An Approximate Minimum Variance Filter for Nonlinear Systems with Randomly Delayed Observations

Seham Allahyani
College of Engineering,
Design and Physical Sciences,
Brunel University, London, UK
Email:seham.allahyani@brunel.ac.uk

Paresh Date
College of Engineering,
Design and Physical Sciences,
Brunel University, London, UK
Email:paresh.date@brunel.ac.uk

***Abstract*—In this paper, we extend our earlier results on minimum variance filter for systems with additive multiplicative noise in two different ways. Firstly, we propose a novel characterization of the linearization error in terms of multiplicative noise. Secondly, we also allow for random delay of up to one time step in the measurement. The delay is modelled by Bernoulli random variables. We derive a closed-form expression for the minimum variance filter for the resulting system with a linearized state transition equation, accounting for both the linearization error as well as the random delay. The utility of the proposed filtering algorithm is demonstrated through numerical experiments.**

Keywords Multiplicative noise, nonlinear filtering, delayed measurements

1. Introduction

In our previous work reported in [1], we considered a *linear* discrete-time system with an interpretation of multiplicative uncertainty as stochastic uncertainty in parameters. It is however an accepted fact that, in many applications of interest, the system dynamics and the observation equation are nonlinear including economics, radar tracking, weather forecasting and navigation systems. In the past few years, considerable attention has also been devoted to the nonlinear filtering problem. A series of suboptimal approaches have been developed in the literature to solve the nonlinear filtering problem, which include the Extended Kalman filter or the EKF [2], the particle filter or the PF [3], the Unscented Kalman filter or the UKF [4], the Ensemble Kalman filter or EnKF [5] and the quadrature Kalman filter or the QKF [6].

Moreover, an extensive theory and a number of algorithms have been developed for filtering in nonlinear systems using linearization, see *e.g.* [7]. If computational time is not an issue, there are powerful sampling based particle filtering algorithms available which are guaranteed to converge and outperform the suboptimal methods if a sufficient number of samples are generated at each time-step. However, linearization-based filters are still often preferred in real time systems where sampling tends to be an expensive operation.

A separate strand of literature considers randomly delayed measurements, which are frequently encountered in many practical application, such as target tracking, communication, signal processing, control, transportation, and so on see, e.g. [8], [9], [10] and [11]. In order to make sure that the filtering error dynamics converge, the possible measurement delays should be taken into account in designing filters. In the past few years, many results have been reported in the literature on the estimation of the state of discrete-time systems with randomly delayed observations see, e.g., [12], [13]. These models have received a considerable research attention and many important results have been reported in the literature. Linear filtering for discrete time systems with finite random measurement delays is investigated in [14]. In [15], an optimal filtering problem for networked systems with random transmission delays is investigated using a multi-state Markov chain model for the delay process. Centralized fusion filters have been designed in [16] for linear systems with multiple sensors with different delay rates. In [17], different approximations of the statistics of a nonlinear transformation of a random vector are used to investigate the filtering problem for a class of nonlinear stochastic systems with randomly delayed observations, where the possible delay is restricted to a single step. This work is extended in [18] to address the case when the measurements might be delayed randomly by one or two sampling steps. For the same model, a Gaussian approximation filter is derived [19]. Furthermore, many researchers have investigated these problems under different assumptions about the possible delays and different filtering approximations; See [20], [21], [22], [23] and [24], for example.

To the best of authors' knowledge, the problem of state estimation in nonlinear systems with random delays has not been fully investigated. This paper is concerned with the problem of the state estimation for a nonlinear discrete time systems with a single random delay in measurement. We use stochastic multiplicative noise as a proxy for the impact of linearization error, and design a closed form minimum variance filter for the resulting linear system with additive multiplicative noise as well as a random delay.

The remainder of the paper is arranged as follows. In section 2, the aforementioned class of systems is described. We outline the problem and derive its solution for a single random delay in this section. In section 3, the proposed filtering algorithm is demonstrated with a numerical example. Some concluding remarks given in section 4. Proof of the main theorem in section 2 is provided in the Appendix.

