Extended Kalman Filtering with Stochastic Nonlinearities and Multiple Missing Measurements ⋆

Jun Hu [a], Zidong Wang [b], Huijun Gao [a], Lampros K. Stergioulas [b]

[a] *Research Institute of Intelligent Control and Systems, Harbin Institute of Technology, Harbin 150001, China.*

[b] *Department of Information Systems and Computing, Brunel University, Uxbridge, Middlesex, UB8 3PH, U.K.*

Abstract

In this paper, the extended Kalman filtering problem is investigated for a class of nonlinear systems with multiple missing measurements over a finite horizon. Both deterministic and stochastic nonlinearities are included in the system model, where the stochastic nonlinearities are described by statistical means that could reflect the multiplicative stochastic disturbances. The phenomenon of measurement missing occurs in a random way and the missing probability for each sensor is governed by an individual random variable satisfying a certain probability distribution over the interval $[0, 1]$. Such a probability distribution is allowed to be any commonly used distribution over the interval $[0, 1]$ with known conditional probability. The aim of the addressed filtering problem is to design a filter such that, in the presence of both the stochastic nonlinearities and multiple missing measurements, there exists an upper bound for the filtering error covariance. Subsequently, such an upper bound is minimized by properly designing the filter gain at each sampling instant. It is shown that the desired filter can be obtained in terms of the solutions to two Riccati-like difference equations that are of a form suitable for recursive computation in online applications. An illustrative example is given to demonstrate the effectiveness of the proposed filter design scheme.

Key words: Nonlinear systems; Extended Kalman filter; Stochastic nonlinearities; Multiple missing measurements; Recursive filter; Riccati-like difference equation.

1 Introduction

In the past few decades, the filtering or state estimation problems for stochastic systems have been extensively investigated. Accordingly, the filter theory has been successfully applied in many branches of practical domains such as computer vision, communications, navigation and tracking systems, econometrics and finance, etc. It is well known that the traditional Kalman filter (KF) serves as an optimal filter in the least mean square sense for *linear* systems with the assumption that the system model is exactly known. In the case that the system model is nonlinear and/or uncertain, there has been an increasing research effort to improve KF with hope to enhance their capabilities of handling nonlinearities and uncertainties. Along this direction, many alternative filtering schemes have been reported in the literature including the H_∞ filtering [15,21,27,30,36], mixed H_2/H_∞ filtering [20, 29], set-value estimation [1, 5, 6, 18] and robust extended Kalman filter (EKF) design [11,12,31,32]. Among them, the EKF has shown to be an effective way for tackling the nonlinear system estimation problems. In fact, EKF has recently gain particular research attention with promising application potentials in various engineering practice. For example, the EKF has been designed in [11, 12] for uncertain systems with quadratic constraints. Moreover, the EKF algorithm has been successfully applied in [25] to identify the parameters and predict the states of a nonlinear stochastic biological network modeled by time series data.

Apart from the stochastcity, the nonlinearity is another ubiquitous feature existing in almost all practical systems that contributes significantly to the complexity of system modeling. Since nonlinearities may cause undesirable dynamic behaviors such as oscillation or even instability, the analysis and synthesis problems for nonlinear systems have long been the main stream of research topics and much effort has been made to deal with the nonlinear stochastic systems, see e.g. [2, 4, 14, 19, 33]. It is worth pointing out that, in most literature, the nonlinearities are assumed to occur in a deterministic way. While this assumption is generally true especially for systems modeled according to physical laws, another kind of nonlinearities, namely, stochastic nonlinearities, deserve particular research attention since they occur randomly due probably to the high manoeuvrability of the tracked target, intermittent network congestion, random failures and repairs of the components, changes in the interconnections of subsystems, sudden environment changes, modification of the operating point of a linearized model of nonlinear systems. In fact, such stochastic nonlinearities include the state-multiplicative noises as spe-

⋆ This work was supported in part by the National 973 Project under Grant 2009CB320600, National Natural Science Foundation of China under Grants 61028008, 61134009 and 60825303, 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. Corresponding author Zidong Wang.

Email address: `Zidong.Wang@brunel.ac.uk` (Zidong Wang).

cial cases. Recently, the filtering problem with stochastic nonlinearities described by statistical means has already stirred some research interests, and some latest results can be found in [26, 35] and the references therein. On the other hand, almost all real-time systems are time-varying and therefore finite-horizon filtering problem is of practical significance. However, so far, there have been very few results in the literature regarding filtering problems with stochastic nonlinearities over a finite horizon due probably to the mathematical complexity and/or the computational difficulty.

In recent years, networked systems have become very prevalent and, accordingly, much work has been done in the literature on the network-induced problems such as missing measurements (also called packet loss or dropout) and random communication delays, see e.g. [3, 8–10, 22, 34]. To be more specific, the optimal estimation problems have been investigated in [8, 22] for linear systems with multiple packet dropouts and the random sensor delays have been taken into account in [9, 34]. It is worth mentioning that, in most reported results, the measurement signal has been assumed to be either completely lost or successfully transferred, and a typical way is to model the missing measurements by the Bernoulli distribution. However, in practical applications, owing to the sensors aging, sensor temporal failure or some of the data coming from a highly noisy environment, the measurement missing might be partial and individual sensor could have different missing probability in the data transmission process [26]. It is noted that most available results with respect to filtering problem with missing measurements have been concentrated on linear systems only, and the corresponding results for nonlinear systems have been very few. It is mentioning that, in [13], the stochastic stability has been analyzed for EKF with intermittent observations. Up to now, to the best of the authors' knowledge, the finite-horizon extended Kalman filtering problem with both stochastic nonlinearities and multiple missing measurements has not been addressed yet, which still remains as a challenging research issue. It is, therefore, the purpose of this paper to shorten such a gap by resorting to a recursive Riccati-like equation approach.

