Error Statistics for Bias-Naïve Filtering in the Presence of Bias

Zachary Chance, Stephen Relyea, and Evan Anderson

MIT Lincoln Laboratory, 244 Wood Street, Lexington, MA 02421

ABSTRACT

In the field of sensing, a typically unavoidable nuisance is the inherent bias of a sensor due to imperfections in timing, calibration, and other sources. The errors incurred by the bias ripple through higher-level processes such as tracking and sensor fusion, causing varying effects to each operation. In many different applications, such as track-to-track correlation, the overall effect of the biases on state estimation is modeled as a constant, translational shift in the position dimension of the track states. This assumption can be appropriate when the required precision of the track states is not stringent. However, in general, sensor bias can not only affect position estimates but also positional derivatives, i.e., velocity, acceleration, in a manner that can change dramatically depending on sensor-target geometry; for situations where high state estimation accuracy is required, these consequences become apparent and need to be handled.

The contribution from measurement bias to state estimation error depends on many different aspects, e.g., measurement uncertainty, dynamic model uncertainty, sensor-target geometry. The focus of this work is the quantification of the relative significance of measurement error and measurement bias in the resultant state estimation error. In short, using the results in this work, it is straightforward to: (i) determine regimes where measurement bias becomes a predominant factor, (ii) bound the impact of the sensor bias on the outputted tracking information, (iii) analyze the dependence of the tracking error on sensor-target geometry, all of which can be of great impact when designing a tracking system architecture.

Keywords: state estimation, tracking, bias, multisensor

1. INTRODUCTION

The error of a state estimator contains contributions from many different mechanisms, e.g., measurement noise, measurement bias, modeling mismatch, and sensor-target geometry. The relative importance of these error sources can vary dramatically from one application to another. For instance, in circumstances with low signal-to-noise ratio, errors may be dominated by measurement noise with relatively negligible impact from measurement bias or model. To address this disparity, a candidate state estimator may require tailoring by adding or removing nuisance parameters to the state to achieve a desired level of precision and accuracy. Thus, it is of primary interest to quantify the relative significance of the constituent factors to state estimation error. The main sources of error considered in this work are process noise, measurement noise, and measurement bias. The objective is to analyze the relative contribution of measurement bias to resultant state estimation error and, subsequently, determine when it is necessary to account for measurement bias during state estimation.

Multisensor, multitarget tracking is a practical application for which error analysis is of particular interest, as successfully combining tracking data from a network of sensors hinges on the ability to characterize the errors of each sensor. More specifically, multisensor tracking architectures at a fusion node are commonly constructed by combining the outputs of a set of bias-naïve (i.e., do not account for measurement bias) multitarget trackers. Upon receiving track states from each sensor, the fusion node *does* consider the effects of measurement bias for each sensor by assuming that the track states from each sensor are shifted in position by a fixed amount that is different

DISTRIBUTION STATEMENT A. Approved for public release. Distribution is unlimited. This material is based upon work supported under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Air Force. Further author information: (Send correspondence to Z.T.C.)
Z.T.C.: E-mail: zachary.chance@ll.mit.edu, Telephone: (781) 981-1993

Signal Processing, Sensor/Information Fusion, and Target Recognition XXVII, edited by Ivan Kadar,
Proc. of SPIE Vol. 10646, 106461K · © 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2303765

from sensor to sensor; this assumption is crucial in many track-to-track association[1–6] algorithms. After track-to-track association is completed, the estimates from each sensor are fused and outputted. If the assumption that all track states are a constant translational shift from the truth is violated, track-to-track correlation may have issues and/or the fused track states might be inconsistent; thus, it is imperative to understand how measurement bias affects positional derivative (i.e., velocity, acceleration) state estimates.

The effects of measurement bias on state estimation error have been studied in varying degrees across a collection of other works. An initial discussion of measurement bias in state estimation was raised in the context of performance analysis[7,8] for suboptimal, reduced order filters. Bias-naïve estimation can be seen as an instance of reduced order filtering in which the filter state has fewer dimensions than the true dynamic state. Further, the implications of measurement and dynamic model mismatch on the estimation error covariance were considered as part of sensitivity analysis[9,10] for optimal linear filtering and smoothing. These analyses produced expressions for the error covariance as functions of the difference between the true and assumed models. Likewise, error covariance analysis[11,12] in the presence of measurement biases was performed, illustrating the evolution of state estimation error covariance from measurement to measurement. In this work, a narrower scope is assumed than previous related treatments. This allows for discussion of both linear and non-linear systems, deterministic and stochastic biases, and simpler alternative derivations for estimation error moments. Additionally, a measure of significance of the measurement bias on the eventual estimation error is defined and illustrated.

The paper is organized using the following structure. Section 2 provides the mathematical background for a bias-naïve state estimation technique, the definitions of the error statistics of interest, and their subsequent expressions in the context of linear systems. Building upon these results, Section 3 extends the error statistic expressions to the class of nonlinear systems using suboptimal filtering and linearization techniques. To measure the impact of bias on the estimation error, a bias significance measure is introduced in Section 4. Finally, a summary of the work and results is presented in Section 5.

2. BIAS-NAÏVE STATE ESTIMATION FOR LINEAR SYSTEMS

To understand the consequences of measurement bias on the output of state estimation, one can analyze the error statistics of an estimator that assumes there is no measurement bias but, in reality, only has access to biased measurements. To begin, consider a sensor observing a target whose dynamics can be described using the time-varying state, $\mathbf{x}_k$, where $k = 1, 2, \dots$ is a discrete time index. The sensor takes measurements of the target state at time instances, t_k, such that the measurement at time t_k can be written as

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{G}_k \mathbf{b}_k + \mathbf{w}_k, \tag{1}$$

where $\mathbf{H}_k$ is a known measurement matrix, $\mathbf{G}_k$ is a known mapping from bias space to measurement space, $\mathbf{b}_k$ is a time-varying measurement bias, and $\mathbf{w}_k$ is additive measurement noise that is distributed as $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$. Accordingly, the target state evolves according to a discrete-time dynamic model given by

$$\mathbf{x}_k = \mathbf{\Phi}_{k,k-1} \mathbf{x}_{k-1} + \mathbf{v}_k, \tag{2}$$

where $\mathbf{\Phi}_{k,k-1}$ is the state transition matrix from time t_{k-1} to time t_k, and $\mathbf{v}_k$ is additive process noise distributed as $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$. The distribution of the initial state, $\mathbf{x}_0$, is Gaussian and known. Both the measurement noise, $\mathbf{w}_k$, and process noise, $\mathbf{v}_k$, are independent from time instance to time instance and of each other. Presently, no assumptions are made on the form of the measurement bias, $\mathbf{b}_k$.

To form the bias-naïve estimator, it is assumed that the bias term in (1) vanishes, i.e., $\mathbf{G}_k \mathbf{b}_k = \mathbf{0}$. Then, the system described in (1) and (2) is linear with Gaussian noise inputs. With this setup, the optimal minimum-mean-square-error estimator is the Kalman filter[13] in which the state estimate, $\hat{\mathbf{x}}_{k\,|\,k}$, and state estimate covariance,

$\mathbf{P}_{k\,|\,k}$, evolve according to the following:

$$\hat{\mathbf{x}}_{k\,|\,k-1} = \mathbf{\Phi}_{k,k-1}\hat{\mathbf{x}}_{k-1\,|\,k-1}, \tag{3a}$$

$$\mathbf{P}_{k\,|\,k-1} = \mathbf{\Phi}_{k,k-1}\mathbf{P}_{k-1\,|\,k-1}\mathbf{\Phi}_{k,k-1}^{T} + \mathbf{Q}_k, \tag{3b}$$

$$\hat{\mathbf{x}}_{k\,|\,k} = \hat{\mathbf{x}}_{k\,|\,k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k\,|\,k-1}\right), \tag{3c}$$

$$\mathbf{P}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k\,|\,k-1}, \tag{3d}$$

$$\mathbf{K}_k = \mathbf{P}_{k\,|\,k-1}\mathbf{H}_k^{T}\left(\mathbf{H}_k\mathbf{P}_{k\,|\,k-1}\mathbf{H}_k^{T} + \mathbf{R}_k\right)^{-1}. \tag{3e}$$

Because of the strict linear structure of the system in (1)-(2), the errors of the estimator in (3) will be a linear function of the bias. In the rest of this section, the linear projection of the bias history into the error mean and its subsequent effect on the covariance is derived.