2. System model and problem formulation

Our aim is to address the problem of the state estimation for nonlinear discrete time systems with additive noise, where the measurement might be delayed randomly by one sample time when the Bernoulli random variables describing the delayed observations, with their values one or zero indicate whether the measurement is delayed or not.

Consider a class of discrete-time nonlinear stochastic systems with additive noise, where the measurement might be delayed. The model is described by the following state and measurement equations:

$$\mathcal{X}(k+1) = f(\mathcal{X}(k)) + U_w\mathcal{W}(k), \tag{1}$$

$$\mathcal{Y}(k) = h(\mathcal{X}(k)) + U_v\mathcal{V}(k), \tag{2}$$

$$\mathcal{Z}(k) = (1-p_k)\mathcal{Y}(k) + p_k\mathcal{Y}(k-1), \tag{3}$$

where $\mathcal{X}(k) \in \mathbb{R}^n$ is the state vector at time k to be estimated, $\mathcal{Y}(k) \in \mathbb{R}^r$ is the measurement vector at time k and $\mathcal{Z}(k) \in \mathbb{R}^r$ is the one step randomly delayed measurement equation, p_k denotes the Bernoulli random variable at each time k (binary switching sequence taking the values 0 or 1) with a known distribution $\mathbb{P}(p_k = 0) = \beta$ and $\mathbb{P}(p_k = 1) = 1-\beta$ and p_k are uncorrelated with other random variables. U_w and U_v are given deterministic matrices. $\mathcal{W}(k) \in \mathbb{R}^n$ and $\mathcal{V}(k) \in \mathbb{R}^r$ are the process noise and the measurement noise, respectively The nonlinear function $f(\mathcal{X}(k))$ and $h(\mathcal{X}(k))$ are analytic everywhere with known form. The work of this paper is carried out based on the following assumptions.

Assumption 1: The noise signals $\mathcal{W}(k)$ and $\mathcal{V}(k)$ are zero mean, i.i.d. random vectors with identity covariance matrix $\mathcal{I}$ and mutually uncorrelated.

Assumption 2: The initial state is a random vector with a known mean and covariance matrix, $\mathbb{E}[\mathcal{X}(0)] = \hat{\mathcal{X}}(0)$ and $\mathbb{E}[(\mathcal{X}(0) - \hat{\mathcal{X}}(0))(\mathcal{X}(0) - \hat{\mathcal{X}}(0))^\top] = P(0)$ respectively. $\mathcal{X}(0)$, $\mathcal{W}(k)$ and $\mathcal{V}(k)$ are mutually independent.

The approximated conditional mean of $\mathcal{X}(k+1)$, which provides the predictor, $\hat{\mathcal{X}}(k+1|k)$, is derived using (1):

$$\hat{\mathcal{X}}(k+1|k) = f(\hat{\mathcal{X}}(k|k)). \tag{4}$$

For further brevity of notation, an expression $LL^\top$ will sometimes be denoted as $(L)(\star)^\top$, where L is a matrix-valued or vector-valued expression and where there is no risk of confusion.

The update equation for a nonlinear filter using one step randomly delayed measurements is

$$\begin{aligned}\hat{\mathcal{X}}(k+1|k+1) &= \hat{\mathcal{X}}(k+1|k)+\\ &\bar{K}(k+1)(\mathcal{Z}(k+1) - \hat{\mathcal{Z}}(k+1|k)),\end{aligned} \tag{5}$$

and the estimation error covariance matrix is given by

$$\begin{aligned}&\bar{P}(k+1|k+1) = \mathbb{E}[(\mathcal{X}(k+1) - \hat{\mathcal{X}}(k+1|k+1))(\star)^\top]\\ &= \mathbb{E}[(f(\mathcal{X}(k)) + U_w\mathcal{W}(k) - (f(\hat{\mathcal{X}}(k|k))\\ &+ \bar{K}(k+1)(\mathcal{Z}(k+1) - \hat{\mathcal{Z}}(k+1|k))))(\star)^\top]\\ &= \mathbb{E}[f(\mathcal{X}(k)) - f(\hat{\mathcal{X}}(k|k))(\star)^\top] + U_wU_w^\top\\ &- \mathbb{E}[\bar{K}(k+1)(\mathcal{Z}(k+1) - \hat{\mathcal{Z}}(k+1|k))(\star)^\top],\end{aligned} \tag{6}$$