Motivated by the above discussion, in this paper, we make a major effort to design the EKF for a class of discrete time-varying systems with stochastic nonlinearities and multiple missing measurements. The considered stochastic nonlinearities are governed by zero mean Gaussian noises. The multiple missing measurements are included to model the randomly intermittent behaviors of the individual sensors. The description of the multiple missing measurements is more general than the commonly used one modeled by Bernoulli distribution. The probability distribution governing the missing measurements from individual sensor is allowed to be any discrete distribution taking values over the interval $[0, 1]$ with known occurrence probability. A recursive approach is developed here to deal with the EKF design problem. An optimized upper bound is guaranteed on the filtering error covariance for both the stochastic nonlinearities and multiple missing measurements. The main contributions of this paper can be summarized (from the aspects of model, problem and algorithm) as follows: 1) the system model is comprehensive that covers stochastic nonlinearities and multiple missing measurements, thereby better reflecting the reality; 2) the addressed extended Kalman filtering problem over a finite horizon is new especially when multiple missing measurements are presented; and 3) the developed filter design algorithm is of a form suitable for recursive computation in online applications.

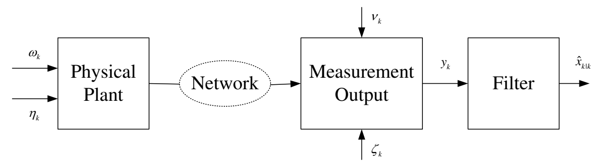

Fig. 1. Schematic structure for the plant and filter over network

The remainder of this paper is organized as follows. Section 2 briefly introduces the problem under consideration. In Section 3, the linearization is firstly enforced to facilitate the filter design. Then, the evolution of one-step prediction error covariance and filtering error covariance are derived for the addressed model. In the same section, an upper bound of the filtering error covariance is obtained and the filter gain is then designed to minimize such an upper bound at each sampling instant. An illustrative example is utilized in Section 4 to show the effectiveness of the proposed algorithm. The paper is concluded in Section 5.

Notation The notations used throughout the paper are standard. $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ denote the n-dimensional Euclidean space and the set of all $n \times m$ matrices, respectively. For a matrix P, P^T and P^{-1} represent its transpose and inverse, respectively. $P > 0$ means that the matrix P is real symmetric and positive definite. $\circ$ 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 random variable x. I and 0 represent the identity matrix and the zero matrix with appropriate dimensions, respectively. $\mathrm{diag}\{X_1, X_2, \ldots, X_n\}$ 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.

2 Problem Formulation and Preliminaries

In this paper, we consider the filtering problem for a general class of discrete time-varying systems with stochastic nonlinearities and multiple missing measurements, where the schematic diagram is shown in Fig. 1. The plant under consideration is of the following form:

$$x_{k+1} = f(x_k) + g(x_k, \eta_k) + D_k\omega_k \tag{1}$$

$$y_k = \Xi_k h(x_k) + s(x_k, \zeta_k) + \nu_k \tag{2}$$

where k is the sampling instant, $x_k \in \mathbb{R}^n$ is the state vector to be estimated, $y_k \in \mathbb{R}^q$ is the measurement output, η_k and ζ_k are zero-mean Gaussian noise sequences, D_k is a known matrix with appropriate dimension, $\omega_k \in \mathbb{R}^m$ is the process noise, and $\nu_k \in \mathbb{R}^q$ is the measurement noise. $\Xi_k := \mathrm{diag}\{\alpha_k^1, \alpha_k^2, \ldots, \alpha_k^q\}$ where α_k^i $(i = 1, 2, \ldots, q)$ are q independent random variables in k as well as i and are independent of all noise signals. It is assumed that

α_k^i has the probability density function $p_k^i(s)$ on the interval $[0,1]$ with mathematical expectation μ_k^i and variance $(\sigma_k^i)^2$ $(i=1,2,\ldots,q)$. Also, the noise signals η_k, ζ_k, ω_k and ν_k are uncorrelated with each other.

The deterministic nonlinearities $f(x_k):\mathbb{R}^n\to\mathbb{R}^n$ and $h(x_k):\mathbb{R}^n\to\mathbb{R}^q$ are known and continuously differentiable with

$$\|h(x_k)\|\leq a_1\|x_k\|+a_2, \tag{3}$$

for some nonnegative scalars a_1 and a_2. On the other hand, the stochastic nonlinearities $g(x_k,\eta_k):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$ and $s(x_k,\zeta_k):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^q$ satisfy $g(0,\eta_k)=0$ and $s(0,\zeta_k)=0$, respectively, and are assumed to have the following first moment for all x_k:

$$\mathbb{E}\left\{\left[\begin{array}{c} g(x_k,\eta_k)\\ s(x_k,\zeta_k)\end{array}\right]\middle| x_k\right\}=0 \tag{4}$$

and the covariance given by

$$\begin{aligned}
&\mathbb{E}\left\{\left[\begin{array}{c} g(x_k,\eta_k)\\ s(x_k,\zeta_k)\end{array}\right]\left[\begin{array}{c} g(x_j,\eta_j)\\ s(x_j,\zeta_j)\end{array}\right]^T\middle| x_k\right\}=0,\quad k\neq j\\
&\mathbb{E}\left\{\left[\begin{array}{c} g(x_k,\eta_k)\\ s(x_k,\zeta_k)\end{array}\right]\left[\begin{array}{c} g(x_k,\eta_k)\\ s(x_k,\zeta_k)\end{array}\right]^T\middle| x_k\right\}=\sum_{i=1}^{r}\Pi_k^i x_k^T\Gamma_k^i x_k
\end{aligned} \tag{5}$$

where r is a known positive integer, $\Pi_k^i=\operatorname{diag}\left\{\Pi_k^{1i},\Pi_k^{2i}\right\}$ and Γ_k^i $(i=1,2,\ldots,r)$ are known matrices with appropriate dimensions.

The initial state x_0, the process noise ω_k and the measurement noise ν_k are mutually uncorrelated and have the following statistical properties:

$$\begin{aligned}
&\mathbb{E}\{x_0\}=\bar{x}_0,\ \ \mathbb{E}\left\{(x_0-\bar{x}_0)(x_0-\bar{x}_0)^T\right\}=P_{0|0},\\
&\mathbb{E}\{\omega_k\}=0,\ \ \mathbb{E}\{\nu_k\}=0,\\
&\mathbb{E}\left\{\omega_k\omega_k^T\right\}=Q_k,\ \ \mathbb{E}\left\{\nu_k\nu_k^T\right\}=R_k,
\end{aligned} \tag{6}$$

where $P_{0|0}>0$, $Q_k>0$ and $R_k>0$ are known matrices with appropriate dimensions.