Before continuing, it is now pertinent to introduce the two error moments that are to be studied. More specifically, the primary concern is the effect that the measurement bias, $\mathbf{b}_k$, plays in the mean and covariance of the estimation error; these are defined respectively as

$$\mathbf{m}_k \triangleq E[\hat{\mathbf{x}}_{k\,|\,k} - \mathbf{x}_k], \tag{4}$$

$$\mathbf{C}_k \triangleq E\left[\left(\hat{\mathbf{x}}_{k\,|\,k} - \mathbf{x}_k - \mathbf{m}_k\right)\left(\hat{\mathbf{x}}_{k\,|\,k} - \mathbf{x}_k - \mathbf{m}_k\right)^{T}\right]. \tag{5}$$

Note that the behavior of these quantities will change depending on linearity of the system, sensor-target geometry, time variation of the measurement bias, and other factors. Because of this, multiple circumstances will be considered. Initially, the measurement bias is assumed to be an unknown, deterministic value, and estimation error for linear systems are considered.

2.1 Unknown, Deterministic Bias

To begin, it is of interest to investigate the estimation error statistics for an unknown, deterministic measurement bias. This provides insight into the consequences on estimation as a function of an unknown measurement bias trajectory. This is important to understand as the performance of the estimator can differ appreciably even within small variations of measurement bias trajectories. The following theorem provides the estimation error statistics that explicitly rely on the measurement bias.

Theorem 1. *The estimation error mean and covariance with an unknown, deterministic bias,* $\mathbf{b}_k$*, can be written for a linear system as*

$$\mathbf{m}_k = \sum_{i=1}^{k}\mathbf{\Lambda}_i^k\mathbf{K}_i\mathbf{G}_i\mathbf{b}_i, \tag{6}$$

$$\mathbf{C}_k = \mathbf{P}_{k\,|\,k}, \tag{7}$$

where

$$\mathbf{\Lambda}_i^k = \begin{cases} \displaystyle\prod_{j=i+1}^{k}\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}_j\right)\mathbf{\Phi}_{j,j-1}, & i < k \\ \mathbf{I}, & i = k. \end{cases}$$

Additionally, the mean and covariance obey the following recursive relations:

$$\mathbf{m}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{\Phi}_{k,k-1}\mathbf{m}_{k-1} + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k,$$

$$\mathbf{C}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\mathbf{\Phi}_{k,k-1}\mathbf{C}_{k-1}\mathbf{\Phi}_{k,k-1}^{T} + \mathbf{Q}_k\right).$$

The proof of Theorem 1 is given in Appendix A.

It is important to note that the mean of the state estimation error in (6) is now dependent on the measurement bias; in fact, the expression in (6) can be viewed as a weighted projection of the measurement bias history into the space of target dynamics. It is also of value to point out that the recursive relationships for the error mean and covariance posed in Theorem 1 agree with expressions in previously published[9] results.

A case of particular interest is a constant bias, which is a common assumption in many bias estimation and track-to-track correlation algorithms. This is a popular assumption as it simplifies system model definition, and, commonly, measurement biases vary extremely slowly in comparison to the rate of observation. To address a constant bias, the following corollary gives the error statistics for a bias that remains the same over the duration of target observation.

Corollary 1. *For a constant bias, i.e.,* $\mathbf{b}_k = \mathbf{b}$*, the estimation error mean and covariance given in Theorem 1 can be simplified as*

$$\begin{aligned} \mathbf{m}_k &= \mathbf{S}_k \mathbf{b}, \\ \mathbf{C}_k &= \mathbf{P}_{k\,|\,k}, \end{aligned} \tag{8}$$

where

$$\mathbf{S}_k = \sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i, \tag{9}$$

Further, the matrices $\mathbf{S}_k$ *satisfy the following:*

$$\mathbf{S}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\, \mathbf{\Phi}_{k,k-1} \mathbf{S}_{k-1} + \mathbf{K}_k.$$

Proof. In the case of a constant bias, the mean can be simplified as

$$\begin{aligned} \mathbf{m}_k &= \sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i \mathbf{b}_i, \\ &= \left(\sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i \right) \mathbf{b}, \\ &= \mathbf{S}_k \mathbf{b}. \end{aligned}$$

As the bias does not play a role in the covariance, $\mathbf{P}_{k\,|\,k}$, the proof is concluded.

The matrix $\mathbf{S}_k$ can be interpreted as the projection of the constant bias vector, $\mathbf{b}$, into the state estimate at time t_k. From the form given in (9), $\mathbf{S}_k$ is a weighted combination of the Kalman gains used at each measurement in the observation history.

To illustrate the distribution of state estimation errors with a biased sensor, an example is now shown using a sensor that observes a target undergoing constant velocity motion. The target state, $\mathbf{x}_k$, is explicitly defined as

$$\mathbf{x}_k = \begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix},$$

where x, y, and z are the position coordinates of the target in a Cartesian coordinate system, and $\dot{x}$, $\dot{y}$, $\dot{z}$ are the respective time derivatives, i.e., velocities. In this case, the state transition matrix, $\mathbf{\Phi}_{k,k-1}$, is given by

$$\mathbf{\Phi}_{k,k-1} = \begin{bmatrix} 1 & 0 & 0 & \Delta_k & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta_k & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta_k \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where $\Delta_k = t_k - t_{k-1}$. The sensor makes position measurements every 0.5 sec directly in the Cartesian space in which the target dynamics are defined. The measurement error standard deviations are 1 m in the x dimension, 100 m in the y dimension, and 100 m in the z dimension. The measurement matrix, $\mathbf{H}_k$, can be written as

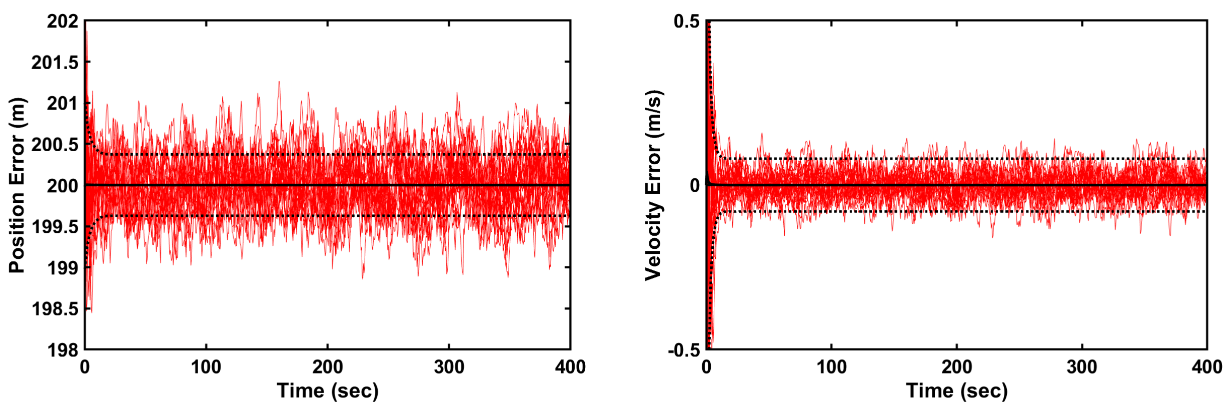
$$\mathbf{H}_k = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Note that the form of $\mathbf{H}_k$ implies that only the position dimension of the target state is observed. Also, the measurement matrix is constant with respect to time; later, this will be seen to have consequences in how the measurement bias affects the estimation error. The measurement bias, $\mathbf{b}_k$, is a constant bias injected directly (i.e., $\mathbf{G}_k = \mathbf{I}$) with shifts of 200 m, 50 m, and -50 m in the x, y, and z dimensions, respectively. The process noise covariance, $\mathbf{Q}_k$, is of the form[13]

$$\mathbf{Q}_k = q \begin{bmatrix} \frac{1}{3}\Delta_k^3 \mathbf{I} & \frac{1}{2}\Delta_k^2 \mathbf{I} \\ \frac{1}{2}\Delta_k^2 \mathbf{I} & \Delta_k \mathbf{I} \end{bmatrix},$$

where $q = 0.001$ m^2/s^3 is the power spectral density of the process noise.

Figure 1 shows the state estimation error resulting from running a Kalman filter on 20 Monte Carlo trials of random measurements. Additionally, the mean and standard deviation of the error as described in (6) and (7) are also shown. As can be observed, the mean and covariance in (6) and (7) match well with the distribution of the errors from the different Monte Carlo trials. Further, the effect of the bias can be clearly observed in the position errors of the state estimates; due to the constant, linear dynamics and measurements, the velocity errors are not evidently affected by the constant position bias. This will notably change in the presence of nonlinear and/or time-dependent measurement mappings.



(a) Position error in x-dimension. (b) Velocity error in x-dimension.

Figure 1: State estimation error (for 20 Monte Carlo runs) in the x dimension of a Cartesian coordinate system. Solid black lines show the mean, $\mathbf{m}_k$ from (6), and dashed black lines show the standard deviation in that dimension, obtained from the covariance $\mathbf{C}_k$ in (7).