$$\begin{aligned}\text{where}\quad \mathcal{Z}(k+1) &= (1-p_k)(h(\mathcal{X}(k+1)) + U_v\mathcal{V}(k+1))\\ &+ p_k(h(\mathcal{X}(k)) + U_v\mathcal{V}(k)).\end{aligned} \tag{7}$$

By using the Taylor series expansion around $\hat{\mathcal{X}}(k|k)$, we linearize $f(\mathcal{X}(k))$ and $h(\mathcal{X}(k))$ as following:

$$f(\mathcal{X}(k)) = f(\hat{\mathcal{X}}(k|k)) + A(k)\tilde{\mathcal{X}}(k|k) + o(|\tilde{\mathcal{X}}(k|k)|),$$

$$\begin{aligned}&h(\mathcal{X}(k+1)) = h(\hat{\mathcal{X}}(k+1|k)) + C(k+1)\tilde{\mathcal{X}}(k+1|k)\\ &+ o(|\tilde{\mathcal{X}}(k+1|k)),\end{aligned}$$

where

$$A(k) = \frac{\partial f(\mathcal{X}(k))}{\partial \mathcal{X}(k)}|_{\mathcal{X}(k)=\hat{\mathcal{X}}(k|k)},$$

$$C(k+1) = \frac{\partial h(\mathcal{X}(k+1))}{\partial \mathcal{X}(k+1)}|_{\mathcal{X}(k+1)=\hat{\mathcal{X}}(k+1|k)}$$

$$\tilde{\mathcal{X}}(k+i|k) = \mathcal{X}(k+i) - \hat{\mathcal{X}}(k+i|k) \quad i = 0, 1. \tag{8}$$

In [25], $o(|\tilde{\mathcal{X}}(k|k)|)$ are characterized as

$$o(|\tilde{\mathcal{X}}(k|k)|) = B(k)N(k)L(k)\tilde{\mathcal{X}}(k|k),$$

where $B(k)$ are bounded problem-dependent scaling matrices, $L(k)$ is for providing an extra degree of freedom to tune the filter and $N(k)$ are unknown time-varying matrices accounting for the linearisation errors of the dynamical model and satisfies

$$N(k)N(k)^\top \leq I.$$

In our work presented here, we characterise the linearization error $o(|\tilde{\mathcal{X}}(k+i|k)|), \quad i=0,1$ as a stochastic perturbation which is linear in $\tilde{\mathcal{X}}(k+i|k)$:

$$o(|\tilde{\mathcal{X}}(k+i|k)|) = Q_j(k+i)\tilde{\mathcal{X}}(k+i|k)R_j(k+i), \quad j=1,2.$$

This gives us an approximate equivalent linear system with additive multiplicative noise:

$$\begin{aligned}\mathcal{X}(k+1) &= f(\hat{\mathcal{X}}(k|k)) + A(k)\tilde{\mathcal{X}}(k|k)+ \\ &Q_1(k)\tilde{\mathcal{X}}(k|k)R_1(k) + U_w\mathcal{W}(k),\end{aligned} \tag{9}$$

$$\begin{aligned}\mathcal{Y}(k+1) &= h(\hat{\mathcal{X}}(k+1|k)) + C(k+1)\tilde{\mathcal{X}}(k+1|k) \\ &+ Q_2(k+1)\tilde{\mathcal{X}}(k+1|k)R_2(k+1) + U_v\mathcal{V}(k),\end{aligned} \tag{10}$$