The recursive filter to be designed is of the following form:

$$\hat{x}_{k+1|k}=f\left(\hat{x}_{k|k}\right), \tag{7}$$

$$\hat{x}_{k+1|k+1}=\hat{x}_{k+1|k}+K_{k+1}\left[y_{k+1}-\bar{\Xi}_{k+1}h\left(\hat{x}_{k+1|k}\right)\right], \tag{8}$$

where $\hat{x}_{k|k}$ is the estimate of x_k at time k with $\hat{x}_{0|0}=\bar{x}_0$, $\hat{x}_{k+1|k}$ is the one-step prediction at time k, K_{k+1} is the filter gain to be determined, and $\bar{\Xi}_{k+1}:=\mathbb{E}\{\Xi_{k+1}\}:=\operatorname{diag}\{\mu_{k+1}^1,\mu_{k+1}^2,\ldots,\mu_{k+1}^q\}$.

The objective of this paper is to design a finite-horizon filter of the structure (7)-(8) such that, for all stochastic nonlinearities and multiple missing measurements, an upper bound for the filtering error covariance is guaranteed, that is, there exists a sequence of positive-definite matrices $\Sigma_{k+1|k+1}$ $(0\leq k\leq N)$ satisfying

$$\mathbb{E}\left\{\left(x_{k+1}-\hat{x}_{k+1|k+1}\right)\left(x_{k+1}-\hat{x}_{k+1|k+1}\right)^T\right\}\leq\Sigma_{k+1|k+1}. \tag{9}$$

Moreover, the designed filter gain K_{k+1} is expected to minimize the upper bound $\Sigma_{k+1|k+1}$ through a recursive scheme.

Remark 1 *In (2), $h(x_k)$ represents the sensor outputs coupled with nonlinearities. In engineering practice, the nonlinearities in the sensor outputs result primarily from the sensor saturations due to finite register-length of digital hardware, and such kind of nonlinearities can be covered by the assumption made in (3). To be more specific, the assumption in (3) could encompass a number of frequently occurred sensor-related nonlinearities such as sector-bounded nonlinearities, quantization, overflow nonlinearities, etc. Note that, under the same norm-bounded assumption, the control and filtering problems have been extensively investigated for nonlinear stochastic systems, see e.g. [16, 17].*

Remark 2 *In recent years, it is quite common that the measurement signals are transmitted through a large number of sensors in a network. Due to the limited bandwidth of a network, the missing measurement phenomenon may occur intermittently and the data-missing probability may be different for individual sensor. In (2), the multiple missing measurements (i.e., data missing with multiple sensors) are taken into account, where the diagonal matrix Ξ_k stands for the missing status for all sensors as a whole and the random variable α_k^i corresponds to the ith sensor $(i=1,2,\ldots,q)$. As discussed in [26], the random variable α_k^i can take any value over the interval $[0,1]$ and the probability for α_k^i to take different values may vary with the sensors. Moreover, α_k^i can obey any discrete probability distributions over the interval $[0,1]$ that includes the Bernoulli (binary) distribution as a special case. By considering the phenomenon of the multiple missing measurements, the new measurement model (2) is capable of describing the actual arrivals of the measured information from multiple sensors especially when only partial data is missing.*

Before proceeding further, we are in a position to introduce the following lemmas which will be used in subsequent developments.

Lemma 1 *[7] Let $A=[a_{ij}]_{p\times p}$ be a real-valued matrix and $B=\operatorname{diag}\{b_1,b_2,\ldots,b_p\}$ be a diagonal random matrix. Then*

$$\mathbb{E}\{BAB^T\}=\left[\begin{array}{cccc}
\mathbb{E}\{b_1^2\} & \mathbb{E}\{b_1b_2\} & \cdots & \mathbb{E}\{b_1b_p\}\\
\mathbb{E}\{b_2b_1\} & \mathbb{E}\{b_2^2\} & \cdots & \mathbb{E}\{b_2b_p\}\\
\vdots & \vdots & \ddots & \vdots\\
\mathbb{E}\{b_pb_1\} & \mathbb{E}\{b_pb_2\} & \cdots & \mathbb{E}\{b_p^2\}
\end{array}\right]\circ A$$

where $\circ$ is the Hadamard product.

Lemma 2 *[28] Given matrices A, H, E and F with appropriate dimensions such that $FF^T\leq I$. Let X be a* *symmetric positive definite matrix and γ be an arbitrary positive constant such that $\gamma^{-1}I - EXE^T > 0$. Then the following inequality holds*

$$\begin{aligned}&(A+HFE)\,X\,(A+HFE)^T\\ \leq &A\left(X^{-1}-\gamma E^TE\right)^{-1}A^T+\gamma^{-1}HH^T.\end{aligned} \tag{10}$$

Lemma 3 *[23] For $0 \leq k \leq N$, suppose that $X = X^T > 0$, $\mathcal{S}_k(X) = \mathcal{S}_k^T(X) \in \mathbb{R}^{n\times n}$ and $\mathcal{H}_k(X) = \mathcal{H}_k^T(X) \in \mathbb{R}^{n\times n}$. If*

$$\mathcal{S}_k(Y) \geq \mathcal{S}_k(X), \ \ \forall \ \ X \leq Y = Y^T \tag{11}$$

and

$$\mathcal{H}_k(Y) \geq \mathcal{S}_k(Y), \tag{12}$$

then the solutions M_k and N_k to the following difference equations

$$M_{k+1} = \mathcal{S}_k(M_k), \ \ N_{k+1} = \mathcal{H}_k(N_k), \ \ M_0 = N_0 > 0 \tag{13}$$

satisfy

$$M_k \leq N_k.$$

3 Main Results

In this section, we aim to establish a unified framework to deal with the addressed filtering problem in the simultaneous presence of stochastic nonlinearities as well as multiple missing measurements. The linearization is firstly enforced to facilitate the later developments. Subsequently, the one-step prediction error covariance and the filtering error covariance are calculated so as to design the finite-horizon EKF, where special effort is made to compensate the effects of multiple missing measurements. Next, the upper bound of the filtering error covariance is presented and the filter gain is designed to guarantee that such an upper bound is minimized.

