Discrete-Time Linear Filtering in Arbitrary Noise*

X. Rong Li
University of New Orleans, xli@uno.edu

Chongzhao Han and Jie Wang
Xi'an Jiaotong University, czhan@xjtu.edu.cn

Abstract—The Kalman filter is a recursive Best Linear Unbiased Estimator (BLUE) for a linear dynamic system with uncorrelated white process and measurement noises. It has been extended to the case where the noises are Markov and/or crosscorrelated for the same time instant. This paper presents optimal batch and semi-recursive filters and a suboptimal recursive filter for a linear discrete-time system with arbitrarily colored (not necessarily Markov) noises that are arbitrarily cross-correlated and correlated with the initial state of the system. They are generalizations of the Kalman filter for the case of arbitrary additive noise of known first two moments. Numerical examples are provided. They demonstrate the superiority in terms of performance and efficiency of the proposed recursive filter.

1 Introduction

The development of the Kalman filter [4] is widely regarded as one of the breakthroughs that marked the beginning of the era of modern control system theory [8]. It has found widespread applications in many areas. The Kalman filter provides an optimal recursive solution to the problem of filtering the state of a linear dynamic system, which has the following two major advantages over the Wiener filter: it is in a very simple recursive form in time domain and is perfectly suitable for nonstationary as well as stationary processes. Although optimal only for linear systems, the Kalman filter plays an important role for many nonlinear problems. For example, it is the backbone of most solutions for nonlinear state estimation where the nonlinearity is mild; it often serves as a building block for many more sophisticated state estimators, such as the multiple-model method, a state-of-the-art approach for many hybrid estimation problems [1, 5].

The original Kalman filter is valid only for systems with white and mutually uncorrelated process and measurement noises. It has been generalized to systems with colored noises that are Markov processes. In many practical situations, however, the noises are not necessarily Markov. For example, many sensor errors are dependent on the random state of the unknown system it is observing and thus are correlated over time in a non-Markov fashion, such as when sensors are on a moving platform without accurate knowledge of the state of the platform. Guidance, navigation, and positioning applications all fall into this category. Another example is that the state of the system is observed by a sensor in a time-correlated noisy environment, such as during noise jamming generated by a target. Also, it is well-known that the process and measurement noise sequences of a discrete-time system sampled from a linear continuous-time system are in general cross-correlated.

The measurement process in the state estimation theory does not have to be the actual physical measurement process of the state of a system. It can be virtually any temporal or spatial process provided that it is related with the state process. In this spirit, many filtering, prediction, and smoothing applications involve the use of "artificial measurements" with colored noise that is correlated with the process noise and with the initial state of the system. One important such application area involving a distributed system of multiple local stations (sensors) is known as track fusion (see, e.g., [3, 2, 7]). In such a system, each local station makes its own estimates of the state of a target based on local measurements and a state estimator at a fusion center estimates the target's state using the received local estimates as the "measurements" of the target [6]. Obviously, the noise of these "measurements" are not only autocorrelated but also cross-correlated with the process noise and the initial state of the system. This noise is not Markov if the process noise or the real measurement noise is not white.

This paper presents optimal batch and semi-recursive and suboptimal recursive filters for linear discrete-time systems in noise of known first two moments. The process and measurement noises considered are not necessarily Markov, mutually uncorrelated, or uncorrelated with the initial state. Various types of arbitrary correlation are accounted for in these filters, including auto- and cross-correlation of the process and measurement noise, and their correlation with the initial state. Computer simulation examples are provided to verify and illustrate the filters presented. While the optimal batch and semi-recursive filters require heavy computation, the recursive filter is more efficient at a price of an insignificant performance degradation.

2 Problem Formulation and Notations

Consider a general discrete-time linear dynamic system with arbitrary additive noise:

$$\begin{cases} x_{k+1} = F_k x_k + w_k \\ z_k = H_k x_k + v_k \end{cases} \tag{1}$$

where k is the time index; $x_k \in \mathbb{R}^{n_x}, z_k \in \mathbb{R}^{n_z}$; F_k and H_k are the system matrices with appropriate dimensions. The

*Supported in part by ONR Grant N00014-00-1-0677 and NSF via Grant ECS-9734285.

process noise $\{w_k\}$ and the measurement noise $\{v_k\}$ are zero mean and **colored** with **arbitrary** covariance matrices

$$\begin{aligned} \text{cov}(w_i, w_j) &= Q_{ij}, \quad Q_i = Q_{ii}, \qquad i,j = 0,1,2,\dots \\ \text{cov}(v_i, v_j) &= R_{ij}, \quad R_i = R_{ii}, \qquad i,j = 1,2,\dots \end{aligned}$$

The initial state x_0 is a random vector with mean $\bar{x}_0$ and covariance matrix P_0. x_0, $\{w_k\}$, and $\{v_k\}$ are **arbitrarily correlated** with each other with known cross-covariances

$$\begin{aligned} &\text{cov}(x_0, v_i) = A_i, \quad \text{cov}(x_0, w_j) = B_j, \\ &\text{cov}(w_j, v_i) = C_{ji}, \quad i = 1,2,\dots, \quad j = 0,1,2,\dots \end{aligned}$$

