Gain-Constrained Recursive Filtering with Stochastic Nonlinearities and Probabilistic Sensor Delays

Jun Hu, Zidong Wang, Bo Shen and Huijun Gao

Abstract— **This paper is concerned with the gain-constrained recursive filtering problem for a class of time-varying nonlinear stochastic systems with probabilistic sensor delays and correlated noises. The stochastic nonlinearities are described by statistical means that cover the multiplicative stochastic disturbances as a special case. The phenomenon of probabilistic sensor delays is modeled by introducing a diagonal matrix composed of Bernoulli distributed random variables taking values of 1 or 0, which means that the sensors may experience randomly occurring delays with individual delay characteristics. The process noise is finite-step autocorrelated. The purpose of the addressed gain-constrained filtering problem is to design a filter such that, for all probabilistic sensor delays, stochastic nonlinearities, gain constraint as well as correlated noises, the cost function concerning the filtering error is minimized at each sampling instant, where the filter gain satisfies a certain equality constraint. A new recursive filtering algorithm is developed that ensures both the local optimality and the unbiasedness of the designed filter at each sampling instant which achieving the pre-specified filter gain constraint. A simulation example is provided to illustrate the effectiveness of the proposed filter design approach.**

Keywords—**Recursive filtering, probabilistic sensor delays, gain constraint, stochastic nonlinearities, correlated noises, time-varying systems.**

I. Introduction

In recent years, the filtering technique has been playing an important role in a variety of application areas including target tracking, computer vision and estimation of structural macroeconomic models. A rich body of results has been reported in the literature with different performance indices [4, 10, 11, 14, 18, 27, 34]. For example, the Kalman filter has proven to be the globally optimal linear estimator in [14] for the linear systems. In [24, 25], the optimal linear filters have been designed for systems with multiple packet dropouts. The extended Kalman filtering algorithm has been successfully applied in [27] to identify both the model parameters and the actual value of gene expression levels for gene regulatory network. Most of the existing results have relied on an implicit assumption that the process noises are uncorrelated, see e.g. [12, 26, 27, 36]. However, such an assumption is not always true in practice since the process noise sequences of a discrete-time system sampled from a continuous-time system are inherently correlated across time. Recently, the filter design problems have been dealt with in [6, 7, 23] for *linear* discrete-time systems with correlated noises.

Most traditional filtering algorithms have been based on the measurement outputs that are supposed to contain information about the current state of the system. However, in engineering practice, the system measurements may be subject to unavoidable sensor delays, which is particularly true in a networked environment. In the past decade, a great number of results have been reported for filtering problems with deterministic/fixed sensor delays, see e.g. [1, 37]. On the other hand, because of limited bandwidth of the communication channel, it is often the case that the sensor delay occurs in a random way when, for example, the information is transmitted through networks in real-time distributed decision-making and multiplexed data communication environment [33]. Accordingly, the filtering problems with *random* sensor delays have recently received much research attention (see e.g. [3, 16, 18, 35]), where all sensors share the same type of delay characteristics as pointed out in [4, 12]. Nevertheless, in reality, the system measurements are usually collected through multiple sensors with different physical constraints. In this case, it is somewhat conservative to assume that all sensors undergo random delays of the same probability distribution law. Rather, it would make more practical sense to consider individual features for randomly occurring sensor delays [4]. As such, in this paper, one of our motivations is to utilize a series of mutually independent Bernoulli-distributed random variables to parameterize the random occurrence of the delays for each individual sensor.

Apart from the measurement delays, it has been well recognized that the existence of nonlinearities may lead to

This work was supported in part by the National Natural Science Foundation of China under Grants 61273156, 61028008, 60825303, 61104125 and 11271103, National 973 Project under Grant 2009CB320600, the Fok Ying Tung Education Fund under Grant 111064, the Special Fund for the Author of National Excellent Doctoral Dissertation of China under Grant 2007B4, the State Key Laboratory of Integrated Automation for the Process Industry (Northeastern University) of China, the Engineering and Physical Sciences Research Council (EPSRC) of the U.K. under Grant GR/S27658/01, the Royal Society of the U.K., and the Alexander von Humboldt Foundation of Germany.

J. Hu and H. Gao are with the Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China. Email: hujun2013@gmail.com; hjgao@hit.edu.cn

Z. Wang is with the Department of Automation, Tsinghua University, Beijing 100084, China, and is also with the Department of Information Systems and Computing, Brunel University, Uxbridge, Middlesex, UB8 3PH, United Kingdom. Email: Zidong.Wang@brunel.ac.uk

B. Shen is with the School of Information Science and Technology, Donghua University, Shanghai 200051, China. Email: shenbodh@gmail.com

undesirable oscillatory behavior and therefore poses great challenges on the filter design. In the past few decades, the nonlinear filtering/control problems have been the focuses of research that have attracted considerable research attention, see e.g. [2, 5, 9, 22, 28, 29, 31]. Parallel to the sensor delay phenomenon, the nonlinearity may also occur in a random fashion especially when the signals are transmitted through networks suffering from limited bandwidth. For example, the intensity of nonlinear disturbances may vary with the network conditions for various reasons including signal congestion, quantization, fading and disorder. Accordingly, the so-called stochastic nonlinearities have started to receive some initial research interest for filtering problems, see e.g. [30, 36]. On the other hand, for practical purposes, the filter design is inevitably subject to certain physical constraints. For example, in many applications, the system states should preserve the positivity, the system outputs experience saturations, and the filter gains may need to be of a specific structure for easy implementation. The filtering problems with constraints have been gaining a recurring research interest in the past decade, see e.g. [13,21,26]. Very recently, in [26], a Kalman filter algorithm has been developed to cope with the constraints on the data injection gain. Unfortunately, up to now, very little research effort has been made on the gain-constrained filtering problem with either stochastic nonlinearities or probabilistic sensor delays, not to mention the case where multiple sensors may undergo varying communication delays with different delay rates. It is, therefore, the main purpose of this paper to shorten such a gap.

Summarizing the above discussion, it is of both theoretical importance and practical significance to investigate the gain-constrained recursive filtering problems with stochastic nonlinearities, probabilistic sensor delays as well as correlated noises. Our aim is to develop an effective recursive filter such that the specified cost function with respect to the filtering error is minimized. The main contribution of this paper lies in the following four aspects. 1) A unified framework is established to solve the gain-constrained filtering problem for discrete time-varying system in the simultaneous presence of probabilistic sensor delays, stochastic nonlinearities and correlated noises. 2) Individual delay rate is introduced to cater for the randomly occurring delay phenomenon for each sensor. 3) The Hadamard product is used to facilitate the algorithm development and intensive stochastic analysis is carried out to obtain the filter parameter for ensuring the desired filtering performance. 4) The presented filter scheme is both unbiased and recursive, hence suitable for online applications.

Notations. The notations used throughout the paper are standard. δ_{i-j} is the Kronecker delta function, which is equal to unity for $i = j$ and zero for $i \neq j$. $\circ$ is the Hadamard product with this product being defined as $[A \circ B]_{ij} = A_{ij} \cdot B_{ij}$. $\mathrm{tr}(\cdot)$ stands for the trace of a matrix. $\mathbb{E}\{x\}$ stands for the expectation of stochastic variable x. $\mathrm{Prob}\{\cdot\}$ represents the occurrence probability of the event "$\cdot$". $\mathrm{diag}\{X_1, X_2, \ldots, X_n\}$ stands for a block-diagonal matrix with matrices $X_1, X_2, \ldots, X_n$ on the diagonal. Matrices, if their dimensions are not explicitly stated, are assumed to be compatible for algebraic operations.

II. Problem Formulation and Preliminaries

Consider the following class of time-varying nonlinear systems:

$$\begin{aligned} \vec{x}_{k+1} &= \vec{A}_k \vec{x}_k + \vec{f}\left(\vec{x}_k, \eta_k\right) + \vec{B}_k \omega_k && (1) \\ \vec{y}_k &= \vec{C}_k \vec{x}_k + \vec{g}\left(\vec{x}_k, \zeta_k\right) + \vec{\nu}_k && (2) \end{aligned}$$

where $\vec{x}_k \in \mathbb{R}^n$ is the state vector to be estimated, $\vec{y}_k \in \mathbb{R}^m$ is the *ideal* output vector, $\eta_k \in \mathbb{R}$ and $\zeta_k \in \mathbb{R}$ are mutually uncorrelated zero-mean Gaussian noise sequences in k, $\omega_k \in \mathbb{R}^q$ is the process noise, $\vec{\nu}_k \in \mathbb{R}^m$ is the measurement noise, $\vec{A}_k$, $\vec{B}_k$ and $\vec{C}_k$ are known matrices with appropriate dimensions.