where $\mathcal{R}_1(k) \in \mathbb{R}^n$ and $\mathcal{R}_2(k) \in \mathbb{R}^r$ are zero mean, i.i.d. random vectors with identity covariance matrix $\mathcal{I}$ and are mutually independent with the initial state and other noise signals. $Q_1(k)$ and $Q_2(k)$ describe the effect of higher order terms in the Taylor series in terms of parameter uncertainties. The justification of characterising deterministic Taylor series truncation error by stochastic multiplicative noise can be given as follows. Firstly, we are typically interested in filter tracking performance over a period of time, e.g. as measured by the root mean squared error, and treating the error as stochastic can be advantageous if it yields a closed-form result (as is the case here). Secondly, as demonstrated in the numerical example presented, the *size* of the stochastic uncertainty representing the linearization error can be used as a tuning parameter for the linearized filter, to improve the filtering performance.

The objective of this section is to find the optimum filter gain $\bar{K}(k+1)$ that minimizes the trace of the covariance matrix $\bar{P}(k+1|k+1)$ of the state estimate $\hat{\mathcal{X}}(k+1|k+1)$ for the approximate linear system given by (7), (8), (9) and (10). Our main result in this section is given in the next theorem.

Theorem 1. For equations (7)-(10), the filter gain $\bar{K}(k+1)$ that minimizes the trace of the covariance $\bar{P}(k+1|k+1)$ is given by

$$\begin{aligned}&\bar{K}(k+1) = (\beta\bar{P}(k+1|k)C(k+1)^\top + (1-\beta)\times \\ &A(k)\bar{P}(k|k)C(k)^\top)[\beta(C(k+1)\bar{P}(k+1|k)C(k+1)^\top \\ &+ Q_2(k+1)\bar{P}(k+1|k)Q_2(k+1)^\top + U_vU_v^\top)+ \\ &(1-\beta)(C(k)\bar{P}(k|k)C(k)^\top + Q_2(k)\bar{P}(k|k)Q_2(k)^\top+ \\ &U_vU_v^\top) + \beta(1-\beta)(\widetilde{\psi}_0(k+1) + \widetilde{\psi}_1(k+1) + C(k+1)\times \\ &A(k)\bar{P}(k|k)C(k)^\top + C(k)\bar{P}(k|k)A(k)^\top C(k+1)^\top)- \\ &\beta(1-\beta)(\psi_0(k+1)\psi_1(k+1)^\top \\ &+ \psi_1(k+1)\psi_0(k+1)^\top)]^{-1},\end{aligned} \tag{11}$$

where $\bar{P}(k+1|k)$ and $\psi_i(k+1)$, $\widetilde{\psi}_i(k+1)$, $\quad i=0,1$ are as defined in (14) and (13), respectively in the Appendix.

Proof: See Appendix for an outline of proof; details are omitted for brevity.

Remark: It can be easily verified that setting $\beta=1$ and $Q_1(k)=Q_2(k)=0$ gives the familiar linear Kalman filter for delay-free case.

3. Numerical Example

To test the accuracy of the new algorithm, the following univariate non-stationary growth model is considered,

$$\mathcal{X}(k+1) = a\mathcal{X}(k) + b\frac{\mathcal{X}(k)}{1+\mathcal{X}(k)^2} + d\cos(1.2k) + U_w\mathcal{W}(k),$$

$$\mathcal{Y}(k) = \frac{\mathcal{X}(k)^2}{20} + U_v\mathcal{V}(k),$$

where $\mathcal{V}(k)$ and $\mathcal{W}(k)$ are i.i.d random variables with zero mean and unit variance. This model has been previously used in [26]. We use the parameters $a=0.5$, $b=1$, $d=8$, $U_w=0.1$ and $U_v=0.1$. Initial conditions are $\mathcal{X}(0)=1$, $\hat{\mathcal{X}}(0)=0$ and $P(0)=0.1$. In equations (8) and (9), we use $Q_1(k)=\gamma\,\mathrm{trace}\,(A(k))$ and $Q_2(k)=\gamma\,\mathrm{trace}\,(C(k))$, where γ is our tuning parameter expressing the linearization error as a percentage of linearized parameters. As the model has strong nonlinearities, we expect that using a large non-zero gamma might improve the performance.