Such process and measurement noises will be referred to as "*arbitrary noise*" or "*arbitrarily correlated noise.*"

Filtering aims at estimating something about an unobservable process $x(t)$ given a related process $z(t)$. $x(t)$ and $z(t)$ and their relationships are completely described by their (finite-dimensional) joint distributions. For the state estimation of the above system, the first two moments of these joint distributions are completely determined given $F_k, H_k, A_i, B_i, C_{ij}, Q_{ij}, R_{ij}, P_0$ and $\bar{x}_0$. As far as filters based only on the first two moments are concerned, such as the best linear unbiased estimator (BLUE)[1], the above correlation model is general, except for its linearity, additivity of the noise, and the assumption of a perfect knowledge of $F_k, H_k, A_i, B_i, C_{ij}, Q_{ij}, R_{ij}$. The zero-mean assumption for the noise is made only for simplicity of presentation.

This paper studies linear filtering in the BLUE sense for the above discrete-time linear system with arbitrary noise. Let

$$\begin{aligned} \hat{x}_{k|k-1} &= E^*[x_k|Z_{k-1}] \quad \text{(prediction)} \\ \tilde{x}_{k|k-1} &= x_k - \hat{x}_{k|k-1} \quad \text{(prediction error)} \\ P_{k|k-1} &= \text{cov}(\tilde{x}_{k|k-1}|Z_{k-1}) \quad \text{(prediction error covariance)} \\ \hat{x}_{k|k} &= E^*[x_k|Z_k] \quad \text{(updated estimate)} \\ \tilde{x}_{k|k} &= x_k - \hat{x}_{k|k} \quad \text{(update error)} \\ P_{k|k} &= \text{cov}(\tilde{x}_{k|k}|Z_k) \quad \text{(error covariance of update)} \end{aligned}$$

where $E^*[y|Z_i]$ denotes the BLUE estimate of y given $Z_i = \{z_1, \dots, z_i\}$.

3 Batch Filtering in Arbitrary Noise

This section presents an optimal filter in the batch form. Let

$$\begin{aligned} z^k &= [z_1', \dots, z_k']', \quad w^k = [w_1', \dots, w_k']' \\ v^k &= [v_1', \dots, v_k']', \quad H^k = [H_1', \dots, H_k']' \end{aligned}$$

Theorem 1 (Batch Optimal Filter). *Assume that F_k^{-1} exists for all k. Then, the following algorithm provides a batch form of the unique BLUE filter for the system (1) with arbitrary process and measurement noise:*

1. *Initialization:* $\hat{x}_{0|0} = \bar{x}_0$, $P_{0|0} = P_0$.
2. *The prediction is done recursively by:*

$$\begin{aligned} \hat{x}_{k|0} &= F_{k-1}\hat{x}_{k-1|0} \\ P_{k|0} &= F_{k-1}P_{k-1|0}F_{k-1}' + Q_{k-1} \\ &\quad + F_{k-1}\Psi_{k-1|0}^{k-1} + (F_{k-1}\Psi_{k-1|0}^{k-1})' \end{aligned}$$

with the initial value

$$\begin{aligned} \hat{x}_{0|0} &= \bar{x}_0 \\ P_{0|0} &= P_0 \end{aligned}$$

and $\Psi_{k-1|0}^{k-1}$ is given by the following recursion

$$\Psi_{i|0}^{k-1} = F_{i-1}\Psi_{i-1|0}^{k-1} + Q_{i-1,k-1}$$

with the initial value $\Psi_{0|0}^{k-1} = B_{k-1}$.

3. *The update is done in a batch form for each time k:*

$$\hat{x}_{k|k} = \hat{x}_{k|0} + K^k[z^k - \tilde{H}^k\hat{x}_{k|0}] \tag{2}$$
$$P_{k|k} = P_{k|0} - K^kS^kK^{k\prime} \tag{3}$$

where

$$S^k = \tilde{H}^kP_{k|0}\tilde{H}^{k\prime} + \tilde{R}^k + \tilde{H}^k\tilde{A}^k + (\tilde{H}^k\tilde{A}^k)' \tag{4}$$
$$K^k = [P_{k|0}\tilde{H}^{k\prime} + \tilde{A}^k](S^k)^{-1} \tag{5}$$

