Reconstructing Estimates from Noisy Transmissions with Serially-Connected Kalman Filters

D. Richard Brown III
Electrical and Computer Engineering Department
Worcester Polytechnic Institute
Worcester, MA 01609 USA
drb@wpi.edu

Yaakov Bar-Shalom
Electrical and Computer Engineering Department
University of Connecticut
Storrs, CT 06269 USA
ybs@ee.uconn.edu

***Abstract*—This paper considers the problem of tracking a time-varying variable with serially-connected Kalman filters. Two nodes are assumed to be serially connected to the target such that only node 1 can directly observe a noisy signal from the target. Node 2 can only observe noisy signals from node 1 corresponding to a linear combination of the current observation and current state estimate at node 1. The objective is to find the linear combination at node 1 that minimizes the mean squared error of the state estimates at node 2 under a transmit power constraint for the signals from node 1 to node 2. An augmented state model is developed to facilitate tracking at node 2. Transmission scaling factors are also derived to satisfy the power constraint. Numerical results are presented for two-node serial tracking in two scenarios: scalar parameter tracking and two-state oscillator phase and frequency tracking. In the scalar parameter tracking example, the results demonstrate that a non-trivial combination of the observation and state estimate at node 1 can improve performance at node 2 with respect to a baseline scenario of simply forwarding scaled observations. In the two-state clock tracking example, an optimal transmission strategy is developed which allows node 2 to achieve the same tracking performance as at node 1.**

TABLE OF CONTENTS

1. INTRODUCTION

Consider a discrete-time linear time-invariant dynamic system with state $x[k] \in \mathbb{R}^n$ and state update equation

$$x[k+1] = Fx[k] + Gu[k] \tag{1}$$

with $F \in \mathbb{R}^{n\times n}$ and $G \in \mathbb{R}^{n\times p}$ known and where $u[k] \in \mathbb{R}^p$ is zero-mean Gaussian process noise with covariance

$$\mathrm{E}[u[k]u^\top[\ell]] = Q\delta[k-\ell] \tag{2}$$

where $Q \in \mathbb{R}^{p\times p}$ and $\delta[k-\ell] = 1$ if $k = \ell$ and is equal to zero otherwise.

As shown in Fig. 1, we consider a scenario in which there are two nodes in the system that wish to estimate/predict the state $x[k]$. We assume observations at each node are scalar. Node 1 observes the state $x[k]$ directly through the equation

$$y_1[k] = H_1x[k] + w_1[k] \tag{3}$$

with $y_1[k] \in \mathbb{R}$, $H_1 \in \mathbb{R}^{1\times n}$, and where $w_1[k] \in \mathbb{R}$ is zero-mean Gaussian measurement noise with covariance

$$\mathrm{E}\left[w_1[k]w_1[\ell]\right] = r_1\delta[k-\ell]. \tag{4}$$

We assume that the measurement noise $w_1[k]$ is independent of the process noise $u[k]$. Node 1 uses a Kalman filter to generate minimum mean squared error (MMSE) estimates and one-step predictions of the state $x[k]$. The MMSE estimate of the state $x[k]$ given observations $\{y_1[0], \ldots, y_1[k]\}$ at node 1 is denoted as $\hat{x}_1[k|k]$. The MMSE prediction of the state $x[k+1]$ given observations $\{y_1[0], \ldots, y_1[k]\}$ at node 1 is denoted as $\hat{x}_1[k+1|k]$. According to the standard Kalman filter recursion [1] with $\Sigma_1[k|k]$ and $\Sigma_1[k+1|k]$ denoting the estimate and one-step prediction covariances at node 1, respectively, we have

$$K_1[k] = \Sigma_1[k|k-1]H_1^\top\left(H_1\Sigma_1[k|k-1]H_1^\top + r_1\right)^{-1} \tag{5}$$

$$\hat{x}_1[k|k] = \hat{x}_1[k|k-1] + K_1[k](y_1[k] - H_1\hat{x}_1[k|k-1]) \tag{6}$$

$$\Sigma_1[k|k] = \Sigma_1[k|k-1] - K_1[k]H_1\Sigma_1[k|k-1] \tag{7}$$

and

$$\hat{x}_1[k+1|k] = F\hat{x}_1[k|k] \tag{8}$$

$$\Sigma_1[k+1|k] = F\Sigma_1[k|k]F^\top + GQG^\top. \tag{9}$$

Node 2 also wants to track the state $x[k]$ but can only observe the state through noisy scalar transmissions from node 1. In other words, node 2 observes

$$y_2[k] = v_1[k] + w_2[k] \tag{10}$$

where $v_1[k]$ is a scalar transmitted by node 1 to node 2 at time k, based only on information known to node 1 at time k, and where $w_2[k]$ is zero-mean Gaussian measurement noise, assumed to be independent of $w_1[k]$ and $u[k]$, with covariance

$$\mathrm{E}[w_2[k]w_2[\ell]] = r_2\delta[k-\ell]. \tag{11}$$

We assume that transmissions from node 1 to node 2 are linear combinations of the current observation and estimate at node 1 of the form

$$v_1[k] = a^\top[k]\begin{bmatrix} y_1[k] \\ \hat{x}_1[k|k] \end{bmatrix} = a^\top[k]z[k] \tag{12}$$

Proc. IEEE Aerospace Conf, Big Sky, MT, March 2015.
This work was supported by the National Science Foundation awards CCF-1302104 and CCF-1319458 and Army Research Office award W991NF-10-1-0369.
978-1-4799-5380-6/15/$31.00 ©2015 IEEE.

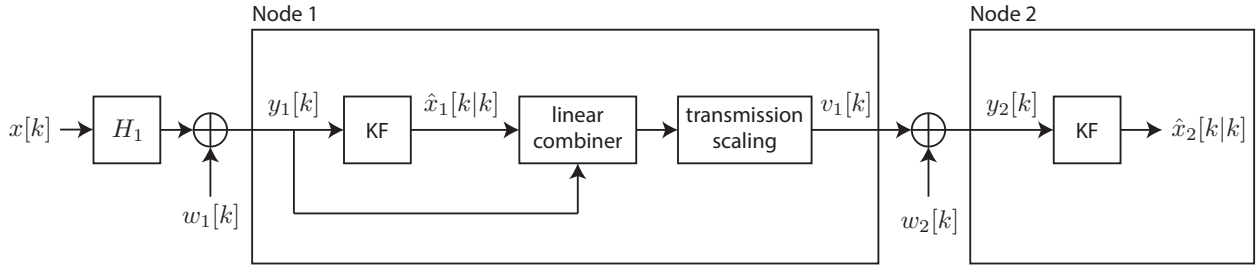

Figure 1. System model.

with $z[k] \in \mathbb{R}^{n+1}$. The goal is to select the linear combiner $a[k]$ to minimize the mean squared state estimation error at node 2 with respect to the original state $x[k]$ while satisfying an average power constraint such that $\text{var}(v_1[k]) \leq P[k]$. More formally, we wish to solve

$$a^*[k] = \arg \min_{a[k]:\, \text{var}(v_1[k]) \leq P[k]} \text{E}\left[|x[k] - \hat{x}_2[k|k]|^2\right] \quad (13)$$

where $\hat{x}_2[k|k]$ are the MMSE estimates of $x[k]$ given observations $\{y_2[0], \ldots, y_2[k]\}$ at node 2. The power constraint $P[k]$ may be fixed or time-varying.