To start with, let us denote the one-step prediction error as $\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k}$ and the filtering error as $\tilde{x}_{k+1|k+1} = x_{k+1} - \hat{x}_{k+1|k+1}$. Subtracting (7) from (1), we have

$$\tilde{x}_{k+1|k} = f(x_k) - f\left(\hat{x}_{k|k}\right) + g(x_k, \eta_k) + D_k\omega_k. \tag{14}$$

By using the Taylor series expansion around $\hat{x}_{k|k}$, we linearize $f(x_k)$ as follows:

$$f(x_k) = f\left(\hat{x}_{k|k}\right) + A_k\tilde{x}_{k|k} + o(\tilde{x}_{k|k}^2) \tag{15}$$

where

$$A_k = \frac{\partial f(x_k)}{\partial x_k}\Big|_{x_k=\hat{x}_{k|k}}$$

and $o(\tilde{x}_{k|k}^2)$ represents the high-order terms of the Taylor series expansion. For presentation convenience, following [6, 31], the high-order terms are transformed into the following easy-to-handle formulation:

$$o(\tilde{x}_{k|k}^2) = B_k\aleph_{1,k}L_k\tilde{x}_{k|k} \tag{16}$$

where B_k is a problem-dependent scaling matrix, L_k is introduced to provide an extra degree of freedom to tune the filter, and $\aleph_{1,k}$ is an unknown time-varying matrix accounting for the linearization errors of the dynamical model that satisfies

$$\aleph_{1,k}\aleph_{1,k}^T \leq I. \tag{17}$$

It follows from (14)-(16) that

$$\tilde{x}_{k+1|k} = (A_k + B_k\aleph_{1,k}L_k)\,\tilde{x}_{k|k} + g(x_k, \eta_k) + D_k\omega_k. \tag{18}$$

Similarly, by applying the Taylor series expansion for $h(x_{k+1})$ around $\hat{x}_{k+1|k}$, the innovation of the filter can be obtained as follows:

$$\begin{aligned}&\tilde{y}_{k+1}\\ =&y_{k+1} - \bar{\Xi}_{k+1}h\left(\hat{x}_{k+1|k}\right)\\ =&\left(\Xi_{k+1} - \bar{\Xi}_{k+1}\right)h(x_{k+1}) + \bar{\Xi}_{k+1}(C_{k+1} + E_{k+1}\\ &\times\aleph_{2,k+1}L_{k+1})\tilde{x}_{k+1|k} + s(x_{k+1}, \zeta_{k+1}) + \nu_{k+1}\end{aligned} \tag{19}$$

where

$$C_{k+1} = \frac{\partial h(x_{k+1})}{\partial x_{k+1}}\Big|_{x_{k+1}=\hat{x}_{k+1|k}},$$

E_{k+1} is a problem-dependent scaling matrix, and $\aleph_{2,k+1}$ is an unknown time-varying matrix representing the linearization errors of the dynamical model that satisfies

$$\aleph_{2,k+1}\aleph_{2,k+1}^T \leq I. \tag{20}$$

In this paper, as in [6], the deterministic matrices $\aleph_{1,k}$, $\aleph_{2,k+1}$ and the scaling matrices B_k, E_{k+1} are employed to account for the linearization errors. For more details we refer the reader to Appendix C of [6].

According to (8) and (19), the filtering error can be written as:

$$\begin{aligned}&\tilde{x}_{k+1|k+1}\\ =&\left[I - K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1} + E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]\tilde{x}_{k+1|k}\\ &- K_{k+1}\left(\Xi_{k+1} - \bar{\Xi}_{k+1}\right)h(x_{k+1})\\ &- K_{k+1}s(x_{k+1}, \zeta_{k+1}) - K_{k+1}\nu_{k+1}.\end{aligned} \tag{21}$$

Subsequently, in the light of (18) and (21), the covariances for the one-step prediction error and filtering error can be derived, respectively, in the following theorems.

Theorem 1 *The one-step prediction error covariance $P_{k+1|k}$ is given by*

$$\begin{aligned}P_{k+1|k} =& (A_k + B_k\aleph_{1,k}L_k)\,P_{k|k}\,(A_k + B_k\aleph_{1,k}L_k)^T\\ &+ \sum_{i=1}^{r}\Pi_k^{1i}\mathrm{tr}\left(\mathbb{E}\left\{x_kx_k^T\right\}\Gamma_k^i\right) + D_kQ_kD_k^T.\end{aligned} \tag{22}$$

Proof: It can be shown that (22) follows directly from (5)-(6) and (18), and therefore the proof is omitted for conciseness.

Theorem 2 *The filtering error covariance $P_{k+1|k+1}$ satisfies*

$$
\begin{aligned}
&P_{k+1|k+1} \\
=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]P_{k+1|k} \\
&\times\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]^T \\
&+K_{k+1}\Bigg[\breve{\Xi}_{k+1}\circ\mathbb{E}\left\{h(x_{k+1})h^T(x_{k+1})\right\} \\
&+\sum_{i=1}^{r}\Pi_{k+1}^{2i}\mathrm{tr}\left(\mathbb{E}\left\{x_{k+1}x_{k+1}^T\right\}\Gamma_{k+1}^i\right)+R_{k+1}\Bigg]K_{k+1}^T
\end{aligned}
\tag{23}
$$

where

$$
\breve{\Xi}_{k+1}=\mathrm{diag}\left\{(\sigma_{k+1}^1)^2,(\sigma_{k+1}^2)^2,\ldots,(\sigma_{k+1}^q)^2\right\}. \tag{24}
$$