The delayed sensor measurement is described by

$$y_k = \left(I - \Gamma_k\right) \vec{y}_k + \Gamma_k \vec{y}_{k-1}. \quad (3)$$

Here, $y_k \in \mathbb{R}^m$ is the *actual* measurement output vector, $\Gamma_k = \mathrm{diag}\{\gamma_{k,1}, \gamma_{k,2}, \ldots, \gamma_{k,m}\}$ accounts for the different delay rate of the individual sensor where the random variables $\gamma_{k,i} \in \mathbb{R}$ $(i = 1, 2, \ldots, m)$ are mutually independent in k and i taking the values of 1 or 0 with

$$\begin{aligned} &\mathrm{Prob}\left\{\gamma_{k,i} = 1\right\} = \mathbb{E}\left\{\gamma_{k,i}\right\} := \beta_{k,i} \\ &\mathrm{Prob}\left\{\gamma_{k,i} = 0\right\} = 1 - \mathbb{E}\left\{\gamma_{k,i}\right\} := 1 - \beta_{k,i} \end{aligned} \quad (4)$$

with $\beta_{k,i} \in [0, 1)$ being a known scalar. $\gamma_{k,i}$ is assumed to be independent of η_k, ζ_k, ω_k, $\vec{\nu}_k$ and $\vec{x}_0$.

The functions $\vec{f}(\vec{x}_k, \eta_k)$ and $\vec{g}(\vec{x}_k, \zeta_k)$ represent the stochastic nonlinearities of the states with $\vec{f}(0, \eta_k) = 0$ and $\vec{g}(0, \zeta_k) = 0$ and have the following first moment for all $\vec{x}_k$:

$$\mathbb{E}\left\{ \left[\begin{array}{c} \vec{f}(\vec{x}_k, \eta_k) \\ \vec{g}(\vec{x}_k, \zeta_k) \end{array} \right] \middle| \vec{x}_k \right\} = 0 \quad (5)$$

and the covariance given by

$$\begin{aligned} &\mathbb{E}\left\{ \left[\begin{array}{c} \vec{f}(\vec{x}_k, \eta_k) \\ \vec{g}(\vec{x}_k, \zeta_k) \end{array} \right] \left[\begin{array}{c} \vec{f}(\vec{x}_j, \eta_j) \\ \vec{g}(\vec{x}_j, \zeta_j) \end{array} \right]^T \middle| \vec{x}_k \right\} \\ &= \left\{ \begin{array}{ll} 0, & k \neq j \\ \sum\limits_{i=1}^{r} \Pi_i \vec{x}_k^T \Omega_i \vec{x}_k, & k = j \end{array} \right. \end{aligned} \quad (6)$$

where $\Pi_i = \mathrm{diag}\left\{\Pi_{1i}, \Pi_{2i}\right\}$ and Ω_i $(i = 1, 2, \ldots, r)$ are known matrices with appropriate dimensions, and r is a known positive integer.

The initial state $\vec{x}_0$ and all the noise signals are uncorrelated with each other while possessing the following statistical properties:

$$\begin{aligned} &\mathbb{E}\left\{\vec{x}_0\right\} = \bar{x}_0, \ \ \mathbb{E}\left\{\left(\vec{x}_0 - \bar{x}_0\right)\left(\vec{x}_0 - \bar{x}_0\right)^T\right\} = \vec{P}_{0|0}, \\ &\mathbb{E}\left\{\omega_k\right\} = 0, \\ &\mathbb{E}\left\{\omega_k \omega_l^T\right\} = Q_k \delta_{k-l} + \sum_{t=1}^{f_k} Q_{k,l} \delta_{k-l-t} + \sum_{t=1}^{d_k} Q_{k,l} \delta_{k-l+t} \\ &\mathbb{E}\left\{\vec{\nu}_k\right\} = 0, \ \ \mathbb{E}\left\{\vec{\nu}_k \vec{\nu}_k^T\right\} = \vec{R}_k \end{aligned} \quad (7)$$

where $\vec{P}_{0|0}$, $\vec{R}_k > 0$, $Q_k > 0$ and $Q_{k,l}$ are known matrices with appropriate dimensions.

Remark 1: Note that the system measurement model (3) was used in [4,17,19]. As pointed out in [4,19], the random variable $\gamma_{k,i}$ accounts for the random varying delay of the i-th sensor and the value $\beta_{k,i}$ represents the probabilities of delay in the measurements of the i-th sensor. The delayed model in [4] considers the case where the measurements from multiple sensors could have different random delay characteristics. Following the standard practice of communication network design [20], the assumption of one-step sensor delay is based on the supposition that the induced data latency from the sensor to the controller is restricted not to exceed the sampling period.

By defining

$$\begin{aligned}
&x_k := \begin{bmatrix} \vec{x}_k \\ \vec{x}_{k-1} \end{bmatrix}, \ A_k := \begin{bmatrix} \vec{A}_k & 0 \\ I & 0 \end{bmatrix}, \ B_k := \begin{bmatrix} \vec{B}_k \\ 0 \end{bmatrix}, \\
&C_k := \begin{bmatrix} \vec{C}_k & 0 \\ 0 & \vec{C}_{k-1} \end{bmatrix}, \ f(x_k, \eta_k) := \begin{bmatrix} \vec{f}(\vec{x}_k, \eta_k) \\ 0 \end{bmatrix}, \\
&\nu_k := \begin{bmatrix} \vec{\nu}_k \\ \vec{\nu}_{k-1} \end{bmatrix}, \ \Upsilon_k := \begin{bmatrix} I-\Gamma_k & \Gamma_k \end{bmatrix}, \\
&g(x_k, \zeta_k, \zeta_{k-1}) := \begin{bmatrix} \vec{g}(\vec{x}_k, \zeta_k) \\ \vec{g}(\vec{x}_{k-1}, \zeta_{k-1}) \end{bmatrix},
\end{aligned} \tag{8}$$

we have the following compact form:

$$x_{k+1} = A_k x_k + f(x_k, \eta_k) + B_k \omega_k, \tag{9}$$

$$y_k = \Upsilon_k \left[C_k x_k + g(x_k, \zeta_k, \zeta_{k-1}) + \nu_k \right] \tag{10}$$

where ν_k is the measurement noise of the augmented system (9)-(10). It follows readily from (7)-(8) that ν_k obeys

$$\begin{aligned}
&\mathbb{E}\{\nu_k\} = 0, \\
&\mathbb{E}\left\{\nu_k \nu_l^T\right\} = R_k \delta_{k-l} + R_{k,k-1}\delta_{k-l-1} + R_{k,k+1}\delta_{k-l+1}
\end{aligned} \tag{11}$$

with

$$\begin{aligned}
R_k &= \begin{bmatrix} \vec{R}_k & 0 \\ 0 & \vec{R}_{k-1} \end{bmatrix}, \ R_{k,k-1} = \begin{bmatrix} 0 & 0 \\ \vec{R}_{k-1} & 0 \end{bmatrix}, \\
R_{k,k+1} &= \begin{bmatrix} 0 & \vec{R}_k \\ 0 & 0 \end{bmatrix}.
\end{aligned}$$

For convenience of later developments, define the following notations:

$$\begin{aligned}
&\bar{\Upsilon}_k := \mathbb{E}\{\Upsilon_k\} = \begin{bmatrix} I-\bar{\Gamma}_k & \bar{\Gamma}_k \end{bmatrix}, \\
&\tilde{\Upsilon}_k := \Upsilon_k - \bar{\Upsilon}_k = \begin{bmatrix} \bar{\Gamma}_k - \Gamma_k & \Gamma_k - \bar{\Gamma}_k \end{bmatrix},
\end{aligned} \tag{12}$$

where $\bar{\Gamma}_k := \text{diag}\{\beta_{k,1}, \beta_{k,2}, \dots, \beta_{k,m}\}$. According to (4), we have an easily accessible result that $\tilde{\Upsilon}_k$ is a zero-mean stochastic matrix sequence.

For system (9)-(10), we are interested in designing a filter of the following form:

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k}, \tag{13}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(y_{k+1} - \bar{\Upsilon}_{k+1} C_{k+1} \times \hat{x}_{k+1|k}) \tag{14}$$

where $\hat{x}_{k|k}$ is the estimate of x_k at time k with $\hat{x}_{0|0} = \begin{bmatrix} \bar{x}_0^T & 0 \end{bmatrix}^T$, $\hat{x}_{k+1|k}$ is the one-step prediction at time k, and K_{k+1} is the filter parameter to be determined.

The criterion for the addressed filtering problem is that the desired filter parameter in (13)-(14) should minimize the following cost function

$$\begin{aligned}
&\aleph_{k+1}(K_{k+1}) \\
:=&\mathbb{E}\left\{ \left(x_{k+1} - \hat{x}_{k+1|k+1}\right)^T W_{k+1} \left(x_{k+1} - \hat{x}_{k+1|k+1}\right) \right\}
\end{aligned} \tag{15}$$

subject to the gain constraint

$$M_{k+1} K_{k+1} N_{k+1} = F_{k+1} \tag{16}$$

where M_{k+1}, N_{k+1} and F_{k+1} are known matrices. As discussed in [26], the symmetric positive-definite weighting matrix W_{k+1} characterizes how much the state elements should be updated relative to each other, which gives a performance index. Furthermore, the matrices M_{k+1} and N_{k+1} are assumed to be of, respectively, full row rank and full column rank.