In order to evaluate the efficiency of the estimators, we use the root mean square error (RMSE) criteria. Consider 100 independent simulations, each with 200 data points. Denoting $\mathcal{X}^{(s)}(k), \quad k=1,...,200$ as the s^{th} set of true values of the state, and $\hat{\mathcal{X}}^{(s)}(k|k)$ as the filtered state estimate at time k for the s^{th} simulation run, the RMSE is calculated by

$$RMSE(s) = \sqrt{\frac{1}{200}\sum_{k=1}^{200}(\mathcal{X}^{(s)}(k) - \hat{\mathcal{X}}^{(s)}(k|k))^2}, \quad s=1,...,100.$$

Then the average of RMSE of the state over 100 simulations is given by

$$AvRMSE = \frac{1}{100}\sum_{s=1}^{100} RMSE(s).$$

We performed two experiments for this system with different levels of of linearization error γ. Firstly, to isolate the improvement in performance even in the absence of delay, we first conduct a delay-free experiment (i.e. with $\beta=1$) and compare the performance of filter with pure linearization (i.e. $\gamma=0$) with filter with different values of γ (i.e.$\gamma=0.25, 0.5, 0.75$). We then use the first experiment with actual delay free filter. Table 1 summarizes the results of this experiment. As can be seen, the filtering algorithm with uncertainties outperforms the filtering algorithm with pure linearization (i.e.

$\gamma = 0$). Further, improvement in the performance of filter becomes more pronounced as γ increases.

TABLE 1. Comparison of $AvRMSE$ for different values of γ

	$\gamma = 0$	$\gamma = 0.25$	$\gamma = 0.5$	$\gamma = 0.75$
$AvRMSE$	0.1169	0.1168	0.1166	0.1164

In the second experiment, we compared a one delay filter with pure linearization (i.e. $\gamma = 0$) with a one delay filter using interpretation of multiplicative noise in terms of linearization error. We use different level of uncertainties (i.e.$\gamma = 0.25, 0.5, 0.75$) and different values of β (i.e.$\beta = 0.3, 0.5, 0.7, 0.9$). Table 2 summarizes the results of this experiment. As can be seen, the filtering algorithm with uncertainties outperforms the filtering algorithm with pure linearization (i.e. $\gamma = 0$) in all the cases.

TABLE 2. Comparison of $AvRMSE$ for different values of β and γ

		$\beta = 0.9$	$\beta = 0.7$	$\beta = 0.5$	$\beta = 0.3$
$AvRMSE$	$\gamma = 0$	0.1170	0.1229	0.1216	0.1154
	$\gamma = 0.25$	0.1169	0.1228	0.1215	0.1153
	$\gamma = 0.5$	0.1168	0.1227	0.1214	0.1152
	$\gamma = 0.75$	0.1167	0.1225	0.1213	0.1151

4. Conclusions

In this paper, an approximate minimum variance filter is discussed for a class of nonlinear discrete time systems with additive noise and a random delay. The paper makes two distinct contributions. Firstly, it generalizes the closed form solution for minimum variance filtering for linear systems in [1]. We have used a novel approach of modelling the linearization error as multiplicative noise, and (in essence) 'de-tuning' the extended Kalman filter to account for this 'noise' or the linearization error. Secondly, we extend the results to cope with a random delay of a single time step. Our numerical experiment indicates that the proposed filtering algorithm can be used to improve the filtering performance (as measured by root mean squared error), when linearized dynamics is used for filter design.

Current research involves exploring the use of this filter in a biomedical application, where the purpose is to detect atherosclerotic coronary stenosis (or *blocked arteries*) by detecting, at the chest surface, the acoustic signals that result from the arterial wall stresses induced by the post-stenotic blood flow disturbance. We plan to adapt our methods, with a computational model of wave transit in the chest, to develop a new tool for non-invasive testing for coronary stenosis.