Proof: Noting (21), we have

$$
\begin{aligned}
&P_{k+1|k+1} \\
=&\mathbb{E}\left\{\tilde{x}_{k+1|k+1}\tilde{x}_{k+1|k+1}^T\right\} \\
=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]P_{k+1|k} \\
&\times\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]^T \\
&+K_{k+1}\mathbb{E}\{\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)h(x_{k+1})h^T(x_{k+1}) \\
&\times\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)\}K_{k+1}^T+K_{k+1}\mathbb{E}\{s\left(x_{k+1},\zeta_{k+1}\right) \\
&\times s^T\left(x_{k+1},\zeta_{k+1}\right)\}K_{k+1}^T+K_{k+1}\mathbb{E}\left\{\nu_{k+1}\nu_{k+1}^T\right\}K_{k+1}^T \\
&-\mathscr{P}_{k+1}-\mathscr{P}_{k+1}^T-\mathscr{Q}_{k+1}-\mathscr{Q}_{k+1}^T-\mathscr{R}_{k+1}-\mathscr{R}_{k+1}^T \\
&+\mathscr{X}_{k+1}+\mathscr{X}_{k+1}^T+\mathscr{Y}_{k+1}+\mathscr{Y}_{k+1}^T+\mathscr{Z}_{k+1}+\mathscr{Z}_{k+1}^T
\end{aligned}
\tag{25}
$$

where

$$
\begin{aligned}
\mathscr{P}_{k+1}=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right] \\
&\times\mathbb{E}\left\{\tilde{x}_{k+1|k}h^T(x_{k+1})\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)\right\}K_{k+1}^T, \\
\mathscr{Q}_{k+1}=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right] \\
&\times\mathbb{E}\left\{\tilde{x}_{k+1|k}s^T\left(x_{k+1},\zeta_{k+1}\right)\right\}K_{k+1}^T, \\
\mathscr{R}_{k+1}=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right] \\
&\times\mathbb{E}\left\{\tilde{x}_{k+1|k}\nu_{k+1}^T\right\}K_{k+1}^T, \\
\mathscr{X}_{k+1}=&K_{k+1}\mathbb{E}\{\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)h(x_{k+1}) \\
&\times s^T\left(x_{k+1},\zeta_{k+1}\right)\}K_{k+1}^T, \\
\mathscr{Y}_{k+1}=&K_{k+1}\mathbb{E}\left\{\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)h(x_{k+1})\nu_{k+1}^T\right\}K_{k+1}^T, \\
\mathscr{Z}_{k+1}=&K_{k+1}\mathbb{E}\left\{s\left(x_{k+1},\zeta_{k+1}\right)\nu_{k+1}^T\right\}K_{k+1}^T.
\end{aligned}
$$

It is not difficult to show that the terms $\mathscr{P}_{k+1}$, $\mathscr{Q}_{k+1}$, $\mathscr{R}_{k+1}$, $\mathscr{X}_{k+1}$, $\mathscr{Y}_{k+1}$ and $\mathscr{Z}_{k+1}$ are all equal to zero. It follows from (5)-(6) that (25) can be rewritten as:

$$
\begin{aligned}
&P_{k+1|k+1} \\
=&\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]P_{k+1|k} \\
&\times\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]^T \\
&+K_{k+1}\mathbb{E}\{\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)h(x_{k+1})h^T(x_{k+1}) \\
&\times\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)\}K_{k+1}^T+K_{k+1}\Bigg[\sum_{i=1}^{r}\Pi_{k+1}^{2i} \\
&\times\mathrm{tr}\left(\mathbb{E}\left\{x_{k+1}x_{k+1}^T\right\}\Gamma_{k+1}^i\right)+R_{k+1}\Bigg]K_{k+1}^T.
\end{aligned}
\tag{26}
$$

By using the property of conditional expectation and applying Lemma 1, the second term of the right-hand side of (26) can be determined as follows:

$$
\begin{aligned}
&\mathbb{E}\left\{\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)h(x_{k+1})h^T(x_{k+1})\left(\Xi_{k+1}-\bar{\Xi}_{k+1}\right)\right\} \\
=&\breve{\Xi}_{k+1}\circ\mathbb{E}\left\{h(x_{k+1})h^T(x_{k+1})\right\}
\end{aligned}
\tag{27}
$$

where $\breve{\Xi}_{k+1}$ is defined in (24). Then, from (26) and (27), it can be concluded that (23) is true. The proof is now complete.

Remark 3 *In Theorem 2, the recursive form of the filtering error covariance has been developed. Note that the linearization is enforced to tackle the nonlinearities $f(\cdot)$ and $h(\cdot)$. As such, (22) and (23) involve $\aleph_{1,k}$ and $\aleph_{2,k+1}$ which add extra computational difficulty for the design of filter gain. Actually, due to the consideration of the linearization errors, it is literally impossible to obtain the accurate value of the filtering error covariance $P_{k+1|k+1}$, and a seemingly natural way is to design appropriate filter gain K_{k+1} in order to guarantee an upper bound for the filtering error covariance that can then be minimized at each sampling instant.*

Motivated by [32], in the following theorem, an upper bound is provided for the filtering error covariance and the filter gain is then designed to minimize such an upper bound.

Theorem 3 *Consider the covariance matrices of the one-step prediction error and the filtering error in (22) and (23). Assume that (17) and (20) are true. Let $\gamma_{1,k}$, $\gamma_{2,k+1}$ and ε_j $(j=1,2)$ be positive scalars. If the following two discrete-time Riccati-like difference equations:*

$$
\begin{aligned}
&\Sigma_{k+1|k} \\
=&A_k\left(\Sigma_{k|k}^{-1}-\gamma_{1,k}L_k^TL_k\right)^{-1}A_k^T+\gamma_{1,k}^{-1}B_kB_k^T+D_kQ_kD_k^T \\
&+\sum_{i=1}^{r}\Pi_k^{1i}\mathrm{tr}\left\{\left[\left(1+\varepsilon_1\right)\Sigma_{k|k}+\left(1+\varepsilon_1^{-1}\right)\hat{x}_{k|k}\hat{x}_{k|k}^T\right]\Gamma_k^i\right\}
\end{aligned}
\tag{28}
$$

$$\begin{aligned}&\Sigma_{k+1|k+1}\\&=\left(I-K_{k+1}\bar{\Xi}_{k+1}C_{k+1}\right)\left(\Sigma_{k+1|k}^{-1}-\gamma_{2,k+1}L_{k+1}^TL_{k+1}\right)^{-1}\\&\quad\times\left(I-K_{k+1}\bar{\Xi}_{k+1}C_{k+1}\right)^T+\gamma_{2,k+1}^{-1}K_{k+1}\bar{\Xi}_{k+1}E_{k+1}\\&\quad\times E_{k+1}^T\bar{\Xi}_{k+1}K_{k+1}^T+K_{k+1}\Big\{\breve{\Xi}_{k+1}\circ[2(a_1^2\mathrm{tr}\left(\Omega_{k+1|k}\right)\\&\quad+a_2^2)I]+\sum_{i=1}^r\Pi_{k+1}^{2i}\mathrm{tr}\left(\Omega_{k+1|k}\Gamma_{k+1}^i\right)+R_{k+1}\Big\}K_{k+1}^T\end{aligned} \tag{29}$$