2.2 Stochastic Bias

Up until now, the bias was assumed to be an unknown, deterministic quantity; thus, the error statistics can be seen as a function of a variable bias. However, it would also be of interest to derive error statistics when the bias can be described by a stochastic process. An appreciable distinction to the unknown, deterministic bias case is that the expected value used in the derivation of the error moments also includes the probability distribution of the bias. The following theorem produces the error statistics for a stochastic bias:

Theorem 2. *Let the bias,* $\mathbf{b}_k$*, be a wide-sense-stationary Gaussian stochastic process*[14] *with constant mean,* $E[\mathbf{b}_k] = \mathbf{u}$*, and autocovariance given for all* k *as* $\mathbf{V}_\ell = E\left[\mathbf{b}_k \mathbf{b}_{k+\ell}^T\right]$*. Also, it is assumed that the bias is drawn* *independently of the measurement noise process,* $\mathbf{w}_k$*. Then, the estimation error mean and covariance for a linear system can be expressed as*

$$\mathbf{m}_k = \mathbf{S}_k \mathbf{u}, \tag{10}$$
$$\mathbf{C}_k = \mathbf{B}_k + \mathbf{P}_{k\,|\,k}, \tag{11}$$

where $\mathbf{S}_k$ *is as defined in Corollary 1 and*

$$\mathbf{B}_k = \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i \mathbf{V}_{j-i} \left(\mathbf{\Lambda}_j^k \mathbf{K}_j \mathbf{G}_j \right)^T. \tag{12}$$

Proof. In this case, the mean of the error statistics can now be written as

$$\begin{aligned} \mathbf{m}_k &= E\left[\sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i \mathbf{b}_i \right], \\ &= \sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i E\left[\mathbf{b}_i\right], \\ &= \mathbf{S}_k \mathbf{u}. \end{aligned}$$

Now, the estimation error covariance can be derived as

$$\begin{aligned} \mathbf{C}_k &= E\left[\left(\left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} \right) + \left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k \right) \right) \left(\left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} \right) + \left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k \right) \right)^T \right], \\ &= E\left[\left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} \right) \left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} \right)^T \right] + \mathbf{P}_{k\,|\,k}, \end{aligned} \tag{13}$$

due to the independence of the measurement noise, $\mathbf{w}_k$, and bias, $\mathbf{b}_k$. Plugging (35) from Appendix A into (13), the result in (11) is obtained. □

The matrix, $\mathbf{B}_k$, is the effective inflation in covariance due to the presence of bias; for largely varying or highly correlated biases, this term can become the predominant factor in the total error covariance.

As with an unknown, deterministic bias, the error statistics for a stochastic bias can be simplified in the case of a constant bias; this is given in the following corollary:

Corollary 2. *If the bias is also constant, i.e.,* $\mathbf{b}_k = \mathbf{b}$*, the estimation error mean and covariance expressions from Theorem 2 can be simplified as*

$$\begin{aligned} \mathbf{m}_k &= \mathbf{S}_k \mathbf{u}, \\ \mathbf{C}_k &= \mathbf{S}_k \mathbf{V} \mathbf{S}_k^T + \mathbf{P}_{k\,|\,k}, \end{aligned} \tag{14}$$

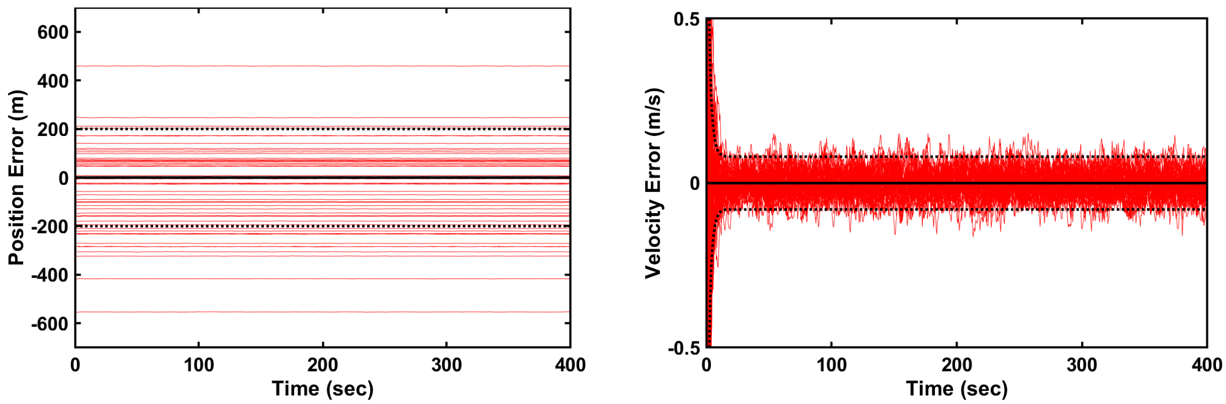
where $\mathbf{S}_k$ *is as defined in Corollary 1 and* $\mathbf{V} = E\left[\mathbf{b}\mathbf{b}^T\right]$.

Proof. In the case of a constant bias, the autocovariance function, $\mathbf{V}_\ell$, no longer depends on the lag parameter, ℓ. Thus, letting $\mathbf{V}_\ell = \mathbf{V}$ in the expression for $\mathbf{B}_k$ in (12), one obtains the result in (14). □

To illustrate the effects of a random bias on the error statistics, the example in Figure 1 is now briefly revisited. Now, for each Monte Carlo trial, the bias is drawn from a known distribution. In Figure 2, the effect of a random bias on state estimation error is shown for the linear system presented originally in Figure 1. For this case, the bias is constant and drawn from a Gaussian distribution with covariance,

$$\mathbf{V} = \begin{bmatrix} 200^2 & 0 & 0 \\ 0 & 50^2 & 0 \\ 0 & 0 & 50^2 \end{bmatrix},$$

where the dimensions are x (meters), y (meters), z (meters), respectively. It can be observed in Figure 2 that the random bias manifests itself as a random positional shift that remains constant over time. Further, the velocity error in Figure 2b is not affected by a random bias as it appears to have the same behavior as in Figure 1b; intuitively, this stems from the net effect of the bias being a constant translational shift over time, which does not alter positional derivatives, e.g., velocity, acceleration.



(a) Position error in x-dimension. (b) Velocity error in x-dimension.

Figure 2: State estimation error (for 50 Monte Carlo runs) in the x dimension of a Cartesian coordinate system. Solid black lines show the mean, $\mathbf{m}_k$ from (10), and dashed black lines show the standard deviation in that dimension, obtained from the covariance $\mathbf{C}_k$ in (7).

3. BIAS-NAÏVE STATE ESTIMATION FOR NONLINEAR SYSTEMS

Many practical systems can only be described using nonlinear functions for the evolution of the state or the mapping from the state space to the measurement space. Due to the nonlinearities, many interesting phenomena can arise in how measurement bias manifests itself within the total estimation error. Accordingly, the focus of this section is to extend the results of linear systems provided in Section 2 to suboptimal filtering for nonlinear systems. The measurements of the target state are now expressed as

$$\mathbf{z}_k = h_k\left(\mathbf{x}_k\right) + \mathbf{G}_k \mathbf{b}_k + \mathbf{w}_k, \tag{15}$$

where $h_k(\cdot)$ is a known measurement function, $\mathbf{G}_k$ is a known mapping from bias space to measurement space, $\mathbf{b}_k$ is a time-varying measurement bias, and $\mathbf{w}_k$ is additive measurement noise which is distributed as $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$. Likewise, the target state evolves according to a discrete-time dynamic model given by

$$\mathbf{x}_k = \phi_{k,k-1}\left(\mathbf{x}_{k-1}\right) + \mathbf{v}_k, \tag{16}$$

where $\phi_{k,k-1}(\cdot)$ is the state transition function from time t_{k-1} to time t_k, and $\mathbf{v}_k$ is additive process noise distributed as $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$. The distribution of the initial state, $\mathbf{x}_0$, is Gaussian and known. Both the measurement noise, $\mathbf{w}_k$, and process noise, $\mathbf{v}_k$, are independent from time instance to time instance and of each other.

