathbf{e}_*^- + \tilde{S}_* \left(\mathbf{e}_{s*}^+ - \mathbf{e}_{s*}^- \right) = \mathbf{e}_{s*}^- - \tilde{S}_* \sum_{i=1}^{N} K_i \left(H_{i,*} \mathbf{e}_*^- + \mathbf{v}_i + \mathbf{u}_{di} \right), \tag{30}$$

where $\tilde{S}_* = \tilde{S}(t_*)$. Now equation (11) is used to partition the state error into *a priori*, measurement, and

process noise components:

$$\mathbf{e}_{a*}^{+} = \left(I_n - \tilde{S}_* \sum_{i=1}^{N} K_i H_{i,*} \right) \mathbf{e}_*^{-} \tag{31a}$$

$$\mathbf{e}_{v*}^{+} = -\tilde{S}_* \sum_{i=1}^{N} K_i \mathbf{v}_i \tag{31b}$$

$$\mathbf{e}_{w*}^{+} = -\tilde{S}_* \sum_{i=1}^{N} K_i \mathbf{u}_{di}. \tag{31c}$$

COVARIANCE ANALYSIS

The function of an estimator is to produce an estimate $\hat{\mathbf{s}}_*^{+}$ given measurements $\mathbf{y}_i$. Error analysis, however, does not actually compute an estimate, but determines how good an estimate would be produced in a given situation. The covariance matrix of the estimation error, defined, assuming $\mathrm{E}\left[\mathbf{e}\right] = \mathbf{0}$, by

$$P(t_*) = \mathrm{E}\left[\mathbf{e}(t_*)\mathbf{e}(t_*)^{\mathbf{T}}\right], \tag{32}$$

provides a statistical measure of the quality of an estimate at the epoch time t_* of a given scenario[‡]. The *a priori* covariance is simply

$$P_*^{-} = \mathrm{E}\left[\mathbf{e}_*^{-}(\mathbf{e}_*^{-})^{\mathbf{T}}\right] \tag{33}$$

The post-update (*a posteriori*) covariance is obtained by substituting equation (31) into

$$P_*^{+} = \mathrm{E}\left[\mathbf{e}_*^{+}(\mathbf{e}_*^{+})^{\mathbf{T}}\right] = \mathrm{E}\left[(\mathbf{e}_{a*}^{+} + \mathbf{e}_{v*}^{+} + \mathbf{e}_{w*}^{+})(\mathbf{e}_{a*}^{+} + \mathbf{e}_{v*}^{+} + \mathbf{e}_{w*}^{+})^{\mathbf{T}}\right] \tag{34}$$

This computation is complicated by correlations between $\mathbf{e}_{a*}^{+}$ and $\mathbf{e}_{w*}^{+}$. To see how these arise, consider a time t_0 (*not* the epoch time) prior to the epoch time t_* and to all of the measurement times t_i. Equation (7) relates the *a priori* error at the epoch time t_* to the error at t_0 by

$$\mathbf{e}_*^{-} = \Phi(t_*, t_0)\mathbf{e}_0^{-} + \mathbf{w}_d(t_*, t_0). \tag{35}$$

Reference 1 assumed that $\mathbf{e}_*^{-}$ is not correlated with $\mathbf{e}_{w*}^{+}$, but it is more natural to assume that $\mathbf{e}_0^{-}$ is not correlated with $\mathbf{e}_{w*}^{+}$, with the result that the second term on the right side of equation (35) will give rise

[‡]It should be kept in mind that if systematic errors cause $\mathrm{E}\left[\mathbf{e}\right] \neq \mathbf{0}$, then $P(t_*)$ will describe the mean square error matrix, rather than the covariance.

to correlations between $\mathbf{e}_*^-$ and $\mathbf{e}_{w*}^+$ unless the epoch time is earlier than all the measurement times.[§] This is where the "arrow of time" enters the picture. The *a priori* error $\mathbf{e}_*^-$ depends on the process noise in the past of t_*, but not on process noise in the future. We will see that the correlations introduced only involve integrals of the process noise power spectral density over the overlaps of the interval between t_* and t_0 and the intervals between t_* and t_i, so the time t_0 will not appear in our final equations.

The covariance matrix can be written as the sum of *a priori*, measurement noise, and dynamic noise contributions. We include the correlation between $\mathbf{e}_{a*}^+$ and $\mathbf{e}_{w*}^+$ in the process noise part, giving

$$P_*^+ = P_{a*}^+ + P_{v*}^+ + P_{w*}^+, \quad \text{where} \tag{36a}$$

$$P_{a*}^+ = \mathrm{E}\left[\mathbf{e}_{a*}^+(\mathbf{e}_{a*}^+)^{\mathbf{T}}\right], \tag{36b}$$

$$P_{v*}^+ = \mathrm{E}\left[\mathbf{e}_{v*}^+(\mathbf{e}_{v*}^+)^{\mathbf{T}}\right], \tag{36c}$$

$$P_{w*}^+ = \mathrm{E}\left[\mathbf{e}_{w*}^+(\mathbf{e}_{w*}^+)^{\mathbf{T}}\right] + \mathrm{E}\left[\mathbf{e}_{w*}^+(\mathbf{e}_{a*}^+)^{\mathbf{T}}\right] + \mathrm{E}\left[\mathbf{e}_{a*}^+(\mathbf{e}_{w*}^+)^{\mathbf{T}}\right]. \tag{36d}$$

The *a priori* error is uncorrelated with the measurement errors because the *a priori* estimate is assumed to contain no information from the measurements, even from those collected prior to the epoch time.

Substituting equation (31) into equation (36) now gives[¶]

$$P_{a*}^+ = \left(I_n - \tilde{S}_* \sum_{i=1}^N K_i H_{i,*}\right) P_*^- \left(I_n - \tilde{S}_* \sum_{i=1}^N K_i H_{i,*}\right)^{\mathbf{T}} \tag{37a}$$

$$P_{v*}^+ = \tilde{S}_* \left(\sum_{i=1}^N K_i R_i K_i^{\mathbf{T}}\right) \tilde{S}_*^{\mathbf{T}}, \tag{37b}$$

$$\begin{aligned} P_{w*}^+ &= \tilde{S}_* \sum_{i,j=1}^N K_i \mathrm{E}\left[\mathbf{u}_{di}\mathbf{u}_{dj}^{\mathbf{T}}\right] K_j^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}} - \tilde{S}_* \sum_{i=1}^N K_i \mathrm{E}\left[\mathbf{u}_{di}\mathbf{w}_d^{\mathbf{T}}(t_*,t_0)\right] \left(I_n - \tilde{S}_* \sum_{j=1}^N K_j H_{j,*}\right)^{\mathbf{T}} \\ &\quad - \left(I_n - \tilde{S}_* \sum_{i=1}^N K_i H_{i,*}\right) \sum_{j=1}^N \mathrm{E}\left[\mathbf{w}_d(t_*,t_0)\mathbf{u}_{dj}^{\mathbf{T}}\right] K_j^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}} \\ &= \tilde{S}_* \sum_{i,j=1}^N K_i \Gamma_{ij} K_j^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}} - \tilde{S}_* \sum_{i=1}^N K_i \mathrm{E}\left[\mathbf{u}_{di}\mathbf{w}_d^{\mathbf{T}}(t_*,t_0)\right] - \sum_{j=1}^N \mathrm{E}\left[\mathbf{w}_d(t_*,t_0)\mathbf{u}_{dj}^{\mathbf{T}}\right] K_j^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}}, \end{aligned} \tag{37c}$$

[§]Correlations between $\mathbf{e}_*^-$ and $\mathbf{e}_{w*}^+$ are absent in Refs. 2 and 3 because these references implicitly assumed the epoch time to be prior to all the measurement times.