A Simple Forwarding Strategy

Since the observations at both nodes are assumed to be scalar, a simple forwarding strategy is for node 1 to directly forward scaled copies its observations to node 2, i.e.,

$$v_1[k] = \alpha[k] y_1[k] = \alpha[k] \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} z[k] \quad (14)$$

with $\alpha[k]$ chosen to satisfy the power constraint $\text{var}(v_1[k]) \leq P[k]$ and with $z[k]$ defined in (12). Node 2 can then track the state at node 1 using the same Kalman filter recursion as in (5) – (9) except replacing r_1 by $\alpha^2[k] r_1 + r_2$ in (5) and scaling H_1 by $\alpha[k]$ in (5) – (7).

This simple forwarding strategy will serve as a benchmark for quantifying the performance gain achieved through optimal forwarding as in (13).

Related Prior Work

The forwarding strategy employed by node 1 is conceptually similar to a constrained measurement strategy at node 2. Measurement strategies for discrete-time state estimators have been considered in [2] where the problem was to optimally select a single sensor measurement from a pool of available sensors. The sensor selection problem has been extensively studied in a variety of contexts under various constraints, e.g., [3], [4], [5], [6]. The problem considered in this paper is somewhat similar to the sensor selection problem in that our goal is to find a measurement strategy at node 2 (equivalently, a forwarding strategy at node 1) that minimizes the mean squared error at node 2. The majority of the sensor selection literature, however, considers combinatorial measurement strategies, e.g., selecting k sensor measurements from a pool of $m > k$ possible sensors. In this paper, the measurement strategy in (12) can be any linear combination of the observation and the state estimate at node 1. Moreover, in the setting considered in this paper, the "sensors" represented by the observation and the state estimate at node 1 are highly correlated.

Another approach toward designing good measurement strategies for parameter tracking involves minimizing the condition number of the observability matrix [7]. This has been considered in the context of the sensor placement problem, e.g., [8], [9], [10], [11]. This approach is heuristic, however, and does not necessarily minimize the mean squared error at node 2 in our setting. In this paper, we show an example of a system in which a minimum condition number strategy does not yield the minimum mean squared error performance at node 2.

2. TRACKING AT NODE 2

From the system model in Section 1, recall that the observation at node 2 is a linear combination of the observation $y_1[k]$ and the state estimate $\hat{x}_1[k|k]$ plus additive white Gaussian noise $w_2[k]$. We define the state estimation error as

$$\tilde{x}_1[k|k] = x[k] - \hat{x}_1[k|k]. \quad (15)$$

and further define the augmented state at node 2 as

$$\bar{x}[k] = \begin{bmatrix} x[k] \\ \tilde{x}_1[k|k] \\ w_1[k] \end{bmatrix} \in \mathbb{R}^{2n+1}. \quad (16)$$

As shown in (17) – (24), the estimation error $\tilde{x}[k|k]$ at node 1 has its own linear but time-varying dynamics of the form

$$\tilde{x}_1[k+1|k+1] = \tilde{F}[k] \tilde{x}_1[k|k] + \tilde{G}[k] \begin{bmatrix} u[k] \\ w_1[k+1] \end{bmatrix} \quad (25)$$

with

$$\tilde{F}[k] = (I - K_1[k+1] H_1) F \quad (26)$$

and

$$\tilde{G}[k] = \begin{bmatrix} (I - K_1[k+1] H_1) G & -K_1[k+1] \end{bmatrix} \quad (27)$$

From these results, we can write the time-varying augmented state model at node 2 as

$$\bar{x}[k+1] = \bar{F}[k] \bar{x}[k] + \bar{G}[k] \bar{u}[k] \quad (28)$$

with

$$\bar{F}[k] = \begin{bmatrix} F & 0 & 0 \\ 0 & \tilde{F}[k] & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad (29)$$

$$\bar{G}[k] = \begin{bmatrix} G & 0 \\ (I - K_1[k+1] H_1) G & -K_1[k+1] \\ 0 & 1 \end{bmatrix}, \quad (30)$$

and

$$\bar{u}[k] = \begin{bmatrix} u[k] \\ w_1[k+1] \end{bmatrix} \quad (31)$$

$$\tilde{x}_1[k+1|k+1] = x[k+1] - \hat{x}_1[k+1|k+1] \tag{17}$$
$$= x[k+1] - \hat{x}_1[k+1|k] - K_1[k+1]\left(y_1[k+1] - H_1\hat{x}_1[k+1|k]\right) \tag{18}$$
$$= Fx[k] + Gu[k] - F\hat{x}_1[k|k] - K_1[k+1]\left(y_1[k+1] - H_1F\hat{x}_1[k|k]\right) \tag{19}$$
$$= F\tilde{x}_1[k|k] + Gu[k] - K_1[k+1]\left(H_1x[k+1] + w_1[k+1] - H_1F\hat{x}_1[k|k]\right) \tag{20}$$
$$= F\tilde{x}_1[k|k] + Gu[k] - K_1[k+1]\left(H_1Fx[k] + H_1Gu[k] + w_1[k+1] - H_1F\hat{x}_1[k|k]\right) \tag{21}$$
$$= F\tilde{x}_1[k|k] + Gu[k] - K_1[k+1]\left(H_1F\tilde{x}_1[k|k] + H_1Gu[k] + w_1[k+1]\right) \tag{22}$$
$$= \left(I - K_1[k+1]H_1\right)\left(F\tilde{x}_1[k|k] + Gu[k]\right) - K_1[k+1]w_1[k+1] \tag{23}$$
$$= \tilde{F}[k]\tilde{x}_1[k|k] + \tilde{G}[k]\begin{bmatrix} u[k] \\ w_1[k+1] \end{bmatrix} \tag{24}$$

where $\bar{u}[k]$ is zero-mean white Gaussian process noise with

$$\bar{Q} = \mathrm{E}[\bar{u}[k]\bar{u}^\top[k]] = \mathrm{diag}(Q, r_1). \tag{32}$$

The observations at node 2 follow as

$$y_2[k] = a^\top[k]\begin{bmatrix} y_1[k] \\ \hat{x}_1[k|k] \end{bmatrix} + w_2[k] \tag{33}$$
$$= \begin{bmatrix} a_1[k] & a_2^\top[k] \end{bmatrix}\begin{bmatrix} H_1x[k] + w_1[k] \\ x[k] - \tilde{x}_1[k|k] \end{bmatrix} + w_2[k] \tag{34}$$
$$= \bar{H}[k]\bar{x}[k] + w_2[k] \tag{35}$$

with $a_1[k] \in \mathbb{R}$, $a_2[k] \in \mathbb{R}^n$, and

$$\bar{H}[k] = \begin{bmatrix} a_1[k]H_1 + a_2^\top[k] & -a_2^\top[k] & a_1[k] \end{bmatrix}. \tag{36}$$