As with the linear system, the bias-naïve estimator is constructed by assuming that the bias term in (15) is negligible, i.e., $\mathbf{G}_k \mathbf{b}_k = \mathbf{0}$. In this case, the system described in (15) and (16) is nonlinear with Gaussian noise inputs. Due to its ubiquity, the suboptimal estimator considered in this paper is the extended Kalman filter[13] (EKF); the state estimate, $\hat{\mathbf{x}}_{k\,|\,k}$, and state estimate covariance, $\mathbf{P}_{k\,|\,k}$, of the bias-naïve estimator follow the equations

$$\hat{\mathbf{x}}_{k\,|\,k-1} = \phi_{k,k-1}\left(\hat{\mathbf{x}}_{k-1\,|\,k-1}\right), \tag{17a}$$

$$\mathbf{P}_{k\,|\,k-1} = \mathbf{\Phi}_{k,k-1}\mathbf{P}_{k-1\,|\,k-1}\mathbf{\Phi}_{k,k-1}^{T} + \mathbf{Q}_k, \tag{17b}$$

$$\hat{\mathbf{x}}_{k\,|\,k} = \hat{\mathbf{x}}_{k\,|\,k-1} + \mathbf{K}_k\left(\mathbf{z}_k - h_k\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right)\right), \tag{17c}$$

$$\mathbf{P}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_{k\,|\,k-1}, \tag{17d}$$

$$\mathbf{K}_k = \mathbf{P}_{k\,|\,k-1}\mathbf{H}_k^{T}\left(\mathbf{H}_k\mathbf{P}_{k\,|\,k-1}\mathbf{H}_k^{T} + \mathbf{R}_k\right)^{-1}, \tag{17e}$$

where

$$\mathbf{\Phi}_{k,k-1} = \left.\frac{\partial}{\partial\mathbf{x}}\phi_{k,k-1}\left(\mathbf{x}\right)\right|_{\mathbf{x}\,=\,\hat{\mathbf{x}}_{k-1\,|\,k-1}}, \tag{18}$$

$$\mathbf{H}_k = \left.\frac{\partial}{\partial\mathbf{x}}h_k\left(\mathbf{x}\right)\right|_{\mathbf{x}\,=\,\hat{\mathbf{x}}_{k\,|\,k-1}}. \tag{19}$$

The matrices in (18) and (19) are the Jacobians for the state transition and measurement functions, respectively.

The Jacobians for the state transition and measurement functions play a key role in the results to follow. In fact, the error statistics for nonlinear systems will appear identical to that of linear systems; however, the ability to express the error distributions as linear functions of the bias requires first order Taylor series approximations—a common approach to dealing with system nonlinearities. These approximations manifest themselves by replacing the measurement matrix, $\mathbf{H}_k$, and state transition matrix, $\mathbf{\Phi}_{k,k-1}$, from linear systems with their Jacobian counterparts.

Something else worth noting is the choice of Jacobian used for the approximation. A critical degree of freedom when using linearization techniques is the point around which a function is linearized. Some commonsense choices for this application include the current unbiased track state, the current biased track state, the true state, or a mixture of the three. Alternatively, one could also use different linearization points for each instance of a nonlinear mapping within the error derivation. To simplify the results and allow for possible online estimation of error statistics, the current biased track state is used for the point of linearization. For systems with extremely large biases and/or highly nonlinear functions, this choice may prove inadequate and, possibly, higher order approximations are necessary. However, the results shown later demonstrate the efficacy of the chosen linearization as the error statistics derived predict the true error behavior very closely.

To see how linearization is incorporated into the previous results for linear systems, the error statistics for an unknown, deterministic bias in a nonlinear system are now given.

Theorem 3. *The estimation error mean and covariance with an unknown, deterministic bias, $\mathbf{b}_k$, can be approximated for a nonlinear system as*

$$\mathbf{m}_k = \sum_{i=1}^{k}\mathbf{\Lambda}_i^k\mathbf{K}_i\mathbf{G}_i\mathbf{b}_i, \tag{20}$$

$$\mathbf{C}_k = \mathbf{P}_{k\,|\,k}, \tag{21}$$

where

$$\mathbf{\Lambda}_i^k = \begin{cases} \prod_{j=i+1}^{k}\left(\mathbf{I} - \mathbf{K}_j\mathbf{H}_j\right)\mathbf{\Phi}_{j,j-1}, & i < k \\ \mathbf{I}, & i = k, \end{cases}$$

and $\mathbf{\Phi}_{k,k-1}$ and $\mathbf{H}_k$ are defined in (18) and (19), respectively. Additionally, the mean and covariance obey the following recursive relations:

$$\mathbf{m}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{\Phi}_{k,k-1}\mathbf{m}_{k-1} + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k,$$

$$\mathbf{C}_k = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\mathbf{\Phi}_{k,k-1}\mathbf{C}_{k-1}\mathbf{\Phi}_{k,k-1}^{T} + \mathbf{Q}_k\right).$$

The proof of Theorem 3 is given in Appendix B. Simplification of the error statistic expressions from a time-invariant bias is described in the follow corollary.

Corollary 3. *If the bias is also constant, i.e.,* $\mathbf{b}_k = \mathbf{b}$*, the estimation error mean and covariance given in Theorem 3 can be simplified as*

$$\begin{aligned}\mathbf{m}_k &= \mathbf{S}_k \mathbf{b}, \\ \mathbf{C}_k &= \mathbf{P}_{k\,|\,k},\end{aligned} \tag{22}$$

where

$$\mathbf{S}_k = \sum_{i=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i,$$

and $\mathbf{\Lambda}_i^k$ *is as defined in Theorem 3. Further, the matrices,* $\mathbf{S}_k$*, satisfy the following:*

$$\mathbf{S}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\, \mathbf{\Phi}_{k,k-1} \mathbf{S}_{k-1} + \mathbf{K}_k.$$

Proof. Result follows from proof of Corollary 1.

To demonstrate the behavior of state estimation error using a biased sensor with nonlinear measurements, an example is now given. More specifically, a sensor with measurements similar to a phased array radar[15] is used to estimate the state of a ballistic target. The state transition function, $\phi_{k,k-1}(\mathbf{x}_{k-1})$, is obtained by numerically integrating a ballistic model[16] previously provided. The measurements are taken in range-projection space (also known as RUV space) with standard deviations of 1 m, 1 msin, and 1 msin in the range, u, and v dimensions, respectively. The measurement function (conversion to range-projection space) can be summarized as

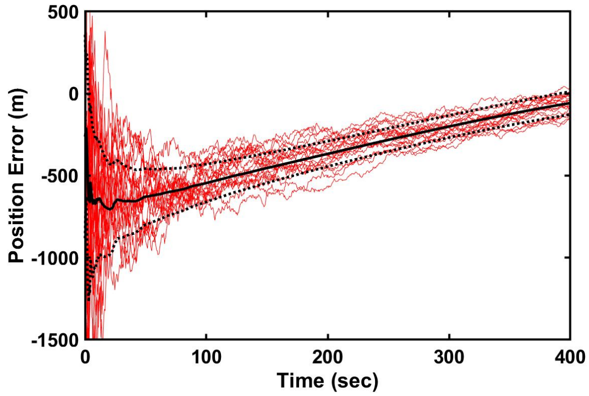
$$h_k\left(\mathbf{x}_k\right) = \begin{bmatrix} \sqrt{x^2 + y^2 + z^2} \\ \frac{x}{\sqrt{x^2+y^2+z^2}} \\ \frac{y}{\sqrt{x^2+y^2+z^2}} \end{bmatrix},$$

where x, y, and z are components in a sensor-centered Cartesian coordinate system, and the output dimensions are the range, u, and v dimensions, respectively. The measurement bias is constant (i.e., $\mathbf{G}_k = \mathbf{I}$) with shifts of 10 m, 500 usin, and -500 usin in the range, u, and v dimensions, respectively. The process noise covariance is defined by the continuous time process noise covariance, $\mathbf{Q}(t)$, given as

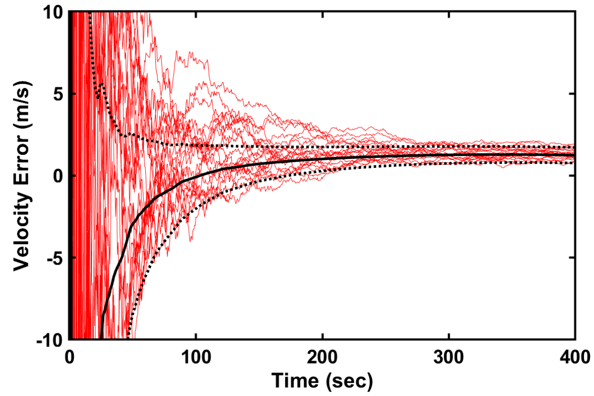
$$\mathbf{Q}(t) = q \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where, as in the linear system, $q = 0.001$ m^2/s^3 is the power spectral density of the process noise. The state estimate covariance is now obtained via numerical integration from time step to time step.

Similar to the previous example with a linear system, Figure 3 shows the behavior of the state estimation errors using bias-naïve filtering. It can be seen that despite the measurement and dynamics both being described using nonlinear functions, the error statistics still follow closely the mean and covariance expressions in (6) and (7). However, there is one main difference from the linear case—the mean of the error in both position and velocity changes with time even though the bias that is injected is constant. This is due to the varying effect of the nonlinear measurement mapping depending on sensor-target geometry.



(a) Position error in x-dimension.



(b) Velocity error in x-dimension.