[¶]We use the notation $\sum_{i,j}(\cdot) \equiv \sum_i \sum_j (\cdot)$.

where

$$\Gamma_{ij} \equiv \mathrm{E}\left[\mathbf{u}_{di}\mathbf{u}_{dj}^{\mathbf{T}}\right] + \mathrm{E}\left[\mathbf{u}_{di}\mathbf{w}_d^{\mathbf{T}}(t_*,t_0)\right] H_{j,*}^{\mathbf{T}} + H_{i,*}\mathrm{E}\left[\mathbf{w}_d(t_*,t_0)\mathbf{u}_{dj}^{\mathbf{T}}\right]. \tag{38}$$

We define the process noise covariance integrals

$$Q_*(t_i,t_j) \equiv \int_{t_j}^{t_i} \Phi(t_*,\tau)Q(\tau)\Phi^{\mathbf{T}}(t_*,\tau)\mathrm{d}\tau, \tag{39a}$$

$$Q_{*,j} \equiv Q_*(t_*,t_j), \tag{39b}$$

the second of which, $Q_{*,j}$, is the usual equivalent discrete-time process noise covariance for the transition from time t_j to t_*. In evaluating these and many subsequent integrals of process noise, care must be taken to observe sign changes from reversing the direction of the integration, especially of integrals involving the Dirac delta function. In particular, we note that $Q_*(t_i,t_j)$ is positive definite if $t_i > t_j$ and negative definite if $t_i < t_j$, assuming that the power spectral density $Q(\tau)$ is positive definite. Using these definitions and the relation $\Phi(t_i,\tau) = \Phi(t_i,t_*)\Phi(t_*,\tau)$, we have

$$\begin{aligned}
\mathrm{E}\left[\mathbf{u}_{di}\mathbf{u}_{dj}^{\mathbf{T}}\right] &= H_i\mathrm{E}\left[\mathbf{w}_d(t_i,t_*)\mathbf{w}_d^{\mathbf{T}}(t_j,t_*)\right] H_j^{\mathbf{T}} \\
&= H_i \int_{t_*}^{t_j}\int_{t_*}^{t_i} \Phi(t_i,\tau)Q(\tau)\delta(\tau-\tau')\Phi^{\mathbf{T}}(t_j,\tau')\mathrm{d}\tau\mathrm{d}\tau' \, H_j^{\mathbf{T}} \\
&= \begin{cases} -H_{i,*}\, Q_{*,\min(i,j)}\, H_{j,*}^{\mathbf{T}} & t_* < t_i, t_* < t_j \\ H_{i,*}\, Q_{*,\max(i,j)}\, H_{j,*}^{\mathbf{T}} & t_i < t_*, t_j < t_*, \\ 0 & \text{otherwise.} \end{cases}
\end{aligned} \tag{40}$$

We also have

$$\begin{aligned}
\mathrm{E}\left[\mathbf{u}_{di}\mathbf{w}_d^{\mathbf{T}}(t_*,t_0)\right] &= H_i\mathrm{E}\left[\mathbf{w}_d(t_i,t_*)\mathbf{w}_d^{\mathbf{T}}(t_*,t_0)\right] \\
&= -H_{i,*}\int_{t_0}^{t_*}\int_{t_i}^{t_*} \Phi(t_*,\tau)Q(\tau)\delta(\tau-\tau')\Phi^{\mathbf{T}}(t_*,\tau')\mathrm{d}\tau\mathrm{d}\tau' \\
&= \begin{cases} 0 & t_* \le t_i \\ -H_{i,*}\, Q_{*,i} & t_i < t_*. \end{cases}
\end{aligned} \tag{41}$$

This expectation is simpler than $\mathrm{E}\left[\mathbf{u}_{di}\mathbf{u}_{dj}^{\mathbf{T}}\right]$ because t_0 is less than t_* and t_i by assumption. We can combine

these to compute Γ_{ij}:

$$\Gamma_{ij} = \begin{cases} -H_{i,*}Q_{*,\min(i,j)}H_{j,*}^{\mathbf{T}} & \text{for } t_* < t_i,\ t_* < t_j, \\ H_{i,*}\left(Q_{*,\max(i,j)} - Q_{*,i} - Q_{*,j}\right)H_{j,*}^{\mathbf{T}} = -H_{i,*}Q_{*,\min(i,j)}H_{j,*}^{\mathbf{T}} & \text{for } t_i < t_*,\ t_j < t_*, \\ -H_{i,*}Q_{*,j}H_{j,*}^{\mathbf{T}} = -H_{i,*}Q_{*,\min(i,j)}H_{j,*}^{\mathbf{T}} & \text{for } t_i > t_*,\ t_j < t_*, \\ -H_{i,*}Q_{*,i}H_{j,*}^{\mathbf{T}} = -H_{i,*}Q_{*,\min(i,j)}H_{j,*}^{\mathbf{T}} & \text{for } t_i < t_*,\ t_j > t_*. \end{cases} \tag{42}$$

Remarkably, Γ_{ij} takes the same form for any ordering of t_i, t_j, and t_*. Putting this together gives

$$P_{w*}^{+} = -\tilde{S}_* \sum_{i,j=1}^{N} K_i H_{i,*} Q_{*,\min(i,j)} H_{j,*}^{\mathbf{T}} K_j^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}} + \sum_{i=1}^{k_*} \left(\tilde{S}_* K_i H_{i,*} Q_{*,i} + Q_{*,i} H_{i,*}^{\mathbf{T}} K_i^{\mathbf{T}} \tilde{S}_*^{\mathbf{T}}\right), \tag{43}$$

where k_* is the index of the latest measurement time earlier than t_*, i.e. $t_{k_*} < t_* \leq t_{k_*+1}$. If $t_* < \min(t_i, t_j)$, then $Q_{*,\min(i,j)}$ is negative. The process noise component of the net covariance is always positive, however, because of the process noise component in P_{a*}^{+} implied by equation (35). Equation (43) reduces to the form in Refs. 1–3 when $t_* \leq t_1$; the second sum is identically zero, and all of the $Q_{*,\min(i,j)}$ terms in the first sum give positive contributions to P_{w*}^{+}.