Remark 2: As pointed out in [26], the gain-constrained filtering problem stems from the data-injection issue arose in practice because 1) the data-injection is restricted to ensure the unbiasedness of the state estimates irrespective to the arbitrary unknown exogenous inputs; 2) the data-injection is restricted to simplify the estimator structure so as to facilitate the multiprocessor implementation for applications or to deal with the partial/complete sensor outage; and 3) the data-injection is restricted to guarantee the state estimates satisfying a linear equality constraint. Note that the gain-constrained filtering problem has been investigated for a broad class of real-time dynamical systems, see e.g. the estimation problem of two state continuous stirred tank reactor [15], the tracking problem of a land based vehicle [21], the tracking problem of a vehicle along circular roads [32] and so on.

Remark 3: In view of (5) and (7), it is reasonable to have the time update equation obeying (13). We will show later that the filter (13)-(14) to be developed is also unbiased. Moreover, due to the stochasticity resulting from multiple sources (stochastic nonlinearities, probabilistic sensor delays as well as correlated noises) and filter gain constraints, we aim to pursue the local optimality of filter design in the sense of minimizing the cost function (15) on the filtering error at each sampling instant.

III. Main Results

To proceed, we introduce the following lemmas which will be helpful in deriving our main results. For presentation clarity, we place all proofs of the results in appendices.

Lemma 1: [8] Let $A = [a_{ij}]_{n\times n}$ be a real matrix and $B = \text{diag}\{b_1, b_2, \dots, b_n\}$ be a diagonal stochastic matrix. Then

$$\mathbb{E}\{BAB^T\} = \begin{bmatrix} \mathbb{E}\{b_1^2\} & \mathbb{E}\{b_1 b_2\} & \cdots & \mathbb{E}\{b_1 b_n\} \\ \mathbb{E}\{b_2 b_1\} & \mathbb{E}\{b_2^2\} & \cdots & \mathbb{E}\{b_2 b_n\} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}\{b_n b_1\} & \mathbb{E}\{b_n b_2\} & \cdots & \mathbb{E}\{b_n^2\} \end{bmatrix} \circ A$$

where $\circ$ is the Hadamard product.

Lemma 2: The state covariance matrix $X_{k+1} = \mathbb{E}\left\{x_{k+1}x_{k+1}^T\right\}$ obeys the following recursion:

$$\begin{aligned} X_{k+1} =& A_k X_k A_k^T + A_k \mathscr{G}_k B_k^T + B_k \mathscr{G}_k^T A_k^T \\ &+ \sum_{i=1}^{r} H_1^T \Pi_{1i} \mathrm{tr}\left(\bar{\Omega}_{1i} X_k\right) H_1 + B_k Q_k B_k^T \end{aligned} \tag{17}$$

with initial value $X_0 = \mathrm{diag}\{\bar{x}_0\bar{x}_0^T, 0\} + \mathrm{diag}\{\vec{P}_{0|0}, 0\}$ and

$$\begin{aligned} &\mathscr{G}_k := B_{k-1}Q_{k-1,k} + \sum_{t=2}^{d_k}\left(\prod_{j=2}^{t} A_{k+1-j}\right) B_{k-t}Q_{k-t,k}, \\ &\bar{\Omega}_{1i} := H_1^T\Omega_i H_1, \\ &H_1 := \left[\begin{array}{cc} I_n & 0 \end{array}\right]. \end{aligned} \tag{18}$$

Proof: See Appendix -A. ■

Lemma 3: The one-step prediction error covariance $P_{k+1|k} = \mathbb{E}\left\{\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T\right\}$ is given by

$$\begin{aligned} P_{k+1|k} =& A_k P_{k|k} A_k^T + A_k \mathscr{Z}_k B_k^T + B_k \mathscr{Z}_k^T A_k^T \\ &+ \sum_{i=1}^{r} H_1^T \Pi_{1i} \mathrm{tr}\left(\bar{\Omega}_{1i} X_k\right) H_1 + B_k Q_k B_k^T \end{aligned} \tag{19}$$

with

$$\begin{aligned} &\mathscr{Z}_k \\ :=& \left(I - K_k\bar{\Upsilon}_k C_k\right)\mathscr{G}_k - \sum_{t=2}^{d_k}\Bigg\{\Bigg[\Bigg(\prod_{j=2}^{t}\Big(I - K_{k+2-j} \\ &\times \bar{\Upsilon}_{k+2-j}C_{k+2-j}\Big)A_{k+1-j}\Bigg)K_{k+1-t}\bar{\Upsilon}_{k+1-t}C_{k+1-t}\Bigg] \\ &\times\left[B_{k-t}Q_{k-t,k} + \sum_{i=t+1}^{d_k}\left(\prod_{l=t+1}^{i} A_{k+1-l}\right)B_{k-i}Q_{k-i,k}\right]\Bigg\} \end{aligned}$$

where $\bar{\Omega}_{1i}$, H_1 and $\mathscr{G}_k$ are defined in (18), and $P_{k|k} := \mathbb{E}\left\{\tilde{x}_{k|k}\tilde{x}_{k|k}^T\right\}$ is the filtering error covariance with $\tilde{x}_{k|k} = x_k - \hat{x}_{k|k}$ being the filtering error.

Proof: See Appendix -B. ■

In Lemma 2 and Lemma 3, similar to [6], the recursions of state covariance and the one-step prediction error covariance have been established. Next, we will proceed to show that 1) the proposed filtering scheme is unbiased; and 2) the cost function (15) with constraint (16) is minimized at each sampling instant by appropriately designing the filter parameter.

Theorem 1: The filter in (13)-(14) is unbiased. Moreover, the filtering error covariance $P_{k+1|k+1}$ obeys the following recursion

$$\begin{aligned} &P_{k+1|k+1} \\ =& \left(I - K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)P_{k+1|k}\left(I - K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)^T \\ &- \left(I - K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)\left(\mathscr{R}_{k+1} + \mathscr{J}_{k+1}\right)\bar{\Upsilon}_{k+1}^T K_{k+1}^T \\ &- K_{k+1}\bar{\Upsilon}_{k+1}\left(\mathscr{R}_{k+1} + \mathscr{J}_{k+1}\right)^T\left(I - K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)^T \\ &+ K_{k+1}[\bar{\Upsilon}_{k+1}\left(\bar{\Omega}_{2,k+1} + R_{k+1}\right)\bar{\Upsilon}_{k+1}^T + \mathscr{K}_{k+1} \\ &+ \mathscr{L}_{k+1} + \mathscr{Q}_{k+1}]K_{k+1}^T \end{aligned} \tag{20}$$

where

$$\begin{aligned} &\bar{\Omega}_{2,k+1} := \mathrm{diag}\Bigg\{\sum_{i=1}^{r}\Pi_{2i}\mathrm{tr}\left(H_1^T\Omega_i H_1 X_{k+1}\right), \\ &\qquad \sum_{i=1}^{r}\Pi_{2i}\mathrm{tr}\left(H_2^T\Omega_i H_2 X_{k+1}\right)\Bigg\}, \\ &H_2 := \left[\begin{array}{cc} 0 & I_n \end{array}\right], \\ &\breve{\Gamma}_{k+1} := \mathrm{diag}\{\beta_{k+1,1}\left(1 - \beta_{k+1,1}\right), \beta_{k+1,2}\left(1 - \beta_{k+1,2}\right), \\ &\qquad \cdots, \beta_{k+1,m}\left(1 - \beta_{k+1,m}\right)\}, \\ &\mathscr{R}_{k+1} := -A_k K_k \bar{\Upsilon}_k R_{k,k+1}, \\ &\mathscr{J}_{k+1} := -A_k K_k \bar{\Upsilon}_k \Psi_{k+1}, \\ &\mathscr{K}_{k+1} := \breve{\Gamma}_{k+1}\circ\left(\bar{H}C_{k+1}X_{k+1}C_{k+1}^T\bar{H}^T\right), \\ &\bar{H} := \left[\begin{array}{cc} I_m & -I_m \end{array}\right], \\ &\mathscr{L}_{k+1} := \breve{\Gamma}_{k+1}\circ\left(\bar{H}\bar{\Omega}_{2,k+1}\bar{H}^T\right), \\ &\mathscr{Q}_{k+1} := \breve{\Gamma}_{k+1}\circ\left(\bar{H}R_{k+1}\bar{H}^T\right), \\ &\Psi_{k+1} := \left[\begin{array}{cc} 0 & \sum_{i=1}^{r}\Pi_{2i}\mathrm{tr}\left(H_2^T\Omega_i H_2 X_{k+1}\right) \\ 0 & 0 \end{array}\right], \end{aligned} \tag{21}$$

and X_{k+1} and $P_{k+1|k}$ are defined, respectively, in (17) and (19).