Figure 3: State estimation error (for 20 Monte Carlo runs) in the x dimension of Earth-centered rotating (ECR) coordinate system. Solid black lines show the mean, $\mathbf{m}_k$ from (20), and dashed black lines show the standard deviation in that dimension, obtained from the covariance $\mathbf{C}_k$ in (21).

3.1 Stochastic Bias

Similar to the discussion on linear systems, the error statistics for a stochastic bias in a nonlinear system are now studied.

Theorem 4. *Let the bias, $\mathbf{b}_k$, be a wide-sense-stationary Gaussian stochastic process[14] with constant mean, $E[\mathbf{b}_k] = \mathbf{u}$, and autocovariance given for all k as $\mathbf{V}_\ell = E\left[\mathbf{b}_k \mathbf{b}_{k+\ell}^T\right]$. Also, it is assumed that the bias is drawn independently of the measurement noise process, $\mathbf{w}_k$. Then, the estimation error mean and covariance for a nonlinear system can be approximated as*

$$\mathbf{m}_k = \mathbf{S}_k \mathbf{u}, \tag{23}$$

$$\mathbf{C}_k = \mathbf{B}_k + \mathbf{P}_{k\,|\,k}, \tag{24}$$

where $\mathbf{\Lambda}_i^k$ is as defined in Theorem 3 and

$$\mathbf{B}_k = \sum_{i=1}^{k} \sum_{j=1}^{k} \mathbf{\Lambda}_i^k \mathbf{K}_i \mathbf{G}_i \mathbf{V}_{j-i} \left(\mathbf{\Lambda}_j^k \mathbf{K}_j \mathbf{G}_j \right)^T .$$

Proof. Result follows from proof of Theorem 2.

As with an unknown, deterministic bias, the error statistics for a stochastic bias can be simplified in the case of a constant bias; this is given in the following corollary:

Corollary 4. *If the bias is also constant, i.e., $\mathbf{b}_k = \mathbf{b}$, the estimation error mean and covariance given in Theorem 4 can be simplified as*

$$\mathbf{m}_k = \mathbf{S}_k \mathbf{u},$$

$$\mathbf{C}_k = \mathbf{S}_k \mathbf{V} \mathbf{S}_k^T + \mathbf{P}_{k\,|\,k}, \tag{25}$$

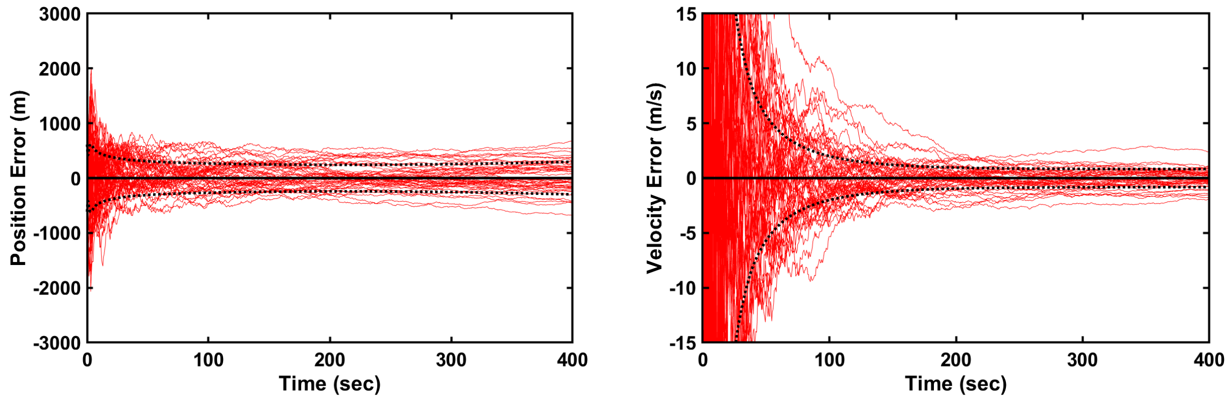
where $\mathbf{S}_k$ is as defined in Corollary 3 and $\mathbf{V} = E\left[\mathbf{b}\mathbf{b}^T\right]$.

Proof. Result follows from proof of Corollary 2.

In Figure 4, a random bias is now injected into the nonlinear system (i.e., ballistic motion, ranging measurements) used for the results in Figure 3. For this example, the bias is drawn from a Gaussian with covariance

$$\mathbf{V} = \begin{bmatrix} 10^2 & 0 & 0 \\ 0 & 0.00025^2 & 0 \\ 0 & 0 & 0.00025^2 \end{bmatrix},$$

where the dimensions are range (meters), u (sin), v (sin), respectively. The behavior of the state estimation error is markedly different from the error trajectories in Figure 3. In Figure 4, the error covariance is dominated by the inclusion of the bias term, $\mathbf{S}_k \mathbf{V} \mathbf{S}_k^T$.



(a) Position error in x-dimension.

(b) Velocity error in x-dimension.

Figure 4: State estimation error (for 50 Monte Carlo runs) in the x dimension of Earth-centered rotating (ECR) coordinate system. Solid black lines show the mean, $\mathbf{m}_k$ from (23), and dashed black lines show the standard deviation in that dimension, obtained from the covariance $\mathbf{C}_k$ in (21).

4. BIAS SIGNIFICANCE

As mentioned in the introduction, it is of primary interest to develop a measure of the relative impact that measurement bias plays in the resultant estimation error. To address this, the concept of bias significance is now given. Let the total bias significance, λ_k, of an estimator for an unknown, deterministic bias be defined as the Mahalanobis distance[17] between the assumed error distribution and the true error distribution. The Mahalanobis distance, d, between two distributions with means $\mathbf{m}_a$ and $\mathbf{m}_b$ and covariances $\mathbf{C}_a$ and $\mathbf{C}_b$ is defined as

$$d = \sqrt{\left(\mathbf{m}_a - \mathbf{m}_b\right)^T \left(\mathbf{C}_a + \mathbf{C}_b\right)^{-1} \left(\mathbf{m}_a - \mathbf{m}_b\right)}.$$

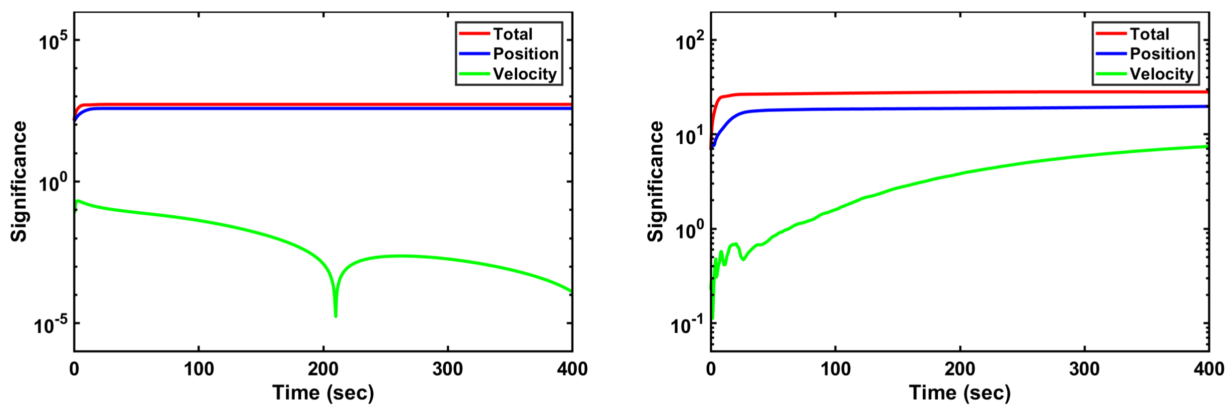
Now, the assumed estimation error distribution for the bias-naïve estimator is $\mathcal{N}\left(\mathbf{0}, \mathbf{P}_{k\,|\,k}\right)$, while its true error distribution is $\mathcal{N}\left(\mathbf{m}_k, \mathbf{C}_k\right)$ from (6) and (7). With those distributions, the total bias significance at time k can be calculated as

$$\begin{aligned} \lambda_k &= \sqrt{\left(\mathbf{m}_k - \mathbf{0}\right)^T \left(\mathbf{P}_{k\,|\,k} + \mathbf{P}_{k\,|\,k}\right)^{-1} \left(\mathbf{m}_k - \mathbf{0}\right)}, \\ &= \sqrt{\frac{\mathbf{m}_k^T \mathbf{P}_{k\,|\,k}^{-1} \mathbf{m}_k}{2}}. \end{aligned} \tag{26}$$

Note that if the bias is identically zero, then, using (6), the true estimation error mean is zero, $\mathbf{m}_k = \mathbf{0}$, and the bias significance is zero, $\lambda_k = 0$. Further, one can separate the contributions from state components by looking at the significance in a subset of state dimensions, e.g., position, velocity.