In the above,

$$\begin{aligned} \tilde{A}^1 &= \tilde{A}_1 = A_1, \quad \tilde{A}_{0|0}^k = A_k \\ \tilde{A}_{i|0}^k &= F_{i-1}\tilde{A}_{i-1|0}^k + C_{i-1,k}, \quad \tilde{A}_k = \tilde{A}_{k|0}^k \\ \tilde{A}^k &= [F_{k-1}\tilde{A}^{k-1} + \tilde{C}_{k-1}^{k-1} - F_{k-1}\Psi_{k-1|0}^{k-1}(\tilde{H}_k^{k-1})' \\ &\quad - Q_{k-1}(\tilde{H}_k^{k-1})', \tilde{A}_k] \\ \bar{R}^{k-1} &= \tilde{R}^{k-1} + \tilde{H}_k^{k-1}Q_{k-1}(\tilde{H}_k^{k-1})' - \tilde{H}_k^{k-1}\tilde{C}_{k-1}^{k-1} \\ &\quad - (\tilde{H}_k^{k-1}\tilde{C}_{k-1}^{k-1})' \\ \tilde{R}^k &= \begin{bmatrix} \bar{R}^{k-1} & \tilde{R}_k^{k-1} \\ (\tilde{R}_k^{k-1})' & R_k \end{bmatrix} \\ \tilde{H}^k &= \begin{bmatrix} \tilde{H}_k^{k-1} \\ H_k \end{bmatrix}, \tilde{H}_k^{k-1} = \begin{bmatrix} H_1(F_{k-1}\cdots F_1)^{-1} \\ \vdots \\ H_{k-1}(F_{k-1})^{-1} \end{bmatrix} \end{aligned}$$

where $\tilde{R}_k^{k-1}$ and $\tilde{C}_{k-1}^{k-1}$ are computed by the recursion:

$$\begin{aligned} \tilde{R}_k^i &= \begin{bmatrix} \tilde{R}_k^{i-1} \\ R_{ik} \end{bmatrix} - \tilde{H}_{i+1}^iC_{ik}, \\ \tilde{C}_{k-1}^i &= [\tilde{C}_{k-1}^{i-1} - Q_{k-1,i-1}(\tilde{H}_i^{i-1})', C_{k-1,i}] \end{aligned}$$

with the initialization

$$\tilde{R}_k^1 = R_{1k} - \tilde{H}_2^1C_{1k}, \qquad \tilde{C}_{k-1}^1 = C_{k-1,1}$$

$P_{k|k}$ and K^k have the following alternative forms

$$\begin{aligned} P_{k|k} &= U^kP_{k|0}U^{k\prime} + K^k\tilde{R}^kK^{k\prime} - U^k\tilde{A}^kK^{k\prime} \\ &\quad - (U^k\tilde{A}^kK^{k\prime})' \end{aligned} \tag{6}$$
$$K^k = (P_{k|k}\tilde{H}^{k\prime} + \tilde{A}^k)(\tilde{R}^k + \tilde{H}^k\tilde{A}^k)^{-1} \tag{7}$$

where $U^k = I - K^k\tilde{H}^k$.

[1] It has many aliases, such as minimum mean square error, (linear) least mean square, linear (unbiased) minimum variance, and (linear) least squares estimation.

All proofs are omitted in this conference version.

If S^k is singular, the unique BLUE filter is given above with $(S^k)^{-1}$ replaced by $(S^k)^+$, the Moore-Penrose pseudoinverse of S^k. However, (7) is no longer valid.

For practical systems where F_k^{-1} does not exist for all k for its discrete-time representation, an equivalent continuous-time representation of the system dynamics may be used to avoid this difficulty. The corresponding mixed-time filter can be derived without much difficulty.

This theorem provides a way, albeit inefficient, of computing the optimal state estimates for system (1) with arbitrarily colored process and measurement noises that are arbitrarily cross-correlated and correlated with the initial state. It will be used to obtain the optimal estimates for comparison with those from other filters. For large k, computation of $(S^k)^{-1}$ amounts to a major burden, which is on the order of $O[(kn_z)^3]$, where $n_z = \dim(z_k)$ is the measurement dimension.

4 Recursive Filtering in Arbitrary Noise

The following theorem presents an optimal filter for linear filtering in arbitrary noise in an alternative form that is computationally more efficient than Theorem 1 because $(S^k)^{-1}$ and $\Xi^k = \mathrm{cov}(\tilde{x}_{k|0}, \tilde{z}^k)$ are computed recursively.

Theorem 2 (Semi-Recursive Optimal Filter). *Assume that F_k^{-1} exists for all k. Then, the following algorithm provides a semi-recursive form of the unique BLUE filter for the system (1) with arbitrary process and measurement noise:*

1. *Initialization:* $\hat{x}_{0|0} = \bar{x}_0$, $P_{0|0} = P_0$.
2. *Prediction: Same as in Theorem 1.*
3. *Update:*

$$
\begin{aligned}
K^k &= \Xi^k (S^k)^{-1} \\
\hat{x}_{k|k} &= \hat{x}_{k|0} + K^k [z^k - \tilde{H}^k \hat{x}_{k|0}] \\
P_{k|k} &= P_{k|0} - \Xi^k (S^k)^{-1} \Xi^{k\prime}
\end{aligned}
$$

where

$$
\begin{aligned}
S_k &= H_k P_{k|0} H_k' + R_k + H_k \tilde{A}_k + (H_k \tilde{A}_k)' \\
\Xi_k^{k-1} &= F_{k-1} \Xi^{k-1} + [(\tilde{H}^{k-1} \Psi_{k-1|0}^{k-1})' + \tilde{C}_{k-1}^{k-1}] \\
\Xi^k &= [\Xi_k^{k-1}, \Xi_k], \quad \Xi_k = P_{k|0} H_k' + \tilde{A}_k \\
S_k^{k-1} &= (H_k \Xi_k^{k-1})' + \tilde{R}_k^{k-1} + \tilde{H}^{k-1} \tilde{A}_k \\
\Delta_k &= S_k - (S_k^{k-1})' (S^{k-1})^{-1} S_k^{k-1} \\
\Lambda_k &= (S^{k-1})^{-1} S_k^{k-1} \Delta_k^{-1} \\
(S^k)^{-1} &= \begin{bmatrix} (S^{k-1})^{-1} [I + S_k^{k-1} \Lambda_k'] & -\Lambda_k \\ -\Lambda_k' & \Delta_k^{-1} \end{bmatrix}
\end{aligned}
$$