Proof: See Appendix -C. ■

Having obtained the filtering error covariance, we are now ready to deal with the optimization issue of the cost function (15) under the constraint (16). Based on [26], the filter parameter is designed to minimize the cost function (15) subject to the constraint (16).

Theorem 2: Let the filter parameter K_{k+1} be

$$\begin{aligned} K_{k+1} =& \mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1} + \mathcal{I}_{k+1}(M_{k+1}\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}N_{k+1} \\ &- F_{k+1})\mathcal{J}_{k+1}. \end{aligned} \tag{22}$$

Then, the cost function $\aleph_{k+1}\left(K_{k+1}\right)$ in (15) with the constraint (16) is minimized by K_{k+1} defined in (22). Moreover, the filtering error covariance $P_{k+1|k+1}$ is given by

$$\begin{aligned} &P_{k+1|k+1} \\ =& P_{k+1|k} - \mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}\mathcal{H}_{k+1}^T + \mathcal{I}_{k+1}(M_{k+1}\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1} \\ &\times N_{k+1} - F_{k+1})\left(N_{k+1}^T\mathcal{S}_{k+1}^{-1}N_{k+1}\right)^{-1} \\ &\times(M_{k+1}\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}N_{k+1} - F_{k+1})^T\mathcal{I}_{k+1}^T \end{aligned} \tag{23}$$

where

$$\begin{aligned} &\mathcal{H}_{k+1} := \left(\mathscr{R}_{k+1} + \mathscr{J}_{k+1} + P_{k+1|k}C_{k+1}^T\right)\bar{\Upsilon}_{k+1}^T, \\ &\mathcal{I}_{k+1} := W_{k+1}^{-1}M_{k+1}^T\left(M_{k+1}W_{k+1}^{-1}M_{k+1}^T\right)^{-1}, \\ &\mathcal{J}_{k+1} := \left(N_{k+1}^T\mathcal{S}_{k+1}^{-1}N_{k+1}\right)^{-1}N_{k+1}^T\mathcal{S}_{k+1}^{-1}, \\ &\mathcal{S}_{k+1} := \bar{\Upsilon}_{k+1}\Big[C_{k+1}P_{k+1|k}C_{k+1}^T + \bar{\Omega}_{2,k+1} + R_{k+1} \\ &\qquad + \left(\mathscr{R}_{k+1} + \mathscr{J}_{k+1}\right)^T C_{k+1}^T + C_{k+1}(\mathscr{R}_{k+1} \\ &\qquad + \mathscr{J}_{k+1})\Big]\bar{\Upsilon}_{k+1}^T + \mathscr{K}_{k+1} + \mathscr{L}_{k+1} + \mathscr{Q}_{k+1}, \end{aligned} \tag{24}$$

and $\bar{\Omega}_{2,k+1}$, $\mathscr{R}_{k+1}$, $\mathscr{J}_{k+1}$, $\mathscr{K}_{k+1}$, $\mathscr{L}_{k+1}$ and $\mathscr{Q}_{k+1}$ are defined in (21).

Proof: See Appendix -D. ■

Remark 4: In this paper, we examine how the probabilistic sensor delays, stochastic nonlinearities, correlated noises and gain constraint influence the performance of the recursive filter for a class of time-varying nonlinear stochastic system. In Theorem 2, all these important aspects are dealt with in a unified yet effective framework. In particular, the proposed filtering algorithm has the following advantages: 1) the filter structure is simple and easy to be implemented; 2) all the system parameters, occurrence probabilities of the sensor delays, statistical characteristics of the stochastic nonlinearities and the moment information of the correlated noises are explicitly reflected in the algorithm; and 3) the algorithm is of a recursive nature suitable for online applications. In the case where global optimality of the recursive filter approach becomes a concern, specific efforts would have to be made for our future research.

Remark 5: It is well known that the traditional Kalman filter serves as an optimal filter in the minimum-variance sense for linear systems with the assumption that the model is exactly known. In order to cope with the network-induced phenomena and the gain constraints, some important filter algorithms have been developed for linear systems, see e.g. [18, 26]. Unfortunately, the existing estimation methods based on the traditional Kalman filtering theory cannot be simply applied to the addressed system (1)-(2) in the simultaneous presence of the probabilistic sensor delays, stochastic nonlinearities, gain constraint and correlated noises. To be specific, the following aspects prevent the existing methods from being applied to the recursive filtering problem considered in this paper: 1) the probabilistic sensor delays are described by a series of random variables, where each sensor is allowed to have individual delay rate, 2) the stochastic nonlinearities described by statistical means are taken into account to better reflect the reality, and 3) the process noises are finite-step auto-correlated and there is a constraint on the filter gain. In conclusion, our developed recursive filtering scheme provides another approach that complements the existing techniques for handling network-induced phenomena and gain constraints.

Remark 6: To deal with the computational complexity of the proposed filtering algorithm, we recall that the length of time-horizon is N, and the variable dimensions can be seen from $x_k \in \mathbb{R}^{2n}$ and $y_k \in \mathbb{R}^m$. It is not difficult to calculate the overall computational complexity of the given algorithm as $O(N(2n)^3)$, which depends linearly on the length of finite time horizon and polynomially on the variable dimension. Generally, the classical Kalman filter and extended Kalman filter have less computation burden than the proposed filter method. However, the new filter scheme has a potential advantage to deal with the complicated problem with multiple constraints addressed in this paper.

IV. An Illustrative Example

Consider the following system:

$$\begin{cases} \vec{x}_{k+1} = \vec{A}_k\vec{x}_k + \vec{f}(\vec{x}_k,\eta_k) + \vec{B}_k\omega_k, \\ \vec{y}_k = \vec{C}_k\vec{x}_k + \vec{g}(\vec{x}_k,\zeta_k) + \vec{\nu}_k, \\ \omega_k = \varsigma_k + \varsigma_{k-1}, \end{cases}$$

with

$$\vec{A}_k = \begin{bmatrix} 1 & T & \frac{T^2}{2} \\ 0 & 1 & T \\ 0 & 0 & 1 \end{bmatrix}, \quad \vec{B}_k = \begin{bmatrix} \frac{T^2}{2} \\ T \\ 1 \end{bmatrix},$$

$$\vec{C}_k = \begin{bmatrix} 1.05 & 0.28 & 0.03 \\ 0 & 1+0.2\sin(3k) & 0.15 \end{bmatrix},$$

where $T = 0.01$ is the sampling period, $\varsigma_k \in \mathbb{R}$ and $\vec{\nu}_k \in \mathbb{R}^2$ are zero-mean Gaussian noises with variances 0.05 and $0.01I_2$, respectively.

The delayed sensor measurement is described by

$$\begin{cases} y_{k,1} = (1-\gamma_{k,1})\vec{y}_{k,1} + \gamma_{k,1}\vec{y}_{k-1,1} \\ y_{k,2} = (1-\gamma_{k,2})\vec{y}_{k,2} + \gamma_{k,2}\vec{y}_{k-1,2} \end{cases}$$

where $\vec{y}_{k,i}$ $(i = 1,2)$ is the i-th element of the ideal output $\vec{y}_k$, $y_{k,i}$ $(i = 1,2)$ is the i-th element of the actual measured output y_k. The random variables $\gamma_{k,i}$ $(i = 1,2)$ satisfy the Bernoulli distribution with $\bar{\Gamma}_k := \text{diag}\{\beta_{k,1},\beta_{k,2}\} = \{0.1, 0.05\}$.

The stochastic nonlinearities $\vec{f}(\vec{x}_k,\eta_k)$ and $\vec{g}(\vec{x}_k,\zeta_k)$ are chosen as follows:

$$\vec{f}(\vec{x}_k,\eta_k) = \begin{bmatrix} 0.5 \\ 0.4 \\ 0.2 \end{bmatrix} [0.5\text{sign}\,(\vec{x}_{k,1})\,\vec{x}_{k,1}\eta_{k,1} + 0.4\text{sign}\,(\vec{x}_{k,2}) \times \vec{x}_{k,2}\eta_{k,2} + 0.3\text{sign}\,(\vec{x}_{k,3})\,\vec{x}_{k,3}\eta_{k,3}]$$

$$\vec{g}(\vec{x}_k,\zeta_k) = \begin{bmatrix} 0.3 \\ 0.6 \end{bmatrix} [0.5\text{sign}\,(\vec{x}_{k,1})\,\vec{x}_{k,1}\zeta_{k,1} + 0.4\text{sign}\,(\vec{x}_{k,2}) \times \vec{x}_{k,2}\zeta_{k,2} + 0.3\text{sign}\,(\vec{x}_{k,3})\,\vec{x}_{k,3}\zeta_{k,3}]$$

where $\vec{x}_{k,i}$ $(i = 1,2,3)$ denotes the i-th element of the system state, $\eta_{k,i}$ and $\zeta_{k,i}$ $(i = 1,2,3)$ stand for zero-mean uncorrelated Gaussian white noises with variance $\Xi_k = 0.2$. It is not difficult to verify that the above stochastic nonlinearities satisfy (5)-(6) with

$$\Pi_{1i} = \begin{bmatrix} 0.25 & 0.20 & 0.10 \\ 0.20 & 0.16 & 0.08 \\ 0.10 & 0.08 & 0.04 \end{bmatrix}, \quad \Pi_{2i} = \begin{bmatrix} 0.09 & 0.18 \\ 0.18 & 0.36 \end{bmatrix},$$

$$\Omega_i = \Xi_k \times \begin{bmatrix} 0.25 & 0 & 0 \\ 0 & 0.16 & 0 \\ 0 & 0 & 0.09 \end{bmatrix}.$$