In Figure 5, the bias significance for the linear and nonlinear systems presented in Sections 2 and 3 are given. For the linear system, the total significance is high, meaning that the presence of bias causes a large deviation in the expected error distribution. However, it can be seen by analyzing the significance in position and velocity separately that the position plays the dominant role in the total bias significance. This is due to the fact that the bias is constant and injected into the position dimensions of the measurement. A negligible amount of velocity error is due to cross-correlation between position and velocity in the initial state estimate. Looking at the nonlinear system, a similar trend can be observed. The total significance in Figure 5b is large due to the high position error incurred by the bias. However, a major difference between the linear and nonlinear systems is that the velocity significance is no longer negligible in the nonlinear system. This is due to the velocity errors incurred by the fact that, even though the measurement bias is constant, the measurement Jacobian changes over the course of the observations.

To see the variation of the bias significance with a stochastic bias, significance trajectories are plotted for a set of Monte Carlo trials (over measurement noise and bias) in Figure 6. It can be seen that the total, position, and velocity significance can vary dramatically based on bias realization.



(a) Significance for linear system. (b) Significance for nonlinear system.

Figure 5: Bias significance for deterministic, unknown bias.

For a stochastic bias, the average squared value of the significance can be obtained in closed form. This can be shown by first noting the following:

$$\begin{aligned} E\left[\lambda_k^2\right] &= \frac{1}{2}E\left[\mathbf{m}_k^T\mathbf{P}_{k\,|\,k}^{-1}\mathbf{m}_k\right], \\ &= \frac{1}{2}E\left[\mathbf{m}_k^T\mathbf{L}_k^{-T}\mathbf{L}_k^{-1}\mathbf{m}_k\right], \\ &= \frac{1}{2}E\left[\operatorname{tr}\left(\mathbf{L}_k^{-1}\mathbf{m}_k\mathbf{m}_k^T\mathbf{L}_k^{-T}\right)\right], \\ &= \frac{1}{2}\operatorname{tr}\left(\mathbf{L}_k^{-1}E\left[\mathbf{m}_k\mathbf{m}_k^T\right]\mathbf{L}_k^{-T}\right), \end{aligned}$$

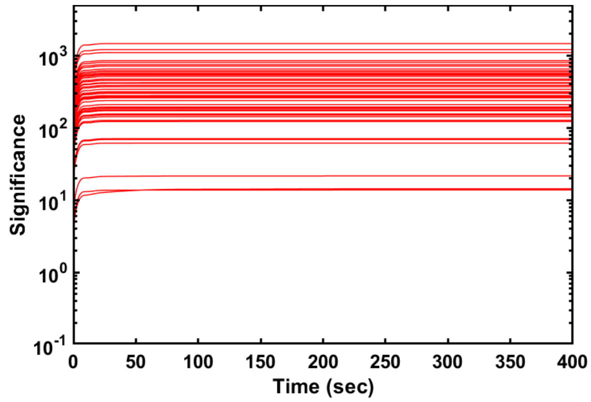
where $\mathbf{L}_k$ is from the Cholesky decomposition[18] of $\mathbf{P}_{k\,|\,k}$ such that $\mathbf{P}_{k\,|\,k} = \mathbf{L}_k\mathbf{L}_k^T$. Finally, using the form of the estimation error mean from (6), the average squared value of the significance can be expressed as

$$E\left[\lambda_k^2\right] = \frac{1}{2}\operatorname{tr}\left(\mathbf{L}_k^{-1}\mathbf{B}_k\mathbf{L}_k^{-T}\right), \tag{27}$$

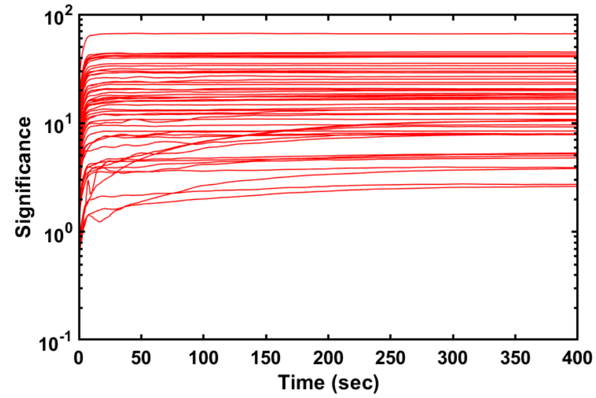
where $\mathbf{B}_k$ is as defined for linear systems in Theorem 2 and nonlinear systems in Theorem 4. Further, if the bias is constant, i.e., $\mathbf{b}_k = \mathbf{b}$ and $\mathbf{V} = E\left[\mathbf{b}\mathbf{b}^T\right]$, then (27) can be simplified to

$$E\left[\lambda_k^2\right] = \frac{1}{2}\operatorname{tr}\left(\mathbf{L}_k^{-1}\mathbf{S}_k\mathbf{V}\mathbf{S}_k^T\mathbf{L}_k^{-T}\right), \tag{28}$$

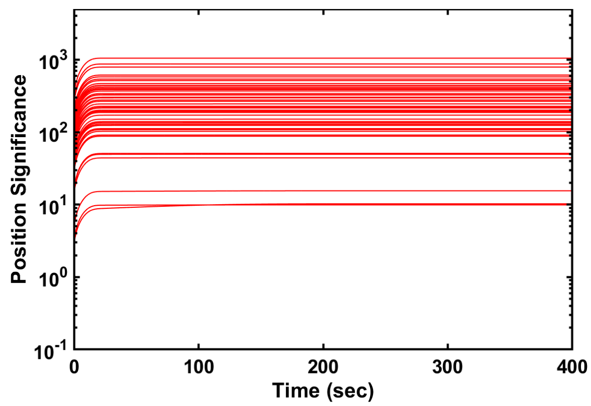
where $\mathbf{S}_k$ is defined above in Corollary 4. In Figure 7, the expression for the mean squared significance in (28) is compared against the Monte Carlo trials in Figure 6b.



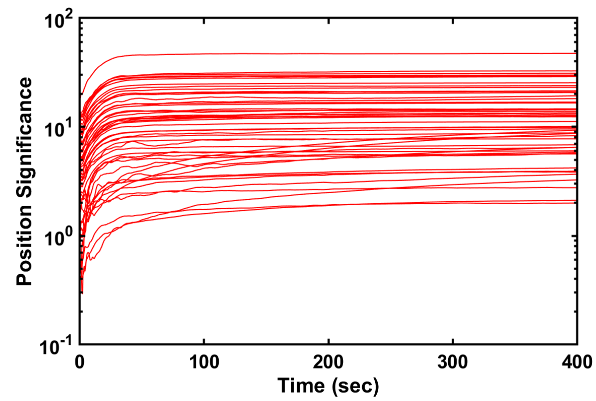
(a) Significance for linear system.



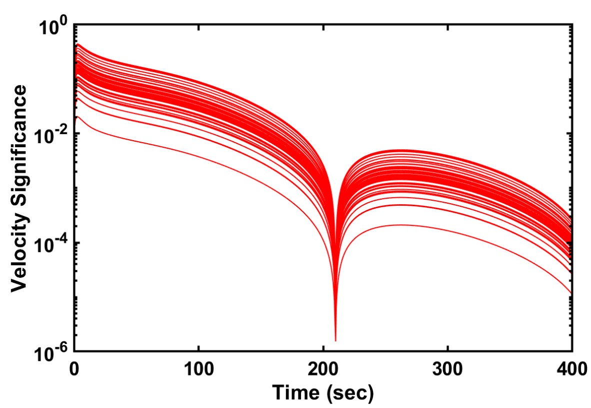
(b) Significance for nonlinear system.



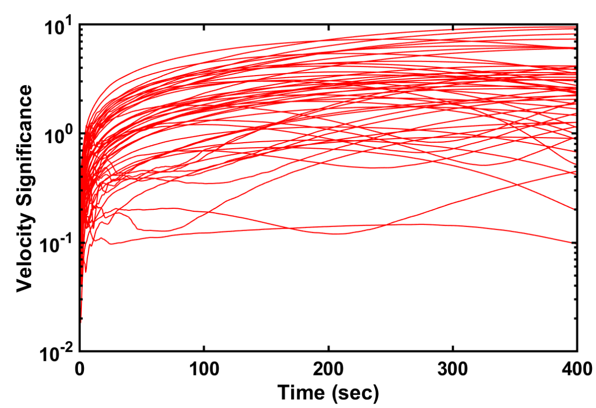
(c) Position significance for linear system.



(d) Position significance for nonlinear system.



(e) Velocity significance for linear system.



(f) Velocity significance for nonlinear system.

Figure 6: Bias significance for stochastic bias.