which can be initialized by

$$
\begin{aligned}
S_1^{-1} &= [H_1 P_{1|0} H_1' + R_1 + H_1 \tilde{A}_1 + (H_1 \tilde{A}_1)']^{-1} \\
\Xi^1 &= P_{1|0} H_1' + \tilde{A}_1
\end{aligned}
$$

In the above, $\tilde{H}^k$, $\tilde{A}_k, \tilde{R}_k^{k-1}$ are given in Theorem 1. The alternative forms of $P_{k|k}$ and K^k are also valid.

This filter is semi-recursive in that $(S^k)^{-1}$ and Ξ^k are computed recursively, but $\hat{x}_{k|k}$ and $P_{k|k}$ depend not only on $\hat{x}_{k-1|k-1}$, $P_{k-1|k-1}$, z_k, and prior information. The following theorem presents a recursive filter.

Theorem 3 (Recursive Filter). *A recursive BLUE filter for system (1) with arbitrary process and measurement noises that satisfy the following equations*

$$
\begin{aligned}
E^*[w_k|Z_k] &= E[w_k] = 0 & (8) \\
E^*[v_k|Z_{k-1}] &= E[v_k] = 0 & (9)
\end{aligned}
$$

is given by the following algorithm, where $E^*[y|Z_k]$ stands for BLUE estimate of y using Z_k.

1. *Initialization:* $\hat{x}_{0|0} = \bar{x}_0$, $P_{0|0} = P_0$.
2. *Prediction:*

$$
\begin{aligned}
\hat{x}_{k|k-1} &= F_{k-1} \hat{x}_{k-1|k-1} \\
P_{k|k-1} &= F_{k-1} P_{k-1|k-1} F_{k-1}' + Q_{k-1} \\
&\quad + F_{k-1} \Psi_{k-1} + (F_{k-1} \Psi_{k-1})'
\end{aligned}
$$

where

$$
\Psi_{k-1} = \Psi_{k-1|k-1}^{k-1}
$$

$\Psi_{i|j}^{k-1}$ *is given recursively by the Kalman filter:*

$$
\begin{aligned}
\Psi_{i|i-1}^{k-1} &= F_{i-1} \Psi_{i-1|i-1}^{k-1} + Q_{i-1,k-1} \\
\Psi_{i|i}^{k-1} &= \Psi_{i|i-1}^{k-1} + K_i (-C_{k-1,i}' - H_i \Psi_{i|i-1}^{k-1})
\end{aligned}
$$

with $-C_{k-1,i}'$ being the new "data" and the initial value $\Psi_{0|0}^{k-1} = B_{k-1}$.

3. *Update:*

$$
\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k [z_k - H_k \hat{x}_{k|k-1}] \\
P_{k|k} &= P_{k|k-1} - K_k S_k K_k'
\end{aligned}
$$

K_k and S_k are given by

$$
\begin{aligned}
K_k &= (P_{k|k-1} H_k' + \Omega_k) S_k^{-1} \\
S_k &= H_k P_{k|k-1} H_k' + H_k \Omega_k + (H_k \Omega_k)' + R_k
\end{aligned}
$$

with

$$
\Omega_k = \Omega_{k|k-1}^k
$$

$\Omega_{i|j}^k$ *is given recursively by the Kalman filter:*

$$
\begin{aligned}
\Omega_{i|i-1}^k &= F_{i-1} \Omega_{i-1|i-1}^k + C_{i-1,k} \\
\Omega_{i|i}^k &= \Omega_{i|i-1}^k + K_i (-R_{ik} - H_i \Omega_{i|i-1}^k)
\end{aligned}
$$

*with $-R_{ik}$ being the new "data" and the initial value $\Omega_{0|0}^k = A_k$. The error covariance (**??**) can also be written in the following generalized Joseph's form*

$$
\begin{aligned}
P_{k|k} &= U_k P_{k|k-1} U_k' + K_k R_k K_k' \\
&\quad - U_k \Omega_k K_k' - (U_k \Omega_k K_k')'
\end{aligned}
$$

with $U_k = I - K_k H_k$, and K_k has an alternative form

$$
K_k = (P_{k|k} H_k' + \Omega_k)(R_k + H_k \Omega_k)^{-1} \qquad (10)
$$

All three filters above are generalizations of the Kalman filter for the case with all five possible types of *arbitrary* correlation under the stated assumption provided that the corresponding covariance matrices are given: (a) the correlation among process noise over time, (b) the correlation among measurement noise over time, (c) the cross-correlation between process and measurement noises, (d) the correlation between the initial state and the process noise, and (e) the correlation between the initial state and the measurement noise. In particular, the process and measurement noise sequences are *arbitrarily* colored, not necessarily Markov processes. Prior to this work, BLUE filters (as extensions of the Kalman filter) are not available in a general setting when any of these *arbitrary* correlations exists.