with initial condition $\Sigma_{0|0}=P_{0|0}>0$ have positive definite solutions $\Sigma_{k+1|k}$ and $\Sigma_{k+1|k+1}$ such that, for all $0\le k\le N$, the following two constraints

$$\gamma_{1,k}^{-1}I-L_k\Sigma_{k|k}L_k^T>0, \tag{30}$$

$$\gamma_{2,k+1}^{-1}I-L_{k+1}\Sigma_{k+1|k}L_{k+1}^T>0 \tag{31}$$

are satisfied where

$$\Omega_{k+1|k}=(1+\varepsilon_2)\,\Sigma_{k+1|k}+\left(1+\varepsilon_2^{-1}\right)\hat{x}_{k+1|k}\hat{x}_{k+1|k}^T, \tag{32}$$

then with the filter gain K_{k+1} given by

$$\begin{aligned}&K_{k+1}\\&=\left(\Sigma_{k+1|k}^{-1}-\gamma_{2,k+1}L_{k+1}^TL_{k+1}\right)^{-1}C_{k+1}^T\bar{\Xi}_{k+1}\Big\{\bar{\Xi}_{k+1}\\&\quad\times C_{k+1}\left(\Sigma_{k+1|k}^{-1}-\gamma_{2,k+1}L_{k+1}^TL_{k+1}\right)^{-1}C_{k+1}^T\bar{\Xi}_{k+1}\\&\quad+\gamma_{2,k+1}^{-1}\bar{\Xi}_{k+1}E_{k+1}E_{k+1}^T\bar{\Xi}_{k+1}\\&\quad+\breve{\Xi}_{k+1}\circ\left[2\left(a_1^2\mathrm{tr}\left(\Omega_{k+1|k}\right)+a_2^2\right)I\right]\\&\quad+\sum_{i=1}^r\Pi_{k+1}^{2i}\mathrm{tr}\left(\Omega_{k+1|k}\Gamma_{k+1}^i\right)+R_{k+1}\Big\}^{-1},\end{aligned} \tag{33}$$

the matrix $\Sigma_{k+1|k+1}$ is an upper bound for $P_{k+1|k+1}$, i.e.,

$$P_{k+1|k+1}\le\Sigma_{k+1|k+1}. \tag{34}$$

Moreover, the filter gain K_{k+1} given by (33) minimizes the upper bound $\Sigma_{k+1|k+1}$.

Proof: To begin with, based on (22) and (23), rewrite the covariance matrices $P_{k+1|k}$ and $P_{k+1|k+1}$ as the functions of $P_{k|k}$ and $P_{k+1|k}$ as follows:

$$\begin{aligned}&P_{k+1|k}\left(P_{k|k}\right)\\&=(A_k+B_k\aleph_{1,k}L_k)\,P_{k|k}\,(A_k+B_k\aleph_{1,k}L_k)^T\\&\quad+\sum_{i=1}^r\Pi_k^{1i}\mathrm{tr}\left(\mathbb{E}\left\{x_kx_k^T\right\}\Gamma_k^i\right)+D_kQ_kD_k^T\end{aligned}$$

$$\begin{aligned}&P_{k+1|k+1}\left(P_{k+1|k}\right)\\&=\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]P_{k+1|k}\\&\quad\times\left[I-K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1}+E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]^T\\&\quad+K_{k+1}\Big[\breve{\Xi}_{k+1}\circ\mathbb{E}\left\{h(x_{k+1})h^T(x_{k+1})\right\}+\sum_{i=1}^r\Pi_{k+1}^{2i}\\&\quad\times\mathrm{tr}\left(\mathbb{E}\left\{x_{k+1}x_{k+1}^T\right\}\Gamma_{k+1}^i\right)+R_{k+1}\Big]K_{k+1}^T.\end{aligned}$$

Then, it is not difficult to verify that the condition (11) in Lemma 3 is satisfied.

Now, we are in a position to tackle the term of the right-hand side of (22). Notice that the following elementary inequality

$$\left(\varepsilon_1^{\frac{1}{2}}\tilde{x}_{k|k}-\varepsilon_1^{-\frac{1}{2}}\hat{x}_{k|k}\right)\left(\varepsilon_1^{\frac{1}{2}}\tilde{x}_{k|k}-\varepsilon_1^{-\frac{1}{2}}\hat{x}_{k|k}\right)^T\ge0$$

yields

$$\tilde{x}_{k|k}\hat{x}_{k|k}^T+\hat{x}_{k|k}\tilde{x}_{k|k}^T\le\varepsilon_1\tilde{x}_{k|k}\tilde{x}_{k|k}^T+\varepsilon_1^{-1}\hat{x}_{k|k}\hat{x}_{k|k}^T \tag{35}$$

where $\varepsilon_1>0$ is a scalar. Based on (35), the second term of the right-hand side of (22) can be rearranged as

$$\begin{aligned}&\sum_{i=1}^r\Pi_k^{1i}\mathrm{tr}\left(\mathbb{E}\left\{x_kx_k^T\right\}\Gamma_k^i\right)\\&=\sum_{i=1}^r\Pi_k^{1i}\mathrm{tr}\left(\mathbb{E}\left\{\left(\hat{x}_{k|k}+\tilde{x}_{k|k}\right)\left(\hat{x}_{k|k}+\tilde{x}_{k|k}\right)^T\right\}\Gamma_k^i\right)\\&\le\sum_{i=1}^r\Pi_k^{1i}\mathrm{tr}\left\{\left[(1+\varepsilon_1)\,P_{k|k}+\left(1+\varepsilon_1^{-1}\right)\hat{x}_{k|k}\hat{x}_{k|k}^T\right]\Gamma_k^i\right\}.\end{aligned} \tag{36}$$