Hence, at node 2 we have a linear time-varying augmented state-space model which satisfies the requirements of a standard filtering problem with white and mutually independent process and measurement noises. This model can be used in conjunction with a Kalman filter to generate MMSE estimates of the original state $x[k]$. The MMSE estimate of the state $\bar{x}[k]$ given observations $\{y_2[0], \ldots, y_2[k]\}$ at node 2 is denoted as $\hat{x}_2[k|k]$. The MMSE prediction of the state $\bar{x}[k+1]$ given observations $\{y_2[0], \ldots, y_2[k]\}$ at node 2 is denoted as $\hat{x}_2[k+1|k]$. With $\Sigma_2[k|k]$ and $\Sigma_2[k+1|k]$ denoting the estimate and one-step prediction covariances at node 2, respectively, we have the Kalman filter recursion at node 2 given as

$$K_2[k] = \Sigma_2[k|k-1]\bar{H}^\top\left(\bar{H}\Sigma_2[k|k-1]\bar{H}^\top + r_2\right)^{-1} \tag{37}$$
$$\hat{x}_2[k|k] = \hat{x}_2[k|k-1] + K_2[k]\left(y_2[k] - \bar{H}\hat{x}_2[k|k-1]\right) \tag{38}$$
$$\Sigma_2[k|k] = \Sigma_2[k|k-1] - K_2[k]\bar{H}\Sigma_2[k|k-1] \tag{39}$$

and

$$\hat{x}_2[k+1|k] = \bar{F}\hat{x}_2[k|k] \tag{40}$$
$$\Sigma_2[k+1|k] = \bar{F}\Sigma_2[k|k]\bar{F}^\top + \bar{G}\bar{Q}\bar{G}^\top. \tag{41}$$

An expression for the initial prediction covariance $\Sigma_2[0|-1]$ at node 2 prior to the first observation $y_2[0]$ is provided in the Appendix.

3. TRANSMISSION SCALING

This section develops an expression for the transmission power $\mathrm{var}(v_1[k])$ to facilitate transmission scaling to satisfy the average transmit power constraint $\mathrm{var}(v_1[k]) \leq P[k]$. From (12), the variance of $v_1[k]$ can be written as

$$\mathrm{var}(v_1[k]) = a^\top[k]\mathrm{cov}\left(z[k], z[k]\right)a[k] \tag{42}$$
$$= a^\top[k]\Gamma[k]a[k] \tag{43}$$

where

$$\Gamma[k] = \begin{bmatrix} \mathrm{cov}\left(y_1[k], y_1[k]\right) & \mathrm{cov}\left(y_1[k], \hat{x}_1[k|k]\right) \\ \mathrm{cov}\left(\hat{x}_1[k|k], y_1[k]\right) & \mathrm{cov}\left(\hat{x}_1[k|k], \hat{x}_1[k|k]\right) \end{bmatrix} \tag{44}$$

is the "open loop" covariance of the observations and state estimates used to compute the appropriate transmission scaling to satisfy the power constraint.

It is straightforward to compute

$$\mathrm{cov}\left(y_1[k], y_1[k]\right) = H_1\Sigma_x[k]H_1^\top + r_1 \tag{45}$$

where $\Sigma_x[k] = \mathrm{cov}\left(x[k], x[k]\right)$ is the covariance of the state at time k which follows the dynamics

$$\Sigma_x[k+1] = F\Sigma_x[k]F^\top + Q \tag{46}$$

with initial covariance $\Sigma_x[0]$. To compute the remaining terms in $\Gamma[k]$, we can use the principle of orthogonality. The principle of orthogonality implies that

$$\mathrm{E}\left[(x[k] - \hat{x}_1[k|k])y_1^\top[k]\right] = 0 \tag{47}$$

which is equivalent to

$$\mathrm{E}\left[x[k]y_1^\top[k]\right] = \mathrm{E}\left[\hat{x}_1[k|k]y_1^\top[k]\right]. \tag{48}$$

Since all quantities are zero-mean, we have

$$\mathrm{cov}\left(\hat{x}_1[k|k], y_1[k]\right) = \mathrm{cov}\left(x[k], y_1[k]\right) \tag{49}$$
$$= \Sigma_x[k]H_1^\top. \tag{50}$$

We now compute an expression for $\mathrm{cov}\left(\hat{x}_1[k|k], \hat{x}_1[k|k]\right)$. From (15), we have

$$\Sigma_1[k|k] = \Sigma_x[k] - 2\mathrm{cov}(x[k], \hat{x}_1[k|k]) + \mathrm{cov}(\hat{x}_1[k|k], \hat{x}_1[k|k]). \tag{51}$$

Observe that

$$\mathrm{cov}(x[k], \hat{x}_1[k|k]) = \mathrm{cov}(\hat{x}_1[k|k] + \tilde{x}_1[k|k], \hat{x}_1[k|k]) \tag{52}$$
$$= \mathrm{cov}(\hat{x}_1[k|k], \hat{x}_1[k|k]) \tag{53}$$

where the second equality follows from the principle of orthogonality. This then implies

$$\mathrm{cov}(\hat{x}_1[k|k], \hat{x}_1[k|k]) = \Sigma_x[k] - \Sigma_1[k|k]. \tag{54}$$

Summarizing these results, we have

$$\Gamma[k] = \begin{bmatrix} H_1\Sigma_x[k]H_1^\top + r_1 & H_1\Sigma_x[k] \\ \Sigma_x[k]H_1^\top & \Sigma_x[k] - \Sigma_1[k|k] \end{bmatrix}. \tag{55}$$

Since there is no advantage to not using all of the available transmission power at each time k, we can set

$$\mathrm{var}(v_1[k]) = a^\top[k]\Gamma[k]a[k] = P[k]. \tag{56}$$

In the case of a fixed mixing vector b satisfying $b^\top\Gamma[k]b > 0$ for all k, we can write $a[k] = \alpha[k]b$ with scale factor $\alpha[k]$ chosen to satisfy the transmit power constraint. The required transmission scaling $\alpha[k]$ can be calculated as

$$\alpha[k] = \sqrt{\frac{P[k]}{b^\top\Gamma[k]b}}. \tag{57}$$

Note that $\Sigma_x[k]$ can grow without bound if F has one or more eigenvalues with magnitude greater than or equal to one. Consequently, elements of $\Gamma[k]$ may also grow without bound. To avoid causing $\alpha[k] \to 0$ in this case, it is necessary either to select b to be orthogonal to the unbounded modes of $\Gamma[k]$ or to allow $P[k]$ to scale at the same rate as $b^\top\Gamma[k]b$. An example with unbounded $\Sigma_x[k]$ is considered in more detail in Section 5.

4. STEADY-STATE ANALYSIS

At node 1, since the state-space model is time-invariant, standard Riccati techniques can be used to solve for the unique positive definite steady-state prediction covariance under standard observability and controllability assumptions [1].

At node 2, the state-space model is time varying. Nevertheless, if a steady-state solution is achieved at node 1, each of the matrices at node 2 approaches a steady-state value as $k \to \infty$. Observe that

$$\bar{F}[k] \to \begin{bmatrix} F & 0 & 0 \\ 0 & (I - K_1^{\mathrm{ss}}H_1)F & 0 \\ 0 & 0 & 0 \end{bmatrix} = \bar{F}^{\mathrm{ss}} \tag{58}$$

$$\bar{G}[k] \to \begin{bmatrix} G & 0 \\ (I - K_1^{\mathrm{ss}}H_1)G & -K_1^{\mathrm{ss}} \\ 0 & 1 \end{bmatrix} = \bar{G}^{\mathrm{ss}} \tag{59}$$

where K_1^{ss} is the steady-state Kalman gain at node 1.