In this filter, the autocorrelation of the process noise $\{w_k\}$ and its correlation with the initial state x_0 only affect Ψ: specifically, the correlation between x_0 and $\{w_k\}$ only affects initial value $\Psi^k_{0|0}$ and the autocorrelation of $\{w_k\}$ only affects additional terms in the recursion of $\Psi^k_{i|j}$; similarly, the effect of the autocorrelation of the measurement noise $\{v_k\}$ and its correlation with x_0 are summarized by Ω: the correlation between x_0 and $\{v_k\}$ only affects initial value $\Omega^k_{0|0}$ and the autocorrelation of $\{v_k\}$ only affects additional terms in the recursion of $\Omega^k_{i|j}$; the effect of crosscorrelation of $\{w_k\}$ and $\{v_k\}$ is summarized by Ω and Ψ jointly. This theorem reduces to the Kalman filter when $\Omega = 0$ and $\Psi = 0$.

The recursive filter is computationally much more efficient than the optimal filters of Theorems 1 and 2: The batch filter involves inversion of an $n_z k \times n_z k$ matrix, where $n_z = \dim(z_k)$; whereas in this recursive filter, the matrix inversion involves only $n_z \times n_z$ matrix S_k or $(R_k + H_k\Omega_k)$ and the major computation comes from the terms Ψ_k and Ω_k for large k. Unfortunately, all old gain matrices $K_i, \forall i < k$, are needed to compute Ψ_k and Ω_k. The computation of Ψ_k and Ω_k increases with k and is inevitable for process and measurement noise of arbitrary correlation since $Q_{ik}, R_{ik}, C_{ik}, C_{ki}, A_k, B_k, \forall i < k$, in general carry new and possibly only new information about the correlation of the new noise terms w_k and v_k. In some special cases, this computation may be reduced significantly. In practice, one may use a moving window by using only recent gain matrices since the correlation between noise terms at different instants generally becomes weaker as the time difference increases.

While the optimal filters of Theorems 1 and 2 are applicable only to systems (1) with nonsingular F_k for all k, the recursive filter of Theorem 3 is applicable to all systems (1), but it is optimal only when the assumptions (8)–(9) are valid. It can be shown that under assumptions (8)–(9), the optimal batch filter gain K^k of Theorems 1 and 2 is related to the recursive filter gain K_k of Theorem 3 by

$$K^k = \left[\; (I - K_k H_k) F_{k-1} K^{k-1}, \quad K_k \; \right] \tag{11}$$

The additional offset of this recursive filter's estimate $\hat{x}$ from the optimal estimate $\hat{x}^*$ in one cycle is given by

$$\begin{aligned} \hat{x}^*_{k|k} - \hat{x}_{k|k} &= K_k(E^*[v_k|Z_{k-1}] - E[v_k]) \\ \hat{x}^*_{k+1|k} - \hat{x}_{k+1|k} &= E^*[w_k|Z_k] - E[w_k] \end{aligned}$$

The above recursion is over the data set $Z_k = \{z_1, \ldots, z_k\}$ with z_k as the basic data block. The recursion can also be applied over each data block $z_k = \{z^1_k, \ldots, z^{n_z}_k\}$, leading to a completely sequential processing of observations. Consequently, this filter uses the recursion $k \times n_z$ times, each time with a scalar-valued measurement.

In the case where S_k is singular, S^{-1}_k in the filter can be replaced by S^+_k, the Moore-Penrose pseudoinverse of S_k. However, (10) is no longer valid when S_k is singular.

This recursive filter has many nice properties of the Kalman filter. For example, it is unbiased and optimal in the sense of having minimum variance (mean-square error) if all the measurements $z_1, z_2, \ldots$, and x_0 are jointly Gaussian and the noises are martingale difference sequences, which is commonly assumed in stochastic systems. Since this theorem gives an optimal recursion for $\{\hat{x}_{k|k}, P_{k|k}\}$, it is clear that $\{\hat{x}_{k|k}, P_{k|k}\}$ is a sufficient statistic for x_k for the jointly Gaussian case whenever the stated assumption is valid. Otherwise it is the best linear filter under the stated assumption. In practice, one can take advantage of the fact that $P_{k|k-1}$, K_k, and $P_{k|k}$ can all be calculated off-line prior to any measurement, just like in the Kalman filter. The computation required by Ψ_k and Ω_k can also be done off-line.

5 Illustrative Simulation Examples

This section presents two illustrative examples of the filters presented above. They serve primarily three purposes: (a) to verify the optimal batch and semi-recursive filters by demonstrating that they yield the same estimates as the Kalman filter does in the case where the noises are Markov; (b) to demonstrate the superiority of the filters presented to the Kalman filter when the noises are non-Markov or cross-correlated; and (c) to demonstrate both computational improvement and performance degradation of the recursive filter relative to the optimal filters.