The state estimation errors propagate to any other time according to equation (12). Therefore, the various components of the error covariances at such other times are given by[‖]

$$P_a(t) = \Phi(t,t_*)P_{a*}^{+}\Phi^{\mathbf{T}}(t,t_*) \tag{44a}$$

$$P_v(t) = \Phi(t,t_*)P_{v*}^{+}\Phi^{\mathbf{T}}(t,t_*) \tag{44b}$$

$$\begin{aligned} P_w(t) = {} & \Phi(t,t_*)P_{w*}^{+}\Phi^{\mathbf{T}}(t,t_*) + \mathrm{E}\left[\mathbf{w}_d(t,t_*)\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right] \\ & + \Phi(t,t_*)\mathrm{E}\left[(\mathbf{e}_{w*}^{+} + \mathbf{e}_{a*}^{+})\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right] + \left\{\Phi(t,t_*)\mathrm{E}\left[(\mathbf{e}_{w*}^{+} + \mathbf{e}_{a*}^{+})\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right]\right\}^{\mathbf{T}}. \end{aligned} \tag{44c}$$

We find that

$$\mathrm{E}\left[\mathbf{w}_d(t,t_*)\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right] = \Phi(t,t_*)Q_*(\max(t,t_*),\min(t,t_*))\Phi^{\mathbf{T}}(t,t_*), \tag{45}$$

[‖]Remember that P_{w*}^{+} contains the correlations with the process noise in the *a priori* error at time t_*.

and, with equation (35),

$$\mathrm{E}\left[(\mathbf{e}_{w*}^{+}+\mathbf{e}_{a*}^{+})\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right] = \left(I_n - \tilde{S}_*\sum_{i=1}^{N}K_iH_{i,*}\right)\mathrm{E}\left[\mathbf{w}_d(t_*,t_0)\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right] - \tilde{S}_*\sum_{i=1}^{N}K_i\mathrm{E}\left[\mathbf{u}_{di}\mathbf{w}_d^{\mathbf{T}}(t,t_*)\right]$$
$$= \begin{cases} \left[-Q_*(t_*,t)+\tilde{S}_*\sum_{i=k_+}^{N}K_iH_{i,*}Q_*(\min(t_*,t_i),t)\right]\Phi^{\mathbf{T}}(t,t_*) & t\le t_*, \\ -\tilde{S}_*\sum_{i=k_+}^{N}K_iH_{i,*}Q_*(\min(t,t_i),t_*)\Phi^{\mathbf{T}}(t,t_*) & t>t_*, \end{cases} \tag{46}$$

where k_+ is the index of the earliest measurement time later than either t or t_*, i.e. $t_{k_+-1}\le\min(t_*,t)<t_{k_+}$. This gives

$$P_w(t) = \Phi(t,t_*)[P_{w*}^{+} - Q_*(t_*,t) + B(t,t_*) + B^{\mathbf{T}}(t,t_*)]\Phi^{\mathbf{T}}(t,t_*), \tag{47}$$

where

$$B(t,t_*) \equiv \tilde{S}_*\sum_{i=k_+}^{N}K_iH_{i,*}Q_*(\min(t_*,t_i),\min(t,t_i)). \tag{48}$$

Note that $Q_*(t_*,t)$ is positive if $t<t_*$ and negative if $t>t_*$.

Formal Covariance and Delta Covariances

The estimation system does not take the consider states into account, possibly out of ignorance but more likely by deliberate choice to simplify its structure, so it updates an n_s-dimensional solve-for state error vector. Denoting the *a priori* estimate of this state error vector at the epoch time by $\hat{\mathbf{e}}_{*|0}$ and the updated value incorporating all N measurements by $\hat{\mathbf{e}}_{*|N}$, we have in parallel with equation (30)

$$\hat{\mathbf{e}}_{*|N} = \hat{\mathbf{e}}_{*|0} - \sum_{i=1}^{N}K_i\left(H_{i,*}\hat{\mathbf{e}}_{*|0}+\hat{\mathbf{v}}_i+\hat{\mathbf{u}}_{di}\right) = \left(I_n - \sum_{i=1}^{N}K_i\hat{H}_{i,*}\right)\hat{\mathbf{e}}_{*|0} - \sum_{i=1}^{N}K_i(\hat{\mathbf{v}}_i+\hat{\mathbf{u}}_{di}). \tag{49}$$

This equation lacks the projection operator onto the solve-for space, because the estimator only deals with the solve-for parameters. In this equation

$$\hat{H}_{i,*} \equiv \hat{H}_i\hat{\Phi}(t_i,t_*). \tag{50}$$

where $\hat{H}_i$ is the $m_i\times n_s$ measurement sensitivity matrix with respect to the solve-for states, $\hat{\Phi}(t_i,t_*)$ an $n_s\times n_s$ transition matrix of the solve-for states, $\hat{\mathbf{v}}_i$ is a Gaussian white measurement noise, and $\hat{\mathbf{u}}_{di}$ is defined analogously to equation (29). These are all the imprecise values modeled by the estimator, and are

not assumed to be equal to the projection onto the solve-for space of their true values.** Of course, an estimator that more closely approximates the truth can be expected to provide better estimates. The same gain matrices appear in the equations for the true and formal covariances, however, because the true gain matrix is the one provided by the estimator.

In addition to providing a solve-for parameter estimate, the estimator will also compute an $n_s \times n_s$ covariance matrix estimate, known as the *formal* error covariance,

$$\hat{P}_{*|N} \equiv \mathrm{E}\left[\hat{\mathbf{e}}_{*|N}\hat{\mathbf{e}}_{*|N}^{\mathbf{T}}\right] = \hat{P}_{a*}^{+} + \hat{P}_{v*}^{+} + \hat{P}_{w*}^{+}, \tag{51}$$

Following a path parallel to that leading to equations (37) and (43) gives

$$\hat{P}_{a*}^{+} = \left(I_{n_s} - \sum_{i=1}^{N} K_i \hat{H}_{i,*}\right) \hat{P}_{*|0} \left(I_{n_s} - \sum_{i=1}^{N} K_i \hat{H}_{i,*}\right)^{\mathbf{T}} \tag{52a}$$

$$\hat{P}_{v*}^{+} = \sum_{i=1}^{N} K_i \hat{R}_i K_i^{\mathbf{T}} \tag{52b}$$

$$\hat{P}_{w*}^{+} = -\sum_{i,j=1}^{N} K_i \hat{H}_{i,*} \hat{Q}_{*,\min(i,j)} \hat{H}_{j,*}^{\mathbf{T}} K_j^{\mathbf{T}} + \sum_{i=1}^{k_*} (K_i \hat{H}_{i,*} \hat{Q}_{*,i} + \hat{Q}_{*,i} \hat{H}_{i,*}^{\mathbf{T}} K_i^{\mathbf{T}}). \tag{52c}$$

In these equations $\hat{P}_{*|0}$ is the *a priori* covariance, $\hat{R}_i \equiv \mathrm{E}\left[\hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^{\mathbf{T}}\right]$ is the measurement error covariance,

$$\hat{Q}_*(t_i, t_j) \equiv \int_{t_j}^{t_i} \hat{\Phi}(t_*, \tau) \hat{Q}(\tau) \hat{\Phi}^{\mathbf{T}}(t_*, \tau) \mathrm{d}\tau, \tag{53a}$$

$$\hat{Q}_{*,j} \equiv \hat{Q}_*(t_*, t_j) \tag{53b}$$

with $\hat{Q}(\tau)$ an $n_s \times n_s$ power spectral density of the process noise assumed by the estimator. As was noted above, none of these quantities is assumed to be equal to the projection onto the solve-for space of its corresponding true value.