In the case of a fixed mixing vector b with $a[k] = \alpha[k]b$, if the limit

$$\alpha^{\mathrm{ss}} = \lim_{k\to\infty} \sqrt{\frac{P[k]}{b^\top\Gamma[k]b}} \tag{60}$$

exists, then the mixing vector $a[k]$ converges to

$$a[k] \to \alpha^{\mathrm{ss}}b = \alpha^{\mathrm{ss}} \begin{bmatrix} b_y \\ b_x \end{bmatrix} \tag{61}$$

with $b_y \in \mathbb{R}$ and $b_x \in \mathbb{R}^n$. Consequently, the observation matrix at node 2 converges to

$$\bar{H}[k] \to \alpha^{\mathrm{ss}} \begin{bmatrix} b_y^\top H_1 + b_x^\top & -b_x^\top & b_y^\top \end{bmatrix} = \bar{H}^{\mathrm{ss}}. \tag{62}$$

If the standard observability and controllability conditions are satisfied, we can solve the Riccati equation for the unique positive definite steady-state prediction covariance at node 2 using the limiting matrix values.

5. NUMERICAL EXAMPLES

This section provides two numerical examples demonstrating the performance of serially connected Kalman filters. In the first example, we consider a simple stable scalar system. This example illustrates the main ideas with minimum notational overhead and shows that a nontrivial combination of the observation and state is optimal for minimizing the mean squared one-step prediction and estimation errors at node 2. We also show that minimizing the condition number of the observability matrix at node 2 does not necessarily result in the minimum MSE at node 2. The second example considers a two-state oscillator tracking problem. The state dynamics are unstable in this problem, but we show under a transmit power constraint proportional to the power of the observations at node 1, node 2 can achieve identical mean squared one-step prediction and estimation error to that at node 1. This result is somewhat surprising since it implies there is no loss of steady-state performance by serially transmitting an appropriate mixture of the observation and state from node 1 to node 2.

Stable Scalar System

We first consider a stable dynamic system governed by

$$x[k+1] = Fx[k] + Gu[k] \tag{63}$$

with $F = 0.95$, $G = 1$, and $Q = \mathrm{E}[u^2[k]] = 0.04$. Node 1 observes

$$y_1[k] = H_1x[k] + w_1[k] \tag{64}$$

with $H_1 = 1$ and $r_1 = \mathrm{E}[w_1^2[k]] = 0.04$. The initial state is distributed as $x[0] \sim \mathcal{N}(0, 1)$. The steady-state variables at node 1 can be computed via the discrete-time algebraic Riccati equation as

$$\lim_{k\to\infty} \Sigma_1[k+1|k] \approx 0.0619 \tag{65}$$

$$\lim_{k\to\infty} K_1[k] = K_1^{\mathrm{ss}} \approx 0.6076 \tag{66}$$

$$\lim_{k\to\infty} \Sigma_1[k|k] \approx 0.0243. \tag{67}$$

Figure 2 shows the one-step prediction and estimation performance of oscillator tracking at node 1 and node 2 in the baseline case with a mixing vector of the form $a[k] = \alpha[k][1, 0]^\top$. This mixing vector corresponds to simply forwarding scaled copies of the noisy observations at node 1 to node 2. We assume a time-varying power constraint

$$P[k] = \mathrm{var}(y_1[k]) = H_1\Sigma_x[k]H_1^\top + r_1 \tag{68}$$

which corresponds to node 2 forwarding observations to node 1 at the same average power that they are received at node 1. Since this system is stable, $P[k]$ will be finite for all k. This power constraint also implies that $\alpha[k] = 1$ for all k and

$$\bar{H}[k] = [1 \quad 0 \quad 1] \tag{69}$$

for all k. We can compute the steady-state state variance by solving the discrete Lyapunov equation

$$FXF^\top - X + Q = 0 \tag{70}$$

which results in steady-state values for the state variance, power constraint, and observation/state covariance as

$$\lim_{k\to\infty} \Sigma_x[k] \approx 0.4103 \tag{71}$$

$$\lim_{k\to\infty} P[k] \approx 0.4503 \tag{72}$$

$$\lim_{k\to\infty} \Gamma[k] \approx \begin{bmatrix} 0.4503 & 0.4103 \\ 0.4103 & 0.3860 \end{bmatrix} \tag{73}$$

respectively. The resulting steady-state variables at node 2 can be computed as

$$\lim_{k\to\infty} \Sigma_2[k+1|k] \approx \begin{bmatrix} 0.0749 & 0.0243 & 0.0000 \\ 0.0243 & 0.0243 & -0.0243 \\ 0.0000 & -0.0243 & 0.0400 \end{bmatrix} \tag{74}$$

$$\lim_{k\to\infty} K_2[k] \approx \begin{bmatrix} 0.4836 \\ 0.0000 \\ 0.2582 \end{bmatrix} \tag{75}$$

and

$$\lim_{k\to\infty} \Sigma_2[k|k] \approx \begin{bmatrix} 0.0387 & 0.0243 & -0.0193 \\ 0.0243 & 0.0243 & -0.0243 \\ -0.0193 & -0.0243 & 0.0297 \end{bmatrix} \tag{76}$$

The (1,1) elements of $\lim_{k\to\infty} \Sigma_2[k+1|k]$ and $\lim_{k\to\infty} \Sigma_2[k|k]$ are plotted as dashed lines in Figure 2.

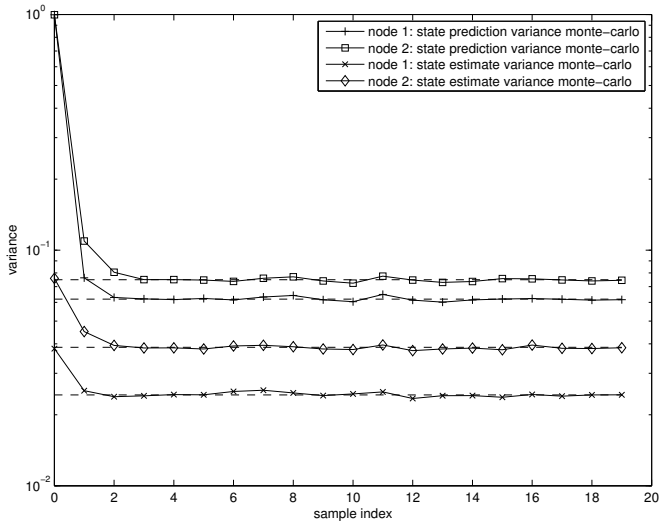

Figure 2. One-step prediction variance and estimation variance in the stable scalar system scenario. Dashed lines correspond to steady-state performance calculations via the discrete-time algebraic Riccati equation.