5. CONCLUSIONS

In this work, several expressions were derived for the mean and covariance of the estimation error of a bias-naïve estimator for both linear and nonlinear systems and deterministic and stochastic measurement biases. These results should prove useful in quickly examining the impact of bias on state estimation error for a fixed trajectory. Additionally, a measure of the relative importance of measurement bias is provided to allow for the quantification of deviations in assumed error statistics caused by measurement bias.

Simulation results were presented that illustrated the efficacy of the mean and covariance formulae and also

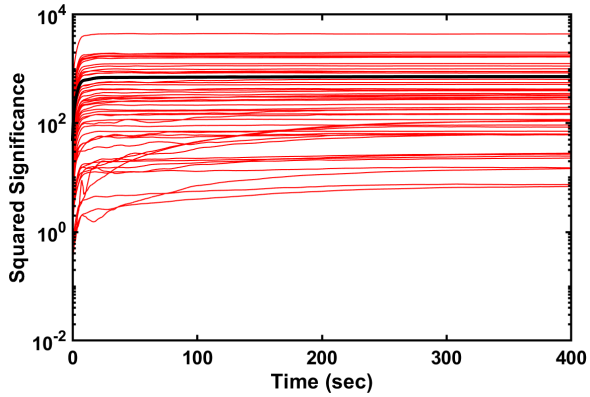


Figure 7: Squared bias significance for stochastic bias with expected value. Significance trajectories for 50 Monte Carlo trials are shown in red; expected value from (28) is shown in black.

illustrated notable phenomena in the estimation error of nonlinear systems. Specifically, a constant measurement bias injected into a nonlinear system can lead to appreciable velocity error biases that need to be considered in filter design. In future work, some topics to be studied are asymptotic behavior of estimation error and the incorporation of target maneuvers and dynamic model mismatch.

APPENDIX A. PROOF OF THEOREM 1

To aid in the decomposition of the mean and covariance of the state estimation error, it is possible to introduce the bias-free estimator $\hat{\mathbf{y}}_{k\,|\,k}$ that has access to the unbiased measurements $\tilde{\mathbf{z}}_k = \mathbf{z}_k - \mathbf{G}_k\mathbf{b}_k$. Note that $\hat{\mathbf{y}}_{k\,|\,k}$ obeys the same set of equations, (3), except $\mathbf{z}_k$ is replaced with $\tilde{\mathbf{z}}_k$. In this case, one can rewrite (4) as

$$\begin{aligned}\mathbf{m}_k &= E\left[\hat{\mathbf{x}}_{k\,|\,k} + (-\hat{\mathbf{y}}_{k\,|\,k} + \hat{\mathbf{y}}_{k\,|\,k}) - \mathbf{x}_k\right], \\ &= E\left[\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k}\right] + E\left[\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right], \qquad (29) \\ &= E\left[\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k}\right]. \qquad (30)\end{aligned}$$

Note that the term $E\left[\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right]$ in (29) vanishes as $\hat{\mathbf{y}}_{0\,|\,0} = E[\mathbf{x}_0]$ and for $k = 1$

$$\begin{aligned}E\left[\hat{\mathbf{y}}_{1\,|\,1} - \mathbf{x}_1\right] &= E\left[(\mathbf{I} - \mathbf{K}_1\mathbf{H}_1)\left(\mathbf{\Phi}_{1,0}\hat{\mathbf{y}}_{0\,|\,0} - \mathbf{x}_0\right) + \mathbf{K}_1\mathbf{w}_1\right], \\ &= (\mathbf{I} - \mathbf{K}_1\mathbf{H}_1)\left(\mathbf{\Phi}_{1,0}\hat{\mathbf{y}}_{0\,|\,0} - E\left[\mathbf{x}_0\right]\right) + \mathbf{K}_1 E\left[\mathbf{w}_1\right], \\ &= \mathbf{0}.\end{aligned}$$

If it is assumed that $E\left[\hat{\mathbf{y}}_{k-1\,|\,k-1} - \mathbf{x}_{k-1}\right] = \mathbf{0}$, then

$$\begin{aligned}E\left[\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right] &= E\left[(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\left(\mathbf{\Phi}_{k,k-1}\hat{\mathbf{y}}_{k-1\,|\,k-1} - \mathbf{x}_k\right) + \mathbf{K}_k\mathbf{w}_k\right], \\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\left(\mathbf{\Phi}_{k,k-1}E\left[\hat{\mathbf{y}}_{k-1\,|\,k-1}\right] - E\left[\mathbf{x}_k\right]\right) + \mathbf{K}_k E\left[\mathbf{w}_k\right], \\ &= \mathbf{0}.\end{aligned}$$

Likewise, the covariance in (5) can be observed to be

$$\begin{aligned}\mathbf{C}_k &= E\left[\left(\hat{\mathbf{x}}_{k\,|\,k} + (-\hat{\mathbf{y}}_{k\,|\,k} + \hat{\mathbf{y}}_{k\,|\,k}) - \mathbf{x}_k - \mathbf{m}_k\right)\left(\hat{\mathbf{x}}_{k\,|\,k} + (-\hat{\mathbf{y}}_{k\,|\,k} + \hat{\mathbf{y}}_{k\,|\,k}) - \mathbf{x}_k - \mathbf{m}_k\right)^T\right], \\ &= E\left[\left(\left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k}\right) + \left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right) - \mathbf{m}_k\right)\left(\left(\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k}\right) + \left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right) - \mathbf{m}_k\right)^T\right]. \qquad (31)\end{aligned}$$

The terms in (30) and (31) can be calculated as

$$\begin{aligned}
\hat{\mathbf{x}}_{k\,|\,k} - \mathbf{x}_k &= \hat{\mathbf{x}}_{k\,|\,k-1} + \mathbf{K}_k\left(\mathbf{H}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{b}_k + \mathbf{w}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k\,|\,k-1}\right) - \mathbf{x}_k, \\
&= \hat{\mathbf{x}}_{k\,|\,k-1} - \mathbf{x}_k + \mathbf{K}_k\left(\mathbf{H}_k\mathbf{x}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k\,|\,k-1}\right) + \mathbf{K}_k\left(\mathbf{G}_k\mathbf{b}_k + \mathbf{w}_k\right), \\
\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k &= \hat{\mathbf{y}}_{k\,|\,k-1} + \mathbf{K}_k\left(\mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k - \mathbf{H}_k\hat{\mathbf{y}}_{k\,|\,k-1}\right) - \mathbf{x}_k, \\
&= \hat{\mathbf{y}}_{k\,|\,k-1} - \mathbf{x}_k + \mathbf{K}_k\left(\mathbf{H}_k\mathbf{x}_k - \mathbf{H}_k\hat{\mathbf{y}}_{k\,|\,k-1}\right) + \mathbf{K}_k\mathbf{w}_k, \\
\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} &= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\hat{\mathbf{x}}_{k\,|\,k-1} - \hat{\mathbf{y}}_{k\,|\,k-1}\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k,
\end{aligned} \tag{32}$$

Rewriting (32) in terms of past state estimates, one obtains

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\mathbf{\Phi}_{k,k-1}\hat{\mathbf{x}}_{k-1\,|\,k-1} - \mathbf{\Phi}_{k,k-1}\hat{\mathbf{y}}_{k-1\,|\,k-1}\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k. \tag{33}$$

The state estimate difference in (33) can now be expressed as

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{\Phi}_{k,k-1}\left(\hat{\mathbf{x}}_{k-1\,|\,k-1} - \hat{\mathbf{y}}_{k-1\,|\,k-1}\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k. \tag{34}$$

Starting the recursion in (34), it can be observed that the initial state estimate difference is

$$\begin{aligned}
\hat{\mathbf{x}}_{1\,|\,1} - \hat{\mathbf{y}}_{1\,|\,1} &= \left(\mathbf{I} - \mathbf{K}_1\mathbf{H}_1\right)\mathbf{\Phi}_{1,0}\left(\hat{\mathbf{x}}_{0\,|\,0} - \hat{\mathbf{y}}_{0\,|\,0}\right) + \mathbf{K}_1\mathbf{G}_1\mathbf{b}_1, \\
&= \mathbf{K}_1\mathbf{G}_1\mathbf{b}_1,
\end{aligned}$$

since $\hat{\mathbf{x}}_{0\,|\,0} = \hat{\mathbf{y}}_{0\,|\,0} = E\left[\mathbf{x}_0\right]$. Continuing on to the next time step,

$$\begin{aligned}
\hat{\mathbf{x}}_{2\,|\,2} - \hat{\mathbf{y}}_{2\,|\,2} &= \left(\mathbf{I} - \mathbf{K}_2\mathbf{H}_2\right)\mathbf{\Phi}_{2,1}\left(\hat{\mathbf{x}}_{1\,|\,1} - \hat{\mathbf{y}}_{1\,|\,1}\right) + \mathbf{K}_2\mathbf{G}_2\mathbf{b}_2, \\
&= \left(\mathbf{I} - \mathbf{K}_2\mathbf{H}_2\right)\mathbf{\Phi}_{2,1}\mathbf{K}_1\mathbf{G}_1\mathbf{b}_1 + \mathbf{K}_2\mathbf{G}_2\mathbf{b}_2.
\end{aligned}$$

Finally, taking the recursion to the k^{th} time step, one can write

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \sum_{i=1}^{k}\mathbf{\Lambda}_i^k\mathbf{K}_i\mathbf{G}_i\mathbf{b}_i. \tag{35}$$

Thus, using (35) and the fact that $\mathbf{b}_k$ is an unknown, deterministic value, the mean and covariance of the state estimation error can be simplified as

$$\mathbf{m}_k = \sum_{i=1}^{k}\mathbf{\Lambda}_i^k\mathbf{K}_i\mathbf{G}_i\mathbf{b}_i, \tag{36}$$

$$\begin{aligned}
\mathbf{C}_k &= E\left[\left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right)\left(\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right)^T\right], \\
&= \mathbf{P}_{k\,|\,k},
\end{aligned} \tag{37}$$

where $\mathbf{P}_{k\,|\,k}$ is the error covariance of the estimator with unbiased measurements, $\hat{\mathbf{y}}_{k\,|\,k}$.