**In the same spirit, $\Phi_{ss}(t_i, t_*)$ and $\tilde{H}_{si}$ in Ref. 1 can be replaced by $\hat{\Phi}(t_i, t_*)$ and $\hat{H}_{i,*}$, respectively.

The formal covariance is propagated to other times by the analogs of equations (44a), (44b), and (47)

$$\hat{P}(t) = \hat{P}_a(t) + \hat{P}_v(t) + \hat{P}_w(t) \tag{54a}$$

$$\hat{P}_a(t) = \hat{\Phi}(t, t_*)\hat{P}^+_{a*}\hat{\Phi}^{\mathbf{T}}(t, t_*) \tag{54b}$$

$$\hat{P}_v(t) = \hat{\Phi}(t, t_*)\hat{P}^+_{v*}\hat{\Phi}^{\mathbf{T}}(t, t_*) \tag{54c}$$

$$\hat{P}_w(t) = \hat{\Phi}(t, t_*)[\hat{P}^+_{w*} - \hat{Q}_*(t_*, t) + \hat{B}(t, t_*) + \hat{B}^{\mathbf{T}}(t, t_*)]\hat{\Phi}^{\mathbf{T}}(t, t_*), \tag{54d}$$

where

$$\hat{B}(t, t_*) \equiv \sum_{i=k_+}^{N} K_i \hat{H}_{i,*} \hat{Q}_*(\min(t_*, t_i), \min(t, t_i)). \tag{55}$$

The full covariance matrix is needed for computation, but in most cases we are only interested in its solve-for part

$$P_{ss}(t) = \mathrm{E}\left[\mathbf{e}_s(t)\mathbf{e}_s^{\mathbf{T}}(t)\right] = SP_a(t)S^{\mathbf{T}} + SP_v(t)S^{\mathbf{T}} + SP_w(t)S^{\mathbf{T}}. \tag{56}$$

This will generally differ from the formal covariance due to estimator mistuning and/or neglect of consider parameters in the formal covariance. We express the difference as

$$P_{ss}(t) = \hat{P}(t) + \Delta P_a(t) + \Delta P_v(t) + \Delta P_w(t), \tag{57}$$

where the delta covariance matrices are

$$\Delta P_a = SP_aS^{\mathbf{T}} - \hat{P}_a, \quad \Delta P_v = SP_vS^{\mathbf{T}} - \hat{P}_v, \quad \text{and} \quad \Delta P_w = SP_wS^{\mathbf{T}} - \hat{P}_w \tag{58}$$

The formal covariance matrices can provide either underestimates or overestimates of the errors, so the delta covariance matrices can be positive or negative.

COMPUTATION OF THE OPTIMAL GAIN MATRIX

The optimal update gain matrix is chosen to minimize the post-update covariance at the epoch time. The total covariance can be written, from equations (51) and (52) as

$$\hat{P}_{*|N} = \hat{P}_{*|0} + \sum_{i,j=1}^{N} K_i W_{ij} K_j^{\mathbf{T}} - \sum_{i=1}^{N} \left[K_i \hat{H}_{i,*}(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,i}) + (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,i})\hat{H}^{\mathbf{T}}_{i,*} K_i^{\mathbf{T}} \right], \tag{59}$$

where

$$W_{ij} \equiv \hat{H}_{i,*}(\hat{P}_{*|0} - \hat{Q}_{*,\min(i,j)})\hat{H}^{\mathbf{T}}_{j,*} + \hat{R}_i\delta_{ij} = W^{\mathbf{T}}_{ji}. \tag{60}$$

The forward process noise covariance is defined as

$$\hat{Q}^{\rightarrow}_{*,i} \equiv \begin{cases} \hat{Q}_{*,i} & t_i < t_* \\ 0 & t_i \geq t_*, \end{cases} \tag{61}$$

and δ_{ij} is the Kronecker delta

$$\delta_{ij} = \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j. \end{cases} \tag{62}$$

We now define the symmetric matrix $\tilde{W}_N$ as an $N \times N$ matrix of $m_i \times m_j$ blocks W_{ij}:

$$\tilde{W}_N = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1N} \\ W_{21} & W_{22} & \dots & W_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1} & W_{N2} & \dots & W_{NN} \end{bmatrix}. \tag{63}$$

The inverse $\tilde{W}_N^{-1}$ can also be written an $N \times N$ matrix of $m_i \times m_j$ blocks,

$$\tilde{W}_N^{-1} = \begin{bmatrix} [\tilde{W}_N^{-1}]_{11} & [\tilde{W}_N^{-1}]_{12} & \dots & [\tilde{W}_N^{-1}]_{1N} \\ [\tilde{W}_N^{-1}]_{21} & [\tilde{W}_N^{-1}]_{22} & \dots & [\tilde{W}_N^{-1}]_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ [\tilde{W}_N^{-1}]_{N1} & [\tilde{W}_N^{-1}]_{N2} & \dots & [\tilde{W}_N^{-1}]_{NN} \end{bmatrix}, \tag{64}$$

and the relations $\tilde{W}_N\tilde{W}_N^{-1} = \tilde{W}_N^{-1}\tilde{W}_N = I$ imply that

$$\sum_{j=1}^{N} W_{ij}[\tilde{W}_N^{-1}]_{jk} = \sum_{j=1}^{N}[\tilde{W}_N^{-1}]_{ij}W_{jk} = \delta_{ik}I_{m_i}. \tag{65}$$

Now equation (51) can be written by completing the square as

$$\hat{P}_{*|N} = \hat{P}_{*|0} - \sum_{i,j=1}^{N} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,i}) \hat{H}^{\mathbf{T}}_{i,*} \, [\tilde{W}_N^{-1}]_{ij} \, \hat{H}_{j,*} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,j})$$
$$+ \sum_{i,j=1}^{N} \left[K_i - \sum_{k=1}^{N} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,k}) \hat{H}^{\mathbf{T}}_{k,*} [\tilde{W}_N^{-1}]_{ki} \right] W_{ij} \left[K_j - \sum_{k=1}^{N} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,k}) \hat{H}^{\mathbf{T}}_{k,*} [\tilde{W}_N^{-1}]_{kj} \right]^{\mathbf{T}} . \quad (66)$$

If A and B are real symmetric matrices, we say that $A \geq B$ if the matrix $A - B$ is positive semidefinite. It is clear that

$$\hat{P}_{*|N} \geq \hat{P}_{*|0} - \sum_{i,j=1}^{N} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,i}) \hat{H}^{\mathbf{T}}_{i,*} \, [\tilde{W}_N^{-1}]_{ij} \, \hat{H}_{j,*} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,j}) \quad (67)$$

in this sense, with equality if the measurement gain matrix is chosen to be

$$K_i = \sum_{k=1}^{N} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,k}) \hat{H}^{\mathbf{T}}_{k,*} [\tilde{W}_N^{-1}]_{ki}. \quad (68)$$