Appendix

The proof follows on the same lines as our earlier work [1] and details are omitted for brevity. The filtering estimates of the state covariance is obtained by combining the equations (1)-(5) as follows. The state covariance matrix at time $k+1$ can be written as

$$\begin{aligned}
&\bar{P}(k+1|k+1) = \mathbb{E}[f(\mathcal{X}(k)) - f(\hat{\mathcal{X}}(k|k))(\star)^\top] + U_w U_w^\top + \\
&\bar{K}(k+1)(\mathbb{E}[(1-p_{k+1})^2]\mathbb{E}[(\mathcal{Y}(k+1) - \hat{\mathcal{Y}}(k+1|k))(\star)^\top] \\
&+ \mathbb{E}[(p_{k+1})^2]\mathbb{E}[(\mathcal{Y}(k) - \hat{\mathcal{Y}}(k|k))(\star)^\top] + \mathbb{E}[((1-p_{k+1}) - \beta)^2] \\
&(\hat{\mathcal{Y}}(k+1|k))(\star)^\top + \mathbb{E}[(p_{k+1} - (1-\beta))^2](\hat{\mathcal{Y}}(k|k))(\star)^\top \\
&+ \mathbb{E}[(1-p_{k+1})(\mathcal{Y}(k+1) - \hat{\mathcal{Y}}(k+1|k))(p_{k+1}(\mathcal{Y}(k) - \hat{\mathcal{Y}}(k|k))^\top] \\
&+ \mathbb{E}[p_{k+1}(\mathcal{Y}(k) - \hat{\mathcal{Y}}(k|k))(1-p_{k+1})(\mathcal{Y}(k+1) - \hat{\mathcal{Y}}(k+1|k))^\top] \\
&+ \mathbb{E}[((1-p_{k+1}) - \beta)(p_{k+1} - (1-\beta))(\hat{\mathcal{Y}}(k+1|k)\hat{\mathcal{Y}}(k|k)^\top + \\
&\hat{\mathcal{Y}}(k|k)\hat{\mathcal{Y}}(k+1|k)^\top)])\bar{K}(k+1)^\top - \bar{K}(k+1)(\mathbb{E}[(f(\mathcal{X}(k) - \\
&f(\hat{\mathcal{X}}(k|k)) + U_w\mathcal{W}(k)((1-p_{k+1})(\mathcal{Y}(k+1) - \hat{\mathcal{Y}}(k+1|k)) + \\
&p_{k+1}(\mathcal{Y}(k) - \hat{\mathcal{Y}}(k|k)) + ((1-p_{k+1}) - \beta)\hat{\mathcal{Y}}(k+1|k) + \\
&(p_{k+1} - (1-\beta))\hat{\mathcal{Y}}(k|k))^\top] + \mathbb{E}[((1-p_{k+1})(\mathcal{Y}(k+1) - \\
&\hat{\mathcal{Y}}(k+1|k)) + p_{k+1}(\mathcal{Y}(k) - \hat{\mathcal{Y}}(k|k)) + ((1-p_{k+1}) - \beta) \\
&\hat{\mathcal{Y}}(k+1|k) + (p_{k+1} - (1-\beta))\hat{\mathcal{Y}}(k|k)) \\
&(f(\mathcal{X}(k)) - f(\hat{\mathcal{X}}(k|k)) + U_w\mathcal{W}(k). \qquad (12)
\end{aligned}$$

where $\hat{\mathcal{Y}}(k+i|k) = h(\hat{\mathcal{X}}(k+i|k)), i = 0, 1$. For further notational brevity, denote

$$\begin{aligned}
&\psi_i(k+1) = \hat{\mathcal{Y}}(k+i|k), i = 0, 1 \\
&\text{and } \widetilde{\psi}_i(k+1) := \psi_i(k+1)\psi_i(k+1)^\top. \qquad (13)
\end{aligned}$$

In addition, let

$$\begin{aligned}
&\bar{P}(k+1|k) = A(k)\bar{P}(k|k)A(k)^\top + \\
&Q_1(k)\bar{P}(k|k)Q_1(k) + U_w U_w^\top, \qquad (14)
\end{aligned}$$