Together with (22) and (36), we have

$$\begin{aligned}&P_{k+1|k}\\&\le(A_k+B_k\aleph_{1,k}L_k)\,P_{k|k}\,(A_k+B_k\aleph_{1,k}L_k)^T+D_kQ_kD_k^T\\&\quad+\sum_{i=1}^r\Pi_k^{1i}\mathrm{tr}\left\{\left[(1+\varepsilon_1)\,P_{k|k}+\left(1+\varepsilon_1^{-1}\right)\hat{x}_{k|k}\hat{x}_{k|k}^T\right]\Gamma_k^i\right\}.\end{aligned} \tag{37}$$

On the other hand, let us handle the terms of the right-hand side of (23). It follows from (3) that

$$\begin{aligned}&\mathbb{E}\left\{h\left(x_{k+1}\right)h^T\left(x_{k+1}\right)\right\}\\&\le\mathbb{E}\left\{\|h\left(x_{k+1}\right)\|^2\right\}I\\&\le\mathbb{E}\left\{\left(a_1\|x_{k+1}\|+a_2\right)^2\right\}I\\&\le\left(2a_1^2\mathbb{E}\left\{\|x_{k+1}\|^2\right\}+2a_2^2\right)I\\&=2\left[a_1^2\mathrm{tr}\left(\mathbb{E}\left\{x_{k+1}x_{k+1}^T\right\}\right)+a_2^2\right]I.\end{aligned} \tag{38}$$

Note that, when deriving (38), we have used the elementary inequality $2ab\le a^2+b^2$. Taking the following inequality into consideration

$$
\begin{aligned}
&\tilde{x}_{k+1|k}\hat{x}_{k+1|k}^T + \hat{x}_{k+1|k}\tilde{x}_{k+1|k}^T \\
\leq &\varepsilon_2 \tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T + \varepsilon_2^{-1}\hat{x}_{k+1|k}\hat{x}_{k+1|k}^T
\end{aligned} \tag{39}
$$

with $\varepsilon_2 > 0$ being a scalar, we obtain

$$
\begin{aligned}
&\mathbb{E}\left\{h\left(x_{k+1}\right)h^T\left(x_{k+1}\right)\right\} \\
\leq &2[a_1^2 \mathrm{tr}(\mathbb{E}\{(1+\varepsilon_2)\,\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T + \left(1+\varepsilon_2^{-1}\right) \\
&\times \hat{x}_{k+1|k}\hat{x}_{k+1|k}^T\}) + a_2^2]I \\
=&2[a_1^2 \mathrm{tr}((1+\varepsilon_2)\,P_{k+1|k} + \left(1+\varepsilon_2^{-1}\right) \\
&\times \hat{x}_{k+1|k}\hat{x}_{k+1|k}^T) + a_2^2]I.
\end{aligned} \tag{40}
$$

Subsequently, by considering (23), (39) and (40), we have

$$
\begin{aligned}
&P_{k+1|k+1} \\
\leq &\left[I - K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1} + E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]P_{k+1|k} \\
&\times \left[I - K_{k+1}\bar{\Xi}_{k+1}\left(C_{k+1} + E_{k+1}\aleph_{2,k+1}L_{k+1}\right)\right]^T \\
&+ K_{k+1}\Bigg[\breve{\Xi}_{k+1}\circ\left(2\left[a_1^2 \mathrm{tr}\left(\Psi_{k+1|k}\right) + a_2^2\right]I\right) \\
&+ \sum_{i=1}^{r}\Pi_{k+1}^{2i}\mathrm{tr}\left(\Psi_{k+1|k}\Gamma_{k+1}^i\right) + R_{k+1}\Bigg]K_{k+1}^T
\end{aligned} \tag{41}
$$

where

$$
\Psi_{k+1|k} = (1+\varepsilon_2)\,P_{k+1|k} + \left(1+\varepsilon_2^{-1}\right)\hat{x}_{k+1|k}\hat{x}_{k+1|k}^T.
$$

Next, according to (28) and (29), we continue to rewrite $\Sigma_{k+1|k}$ and $\Sigma_{k+1|k+1}$ as the function of $\Sigma_{k|k}$ and $\Sigma_{k+1|k}$ as follows:

$$
\begin{aligned}
&\Sigma_{k+1|k}\left(\Sigma_{k|k}\right) \\
=&A_k\left(\Sigma_{k|k}^{-1} - \gamma_{1,k}L_k^T L_k\right)^{-1}A_k^T + \gamma_{1,k}^{-1}B_k B_k^T + D_k Q_k D_k^T \\
&+ \sum_{i=1}^{r}\Pi_k^{1i}\mathrm{tr}\left\{\left[(1+\varepsilon_1)\,\Sigma_{k|k} + \left(1+\varepsilon_1^{-1}\right)\hat{x}_{k|k}\hat{x}_{k|k}^T\right]\Gamma_k^i\right\}
\end{aligned} \tag{42}
$$

$$
\begin{aligned}
&\Sigma_{k+1|k+1}\left(\Sigma_{k+1|k}\right) \\
=&\left(I - K_{k+1}\bar{\Xi}_{k+1}C_{k+1}\right)\left(\Sigma_{k+1|k}^{-1} - \gamma_{2,k+1}L_{k+1}^T L_{k+1}\right)^{-1} \\
&\times \left(I - K_{k+1}\bar{\Xi}_{k+1}C_{k+1}\right)^T + \gamma_{2,k+1}^{-1}K_{k+1}\bar{\Xi}_{k+1}E_{k+1} \\
&\times E_{k+1}^T\bar{\Xi}_{k+1}K_{k+1}^T + K_{k+1}\Bigg\{\breve{\Xi}_{k+1}\circ[2(a_1^2 \mathrm{tr}\left(\Omega_{k+1|k}\right) \\
&+ a_2^2)I] + \sum_{i=1}^{r}\Pi_{k+1}^{2i}\mathrm{tr}\left(\Omega_{k+1|k}\Gamma_{k+1}^i\right) + R_{k+1}\Bigg\}K_{k+1}^T
\end{aligned} \tag{43}
$$

where $\breve{\Xi}_{k+1}$ and $\Omega_{k+1|k}$ are defined in (24) and (32), respectively. Combining (37), (41), (42) and (43), we can show that the condition (12) in Lemma 3 is satisfied. Therefore, it follows from Lemmas 2-3 that