We will use this gain matrix to achieve the minimum $\hat{P}_{*|N}$, which can be expressed more compactly as

$$\hat{P}_{*|N} = \hat{P}_{*|0} - \sum_{j=1}^{N} K_j \, \hat{H}_{j,*} (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,j}) \quad (69)$$

Batch Least-Squares Case

We want to show that our formulation reduces to the usual batch least-squares estimator if the assumed process noise is zero. In that case

$$\tilde{W}_N = \hat{H}_* \, \hat{P}_{*|0} \, \hat{H}^{\mathbf{T}}_* + \hat{R}, \quad (70)$$

where

$$\hat{H}_* \equiv \begin{bmatrix} \hat{H}_{1,*} \\ \vdots \\ \hat{H}_{N,*} \end{bmatrix} \quad (71)$$

and $\hat{R}$ is block-diagonal with the matrices $\hat{R}_i$ along its main diagonal. The Sherman-Morrison-Woodbury matrix inversion lemma[13] gives

$$\tilde{W}_N^{-1} = \hat{R}^{-1} - \hat{R}^{-1} \hat{H}_* \left(\hat{P}^{-1}_{*|0} + \hat{H}^{\mathbf{T}}_* \, \hat{R}^{-1} \, \hat{H}_* \right)^{-1} \hat{H}^{\mathbf{T}}_* \hat{R}_j^{-1} \quad (72)$$

or

$$[\tilde{W}_N^{-1}]_{ki} = \hat{R}_k^{-1}\delta_{ki} - \hat{R}_k^{-1}\hat{H}_{k,*}\left(\hat{P}_{*|0}^{-1} + \sum_{j=1}^{N}\hat{H}_{j,*}^{\mathbf{T}}\,\hat{R}_j^{-1}\,\hat{H}_{j,*}\right)^{-1}\hat{H}_{i,*}^{\mathbf{T}}\hat{R}_i^{-1} \tag{73}$$

Substituting into equation (68) with $\hat{Q}_{*,k}^{\rightarrow} = 0$ gives

$$K_i = \hat{P}_{*|0}\left[I_{n_s} - \sum_{k=1}^{N}\hat{H}_{k,*}^{\mathbf{T}}\hat{R}_k^{-1}\hat{H}_{k,*}\left(\hat{P}_{*|0}^{-1} + \sum_{j=1}^{N}\hat{H}_{j,*}^{\mathbf{T}}\,\hat{R}_j^{-1}\,\hat{H}_{j,*}\right)^{-1}\right]\hat{H}_{i,*}^{\mathbf{T}}\hat{R}_i^{-1} \tag{74}$$

Now writing $I_{n_s} = \left(\hat{P}_{*|0}^{-1} + \sum_{k=1}^{N}\hat{H}_{k,*}^{\mathbf{T}}\,\hat{R}_k^{-1}\,\hat{H}_{k,*}\right)\left(\hat{P}_{*|0}^{-1} + \sum_{j=1}^{N}\hat{H}_{j,*}^{\mathbf{T}}\,\hat{R}_j^{-1}\,\hat{H}_{j,*}\right)^{-1}$ and performing appropriate cancellations gives

$$K_i = \left(\hat{P}_{*|0}^{-1} + \sum_{j=1}^{N}\hat{H}_{j,*}^{\mathbf{T}}\,\hat{R}_j^{-1}\,\hat{H}_{j,*}\right)^{-1}\hat{H}_{i,*}^{\mathbf{T}}\hat{R}_i^{-1}, \tag{75}$$

which is just the usual result obtained by minimizing a quadratic loss function with weighted contributions from the measurement innovations and the *a priori* estimate, as in Refs. 1 – 3.

Sequential Implementation

Inverting the very large matrix $\tilde{W}_N$ is a formidable problem in the general case, and the Sherman-Morrison-Woodbury lemma is not applicable if process noise is present. The inversion of this large matrix can be avoided even when process noise is present by implementing a sequential estimation procedure. In developing a sequential estimator, it will be convenient to denote the gain matrices by $K_{i|k}$, where the index k is the number of measurements included in the update. The estimation errors after processing $k-1$ and k measurements are given by equation (49) as

$$\hat{\mathbf{e}}_{*|k-1} = \left(I_{n_s} - \sum_{i=1}^{k-1}K_{i|k-1}\hat{H}_{i,*}\right)\hat{\mathbf{e}}_{*|0} - \sum_{i=1}^{k-1}K_{i|k-1}(\hat{\mathbf{v}}_i + \hat{\mathbf{u}}_{di}), \tag{76}$$

$$\hat{\mathbf{e}}_{*|k} = \left(I_{n_s} - \sum_{i=1}^{k}K_{i|k}\hat{H}_{i,*}\right)\hat{\mathbf{e}}_{*|0} - \sum_{i=1}^{k}K_{i|k}(\hat{\mathbf{v}}_i + \hat{\mathbf{u}}_{di}). \tag{77}$$

For a sequential method, we need to be able to write

$$\hat{\mathbf{e}}_{*|k} = (I_{n_s} - K_{k|k}\hat{H}_{k,*})\hat{\mathbf{e}}_{*|k-1} - K_{k|k}(\hat{\mathbf{v}}_k + \hat{\mathbf{u}}_{dk}). \tag{78}$$

Substituting equation (76) into equation (78) and comparing with equation (77) gives the consistency requirement

$$K_{i|k} = (I_{n_s} - K_{k|k}\hat{H}_{k,*})K_{i|k-1} \quad i < k. \tag{79}$$

It is not difficult to see that complete agreement among all the terms of equations (76)−(78) follows if the gain matrices satisfy this condition. It is shown in the Appendix that equation (79) is always satisfied if the estimator ignores process noise, but it can only be satisfied in the presence of process noise if $\hat{Q}^{\rightarrow}_{*,i} = \hat{Q}_{*,i}$ for $i < k$, which means that $t_i \leq t_*$. We will thus assume that t_* is in the future of all of the measurements, which means that we can remove the arrow from every appearance of $\hat{Q}^{\rightarrow}_{*,i}$ in the following.