We now consider optimizing the mixing vector to minimize the mean squared one-step prediction and estimation error at node 2. We consider a mixing vector of the form

$$a[k] = \alpha[k] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \in \mathbb{R}^2 \tag{77}$$

with $b_1 > 0$ and $\alpha[k]$ calculated to satisfy the power constraint $P[k]$. This leads to an observation matrix at node 2 of the form

$$\bar{H}[k] = \alpha[k] \begin{bmatrix} b_1 + b_2 & -b_2 & b_1 \end{bmatrix}. \tag{78}$$

Figure 3 shows the steady-state one-step prediction variance and estimation variance at node 2 for this stable scalar system example as a function of the mixing parameter ratio b_2/b_1. The minimum occurs at $b_2/b_1 \approx -0.787$ and the resulting steady-state variables at node 2 can be computed as

$$\lim_{k\to\infty} \Sigma_2[k+1|k] \approx \begin{bmatrix} 0.0711 & 0.0243 & 0.0000 \\ 0.0243 & 0.0243 & -0.0243 \\ 0.0000 & -0.0243 & 0.0400 \end{bmatrix} \tag{79}$$

$$\lim_{k\to\infty} K_2[k] \approx \begin{bmatrix} 0.3326 \\ 0 \\ 0.2026 \end{bmatrix} \tag{80}$$

and

$$\lim_{k\to\infty} \Sigma_2[k|k] \approx \begin{bmatrix} 0.0344 & 0.0243 & -0.0223 \\ 0.0243 & 0.0243 & -0.0243 \\ -0.0223 & -0.0243 & 0.0264 \end{bmatrix} \tag{81}$$

By optimally mixing the observation and the estimate, we observe approximately a 5% reduction in the one-step state prediction variance and an 11% reduction in the state estimation variance at node 2 with respect to the baseline case of simply forwarding noisy observations.

There is also a sharp maximum in the prediction/estimation variances at $b_2/b_1 \approx -1.063$. With this mixing parameter, the steady-state variables at node 2 can be computed as

$$\lim_{k\to\infty} \Sigma_2[k+1|k] \approx \begin{bmatrix} 0.4103 & 0.0243 & -0.0000 \\ 0.0243 & 0.0243 & -0.0243 \\ -0.0000 & -0.0243 & 0.0400 \end{bmatrix} \tag{82}$$

$$\lim_{k\to\infty} K_2[k] \approx \begin{bmatrix} -0.0001 \\ 0 \\ 0.1629 \end{bmatrix} \tag{83}$$

and

$$\lim_{k\to\infty} \Sigma_2[k|k] \approx \begin{bmatrix} 0.4103 & 0.0243 & 0.0000 \\ 0.0243 & 0.0243 & -0.0243 \\ 0.0000 & -0.0243 & 0.0270 \end{bmatrix}. \tag{84}$$

In this case, the estimation and prediction covariances are almost identical except with regards to the third state corresponding to $w_1[k]$ and are considerably worse than the baseline case of simply forwarding noisy observations.

Figure 4 shows the condition number of the observability matrix resulting from the steady-state matrices $\bar{F}^{\text{ss}}$ and $\bar{H}^{\text{ss}}$ at node 2 as a function of the fixed mixing vector $b = [b_1, b_2]^\top$ with $b_1 > 0$ and $a[k] = \alpha[k]b$. While the condition number attains a local minimum at the MMSE-optimal operating point $b_2/b_1 \approx -0.787$, the global minimum condition number occurs at $b_2/b_1 \approx -1.5$. As shown in Fig. 3, setting $b_2/b_1 \approx -1.5$ results in a significantly worse MMSE at node 2 than even the simple forwarding strategy of retransmitting noisy observations.

Oscillator Tracking System

In this section, we consider an oscillator tracking scenario with discrete-time state

$$x[k] = \begin{bmatrix} \phi[k] \\ \dot{\phi}[k] \end{bmatrix} \tag{85}$$

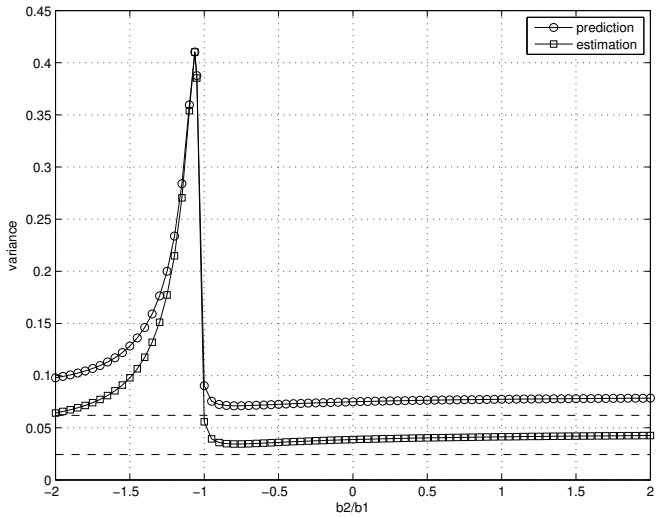

Figure 3. Steady-state one-step prediction variance in the stable scalar system scenario as a function of the fixed mixing vector $b = [b_1, b_2]^\top$ with $b_1 > 0$ and $a[k] = \alpha[k]b$. The dashed lines are the steady-state one-step prediction variance and estimation variance at node 1.

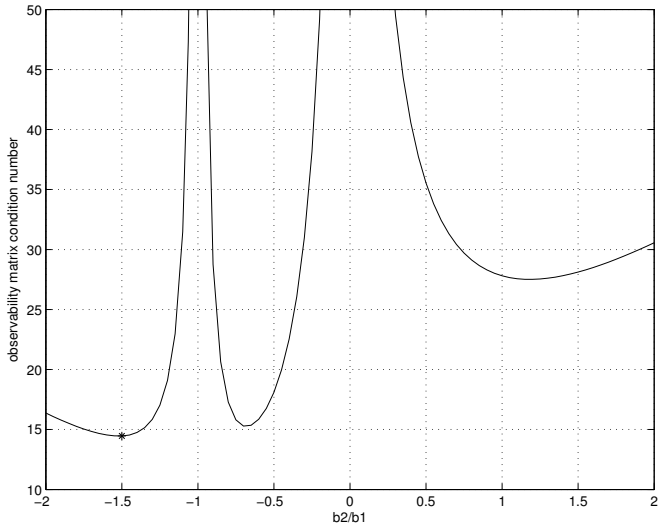

Figure 4. Condition number of the observability matrix resulting from the steady-state matrices $\bar{F}^{\rm ss}$ and $\bar{H}^{\rm ss}$ as a function of the fixed mixing vector $b = [b_1, b_2]^\top$ with $b_1 > 0$ and $a[k] = \alpha[k]b$. Observe the minimum condition number occurs at $b_2/b_1 \approx -1.5$.