In the simulation, set the initial value of estimation as $\hat{\vec{x}}_{0|0} = \begin{bmatrix} 400 & 100 & 9 \end{bmatrix}^T$ and $\vec{P}_{0|0} = 0.01I_3$. Other parameters are chosen as $W_{k+1} = 0.15I_6$ and

$$M_{k+1} = \begin{bmatrix} 1 & 0 & 0.35 & 1 & 0 & 0.35 \\ 0 & 1 & 0 & 0.62 & 1 & 0.01 \end{bmatrix},$$

$$N_{k+1} = \begin{bmatrix} 1 \\ 0.82 \end{bmatrix} \quad F_{k+1} = \begin{bmatrix} 0.35 \\ 0.6 \end{bmatrix}.$$

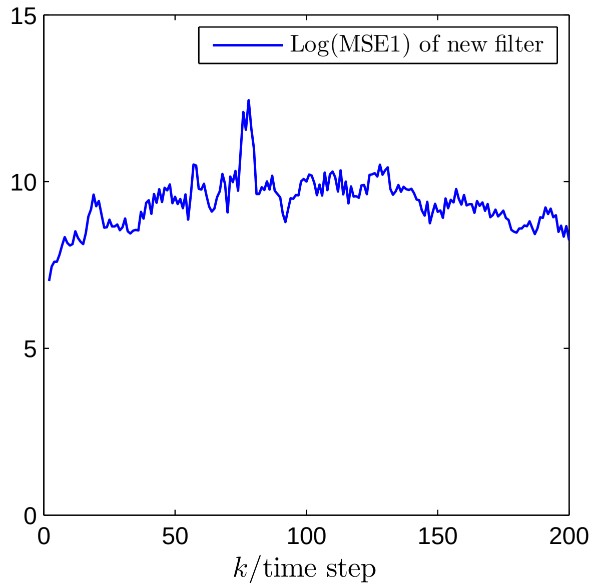

Fig. 1. Log(MSE1)

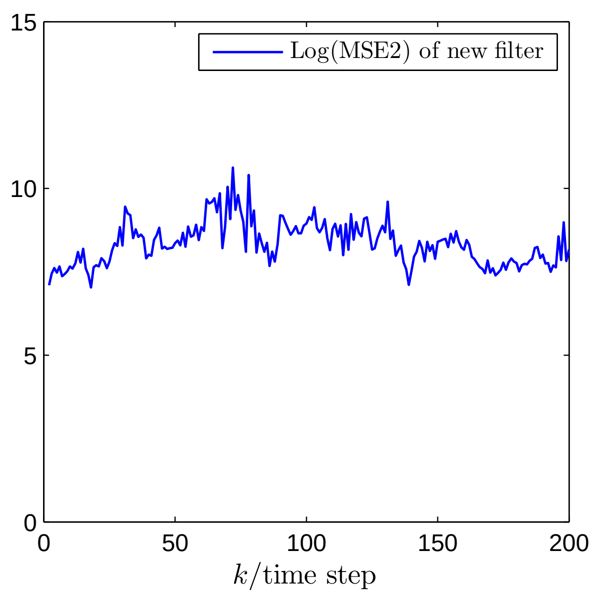

Fig. 2. Log(MSE2)

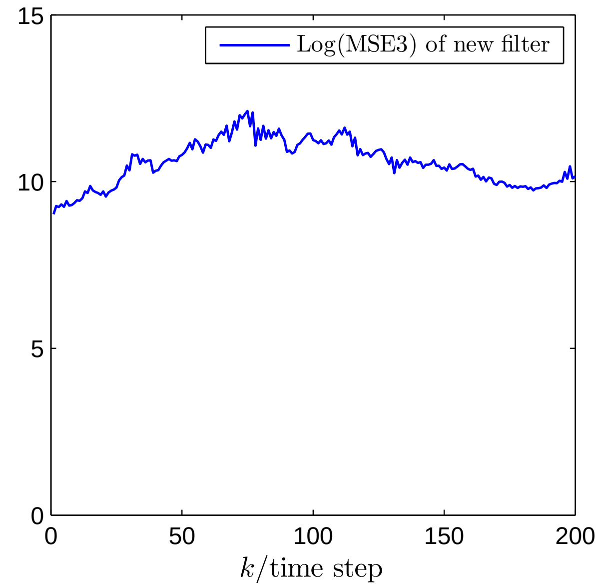

Fig. 3. Log(MSE3)

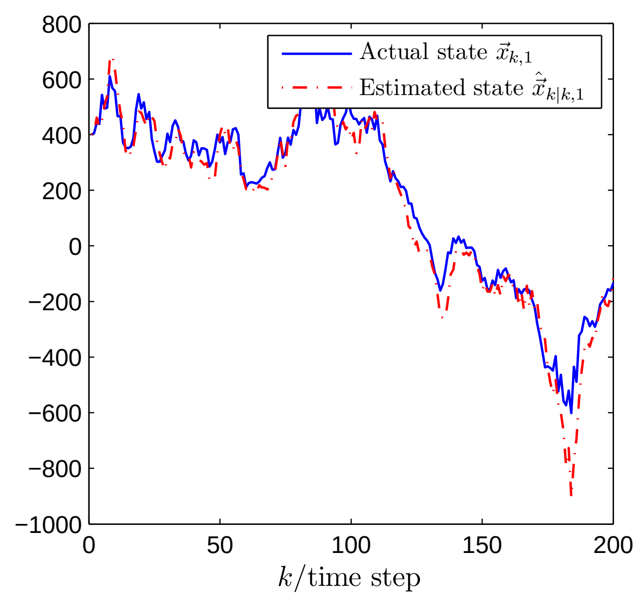

Fig. 4. Actual state $\vec{x}_{k,1}$ and estimated state $\hat{\vec{x}}_{k|k,1}$

Let MSEi denote the mean square error (MSE) for the estimation of $\vec{x}_{k,i}$, i.e., $(1/M)\sum_{j=1}^{M}\left(\vec{x}_{k,i}^{(j)}-\hat{\vec{x}}_{k|k,i}^{(j)}\right)^2$ $(i = 1,2,3)$, where $M = 100$ denotes the number of simulation test. The simulation results are shown in Figs. 1-6. Among them, the Log(MSE)i for the estimation of $\vec{x}_{k,i}$ $(i = 1,2,3)$ are shown in Figs. 1-3. Moreover, the trajectories of the actual states $\vec{x}_{k,i}$ and their estimates $\hat{\vec{x}}_{k|k,i}$ $(i = 1,2,3)$ are plotted in Figs. 4-6. The simulation results illustrate that the presented scheme performs well in estimating the system states, which is due to the fact that we have made specific efforts to compensate the effects of the probabilistic sensor delays, stochastic nonlinearities as well as correlated noises.

V. Conclusions

In this paper, the recursive filtering problem has been investigated for a class of nonlinear systems in the simultaneous presence of probabilistic sensor delays, stochastic nonlinearities, gain constraint and correlated noises. The filter parameter has been designed such that the specified cost function with gain constraint is minimized at each sampling instant. It has been shown that the proposed algorithm is of a recursive form suitable for online applications. A simulation example has been given to illustrate the effectiveness of the presented filtering scheme, where the pro-