To make further progress, let $\tilde{W}_k$ denote the matrix defined by equation (63), but only including k measurements. We write this in partitioned form as

$$\tilde{W}_k = \begin{bmatrix} \tilde{W}_{k-1} & V_k \\ V_k^{\mathbf{T}} & W_{kk} \end{bmatrix}, \tag{80}$$

where

$$V_k \equiv \begin{bmatrix} W_{1k} \\ W_{2k} \\ \vdots \\ W_{(k-1)k} \end{bmatrix}. \tag{81}$$

Letting $\hat{P}_{*|k}$ denote the formal covariance incorporating k measurements, equations (60) and (67) (with equality) can be used to show that

$$\hat{H}_{k,*}(\hat{P}_{*|0} - \hat{P}_{*|k-1})\hat{H}^{\mathbf{T}}_{k,*} = \sum_{i,j=1}^{k-1} W_{ki}[\tilde{W}^{-1}_{k-1}]_{ij}W^{\mathbf{T}}_{jk} = V_k^{\mathbf{T}}\tilde{W}^{-1}_{k-1}V_k, \tag{82}$$

from which it follows that

$$W_{kk} - V_k^{\mathbf{T}}\tilde{W}^{-1}_{k-1}V_k = \hat{H}_{k,*}(\hat{P}_{*|k-1} - \hat{Q}_{*,k})\,\hat{H}^{\mathbf{T}}_{k,*} + \hat{R}_k. \tag{83}$$

The standard equation for the inverse of a partitioned matrix gives

$$\tilde{W}_k^{-1} = \begin{bmatrix} (\tilde{W}_{k-1} - V_k W_{kk}^{-1} V_k^{\mathbf{T}})^{-1} & U_k \\ U_k^{\mathbf{T}} & (W_{kk} - V_k^{\mathbf{T}} \tilde{W}_{k-1}^{-1} V_k)^{-1} \end{bmatrix}, \tag{84}$$

where

$$U_k = -\tilde{W}_{k-1}^{-1} V_k (W_{kk} - V_k^{\mathbf{T}} \tilde{W}_{k-1}^{-1} V_k)^{-1} = -\tilde{W}_{k-1}^{-1} V_k [\tilde{W}_k^{-1}]_{kk}. \tag{85}$$

These relations give the gain matrix $K_{k|k}$, which is the only one needed for the sequential estimator, as

$$\begin{aligned} K_{k|k} &= \sum_{i=1}^{k} (\hat{P}_{*|0} - \hat{Q}_{*,i})\, \hat{H}_{i,*}^{\mathbf{T}} [\tilde{W}_k^{-1}]_{ik} = \sum_{i=1}^{k-1} (\hat{P}_{*|0} - \hat{Q}_{*,i})\, \hat{H}_{i,*}^{\mathbf{T}} [\tilde{W}_k^{-1}]_{ik} + (\hat{P}_{*|0} - \hat{Q}_{*,k})\, \hat{H}_{k,*}^{\mathbf{T}} [\tilde{W}_k^{-1}]_{kk} \\ &= (\hat{P}_{*|k-1} - \hat{Q}_{*,k})\, \hat{H}_{k,*}^{\mathbf{T}} [\tilde{W}_k^{-1}]_{kk} = (\hat{P}_{*|k-1} - \hat{Q}_{*,k})\, \hat{H}_{k,*}^{\mathbf{T}} \left[\hat{H}_{k,*} (\hat{P}_{*|k-1} - \hat{Q}_{*,k})\, \hat{H}_{k,*}^{\mathbf{T}} + \hat{R}_k \right]^{-1} \end{aligned} \tag{86}$$

The third equality uses, for $i < k$,

$$\begin{aligned} [\tilde{W}_k^{-1}]_{ik} = [U_k]_i &= -\sum_{j=1}^{k-1} [\tilde{W}_{k-1}^{-1}]_{ij} [V_k]_j [\tilde{W}_k^{-1}]_{kk} = -\sum_{j=1}^{k-1} [\tilde{W}_{k-1}^{-1}]_{ij} W_{jk} [\tilde{W}_k^{-1}]_{kk} \\ &= -\sum_{j=1}^{k-1} [\tilde{W}_{k-1}^{-1}]_{ij} \hat{H}_{j,*} (\hat{P}_{*|0} - \hat{Q}_{*,j})\, \hat{H}_{k,*}^{\mathbf{T}} [\tilde{W}_k^{-1}]_{kk} \end{aligned} \tag{87}$$

and then

$$\sum_{i,j=1}^{k-1} (\hat{P}_{*|0} - \hat{Q}_{*,i})\, \hat{H}_{i,*}^{\mathbf{T}} [\tilde{W}_{k-1}^{-1}]_{ij} \hat{H}_{j,*} (\hat{P}_{*|0} - \hat{Q}_{*,j}) = \hat{P}_{*|0} - \hat{P}_{*|k-1}, \tag{88}$$

and the fourth equality uses equation (83). Equation (86) expresses $K_{k|k}$ in terms of only $\hat{P}_{*|k-1}$ and quantities available at the measurement time t_k. Completing the development of the recursive estimator requires a recursion relation for $\hat{P}_{*|k}$. Using the consistency condition on the gain matrix, equation (79), in

equation (69) gives

$$\hat{P}_{*|k} = \hat{P}_{*|0} - \sum_{j=1}^{k} K_{j|k}\,\hat{H}_{j,*}(\hat{P}_{*|0} - \hat{Q}_{*,j}) \;= \hat{P}_{*|0} - K_{k|k}\,\hat{H}_{k,*}(\hat{P}_{*|0} - \hat{Q}_{*,k}) \;- \sum_{j=1}^{k-1} K_{j|k}\,\hat{H}_{j,*}(\hat{P}_{*|0} - \hat{Q}_{*,j})$$

$$= (I_{n_s} - K_{k|k}\,\hat{H}_{k,*}) \left[\hat{P}_{*|0} - \sum_{j=1}^{k-1} K_{j|k-1}\,\hat{H}_{j,*}(\hat{P}_{*|0} - \hat{Q}_{*,j})\right] + K_{k|k}\,\hat{H}_{k,*}\hat{Q}_{*,k}$$

$$= (I_{n_s} - K_{k|k}\,\hat{H}_{k,*})\hat{P}_{*|k-1} + K_{k|k}\,\hat{H}_{k,*}\hat{Q}_{*,k} = \hat{P}_{*|k-1} - K_{k|k}\,\hat{H}_{k,*}(\hat{P}_{*|k-1} - \hat{Q}_{*,k}). \tag{89}$$

It is not surprising that the final form strongly resembles equation (69) with an update from only the last measurement and with $\hat{P}_{*|k-1}$ as the *a priori* covariance.

Relation to the Kalman Filter

If we define

$$\hat{P}_{*,k} \equiv \hat{P}_{*|k-1} - \hat{Q}_{*,k}, \tag{90}$$

then equation (86) becomes

$$K_{k|k} = \hat{P}_{*,k}\,\hat{H}^{\mathbf{T}}_{k,*}(\hat{H}_{k,*}\hat{P}_{*,k}\,\hat{H}^{\mathbf{T}}_{k,*} + \hat{R}_k)^{-1} \tag{91}$$

and equation (89) becomes

$$\hat{P}_{*,k+1} = (I_{n_s} - K_{k|k}\,\hat{H}_{k,*})\hat{P}_{*,k} + \hat{Q}_{*,k} - \hat{Q}_{*,k+1} = (I_{n_s} - K_{k|k}\,\hat{H}_{k,*})\hat{P}_{*,k} + \hat{Q}_{*}(t_{k+1}, t_k), \tag{92}$$

which are equivalent to the equations of the epoch-state estimator used by Montenbruck.[14] It should be pointed out, though, that $\hat{P}_{*,k}$ is not the covariance of any obvious error vector.