Consider the following constant-velocity dynamic system with position-only measurements:

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_k + \begin{bmatrix} \frac{1}{2}T^2 \\ T \end{bmatrix} w_k \\ z_k &= [1, 0]\, x_k + v_k \end{aligned} \tag{12}$$

where $x = [\text{position,velocity}]'$ and $T = 5$s is the sampling interval. All results were obtained using Matlab, a platform that is quite efficient for matrix operations, e.g., matrix inverse, a key operation for the batch forms.

Results from both single and multiple Monte Carlo runs are presented. The average normalized estimation error squared (ANEES) is computed by

$$\text{ANEES}(k) = \frac{1}{\dim(x)} \frac{1}{M} \sum_{i=1}^{M} (\tilde{x}^i_{k|k})'(P^i_{k|k})^{-1} \tilde{x}^i_{k|k}$$

where $\tilde{x}^i_{k|k}$ denotes the state estimation error in the ith run, $P^i_{k|k} = \text{cov}(\tilde{x}^i_{k|k})$, $\dim(x)$ is the dimension of x_k, and M is the number of runs. It is used to check how credible the filter is. For a fixed k, its expected value and variance are 1 and $2/[\dim(x)M]$ if $\tilde{x}^i_{k|k}$ is Gaussian distributed.

Example 1 (Markov noise). Here $\{v_k\}$ is zero-mean stationary Gaussian white noise with variance $R_k = (100\text{m})^2$; $\{w_k\}$ is a zero-mean stationary Gauss-Markov sequence with autocovariance Q_{ij} =cov$(w_i, w_j) = \sigma_w^2 e^{-\alpha T|i-j|}$, where $\sigma_w = 0.4\text{m/s}^2, \alpha = 1/100$, which can be generated by

$$w_{k+1} = e^{-\alpha T} w_k + (\sigma_w \sqrt{1 - e^{-2\alpha T}})\eta_k$$

where $w_0 \sim N(0, \sigma_w^2)$ and $\{\eta_k\}$ is zero-mean Gaussian white noise with unity variance and is independent of everything else. The initial state $x_0 \sim N(\bar{x}_0, P_0)$ is random with $\bar{x}_0 = [1000, 100]'$ and P_0 =diag$(100^2, 10^2)$. $\{w_k\}$, $\{v_k\}$, x_0 are uncorrelated. The optimal filter for this example can be obtained by the Kalman filter with state augmentation:

$$x^a_{k+1} = \begin{bmatrix} 1 & T & \frac{1}{2}T^2 \\ 0 & 1 & T \\ 0 & 0 & e^{-\alpha T} \end{bmatrix} x^a_k + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} (\sigma_w \sqrt{1 - e^{-2\alpha T}})\eta_k$$
$$z_k = [1, 0, 0]\, x^a_k + v_k, \qquad x^a_k = [x'_k, w_k]'$$

It is also given by Theorems 1 and 2. Note that $\tilde{R}^k$ here is block-diagonal and thus the computational advantage of Theorem 2 is basically lost. The implementation of Theorems 1 and 2 used the formulas corresponding to zero $\Omega^{k-1}_k, \tilde{R}^{k-1}_k, \tilde{A}_k$. Similarly, the recursive filter of Theorem 3 did not compute Ω_k since it is always equal to zero here.

Example 2 (Non-Markov noise). In this example, $\{w_k\}$ is the same as in Example 1 and $\{v_k\}$ is a zero-mean nonstationary non-Markov noise sequence, generated by $v_k = \beta v_{k-1}\xi_k = \beta^k \xi_1 \cdots \xi_k v_0$, $\beta = \sqrt{3}, \xi_k \sim U(0, 1), \forall k$, independent of everything else, and $v_0 \sim N(0, \sigma_0^2)$ with $\sigma_0 = 100$. It can be shown that, for $i, j \geq 0$,

$$R_{ij} = \text{cov}(v_i, v_j) = \sigma_0^2 \beta^{i+j} \times 3^{-\min\{i,j\}} \times 0.5^{|j-i|}$$
$$C_{ij} = \text{cov}(w_i, v_j) = 0.5\sigma_0\sigma_w e^{-\alpha T i}(\beta/2)^j$$

$\{w_k\}$ and $\{v_k\}$ are correlated through w_0 and v_0 by $E[w_0 v_0] = 50\sigma_w$. $x_0 = \bar{x}_0 + [-v_0, w_0]'$ is a random variable with $\bar{x}_0 = [1000, 100]'$ and $P_0 = \begin{bmatrix} \sigma_0^2 & -0.5\sigma_0\sigma_w \\ -0.5\sigma_0\sigma_w & \sigma_w^2 \end{bmatrix}$. Note that $\{w_k\}$ and $\{v_k\}$ are correlated with x_0:

$$A_k = \text{cov}(x_0, v_k) = [-R_{0k}, C_{0k}]'$$
$$B_k = \text{cov}(x_0, w_k) = [-C_{k0}, Q_{0k}]'$$

The optimal filter for this example is given by Theorems 1 and 2. The Kalman filter implemented is the same as in Example 1, which ignores the autocorrelation of $\{v_k\}$ and the crosscorrelation among $\{w_k\}$, $\{v_k\}$, and x_0 since it is not known how to account for these types of correlation.