where $\phi[k]$ and $\dot{\phi}[k]$ correspond to the phase offset in radians and frequency offset in radians per second of an oscillator with respect to a ideal reference. Based on the two-state models in [12], [13], the state update is given as

$$x[k+1] = F(T)x[k] + u[k] \tag{86}$$

with

$$F(T) = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{87}$$

where $T > 0$ is an arbitrary sampling period. The process noise vector $u[k] \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, Q(T))$ corresponds to the white frequency and random walk frequency process noises that cause the the local oscillator to deviate from an ideal linear phase trajectory. The covariance of the discrete-time process noise is derived from a continuous-time model in [12] and is given as

$$Q(T) = \omega_0^2 \begin{bmatrix} q_1^2 T + q_2^2 \frac{T^3}{3} & q_2^2 \frac{T^2}{2} \\ q_2^2 \frac{T^2}{2} & q_2^2 T \end{bmatrix} \tag{88}$$

where ω_0 is the nominal oscillator frequency in radians per second and q_1^2 (units of seconds) and q_2^2 (units of Hertz) are the process noise parameters corresponding to white frequency noise and random walk frequency noise, respectively. The process noise parameters q_1^2 and q_2^2 can be estimated by fitting the theoretical Allan variance

$$\sigma_y^2(\tau) = \frac{q_1^2}{\tau} + \frac{q_2^2 \tau}{3} \tag{89}$$

to experimental measurements of the Allan variance over a range of τ values. For example, the Allan variance specifications for a Rakon RPFO45 oven-controlled oscillator [14] are given in Table 1. A least squares fit of (89) to these specifications yields $q_1^2 = 2.31 \times 10^{-21}$ and $q_2^2 = 6.80 \times 10^{-23}$.

Table 1. Allan variance specifications for the Rakon RPFO45 oven-controlled oscillator.

Elapsed time τ (sec)	Allan variance specification
0.1	2.25×10^{-20}
1	0.81×10^{-20}
10	0.36×10^{-20}
100	0.36×10^{-20}
1000	2.25×10^{-20}

Figures 5 and 6 show the one-step prediction and estimation performance of oscillator tracking at node 1 and node 2 assuming $q_1^2 = 2.31 \times 10^{-21}$ sec, $q_2^2 = 6.80 \times 10^{-23}$ Hz, $\omega_0 = 2\pi \cdot 900 \times 10^6$ rad/sec, and $T = 1$ sec. The observation matrix at node 1 was set to

$$H_1 = \begin{bmatrix} 1 & 0 \end{bmatrix} \tag{90}$$

corresponding to phase-only measurements at node 1. The measurement noise variances were set to

$$r_1 = r_2 = \left(\frac{2\pi \cdot 10}{360} \right)^2 \tag{91}$$

corresponding to a 10 degrees standard deviation measurement noise. The power constraint $P[k]$ was set as in (68) and the mixing vector was set to

$$b = [1, 0, 0]^\top \tag{92}$$

with $a[k] = \alpha[k]b$ and $\alpha[k]$ chosen according to (57) to satisfy the power constraint. This mixing vector corresponds to the baseline scenario where node 1 simply forwards scaled noisy observations to node 2.

The results in Figures 5 and 6 show that the phase and frequency estimates and predictions at both nodes converge quickly to their steady-state values as calculated by solving the discrete-time algebraic Riccati equation as discussed in

Section 4. The phase predictions/estimates at node 2 suffer more from the additional noise incurred in the forwarding process from node 1 to node 2 whereas the frequency predictions/estimates at node 2 have nearly identical performance to those at node 1.

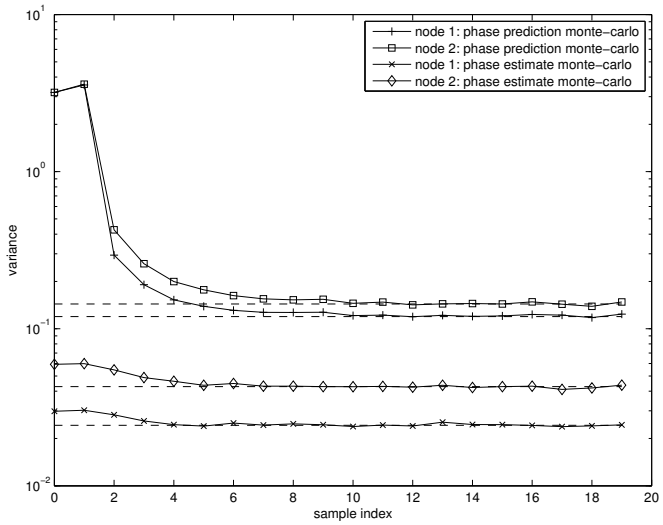

Figure 5. One-step prediction variance and estimation variance for the phase state in the oscillator tracking scenario. Dashed lines correspond to steady-state performance calculations via the discrete-time algebraic Riccati equation.

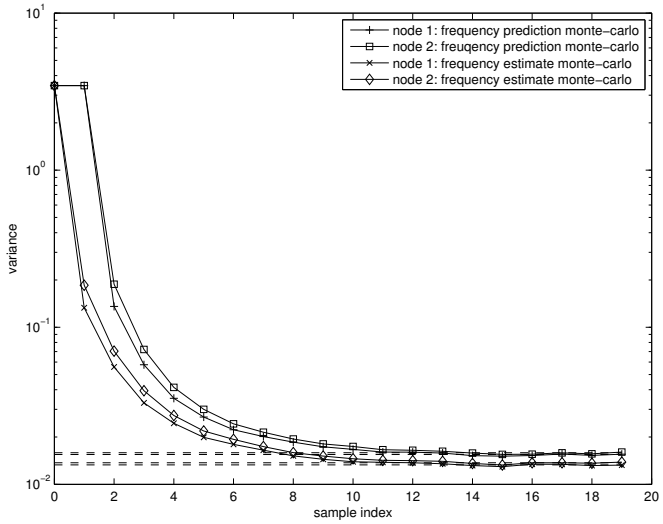

Figure 6. One-step prediction variance and estimation variance for the frequency state in the oscillator tracking scenario. Dashed lines correspond to steady-state performance calculations via the discrete-time algebraic Riccati equation.

Figure 7 shows the steady-state phase prediction error at node 2 as a function of the elements in the mixing vector $a[k]$. For an arbitrary fixed mixing vector $b = [b_1, b_2, b_3]^\top \in \mathbb{R}^3$ with $b_1 > 0$ and $a[k] = \alpha[k]b$, these results show that the best steady-state phase prediction error performance is achieved as $b_2 \to -b_1$ from above with $b_3 \geq 0$. In this case, the steady-state phase prediction error variance at node 2 approaches that of node 1 and is significantly improved with respect to the baseline case of simply forwarding scaled noisy observations, i.e., $b_2 = b_3 = 0$. Figure 8 shows similar results for steady-state frequency prediction errors.

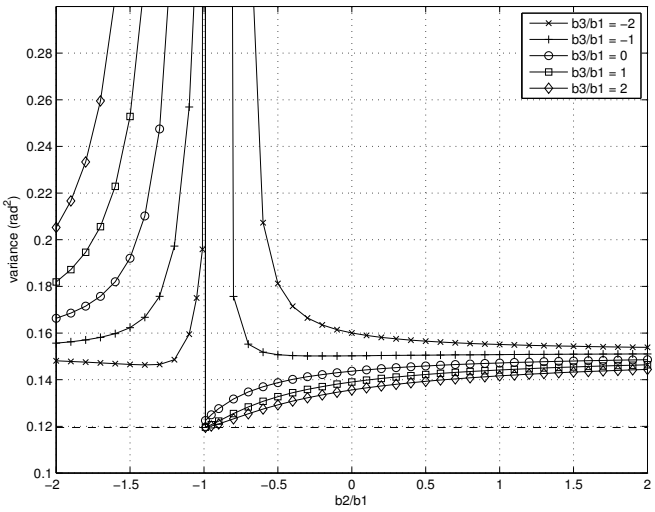

Figure 7. Steady-state one-step prediction variance for the phase state as a function of the fixed mixing vector $b = [b_1, b_2, b_3]^\top$ with $b_1 > 0$ and $a[k] = \alpha[k]b$ in the oscillator tracking scenario. The dashed line is the steady-state one-step prediction variance for the phase state at node 1.