If we further make the identification

$$\hat{P}_{*,k} = \hat{\Phi}(t_*, t_k)\hat{P}_{k|k-1}\hat{\Phi}^{\mathbf{T}}(t_*, t_k) \tag{93}$$

and use

$$\hat{\Phi}(t', t) = \hat{\Phi}(t', \tau)\hat{\Phi}(\tau, t), \tag{94}$$

then equations (91) and (92) become

$$K_{k|k} = \hat{\Phi}(t_*, t_k)\hat{K}_k \tag{95}$$

and

$$\hat{P}_{k+1|k} = \hat{\Phi}(t_{k+1}, t_k)(I_{n_s} - \hat{K}_k\,\hat{H}_k)\hat{P}_{k|k-1}\hat{\Phi}^{\mathbf{T}}(t_{k+1}, t_k) + \hat{\Phi}(t_{k+1}, t_*)\hat{Q}_*(t_{k+1}, t_k)\hat{\Phi}^{\mathbf{T}}(t_{k+1}, t_*), \tag{96}$$

where

$$\hat{K}_k \equiv \hat{P}_{k|k-1}\hat{H}_k^{\mathbf{T}}(\hat{H}_k\hat{P}_{k|k-1}\,\hat{H}_k^{\mathbf{T}} + \hat{R}_k)^{-1}. \tag{97}$$

Notice that $\hat{H}_k$ appears in these equations rather than $\hat{H}_{k,*}$. Equation (96) can be written as the usual two-step Kalman filter process

$$\hat{P}_{k|k} = (I_{n_s} - \hat{K}_k\,\hat{H}_k)\hat{P}_{k|k-1} \tag{98}$$

$$\hat{P}_{k+1|k} = \hat{\Phi}(t_{k+1}, t_k)\hat{P}_{k|k}\hat{\Phi}^{\mathbf{T}}(t_{k+1}, t_k) + \int_{t_k}^{t_{k+1}} \hat{\Phi}(t_{k+1}, \tau)\hat{Q}(\tau)\hat{\Phi}^{\mathbf{T}}(t_{k+1}, \tau)\mathrm{d}\tau, \tag{99}$$

where the expression for the process noise follows from equation (53). It is now easy to recognize $\hat{P}_{k|k-1}$ and $\hat{P}_{k|k}$ as the pre- and post-update covariances of the estimation error at the measurement time t_k.

SUMMARY AND CONCLUSION

We have updated and extended previous work on consider covariance analysis to improve the treatment of process noise when the update time is not prior to all of the measurements. Consider covariance analysis treats estimators that account for only a subset of the parameters, possibly out of ignorance but more likely by deliberate choice to simplify their structure. The covariance analysis, however, considers the effect on the covariance of all the variables, including those ignored by the estimator. Our analysis explicitly partitions the full covariance and solve-for covariance into three subspaces containing the influence of measurement noise, process noise, and *a priori* errors, respectively. Explicitly exhibiting the *a priori* component of the covariance is useful because *a priori* information about observable parameters is "forgotten" as more data are processed, providing a useful test of observability.

We have found an expression for the gain of an optimal epoch-state estimator incorporating a model of process noise, and showed that this reduces to the usual batch estimator if process noise is ignored when computing the measurement update gains. Our formulation agrees with our previous analysis of batch filtering methods if the epoch time is prior to all measurement times, but differs if the epoch time is amid

the data or in the future of all data. If the epoch time is no earlier than the time of the last measurement processed, a sequential implementation of the filter is possible, and we show that this is basically equivalent to an extended Kalman filter.

APPENDIX: CONSISTENCY CONDITION FOR SEQUENTIAL IMPLEMENTATION

We will first show that the consistency condition requires that either the estimator ignores process noise or that $t_i \leq t_*$ for $i < k$. We will use a very simple model with a one-component state vector and two scalar measurements with $\hat{H}_{1,*} = \hat{H}_{2,*} = 1$. The blocks

$$W_{ij} = (\hat{P}_{*|0} - \hat{Q}_{*,\min(i,j)}) + \hat{R}_i \delta_{ij} \tag{100}$$

of $\tilde{W}_k$ are scalars, so we have

$$\tilde{W}_1^{-1} = \frac{1}{W_{11}} \tag{101a}$$

$$\tilde{W}_2^{-1} = \frac{1}{W_{11}W_{22} - W_{12}^2} \begin{bmatrix} W_{22} & -W_{12} \\ -W_{12} & W_{11} \end{bmatrix}. \tag{101b}$$

The gains are given by equation (68) as

$$K_{1|1} = \frac{\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1}}{W_{11}} \tag{102a}$$

$$K_{1|2} = \frac{(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{22} - (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,2})W_{12}}{W_{11}W_{22} - W_{12}^2} \tag{102b}$$

$$K_{2|2} = \frac{(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,2})W_{11} - (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{12}}{W_{11}W_{22} - W_{12}^2} \tag{102c}$$

The consistency condition, equation (79) then gives

$$\begin{aligned}(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})[W_{11}W_{22} - W_{12}^2 - (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,2})W_{11} + (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{12}] = \\ W_{11}[(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{22} - (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,2})W_{12}]\end{aligned} \tag{103}$$

Canceling the common term of $(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{11}W_{22}$ from the two sides of the equation and rearranging gives

$$[(\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,2})W_{11} - (\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1})W_{12}](\hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1} - W_{12}) = 0 \tag{104}$$

The quantity in square brackets is certainly not zero, because there is nothing to cancel the term R_1 in W_{11}. Thus we must have

$$0 = \hat{P}_{*|0} - \hat{Q}^{\rightarrow}_{*,1} - W_{12} = \hat{Q}_{*,1} - \hat{Q}^{\rightarrow}_{*,1} \tag{105}$$

This equation is satisfied if either the estimator ignores process noise entirely or if $t_1 \leq t_*$. In the latter case it follows by by induction that we must have $t_i \leq t_*$ for all $i < k$.

The next task is to show that the consistency condition holds if $\hat{Q}^{\rightarrow}_{*,i} = \hat{Q}_{*,i}$ for $i \leq k$. We will show this in the general case. Equation (68) gives, with the subscripting notation used for the sequential implementation,

$$K_{i|k} = \sum_{j=1}^{k} (\hat{P}_{*|0} - \hat{Q}_{*,j}) \hat{H}^{\mathbf{T}}_{j,*} [\tilde{W}_k^{-1}]_{ji} = \sum_{j=1}^{k-1} (\hat{P}_{*|0} - \hat{Q}_{*,j}) \hat{H}^{\mathbf{T}}_{j,*} [\tilde{W}_k^{-1}]_{ji} + (\hat{P}_{*|0} - \hat{Q}_{*,k}) \hat{H}^{\mathbf{T}}_{k,*} [\tilde{W}_k^{-1}]_{ki}. \tag{106}$$

We use equation (84) and the Sherman-Morrison-Woodbury matrix inversion lemma write the upper left-hand corner of $\tilde{W}_k^{-1}$ in the form

$$\begin{aligned} (\tilde{W}_{k-1} - V_k W_{kk}^{-1} V_k^{\mathbf{T}})^{-1} &= \tilde{W}_{k-1}^{-1} + \tilde{W}_{k-1}^{-1} V_k (W_{kk} - V_k^{\mathbf{T}} \tilde{W}_{k-1}^{-1} V_k)^{-1} V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1} \\ &= \tilde{W}_{k-1}^{-1} + \tilde{W}_{k-1}^{-1} V_k \, [\tilde{W}_k^{-1}]_{kk} V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1}, \end{aligned} \tag{107}$$

which is made up of the blocks

$$[\tilde{W}_k^{-1}]_{ji} = [\tilde{W}_{k-1}^{-1}]_{ji} + [\tilde{W}_{k-1}^{-1} V_k]_j \, [\tilde{W}_k^{-1}]_{kk} \, [V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1}]_i, . \tag{108}$$