Note that the system (12) is of the form $x_{k+1} = F_k x_k + \Gamma_k w_k$ rather than the assumed form (1) in the theorems. Thus, one needs to replace Q_{ij}, B_j, C_{ji} in the theorems with $\Gamma Q_{ij}\Gamma', B_j\Gamma', \Gamma C_{ji}$, respectively, where $\Gamma = \left[T^2/2, T\right]'$.

Figs. 1 and 2 show the results of the proposed batch (BF), semi-recursive (SF), and recursive (RF) filters, as well as the augmented-state Kalman filter (KF). Tables 1 and 2 give the corresponding relative ratios of CPU times and floating point operation (FLOP) counts, obtained from 500 runs. As expected, BF and SF always coincide; they coincide with the optimal KF in Example 1; and RF exhibits a slightly difference as time goes. In Example 2, RF and the optimal BF and SF have virtually the same performance, along with ANEES curves that are not far away from 1, while the suboptimal KF is significantly worse, along with an ANEES much larger than 1, which indicates a model mismatch.

6 Conclusion

The problem of filtering for a linear discrete-time system in arbitrary additive noise (not necessarily Markov) has been considered where the noise sequences have arbitrary autocorrelations, crosscorrelation, and correlations with the initial state. Optimal batch and a semi-recursive filters and a suboptimal recursive filter have been presented. They are generalizations of the Kalman filter for the case of arbitrary additive noise of known first two moments in the most general setting. The recursive filter resembles the Kalman filter. Numerical examples have demonstrated that the proposed recursive filter is efficient and has insignificant performance degradation.

References

[1] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques, and Software*. Boston, MA: Artech House, 1993. (Reprinted by YBS Publishing, 1998).

[2] Y. Bar-Shalom and X. R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[3] C. Y. Chong, S. Mori, and K. C. Chang, "Distributed Multitarget Multisensor Tracking," in *Multitarget-Multisensor Tracking: Advanced Applications*, (Y. Bar-Shalom, ed.), pp. 247–295, Norwood, MA: Artech House, 1990.

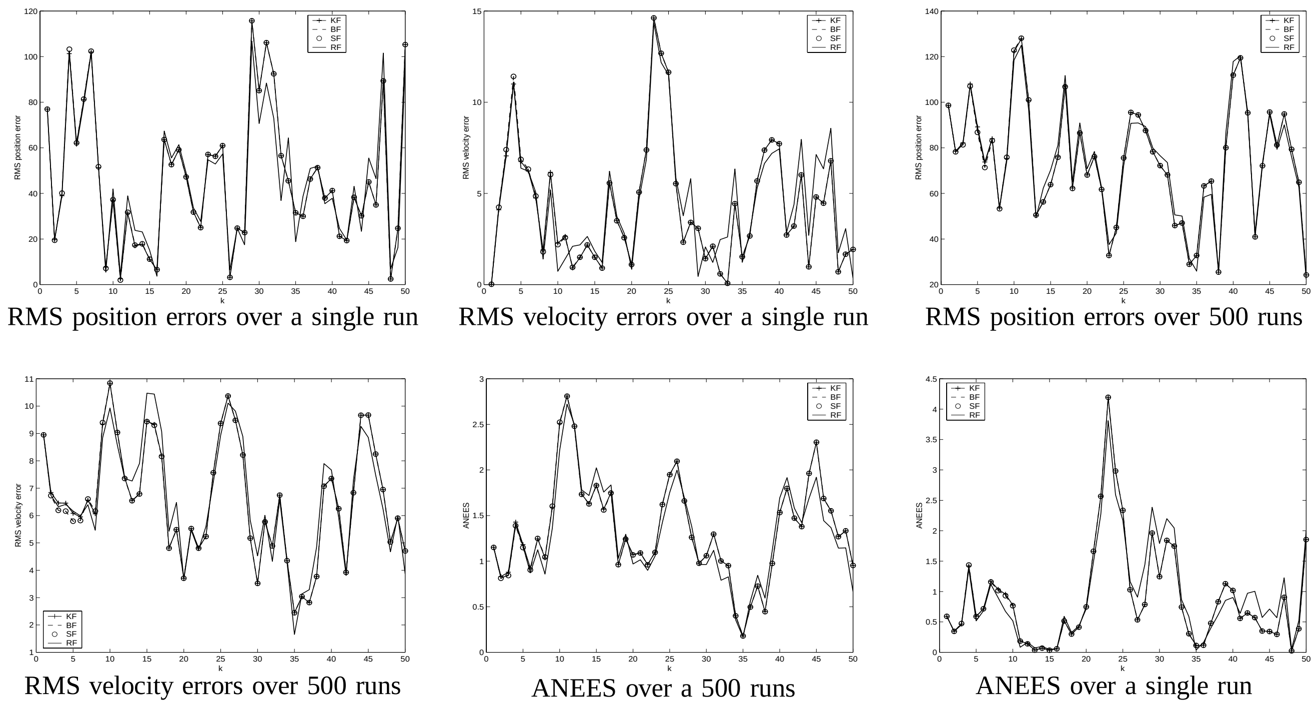
[4] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *ASME J. Basic Engineering*, vol. Series 82D, pp. 35–40, Mar. 1960.