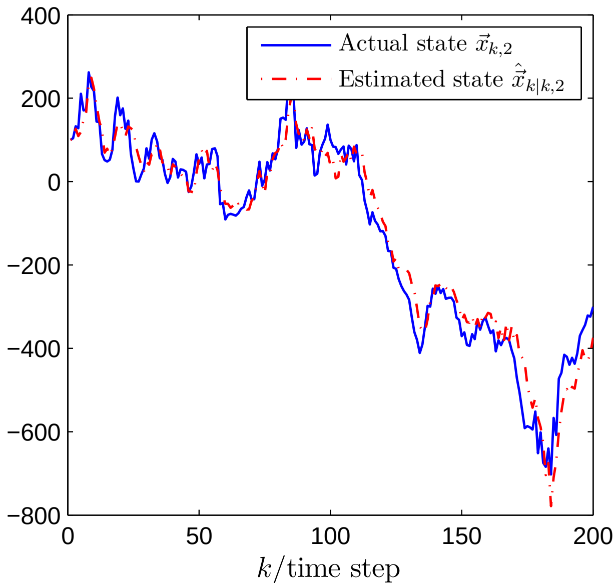

Fig. 5. Actual state $\vec{x}_{k,2}$ and estimated state $\hat{\vec{x}}_{k|k,2}$

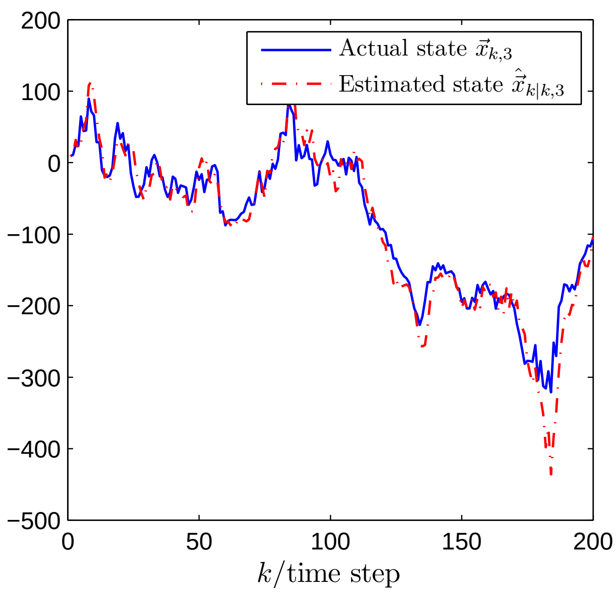

Fig. 6. Actual state $\vec{x}_{k,3}$ and estimated state $\hat{\vec{x}}_{k|k,3}$

posed method has been applied to estimate the states for the addressed time-varying system involving the stochastic nonlinearities, the randomly varying delayed observations coming from two sensors with different delay characteristics and the correlated process noises. The results are potentially applicable to state estimation for stabilizing the output feedback control systems. Further research topics include the extension of the main results to the filtering problem for a more general class of time-varying nonlinear systems and the derivation of globally optimal filter for the considered time-varying nonlinear systems.

Appendices

A. Proof of Lemma 2

Proof: By considering (5)-(7) and (9), the recursion of X_{k+1} can be obtained as follows:

$$\begin{aligned} X_{k+1} =& A_k X_k A_k^T + A_k \mathscr{G}_k B_k^T + B_k \mathscr{G}_k^T A_k^T \\ &+ \mathbb{E}\left\{ f\left(x_k,\eta_k\right) f^T\left(x_k,\eta_k\right)\right\} + B_k Q_k B_k^T \end{aligned} \tag{25}$$

where $\mathscr{G}_k := \mathbb{E}\left\{x_k\omega_k^T\right\}$. Together with (6) and (8), we have

$$\mathbb{E}\left\{ f\left(x_k,\eta_k\right) f^T\left(x_k,\eta_k\right)\right\} = \sum_{i=1}^{r} H_1^T \Pi_{1i} \mathrm{tr}\left(\bar{\Omega}_{1i} X_k\right) H_1 \tag{26}$$

where $\bar{\Omega}_{1i}$ and H_1 are defined in (18).

By (7), the term $\mathscr{G}_k$ can be calculated as follows:

$$\begin{aligned} \mathscr{G}_k =& A_{k-1}\mathbb{E}\left\{x_{k-1}\omega_k^T\right\} + B_{k-1}Q_{k-1,k} \\ &\vdots \\ =& B_{k-1}Q_{k-1,k} + \sum_{t=2}^{d_k}\left(\prod_{j=2}^{t} A_{k+1-j}\right) B_{k-t}Q_{k-t,k}. \end{aligned} \tag{27}$$

Note that, in deriving (27), we have used the fact that η_k is uncorrelated with ω_k. Substituting (26) and (27) into (25) yields (17). ■

B. Proof of Lemma 3

Proof: It follows from (9) and (13) that

$$\tilde{x}_{k+1|k} = A_k \tilde{x}_{k|k} + f(x_k,\eta_k) + B_k\omega_k,$$

and then the one-step prediction error covariance can be determined as

$$\begin{aligned} P_{k+1|k} =& A_k P_{k|k} A_k^T + A_k \mathscr{Z}_k B_k^T + B_k \mathscr{Z}_k^T A_k^T \\ &+ \sum_{i=1}^{r} H_1^T \Pi_{1i} \mathrm{tr}\left(\bar{\Omega}_{1i} X_k\right) H_1 + B_k Q_k B_k^T \end{aligned} \tag{28}$$

where $\mathscr{Z}_k := \mathbb{E}\left\{\tilde{x}_{k|k}\omega_k^T\right\}$.

From (10), (14) and (27), the term $\mathscr{Z}_k$ can be calculated as:

$$\begin{aligned} &\mathscr{Z}_k \\ =& \left(I - K_k\bar{\Upsilon}_k C_k\right)\mathscr{G}_k - \left(I - K_k\bar{\Upsilon}_k C_k\right)\mathbb{E}\left\{\hat{x}_{k|k-1}\omega_k^T\right\} \\ =& \left(I - K_k\bar{\Upsilon}_k C_k\right)\mathscr{G}_k - \left(I - K_k\bar{\Upsilon}_k C_k\right)A_{k-1} \\ &\times \left(I - K_{k-1}\bar{\Upsilon}_{k-1} C_{k-1}\right)\mathbb{E}\left\{\hat{x}_{k-1|k-2}\omega_k^T\right\} \\ &- \left(I - K_k\bar{\Upsilon}_k C_k\right)A_{k-1}K_{k-1}\bar{\Upsilon}_{k-1}C_{k-1}\mathbb{E}\left\{x_{k-1}\omega_k^T\right\} \\ &\vdots \\ =& \left(I - K_k\bar{\Upsilon}_k C_k\right)\mathscr{G}_k - \sum_{t=2}^{d_k}\left\{\left[\left(\prod_{j=2}^{t}(I - K_{k+2-j}\right.\right.\right. \\ &\left.\left.\times\bar{\Upsilon}_{k+2-j}C_{k+2-j})A_{k+1-j}\right)K_{k+1-t}\bar{\Upsilon}_{k+1-t}C_{k+1-t}\right] \\ &\left.\times\left[B_{k-t}Q_{k-t,k} + \sum_{i=t+1}^{d_k}\left(\prod_{l=t+1}^{i} A_{k+1-l}\right)B_{k-i}Q_{k-i,k}\right]\right\} \end{aligned}$$

(29)

where $\mathscr{G}_k$ is defined in (27). It follows from (28)-(29) that (19) holds. ∎

C. Proof of Theorem 1

Proof: To begin with, let us show the unbiasedness of the filter in (13)-(14). According to (10) and (14), the filtering error can be rewritten as

$$\begin{aligned}&\tilde{x}_{k+1|k+1}\\=&\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)\tilde{x}_{k+1|k}-K_{k+1}[\tilde{\Upsilon}_{k+1}C_{k+1}\\&\times x_{k+1}+\Upsilon_{k+1}g\left(x_{k+1},\zeta_{k+1},\zeta_k\right)+\Upsilon_{k+1}\nu_{k+1}].\end{aligned} \tag{30}$$

Taking mathematical expectation of both sides of (30), it follows from (9) and (13) that

$$\mathbb{E}\left\{\tilde{x}_{k+1|k+1}\right\}=\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)A_k\mathbb{E}\left\{\tilde{x}_{k|k}\right\}. \tag{31}$$

With the given initial condition, it is not difficult to show that $\mathbb{E}\left\{\tilde{x}_{k|k}\right\}=0$ for all $k\geq 0$, which confirms the unbiasedness of the filter (13)-(14). Subsequently, the filtering error covariance $P_{k+1|k+1}$ can be obtained as follows:

$$\begin{aligned}&P_{k+1|k+1}\\=&\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)P_{k+1|k}\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)^T\\&-\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)\mathscr{R}_{k+1}\bar{\Upsilon}_{k+1}^TK_{k+1}^T\\&-K_{k+1}\bar{\Upsilon}_{k+1}\mathscr{R}_{k+1}^T\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)^T\\&-\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)\mathscr{J}_{k+1}\bar{\Upsilon}_{k+1}^TK_{k+1}^T\\&-K_{k+1}\bar{\Upsilon}_{k+1}\mathscr{J}_{k+1}^T\left(I-K_{k+1}\bar{\Upsilon}_{k+1}C_{k+1}\right)^T\\&+K_{k+1}\bar{\Upsilon}_{k+1}\left(\bar{\Omega}_{2,k+1}+R_{k+1}\right)\bar{\Upsilon}_{k+1}^TK_{k+1}^T\\&+K_{k+1}\left(\mathscr{K}_{k+1}+\mathscr{L}_{k+1}+\mathscr{Q}_{k+1}\right)K_{k+1}^T,\end{aligned} \tag{32}$$

where

$$\begin{aligned}\mathscr{R}_{k+1}:=&\mathbb{E}\left\{\tilde{x}_{k+1|k}\nu_{k+1}^T\right\},\\\mathscr{J}_{k+1}:=&\mathbb{E}\left\{\tilde{x}_{k+1|k}g^T\left(x_{k+1},\zeta_{k+1},\zeta_k\right)\right\},\\\mathscr{K}_{k+1}:=&\mathbb{E}\left\{\tilde{\Upsilon}_{k+1}C_{k+1}x_{k+1}x_{k+1}^TC_{k+1}^T\tilde{\Upsilon}_{k+1}^T\right\},\\\mathscr{L}_{k+1}:=&\mathbb{E}\{\tilde{\Upsilon}_{k+1}g\left(x_{k+1},\zeta_{k+1},\zeta_k\right)\\&\times g^T\left(x_{k+1},\zeta_{k+1},\zeta_k\right)\tilde{\Upsilon}_{k+1}^T\},\\\mathscr{Q}_{k+1}:=&\mathbb{E}\left\{\tilde{\Upsilon}_{k+1}\nu_{k+1}\nu_{k+1}^T\tilde{\Upsilon}_{k+1}^T\right\},\end{aligned} \tag{33}$$

and $\bar{\Omega}_{2,k+1}$ is defined in (21).