We also have

$$[\tilde{W}_k^{-1}]_{ki} = [U_k^{\mathbf{T}}]_i = -[\tilde{W}_k^{-1}]_{kk} \, [V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1}]_i \tag{109}$$

With these substitutions,

$$\begin{aligned} K_{i|k} = \sum_{j=1}^{k-1} (\hat{P}_{*|0} - \hat{Q}_{*,j}) \hat{H}^{\mathbf{T}}_{j,*} [\tilde{W}_{k-1}^{-1}]_{ji} + \sum_{j=1}^{k-1} (\hat{P}_{*|0} - \hat{Q}_{*,j}) \hat{H}^{\mathbf{T}}_{j,*} [\tilde{W}_{k-1}^{-1} V_k]_j \, [\tilde{W}_k^{-1}]_{kk} \, [V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1}]_i \\ - (\hat{P}_{*|0} - \hat{Q}_{*,k}) \hat{H}^{\mathbf{T}}_{k,*} [\tilde{W}_k^{-1}]_{kk} \, [V_k^{\mathbf{T}} \, \tilde{W}_{k-1}^{-1}]_i. \end{aligned} \tag{110}$$

The first term on the right side of this equation is easily recognized to be $K_{i|k-1}$. The sum over j in the

second term can be written as

$$\begin{aligned}
\sum_{j=1}^{k-1}(\hat{P}_{*|0}-\hat{Q}_{*,j})\hat{H}^{\mathbf{T}}_{j,*}[\tilde{W}^{-1}_{k-1}V_k]_j &= \sum_{j,m=1}^{k-1}(\hat{P}_{*|0}-\hat{Q}_{*,j})\hat{H}^{\mathbf{T}}_{j,*}[\tilde{W}^{-1}_{k-1}]_{jm}[V_k]_m \\
&= \sum_{j,m=1}^{k-1}(\hat{P}_{*|0}-\hat{Q}_{*,j})\hat{H}^{\mathbf{T}}_{j,*}[\tilde{W}^{-1}_{k-1}]_{jm}W_{mk} \\
&= \sum_{j,m=1}^{k-1}(\hat{P}_{*|0}-\hat{Q}_{*,j})\hat{H}^{\mathbf{T}}_{j,*}[\tilde{W}^{-1}_{k-1}]_{jm}\hat{H}_{m,*}(\hat{P}_{*|0}-\hat{Q}_{*,m})\hat{H}^{\mathbf{T}}_{k,*} = (\hat{P}_{*|0}-\hat{P}_{*|k-1})\hat{H}^{\mathbf{T}}_{k,*}, \qquad (111)
\end{aligned}$$

where equation (67) (with equality) gives the last line. Putting all this together gives

$$\begin{aligned}
K_{i|k} &= K_{i|k-1} + (\hat{P}_{*|0}-\hat{P}_{*|k-1})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\,[V_k^{\mathbf{T}}\,\tilde{W}^{-1}_{k-1}]_i - (\hat{P}_{*|0}-\hat{Q}_{*,k})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\,[V_k^{\mathbf{T}}\,\tilde{W}^{-1}_{k-1}]_i \\
&= K_{i|k-1} - (\hat{P}_{*|k-1}-\hat{Q}_{*,k})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\sum_{j=1}^{k-1}[V_k^{\mathbf{T}}]_j[\tilde{W}^{-1}_{k-1}]_{ji} \\
&= K_{i|k-1} - (\hat{P}_{*|k-1}-\hat{Q}_{*,k})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\sum_{j=1}^{k-1}W_{kj}[\tilde{W}^{-1}_{k-1}]_{ji} \\
&= K_{i|k-1} - (\hat{P}_{*|k-1}-\hat{Q}_{*,k})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\sum_{j=1}^{k-1}\hat{H}_{k,*}(\hat{P}_{*|0}-\hat{Q}_{*,j})\hat{H}^{\mathbf{T}}_{j,*}[\tilde{W}^{-1}_{k-1}]_{ji} \\
&= [I_{n_s} - (\hat{P}_{*|k-1}-\hat{Q}_{*,k})\hat{H}^{\mathbf{T}}_{k,*}[\tilde{W}^{-1}_k]_{kk}\hat{H}_{k,*}]K_{i|k-1}, = (I_{n_s} - K_{k|k}\hat{H}_{k,*})K_{i|k-1} \qquad (112)
\end{aligned}$$

recognizing equation (86) in the last line. This is the desired relation.

REFERENCES

[1] F. L. Markley and J. R. Carpenter, "Generalized Linear Covariance Analysis," *Journal of the Astronautical Sciences*, Vol. 57, No. 1,2, Jan–Jun 2009, pp. 233–260.

[2] F. L. Markley, E. Seidewitz, and M. Nicholson, "A General Model for Attitude Determination Error Analysis," *NASA Conference Publication 3011: Flight Mechanics/Estimation Theory Symposium*, May 1988, pp. 3–25.

[3] F. L. Markley, E. Seidewitz, and J. Deutschmann, "Attitude Determination Error Analysis: General Model and Specific Application," *Proceedings of the CNES Space Dynamics Conference*, Toulouse, France, November 1989, pp. 251–266.

[4] K. F. Gauss, *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections*, Dover, New York, 1963.

[5] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME – Journal of Basic Engineering*, Vol. 82, 1960, pp. 35–45.

[6] R. G. Brown and P. Y. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, John Wiley and Sons, Inc., New York, NY, 3rd ed., 1997.

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

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

[9] S. McReynolds, "Covariance Projection Methods for Estimators," *AIAA Guidance, Navigation and Control Conference*, No. A9635623, AIAA Paper 96–3739 in AIAA Meeting Papers on Disk, American Institute of Aeronautics and Astronautics, Inc., July 1996.

[10] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, Vol. 1, Academic Press, New York, 1979.

[11] J. R. Carpenter, F. L. Markley, K. T. Alfriend, C. Wright, and J. Arcido, "Sequential Probability Ratio Test for Collision Avoidance Maneuver Decisions Based on a Bank of Norm-Inequality-Constrained Epoch-State Filters," *AAS/AIAA Astrodynamics Specialist Conference*, No. AAS 11-437, Alyeska, AK, August 2011.

[12] R. Zanetti, M. Majii, R. H. Bishop, and D. Mortari, "Norm-Constrained Kalman Filtering," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 5, September–October 2009, pp. 1458–1465.

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

[14] O. Montenbruck, "An Epoch State Filter for Use with Analytical Orbit Models of Low Earth Satellites," *Aerosp. Sci. Technol.*, Vol. 4, 2000, pp. 277–287.