[5] X. R. Li, "Hybrid Estimation Techniques," in *Control and Dynamic Systems: Advances in Theory and Applications*, vol. 76, (C. T. Leondes, ed.), pp. 213–287, New York: Academic Press, 1996.

[6] X. R. Li, Y. M. Zhu, and C. Z. Han, "Unified Optimal Linear Estimation Fusion—Part I: Unified Models and Fusion Rules," in *Proc. 2000 International Conf. on Information Fusion*, (Paris, France), July 2000.

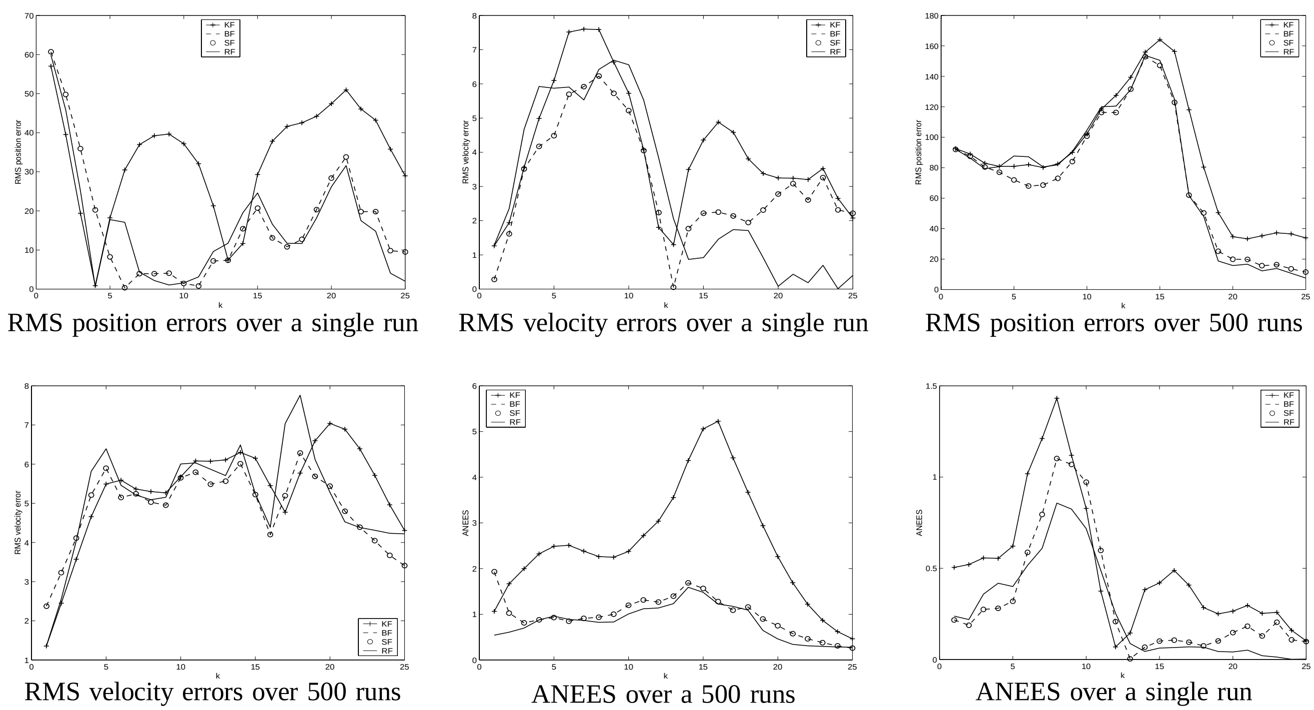
[7] M. Liggins, C. Y. Chong, I. Kadar, M. G. Alford, V. Vannicola, and S. Thomopoulos, "Distributed Fusion Architectures and Algorithms for Target Tracking," *Proceedings of the IEEE*, vol. 85, pp. 95–107, Jan. 1997.

[8] H. W. Sorenson, ed., *Kalman Filtering: Theory and Application*. IEEE Press, 1985.



RMS position errors over a single run
RMS velocity errors over a single run
RMS position errors over 500 runs
RMS velocity errors over 500 runs
ANEES over a 500 runs
ANEES over a single run

Fig. 1: RMS errors and the ANEES of Example 1.



RMS position errors over a single run
RMS velocity errors over a single run
RMS position errors over 500 runs
RMS velocity errors over 500 runs
ANEES over a 500 runs
ANEES over a single run

Fig. 2: RMS errors and the ANEES of Example 2.

Table 1. Computational complexities of four filters (Example 1)

	batch	semi-recursive	recursive	Kalman
cpu time	1	0.847	0.256	0.042
FLOPs	1	0.777	0.01	0.004

Table 2. Computational complexities of four filters (Example 2)

	batch	semi-recursive	recursive	Kalman
cpu time	1	0.883	0.370	0.033
FLOPs	1	0.780	0.014	0.004