By using the property of conditional expectation and applying Lemma 1, we have

$$\begin{aligned}\mathscr{K}_{k+1}&=\mathbb{E}\left\{\tilde{\Upsilon}_{k+1}C_{k+1}x_{k+1}x_{k+1}^TC_{k+1}^T\tilde{\Upsilon}_{k+1}^T\right\}\\&=\breve{\Gamma}_{k+1}\circ\left(\bar{H}C_{k+1}X_{k+1}C_{k+1}^T\bar{H}^T\right)\end{aligned} \tag{34}$$

where $\bar{H}$ and $\breve{\Gamma}_{k+1}$ are defined in (21). Following the same line of the derivation for (34), the terms of $\mathscr{L}_{k+1}$ and $\mathscr{Q}_{k+1}$ can be obtained as

$$\mathscr{L}_{k+1}=\breve{\Gamma}_{k+1}\circ\left(\bar{H}\bar{\Omega}_{2,k+1}\bar{H}^T\right), \tag{35}$$

$$\mathscr{Q}_{k+1}=\breve{\Gamma}_{k+1}\circ\left(\bar{H}R_{k+1}\bar{H}^T\right). \tag{36}$$

Next, let us determine the term $\mathscr{R}_{k+1}$ in (33) as follows:

$$\begin{aligned}&\mathscr{R}_{k+1}\\=&-\mathbb{E}\left\{A_k\left[\hat{x}_{k|k-1}+K_k\left(y_k-\bar{\Upsilon}_kC_k\hat{x}_{k|k-1}\right)\right]\nu_{k+1}^T\right\}\\=&-A_kK_k\bar{\Upsilon}_kR_{k,k+1}\end{aligned} \tag{37}$$

Note that, when deriving (37), we have used the facts that (i) x_{k+1} is uncorrelated with the measurement noise ν_{k+1}; and (ii) $\hat{x}_{k|k-1}$ is uncorrelated with the measurement noise ν_{k+1}. Similarly, the term $\mathscr{J}_{k+1}$ in (32) can be calculated as

$$\mathscr{J}_{k+1}=-A_kK_k\bar{\Upsilon}_k\Psi_{k+1} \tag{38}$$

where Ψ_{k+1} is defined in (21). Then, from (32) and (34)-(38), it can be concluded that (20) is true. ∎

D. Proof of Theorem 2

Proof: Define the Lagrangian

$$\begin{aligned}\Im_{k+1}\left(K_{k+1}\right):=&\mathrm{tr}\left[\aleph_{k+1}\left(K_{k+1}\right)\right]+2\mathrm{tr}[(M_{k+1}K_{k+1}\\&\times N_{k+1}-F_{k+1})\Lambda_{k+1}^T]\end{aligned} \tag{39}$$

where Λ_{k+1} is the Lagrange multiplier. Take the partial derivative of (39) with respect to K_{k+1} and Λ_{k+1}, respectively. Letting the derivative be zero yields

$$\begin{aligned}&\frac{\partial\Im_{k+1}}{\partial K_{k+1}}\\=&2\Bigg\{-W_{k+1}P_{k+1|k}C_{k+1}^T\bar{\Upsilon}_{k+1}^T-W_{k+1}(\mathscr{R}_{k+1}\\&+\mathscr{J}_{k+1})\bar{\Upsilon}_{k+1}^T+W_{k+1}K_{k+1}\Bigg[\bar{\Upsilon}_{k+1}\Bigg(C_{k+1}P_{k+1|k}C_{k+1}^T\\&+\bar{\Omega}_{2,k+1}+R_{k+1}+(\mathscr{R}_{k+1}+\mathscr{J}_{k+1})^TC_{k+1}^T\\&+C_{k+1}(\mathscr{R}_{k+1}+\mathscr{J}_{k+1})\Bigg)\bar{\Upsilon}_{k+1}^T+\mathscr{K}_{k+1}\\&+\mathscr{L}_{k+1}+\mathscr{Q}_{k+1}\Bigg]+M_{k+1}^T\Lambda_{k+1}N_{k+1}^T\Bigg\}\\=&2\left(-W_{k+1}\mathcal{H}_{k+1}+W_{k+1}K_{k+1}\mathcal{S}_{k+1}+M_{k+1}^T\Lambda_{k+1}N_{k+1}^T\right)\\=&0\end{aligned} \tag{40}$$

and

$$\frac{\partial\Im_{k+1}}{\partial\Lambda_{k+1}}=2\left(M_{k+1}K_{k+1}N_{k+1}-F_{k+1}\right)=0 \tag{41}$$

where $\mathcal{H}_{k+1}$ and $\mathcal{S}_{k+1}$ are defined in (24). According to (41), it can be concluded that (16) is satisfied.

On the other hand, it follows from (40) that

$$-W_{k+1}\mathcal{H}_{k+1}+W_{k+1}K_{k+1}\mathcal{S}_{k+1}+M_{k+1}^T\Lambda_{k+1}N_{k+1}^T=0 \tag{42}$$

Pre-multiplying and post-multiplying (42) by W_{k+1}^{-1} and $\mathcal{S}_{k+1}^{-1}$, we have

$$-\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}+K_{k+1}+W_{k+1}^{-1}M_{k+1}^T\Lambda_{k+1}N_{k+1}^T\mathcal{S}_{k+1}^{-1}=0 \tag{43}$$

Subsequently, pre-multiply and post-multiply (43) by M_{k+1} and N_{k+1}, respectively. Then, by considering (16), we can get

$$\begin{aligned}&-M_{k+1}\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}N_{k+1}+F_{k+1}+M_{k+1}W_{k+1}^{-1}M_{k+1}^{T}\\&\times\Lambda_{k+1}N_{k+1}^{T}\mathcal{S}_{k+1}^{-1}N_{k+1}=0.\end{aligned}\tag{44}$$

According to (44), we obtain

$$\begin{aligned}\Lambda_{k+1}=&\left(M_{k+1}W_{k+1}^{-1}M_{k+1}^{T}\right)^{-1}\left(M_{k+1}\mathcal{H}_{k+1}\mathcal{S}_{k+1}^{-1}N_{k+1}\right.\\&\left.-F_{k+1}\right)\left(N_{k+1}^{T}\mathcal{S}_{k+1}^{-1}N_{k+1}\right)^{-1}.\end{aligned}\tag{45}$$

From (43) and (45), the filter parameter can be determined as in (22). Subsequently, substituting (22) into (20) and after tedious algebraic manipulations, we can obtain the recursion (23). ■

References

[1] M. Basin, P. Shi, and D. Calderon-Alvarez, "Joint state filtering and parameter estimation for linear stochastic time-delay systems," *Signal Process.*, vol. 91, no. 4, pp. 782–792, Apr. 2011.

[2] M. Basin, P. Shi, and P. Soto, "Central suboptimal H_∞ filtering for nonlinear polynomial systems with multiplicative noise," *Journal of The Franklin Institute*, vol. 347, no. 9, pp. 1740–1754, Nov. 2010.

[3] R. Caballero-Águila, A. Hermoso-Carazo, J. D. Jiménez-López, J. Linares-Pérez, and S. Nakamori, "Recursive estimation of discrete-time signals from nonlinear randomly delayed observations," *Computers & Mathematics with Applications*, vol. 58, no. 6, pp. 1160–1168, Sept. 2009.

[4] R. Caballero-Águila, A. Hermoso-Carazo, J. D. Jiménez-López, J. Linares-Pérez, and S. Nakamori, "Signal estimation with multiple delayed sensors using covariance information," *Digital Signal Processing*, vol. 20, no. 2, pp. 528–540, Mar. 2010.

[5] H. Dong, Z. Wang, and H. Gao, "Distributed filtering for a class of time-varying systems over sensor networks with quantization errors and successive packet dropouts," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3164–3173, Jun. 2012.

[6] J. Feng, Z. Wang, and M. Zeng, "Recursive robust filtering with finite-step correlated process noises and missing measurements," *Circuits Syst. Signal Process.*, vol. 30, no. 6, pp. 1355–1368, Dec. 2011.