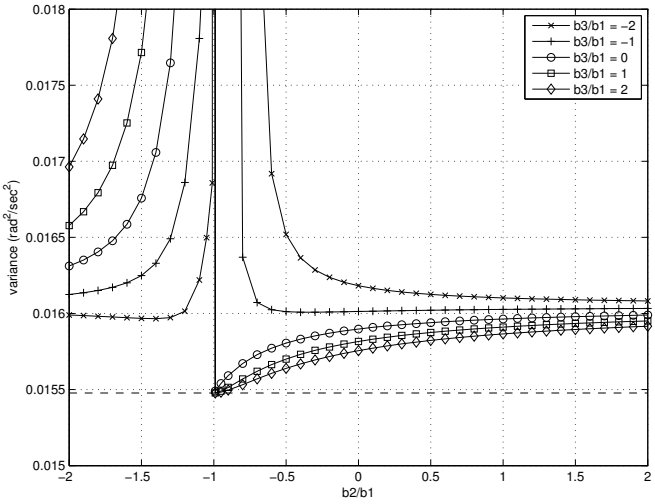

Figure 8. Steady-state one-step prediction variance for the frequency state as a function of the fixed mixing vector $b = [b_1, b_2, b_3]^\top$ with $b_1 > 0$ and $a[k] = \alpha[k]b$ in the oscillator tracking scenario. The dashed line is the steady-state one-step prediction variance for the frequency state at node 1.

The steady-state behavior as $b_2 \to -b_1$ from above deserves additional explanation. Suppose we have a mixing vector of the form

$$a[k] = \alpha[k] \begin{bmatrix} 1 \\ -1+\epsilon \\ b_3 \end{bmatrix}^\top \tag{93}$$

with $\epsilon > 0$ and $\alpha[k]$ the scale factor computed to satisfy the power constraint. From (36), we can write

$$\bar{H}[k] = \begin{bmatrix} a_1[k]H_1 + a_2^\top[k] & -a_2^\top[k] & a_1[k] \end{bmatrix} \tag{94}$$

$$= \alpha[k] \begin{bmatrix} \epsilon & b_3 & 1-\epsilon & -b_3 & 1 \end{bmatrix}. \tag{95}$$

We can calculate the steady-state scale factor $\alpha^{\text{ss}} =$

$\lim_{k\to\infty}\alpha[k]$ explicitly in the oscillator tracking scenario by first computing

$$\lim_{k\to\infty}\frac{\Sigma_x[k]}{P[k]}=\lim_{k\to\infty}\frac{\Sigma_x[k]}{H_1\Sigma_x[k]H_1^\top+r_1}=\begin{bmatrix}1&0\\0&0\end{bmatrix} \quad (96)$$

since the (1,1) element of $\Sigma_x[k]$ grows more quickly than the other elements in $\Sigma_x[k]$. This implies

$$\lim_{k\to\infty}\frac{\Gamma[k]}{P[k]}=\begin{bmatrix}1&1&0\\1&1&0\\0&0&0\end{bmatrix} \quad (97)$$

using (55) and the fact that $\Sigma_1[k|k]$ is bounded. From (57), we have

$$\lim_{k\to\infty}\alpha[k]=\lim_{k\to\infty}\sqrt{\frac{P[k]}{b^\top\Gamma[k]b}} \quad (98)$$

$$=\frac{1}{|b_1+b_2|} \quad (99)$$

$$=\frac{1}{\epsilon} \quad (100)$$

Hence, in the steady-state, we have

$$\lim_{k\to\infty}\bar{H}[k]=\begin{bmatrix}1&\frac{b_3}{\epsilon}&\frac{1-\epsilon}{\epsilon}&\frac{-b_3}{\epsilon}&\frac{1}{\epsilon}\end{bmatrix} \quad (101)$$

Although the unbounded elements in $\lim_{k\to\infty}\bar{H}[k]$ cause numerical problems when solving the discrete-time algebraic Riccati equation for very small values of ϵ, (101) implies that a linear combination of some of the states can be effectively observed at node 2 without any additional measurement noise when ϵ is small. It is easy to verify the observability of the pair $[\bar{F}^{\rm ss},\bar{H}^{\rm ss}]$, hence this allows the steady-state phase and frequency prediction variances at node 2 to approach those at node 1 when $b_2\to -b_1$ from above.

6. Conclusion

This paper studied the problem of tracking a time-varying variable with serially-connected Kalman filters. Our focus was on a two-node scenario in which node 1 can directly measure the state while node 2 can only observe the state through power-constrained transmissions from node 1. A time-varying augmented state model was developed for tracking at node 2. Detailed numerical results were presented for a scenario with scalar variable tracking and stable state dynamics and a scenario with two-state oscillator phase and frequency tracking and unstable state dynamics. In both scenarios, the results demonstrate that a non-trivial combination of the observation and state estimate at node 1 can improve performance at node 2 with respect to a baseline scenario of simply forwarding scaled observations. We also demonstrate that minimizing the condition number of the observability matrix does not necessarily result in MMSE-optimal tracking at node 2.

Appendix: Initial Prediction Covariance at Node 2

Prior to the first observation, the state $\bar{x}[0]\sim\mathcal{N}(0,\Sigma_2[0|-1])$. In this appendix, we derive an expression for the initial prediction covariance $\Sigma_2[0|-1]$. The diagonal elements

$$\mathrm{E}[x[0]x^\top[0]]=\Sigma_1[0|-1] \quad (102)$$

$$\mathrm{E}[\tilde{x}[0|0]\tilde{x}^\top[0|0]]=\Sigma_1[0|0] \quad (103)$$

$$\mathrm{E}[w_1^2[0]]=r_1. \quad (104)$$

are all given by definition. The off-diagonal element

$$\mathrm{E}[x[0]w_1[0]]=0 \quad (105)$$

under our assumption that the process and measurement noises are zero-mean and independent.