After some straightforward, but tedious simplification, we get

$$\begin{aligned}
&\bar{P}(k+1|k+1) = A(k)\bar{P}(k|k)A(k)^\top + Q_1(k)\bar{P}(k|k)Q_1(k)^\top + \\
&U_w U_w^\top + \bar{K}(k+1)(\beta(C(k+1)\bar{P}(k+1|k)C(k+1)^\top + \\
&Q_2(k+1)\bar{P}(k+1|k)Q_2(k+1)^\top + U_v U_v^\top) + \\
&(1-\beta)(C(k)\bar{P}(k|k)C(k)^\top + Q_2(k)\bar{P}(k|k)Q_2(k)^\top + U_v U_v^\top + \\
&\beta(1-\beta)(\widetilde{\psi}_0(k+1) + \widetilde{\psi}_1(k+1) + C(k+1)A(k)\bar{P}(k|k)C(k)^\top + \\
&C(k)\bar{P}(k|k)A(k)^\top C(k+1)^\top) - \beta(1-\beta)(\psi_0(k+1)\psi_1(k+1)^\top \\
&+ \psi_1(k+1)\psi_0(k+1)^\top))\bar{K}(k+1)^\top - (\beta\bar{P}(k+1|k)C(k+1)^\top \\
&+ (1-\beta)A(k)\bar{P}(k|k)C(k)^\top)\bar{K}(k+1)^\top - \\
&\bar{K}(k+1)(\beta C(k+1)\bar{P}(k+1|k) + (1-\beta)C(k)\bar{P}(k|k)A(k)^\top)
\end{aligned}$$

To find the value of $\bar{K}(k+1)$ that minimizes the trace of the covariance $\bar{P}(k+1|k+1)$ we differentiate the trace of the above expression with respect to the filter gain matrix $\bar{K}(k+1)$and set the derivative to zero.

$$\bar{K}(k+1) = (\beta\bar{P}(k+1|k)C(k+1)^\top + (1-\beta)A(k)\bar{P}(k|k)C(k)^\top)$$
$$\times [\beta(C(k+1)\bar{P}(k+1|k)C(k+1)^\top +$$
$$Q_2(k+1)\bar{P}(k+1|k)Q_2(k+1)^\top + U_vU_v^\top)+$$
$$(1-\beta)(C(k)\bar{P}(k|k)C(k)^\top + Q_2(k)\bar{P}(k|k)Q_2(k)^\top + U_vU_v^\top)+$$
$$\beta(1-\beta)(\widetilde{\psi}_0(k+1) + \widetilde{\psi}_1(k+1) + C(k+1)A(k)\bar{P}(k|k)C(k)^\top +$$
$$C(k)\bar{P}(k|k)A(k)^\top C(k+1)^\top) - \beta(1-\beta)(\psi_0(k+1)\psi_1(k+1)^\top$$
$$+ \psi_1(k+1)\psi_0(k+1)^\top)]^{-1}, \tag{15}$$

which is the required expression. ■

References

[1] S. Allahyani and P. Date, "A minimum variance filter for discrete time linear systems with parametric uncertainty," in *MED'16: The 24th Mediterranean Conference on Control and Automation.* Mediterranean Control Association, 2016.

[2] S. S. Haykin, *Kalman filtering and neural networks.* Wiley, 2001.

[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking," *Signal Processing, IEEE Transactions on*, vol. 50, no. 2, pp. 174–188, 2002.

[4] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *IEEE*, vol. 92, pp. 401–422, 2004.

[5] P. L. Houtekamer and H. L. Mitchell, "Data assimilation using an ensemble Kalman filter technique," *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, 1997.

[6] I. Arasaratnam and S. Haykin, "Square-root quadrature kalman filtering," *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2589–2593, 2008.

[7] A. H. Jazwinski, *Stochastic processes and filtering theory.* Courier Corporation, 2007.

[8] H. Zhang and L. Xie, *Control and estimation of systems with input/output delays.* Springer Publishing Company, Incorporated, 2007.