APPENDIX B. PROOF OF THEOREM 3

To begin, the terms in (30) and (31) are examined. However, it should be noted that due to the nonlinearities in the measurement and state transition functions, the expected value of the unbiased track state error can no longer be proven to be identically zero, i.e., $E\left[\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k\right] \neq \mathbf{0}$; this is due to typically small perturbations that arise from the initial state estimate error propagating through the suboptimal filter. However, in most cases, this term will provide negligible contribution to the error after a few measurements and is therefore ignored. With this, the terms in the mean and covariance expressions in (30) and (31) are

$$\begin{aligned}
\hat{\mathbf{x}}_{k\,|\,k} - \mathbf{x}_k &= \hat{\mathbf{x}}_{k\,|\,k-1} + \mathbf{K}_k\left(h_k\left(\mathbf{x}_k\right) + \mathbf{G}_k\mathbf{b}_k + \mathbf{w}_k - h_k\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right)\right) - \mathbf{x}_k, \\
&= \hat{\mathbf{x}}_{k\,|\,k-1} - \mathbf{x}_k + \mathbf{K}_k\left(h_k\left(\mathbf{x}_k\right) - h_k\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right)\right) + \mathbf{K}_k\left(\mathbf{G}_k\mathbf{b}_k + \mathbf{w}_k\right), \\
\hat{\mathbf{y}}_{k\,|\,k} - \mathbf{x}_k &= \hat{\mathbf{y}}_{k\,|\,k-1} + \mathbf{K}_k\left(h_k\left(\mathbf{x}_k\right) + \mathbf{w}_k - h_k\left(\hat{\mathbf{y}}_{k\,|\,k-1}\right)\right) - \mathbf{x}_k, \\
&= \hat{\mathbf{y}}_{k\,|\,k-1} - \mathbf{x}_k + \mathbf{K}_k\left(h_k\left(\mathbf{x}_k\right) - h_k\left(\hat{\mathbf{y}}_{k\,|\,k-1}\right)\right) + \mathbf{K}_k\mathbf{w}_k, \\
\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} &= \hat{\mathbf{x}}_{k\,|\,k-1} - \hat{\mathbf{y}}_{k\,|\,k-1} - \mathbf{K}_k\left(h_k\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right) - h_k\left(\hat{\mathbf{y}}_{k\,|\,k-1}\right)\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k.
\end{aligned} \tag{38}$$

To continue, the first of two linearization steps is taken—replacement of the measurement function, $h_k(\cdot)$, using its first-order Taylor series approximation around the current biased track state. Specifically, the following replacement is made:

$$h_k\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right) - h_k\left(\hat{\mathbf{y}}_{k\,|\,k-1}\right) \approx \mathbf{H}_k\hat{\mathbf{x}}_{k\,|\,k-1} - \mathbf{H}_k\hat{\mathbf{y}}_{k\,|\,k-1}, \tag{39}$$

where $\mathbf{H}_k$ is as defined in (19). Using (39) in (32), the state estimate difference can now be written as

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\hat{\mathbf{x}}_{k\,|\,k-1} - \hat{\mathbf{y}}_{k\,|\,k-1}\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k. \tag{40}$$

Rewriting (40) in terms of past state estimates, one obtains

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left(\phi_{k,k-1}\left(\hat{\mathbf{x}}_{k-1\,|\,k-1}\right) - \phi_{k,k-1}\left(\hat{\mathbf{y}}_{k-1\,|\,k-1}\right)\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k. \tag{41}$$

Similar to the linearization step taken with the measurement function, the state transition function, $\phi_{k,k-1}(\cdot)$, will now be replaced with its first-order Taylor series approximation around the previous biased track state such that

$$\phi_{k,k-1}\left(\hat{\mathbf{x}}_{k\,|\,k-1}\right) - \phi_{k,k-1}\left(\hat{\mathbf{y}}_{k\,|\,k-1}\right) \approx \mathbf{\Phi}_{k,k-1}\hat{\mathbf{x}}_{k\,|\,k-1} - \mathbf{\Phi}_{k,k-1}\hat{\mathbf{y}}_{k\,|\,k-1},$$

where $\mathbf{\Phi}_{k,k-1}$ is as defined in (18). The state estimate difference in (41) can now be expressed as

$$\hat{\mathbf{x}}_{k\,|\,k} - \hat{\mathbf{y}}_{k\,|\,k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{\Phi}_{k,k-1}\left(\hat{\mathbf{x}}_{k-1\,|\,k-1} - \hat{\mathbf{y}}_{k-1\,|\,k-1}\right) + \mathbf{K}_k\mathbf{G}_k\mathbf{b}_k. \tag{42}$$

The rest of the proof follows from the proof of Theorem 1, starting with (34).

REFERENCES

[1] Lin, X., Kirubarajan, T., and Bar-Shalom, Y., "Exact multi-sensor dynamic bias estimation with local tracks," in [*Proceedings of 6th International Conference on Information Fusion*], (July 2003).

[2] Lin, X., Kirubarajan, T., and Bar-Shalom, Y., "Multisensor bias estimation with local tracks without a priori association," in [*Signal and Data Processing of Small Targets*], Drummond, O. E., ed., *Proceedings of SPIE* **5204**, 334–345 (December 2003).

[3] Bar-Shalom, Y., Willett, P. K., and Tan, X., [*Tracking and Data Fusion: A Handbook of Algorithms*], YBS Publishing (2011).

[4] Blackman, S. S. and Banh, N. D., "Track association using correction for bias and missing data," *Proceedings of SPIE* **2235**, 529–539 (July 1994).

[5] Blackman, S. S. and Popoli, R., [*Design and Analysis of Modern Tracking Systems*], ch. 6, Artech House (1999).

[6] Humke, C., "Bias removal techniques for the target-object mapping problem," (July 2002). Technical Report 1060, M.I.T. Lincoln Laboratory, Lexington, MA.

[7] Warren, A., "Bias and covariance formula for filter error analysis and model sensitivity," in [*Proceedings of Joint Automatic Control Conference*], **11**, 53–60 (1973).

[8] Asher, R. and Reeves, R., "Performance evaluation of suboptimal filters," *IEEE Transactions on Aerospace and Electronic Systems* **11**, 400–405 (May 1975).

[9] Brown, R. and Sage, A., "Analysis of modeling and bias errors in discrete-time state estimation," *IEEE Transactions on Aerospace and Electronic Systems* **7**, 340–354 (March 1971).

[10] Gelb, A., ed., [*Applied Optimal Estimation*], The M.I.T. Press (1974).

[11] Asher, R., Herring, K., and Ryles, J., "Bias, variance, and estimation error in reduced order filters," *Automatica* **12**, 589–600 (May 1976).

[12] Fitzgerald, R., "Divergence of the Kalman filter," *IEEE Transactions on Automatic Control* **16**, 736–747 (December 1971).

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

[14] Papoulis, A. and Pillai, S., [*Probability, Random Variables, and Stochastic Processes*], McGraw-Hill (2002).

[15] Mailloux, R., [*Phased Array Antenna Handbook*], Artech House, Inc. (1994).

[16] Li, X. and Jilkov, V., "A survey of maneuvering target tracking—Part II: Motion models of ballistic and space targets," *IEEE Transactions on Aerospace and Electronic Systems* **46**, 96–119 (January 2010).
[17] Mahalanobis, P., "On the generalised distance in statistics," *Proceedings of the National Institute of Sciences of India* **2**, 49–55 (April 1936).
[18] Golub, G. H. and van Loan, C. F., [*Matrix Computations*], Johns Hopkins University Press (1996).