We now consider the off-diagonal element $\mathrm{E}[x[0]\tilde{x}^\top[0|0]]$. Since $\hat{x}_1[0|0]=K_1[0]y_1[0]$, we can write

$$\tilde{x}[0|0]=x[0]-\hat{x}_1[0|0] \quad (106)$$

$$=x[0]-K_1[0]y_1[0] \quad (107)$$

$$=x[0]-K_1[0]hx[0]-K_1[0]w_1[0] \quad (108)$$

$$=(I-K_1[0]h)x[0]-K_1[0]w_1[0]. \quad (109)$$

It follows that

$$\mathrm{E}[x[0]\tilde{x}^\top[0|0]]=\mathrm{E}\left[x[0]x^\top[0]\right](I-K_1[0]h)^\top \quad (110)$$

$$=\Sigma_1[0|-1](I-K_1[0]h)^\top. \quad (111)$$

Finally, we consider the off-diagonal element $\mathrm{E}[w_1[0]\tilde{x}^\top[0|0]]$. Using (109), this can be computed as

$$\mathrm{E}[w_1[0]\tilde{x}^\top[0|0]]=-\mathrm{E}[w_1[0]w_1^\top[0]]K_1^\top[0] \quad (112)$$

$$=-r_1K_1^\top[0]. \quad (113)$$

Putting these results together, we have

$$\bar{\Sigma}[0|-1]=\begin{bmatrix}\Sigma_1[0|-1]&A^\top&0\\A&\Sigma_1[0|0]&-K_1[0]r_1\\0&-r_1K_1^\top[0]&r_1\end{bmatrix} \quad (114)$$

with $A=(I-K_1[0]h)\Sigma_1[0|-1]$.

References

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

[2] Y. Oshman, "Optimal sensor selection strategy for discrete-time state estimators," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 30, no. 2, pp. 307–314, 1994.

[3] Y. He and E. K. Chong, "Sensor scheduling for target tracking in sensor networks," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 1. IEEE, 2004, pp. 743–748.

[4] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, 2006.

[5] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *Signal Processing, IEEE Transactions on*, vol. 57, no. 2, pp. 451–462, 2009.

[6] Y. Mo, R. Ambrosino, and B. Sinopoli, "Sensor selection strategies for state estimation in energy constrained wireless sensor networks," *Automatica*, vol. 47, no. 7, pp. 1330–1338, 2011.

[7] J. Wilson and S. Guhe, "Observability matrix condition number in design of measurement strategies," *Computer Aided Chemical Engineering*, vol. 20, pp. 397–402, 2005.

[8] N. Tali-Maamar, J. Babary, and D. Dochain, "Influence of the sensor location on the practical observability of distributed parameter bioreactors," in *Control, 1994. Control '94. International Conference on*, vol. 1, March 1994, pp. 255–260 vol.1.

[9] K. Hiramoto, H. Doki, and G. Obinata, "Optimal sensor/actuator placement for active vibration control using explicit solution of algebraic riccati equation," *Journal of Sound and Vibration*, vol. 229, no. 5, pp. 1057–1075, 2000.

[10] M. Meo and G. Zumpano, "On the optimal sensor placement techniques for a bridge structure," *Engineering Structures*, vol. 27, no. 10, pp. 1488–1497, 2005.

[11] N. Debnath, A. Dutta, and S. Deb, "Placement of sensors in operational modal analysis for truss bridges," *Mechanical Systems and Signal Processing*, vol. 31, pp. 196–216, 2012.

[12] L. Galleani, "A tutorial on the 2-state model of the atomic clock noise," *Metrologia*, vol. 45, no. 6, pp. S175–S182, Dec. 2008.

[13] G. Giorgi and C. Narduzzi, "Performance analysis of kalman filter-based clock synchronization in IEEE 1588 networks," in *International IEEE Symposium on Precision Clock Synchronization for Measurement, Control, and Communication*, October 12-16 2009, pp. 1–6.

[14] "Rakon RFPO45 SMD oven controlled crystal oscillator datasheet," 2009. [Online]. Available: http://www.rakon.com/Products/Public/Documents/Specifications/RFPO45.pdf

BIOGRAPHY

D. Richard Brown III *received the B.S. and M.S. degrees in Electrical Engineering from The University of Connecticut in 1992 and 1996, respectively, and received the Ph.D. degree in Electrical Engineering from Cornell University in 2000. From 1992-1997, he was with General Electric Electrical Distribution and Control. He joined the faculty at Worcester Polytechnic Institute (WPI) in Worcester, Massachusetts in 2000 and currently is an Associate Professor. He also held an appointment as a Visiting Associate Professor at Princeton University from August 2007 to June 2008. His research interests are currently in coordinated wireless transmission and reception, synchronization, distributed computing, and game-theoretic analysis of communication networks.*

Yaakov Bar-Shalom *Yaakov Bar-Shalom was born on May 11, 1941. He received the B.S. and M.S. degrees from the Technion, Israel Institute of Technology, in 1963 and 1967 and the Ph.D. degree from Princeton University in 1970, all in electrical engineering. From 1970 to 1976 he was with Systems Control, Inc., Palo Alto, California. Currently he is Board of Trustees Distinguished Professor in the Dept. of Electrical and Computer Engineering and Marianne E. Klewin Professor in Engineering at the University of Connecticut. He is also Director of the ESP (Estimation and Signal Processing) Lab. His current research interests are in estimation theory and target tracking. While he started his career in stochastic control, currently he is out of control, except for the (suboptimal) multivariable (12) control of his sloop "Syrah" (see the cap in picture) when racing her. He has published over 500 papers and book chapters in these areas and in stochastic adaptive control. He coauthored the monograph Tracking and Data Association (Academic Press, 1988), the graduate texts Estimation and Tracking: Principles, Techniques and Software (Artech House, 1993), Estimation with Applications to Tracking and Navigation: Algorithms and Software for Information Extraction (Wiley, 2001), the advanced graduate text MultitargetMultisensor Tracking: Principles and Techniques (YBS Publishing, 1995), Tracking and Data Fusion (YBS Publishing, 2011), and edited the books Multitarget-Multisensor Tracking: Applications and Advances (Artech House, Vol. I, 1990; Vol. II, 1992; Vol. III, 2000). He has been elected Fellow of IEEE for "contributions to the theory of stochastic systems and of multi target tracking". He has been consulting to numerous companies and government agencies, and originated the series of MultitargetMultisensor Tracking short courses offered via UCLA Extension, at Government Laboratories, private companies and overseas. During 1976 and 1977 he served as Associate Editor of the IEEE Transactions on Automatic Control and from 1978 to 1981 as Associate Editor of Automatica. He was Program Chairman of the 1982 American Control Conference, General Chairman of the 1985 ACC, and CoChairman of the 1989 IEEE International Conference on Control and Applications. During 1983–1987 he served as Chairman of the Conference Activities Board of the IEEE Control Systems Society and during 1987–1989 was a member of the Board of Governors of the IEEE CSS. He is a member of the Board of Directors of the International Society of Information Fusion (1999-2007) and served as General Chairman of FUSION 2000, President of ISIF in 2000 and 2002 and Vice President for Publications in 2004-07. In 1987 he received the IEEE CSS Distinguished Member Award. Since 1995 he is a Distinguished Lecturer of the IEEE AESS and has given numerous keynote addresses at major national and international conferences. He is co-recipient of the M. Barry Carlton Award for the best paper in the IEEE Transactions on Aerospace and Electronic Systems in 1995 and 2000 and the 1998 University of Connecticut AAUP Excellence Award for Research. In 2002 he received the J. Mignona Data Fusion Award from the DoD JDL Data Fusion Group. He is a member of the Connecticut Academy of Science and Engineering. In 2008 he was awarded the IEEE Dennis J. Picard Medal for Radar Technologies and Applications. He has been listed by academic.research.microsoft.com "Top Authors in Engineering" as #1 among the researchers in Aeronautics & Aerospace Engineering, based on the citations of his work.*