[9] W.-W. Che, Y.-P. Li, and Y.-L. Wang, "H_∞ tracking control for ncs with packet losses in multiple channels case," *International Journal of Innovative Computing Information and Control*, vol. 2011, no. 7, pp. 6507–6522, 2011.

[10] J. Nilsson, B. Bernhardsson, and B. Wittenmark, "Stochastic analysis and control of real-time systems with random time delays," *Automatica*, vol. 34, no. 1, pp. 57–64, 1998.

[11] H. Dong, Z. Wang, and H. Gao, "Robust H_∞ filtering for a class of nonlinear networked systems with multiple stochastic communication delays and packet dropouts," *IEEE Transactions on Signal Processing*, vol. 58, no. 4, pp. 1957–1966, 2010.

[12] S. Wang, H. Fang, and X. Tian, "Recursive estimation for nonlinear stochastic systems with multi-step transmission delays, multiple packet dropouts and correlated noises," *Signal Processing*, vol. 115, pp. 164–175, 2015.

[13] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Fusion estimation using measured outputs with random parameter matrices subject to random delays and packet dropouts," *Signal Processing*, vol. 127, pp. 12–23, 2016.

[14] S. Sun, "Linear minimum variance estimators for systems with bounded random measurement delays and packet dropouts," *Signal Processing*, vol. 89, no. 7, pp. 1457–1466, 2009.

[15] C. Han, H. Zhang, and M. Fu, "Optimal filtering for networked systems with Markovian communication delays," *Automatica*, vol. 49, no. 10, pp. 3097–3104, 2013.

[16] F. O. Hounkpevi and E. E. Yaz, "Minimum variance generalized state estimators for multiple sensors with different delay rates," *Signal Processing*, vol. 87, no. 4, pp. 602–613, 2007.

[17] A. Hermoso-Carazo and J. Linares-Pérez, "Extended and unscented filtering algorithms using one-step randomly delayed observations," *Applied Mathematics and Computation*, vol. 190, no. 2, pp. 1375–1393, 2007.

[18] ——, "Unscented filtering algorithm using two-step randomly delayed observations in nonlinear systems," *Applied Mathematical Modelling*, vol. 33, no. 9, pp. 3705–3717, 2009.

[19] X. Wang, Y. Liang, Q. Pan, and C. Zhao, "Gaussian filter for nonlinear systems with one-step randomly delayed measurements," *Automatica*, vol. 49, no. 4, pp. 976–986, 2013.

[20] C.-L. Su and C.-N. Lu, "Interconnected network state estimation using randomly delayed measurements," *Power Systems, IEEE Transactions on*, vol. 16, no. 4, pp. 870–878, 2001.

[21] A. Hermoso-Carazo and J. Linares-Pérez, "Linear and quadratic least-squares estimation using measurements with correlated one-step random delay," *Digital Signal Processing*, vol. 18, no. 3, pp. 450–464, 2008.

[22] S. Nakamori, R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Recursive estimators of signals from measurements with stochastic delays using covariance information," *Applied Mathematics and Computation*, vol. 162, no. 1, pp. 65–79, 2005.

[23] Z. Wang, J. Lam, and X. Liu, "Filtering for a class of nonlinear discrete-time stochastic systems with state delays," *Journal of Computational and Applied Mathematics*, vol. 201, no. 1, pp. 153–163, 2007.

[24] M. Choi, J. Choi, and W. Chung, "State estimation with delayed measurements incorporating time-delay uncertainty," *Control Theory & Applications, IET*, vol. 6, no. 15, pp. 2351–2361, 2012.

[25] J. Hu, Z. Wang, B. Shen, and H. Gao, "Quantised recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements," *International Journal of Control*, vol. 86, no. 4, pp. 650–663, 2013.

[26] P. Date, L. Jalen, and R. Mamon, "A new algorithm for latent state estimation in non-linear time series models," *Applied Mathematics and Computation*, vol. 203, no. 1, pp. 224–232, 2008.