[7] J. A. Guu and C. H. Wei, "Tracking technique for manoeuvring target with correlated measurement noises and unknown parameters," *IEE Proceedings F-Radar and Signal Processing*, vol. 138, no. 3, pp. 278–288, Jun. 1991.

[8] R. A. Horn and C. R. Johnson, *Topic in Matrix Analysis*, New York: Cambridge University Press, 1991.

[9] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Robust sliding mode control for discrete stochastic systems with mixed time delays, randomly occurring uncertainties, and randomly occurring nonlinearities," *IEEE Trans. Ind. Electron.*, vol. 59, no. 7, pp. 3008–3015, Jul. 2012.

[10] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 48, no. 9, pp. 2007–2015, Sept. 2012.

[11] J. Hu, Z. Wang, H. Gao, and L. K. Stergioulas, "Probability-guaranteed H_∞ finite-horizon filtering for a class of nonlinear time-varying systems with sensor saturations," *Syst. & Control Lett.*, vol. 61, no. 4, pp. 477–484, Apr. 2012.

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

[13] S. J. Julier and J. J. LaViola, "On Kalman filtering with nonlinear equality constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2774–2784, Jun. 2007.

[14] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[15] S. Kolas, B. A. Foss, and T. S. Schei, "Constrained nonlinear state estimation based on the UKF approach," *Computers and Chemical Engineering*, vol. 33, no. 8, pp. 1386–1401, Aug. 2009.

[16] J. Linares-Pérez, A. Hermoso-Carazo, R. Caballero-Águila, and J. D. Jiménez-López, "Least-squares linear filtering using observations coming from multiple sensors with one- or two-step random delay," *Signal Process.*, vol. 89, no. 10, pp. 2045–2052, Oct. 2009.

[17] L. Liou and A. Ray, "A stochastic regulator for integrated communication and control systems: Parts I and II," *ASME Journal of Dynamic Systems, Measurement and Control*, vol. 113, pp. 604–619, Dec. 1991.

[18] J. Ma and S. Sun, "Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5181–5192, Nov. 2011.

[19] 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, Mar. 2005.

[20] A. Ray, "Performance analysis of medium access control protocols for distributed digital avionics," *ASME Journal of Dynamic Systems, Measurement and Control*, vol. 109, pp. 370–377, Dec. 1987.

[21] D. Simon, *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*, John Wiley & Sons, 2006.

[22] B. Shen, Z. Wang, H. Shu, and G. Wei, "On nonlinear H_∞ filtering for discrete-time stochastic systems with missing measurements," *IEEE Trans. Autom. Control*, vol. 53, no. 9, pp. 2170–2180, Oct. 2008

[23] E. Song, Y. Zhu, and Z. You, "The Kalman type recursive state estimator with a finite-step correlated process noises," In *Proceedings of the IEEE International Conference on Automation and Logistics*, Qingdao, China, September, 2008.

[24] S. Sun, L. Xie, and W. Xiao, "Optimal full-order and reduced-order estimators for discrete-time systems with multiple packet dropouts," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 4031–4038, Aug. 2008.

[25] S. Sun, L. Xie, W. Xiao, and Y. C. Soh, "Optimal linear estimation for systems with multiple packet dropouts," *Automatica*, vol. 44, no. 5, pp. 1333–1342, May 2008.

[26] B. O. S. Teixeira, J. Chandrasekar, H. J. Palanthandalam-Madapusi, L. A. B. Tôrres, L. A. Aguirre, and D. S. Bernstein, "Gain-constrained Kalman filtering for linear and nonlinear systems," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4113–4123, Sept. 2008.

[27] Z. Wang, X. Liu, Y. Liu, J. Liang, and V. Vinciotti, "An extended Kalman filtering approach to modelling nonlinear dynamic gene regulatory networks via short gene expression time series," *IEEE/ACM Trans. Computational Biology and Bioinformatics*, vol. 6, no. 3, pp. 410–419, Jul-Sept. 2009.

[28] Z. Wang, B. Shen, and X. Liu, "H_∞ filtering with randomly occurring sensor saturations and missing measurements," *Automatica*, vol. 48, no. 3, pp. 556–562, Mar. 2012.

[29] Z. Wang, B. Shen, H. Shu, and G. Wei, "Quantized H_∞ control for nonlinear stochastic time-delay systems with missing measurements," *IEEE Trans. Autom. Control*, vol. 57, no. 6, pp. 1431–1444, Jun. 2012.

[30] G. Wei, Z. Wang, and H. Shu, "Robust filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 45, no. 3, pp. 836–841, Mar. 2009.

[31] L. Wu and D. W. C. Ho, "Sliding mode control of singular stochastic hybrid systems," *Automatica*, vol. 46, no. 4, pp. 779–783, Apr. 2010.

[32] C. Yang and E. Blasch, "Kalman filtering with nonlinear state constraints," *Proc. of the 9th International Conference on Information Fusion*, Seattle, pp. 1–8, 2006.

[33] F. Yang, Z. Wang, G. Feng, and X. Liu, "Robust filtering with randomly varying sensor delay: the finite-horizon case," *IEEE Trans. Ciruits Syst. I, Reg. Papers*, vol. 56, no. 3, pp. 664–672, Mar. 2009.

[34] X. Yao, L. Wu, and W. Zheng, "Fault detection filter design for Markovian jump singular systems with intermittent measurements," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3099–3109, Jul. 2011.

[35] Y. Yaz and E. Yaz, "A new formulation of some discrete-time stochastic-parameter state estimation problems," *Appl. Math. Lett.*, vol. 10, no. 6, pp. 13–19, Nov. 1997.

[36] E. Yaz and Y. Yaz, "State estimation of uncertain nonlinear

stochastic systems with general criteria," *Appl. Math. Lett.*, vol. 14, no. 5, pp. 605–610, Aug. 2001.

[37] H. Zhang, G. Feng, G. Duan, and X. Lu, "H_∞ filtering for multiple-time-delay measurements," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1681–1688, Oct. 2006.

Jun Hu received the B.Sc. degree in information and computation science and M.Sc. degree in applied mathematics from Harbin University of Science and Technology, Harbin, China, in 2006 and 2009, respectively.

He is currently working toward the Ph.D. degree in control science and engineering with the Harbin Institute of Technology, Harbin, China. From September 2010 to September 2012, he was a Visiting Ph.D. Student in the Department of Information Systems and Computing, Brunel University, U.K. His research interests include nonlinear control and filtering, time-varying systems, stochastic systems.

He is an active reviewer for many international journals.

Zidong Wang (SM'03) was born in Jiangsu, China, in 1966. He received his B.Sc. degree in Mathematics in 1986 from Suzhou University, Suzhou, China, the M.Sc. degree in Applied Mathematics in 1990 and the Ph.D. degree in Electrical and Computer Engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently a Professor of Dynamical Systems and Computing at Brunel University, U.K. His research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 200 papers in refereed international journals.

Dr. Wang is currently serving as an Associate Editor for 12 international journals including *IEEE Transactions on Automatic Control*, *IEEE Transactions on Neural Networks*, *IEEE Transactions on Signal Processing*, *IEEE Transactions on Systems, Man, and Cybernetics - Part C* and *IEEE Transactions on Control Systems Technology*.

Bo Shen (M'12) received the B.Sc. degree in mathematics from Northwestern Polytechnical University, Xi'an, China, in 2003, and the Ph.D. degree in control theory and control engineering from Donghua University, Shanghai, China, in 2011.

He is currently with the School of Information Science and Technology, Donghua University. From 2009 to 2010, he was a Research Assistant with the Department of Electrical and Electronic Engineering, University of Hong Kong, Hong Kong. From 2010 to 2011, he was a Visiting Ph.D. Student with the Department of Information Systems and Computing, Brunel University, Uxbridge, U.K. His current research interests include nonlinear control and filtering, stochastic control and filtering, as well as complex networks and genetic regulatory networks.

He is an active reviewer for many international journals.

Huijun Gao (SM'09) received the Ph.D. degree in control science and engineering from Harbin Institute of Technology, China, in 2005.

He was a Research Associate with the Department of Mechanical Engineering, The University of Hong Kong, from November 2003 to August 2004. From October 2005 to October 2007, he carried out his postdoctoral research with the Department of Electrical and Computer Engineering, University of Alberta, Canada. Since November 2004, he has been with Harbin Institute of Technology, where he is currently a Professor and director of the Research Institute of Intelligent Control and Systems. His research interests include network-based control, robust control/filter theory, time-delay systems and their engineering applications.

Dr. Gao is an Associate Editor for *Automatica*, *IEEE Transactions on Industrial Electronics*, *IEEE Transactions on Systems Man and Cybernetics Part B: Cybernetics*, *IEEE Transactions on Fuzzy Systems*, *IEEE Transactions on Circuits and Systems - I*, *IEEE Transactions on Control Systems Technology* etc.