$$
P_{k+1|k+1} \leq \Sigma_{k+1|k+1}.
$$

Next, we are ready to show that the filter gain given by (33) is optimal in the sense that it minimizes the upper bound $\Sigma_{k+1|k+1}$. Taking the partial derivative of $\Sigma_{k+1|k+1}$ with respect to K_{k+1} and letting the derivative be zero, we have

$$
\begin{aligned}
&\frac{\partial \mathrm{tr}\left(\Sigma_{k+1|k+1}\right)}{\partial K_{k+1}} \\
=&-2\left(I - K_{k+1}\bar{\Xi}_{k+1}C_{k+1}\right)(\Sigma_{k+1|k}^{-1} - \gamma_{2,k+1}L_{k+1}^T \\
&\times L_{k+1})^{-1}C_{k+1}^T\bar{\Xi}_{k+1} + 2K_{k+1}\Bigg\{\gamma_{2,k+1}^{-1}\bar{\Xi}_{k+1}E_{k+1} \\
&\times E_{k+1}^T\bar{\Xi}_{k+1} + \breve{\Xi}_{k+1}\circ\left[2\left(a_1^2 \mathrm{tr}\left(\Omega_{k+1|k}\right) + a_2^2\right)I\right] \\
&+ \sum_{i=1}^{r}\Pi_{k+1}^{2i}\mathrm{tr}\left(\Omega_{k+1|k}\Gamma_{k+1}^i\right) + R_{k+1}\Bigg\} = 0.
\end{aligned}
$$

Based on the above equation, the optimal filter gain K_{k+1} can be determined as

$$
\begin{aligned}
&K_{k+1} \\
=&\left(\Sigma_{k+1|k}^{-1} - \gamma_{2,k+1}L_{k+1}^T L_{k+1}\right)^{-1}C_{k+1}^T\bar{\Xi}_{k+1}\Bigg\{\bar{\Xi}_{k+1} \\
&\times C_{k+1}\left(\Sigma_{k+1|k}^{-1} - \gamma_{2,k+1}L_{k+1}^T L_{k+1}\right)^{-1}C_{k+1}^T\bar{\Xi}_{k+1} \\
&+ \gamma_{2,k+1}^{-1}\bar{\Xi}_{k+1}E_{k+1}E_{k+1}^T\bar{\Xi}_{k+1} \\
&+ \breve{\Xi}_{k+1}\circ\left[2\left(a_1^2 \mathrm{tr}\left(\Omega_{k+1|k}\right) + a_2^2\right)I\right] \\
&+ \sum_{i=1}^{r}\Pi_{k+1}^{2i}\mathrm{tr}\left(\Omega_{k+1|k}\Gamma_{k+1}^i\right) + R_{k+1}\Bigg\}^{-1}
\end{aligned}
$$

which is identical to (33). It is clear that the filter gain given by (33) is optimal that minimizes the upper bound $\Sigma_{k+1|k+1}$ for the filtering error covariance. This completes the proof.

Remark 4 *The recursive EKF problem is solved in Theorems 1-3 for a general class of discrete time-varying nonlinear systems with stochastic nonlinearities and multiple missing measurements. Unlike most existing literature, the EKF scheme presented in this paper has an advantage to cope with the multiple missing measurements where each sensor is allowed to have individual data missing probability especially when only partial information is missing. Note that such a missing measurement phenomenon is typically encountered in practical engineering systems including networked control systems. To handle the emergence of multiple missing measurements, we have made specific efforts to design a recursive filter and derive the upper bound for the filtering error covariance that are dependent on the individual missing probability. Specifically, the Hadamard product has been applied to facilitate the algorithm development. It is worth pointing out that the related (first to third) terms in (29) caused by multiple missing measurements and the fourth term in (28)-(29) due to the consideration of stochastic nonlinearities constitute the main difference between our work and the work of [32].*

Remark 5 *In this paper, our focus is on the recursive filter design problem for time-varying systems with stochastic nonlinearities and multiple missing measurements. Due to such a complicated time-varying nature, we carry out the research for the finite-horizon case, that is, we wish the filtering criteria to be satisfied over a finite-horizon. Instead of the asymptotic behavior (over an infinite-horizon), in this paper, we are only interested in the transient property over the finite-horizon* $k \in [0, N]$*, i.e., the upper bound for the filtering error covariance is obtained at every sampling instant* $k \in [0, N]$*, and such an upper bound is minimized by properly designing the filter gain at each sampling instant. Nevertheless, in case that the convergence analysis of the proposed filter approach becomes a concern, as discussed in [13], some additional assumptions can be made on the system parameters in order to ensure the global boundedness of the estimation errors, which constitutes one of our future research topics.*

Remark 6 *At each sampling instant, the filter gain* K_{k+1} *is designed in Theorem 3 to guarantee that the upper bound for the filtering error covariance is minimized. The system (1)-(2) under consideration is comprehensive that includes two phenomena of the stochastic nonlinearities and the multiple missing measurements, hence reflects the reality more closely especially in a networked environment. In our main results, these two phenomena are dealt with in a unified yet effective framework and are explicitly reflected in the design procedure. Specifically, the matrices* Π_k^{ij} *and* Γ_k^j *(*$i = 1, 2; j = 1, 2, \ldots, r$*) quantify the effects of the stochastic nonlinearities, and the constants* μ_k^i *and* σ_k^i *(*$i = 1, 2, \ldots, q$*) are there to account for the multiple missing measurements. Furthermore, the proposed filter is derived in terms of two discrete Riccati-like difference equations, which are suitable for recursive computation in online applications. In the next section, a simulation example is employed to show the usefulness of the proposed filter scheme.*

4 A Numerical Example

In this section, the effectiveness of the filtering algorithm developed in this paper is demonstrated. A target tracking scenario is used to justify the potential applicability of the designed filter scheme.

As analyzed in [24], consider a maneuvering target that is accelerating with random bursts of gas from its reaction control system thrusters. The state vector could consist of the position and velocity of the target. When tracking a maneuvering target through a radar system equipped with an array of sensors communicating through a (possibly wireless) network, the multiple missing phenomenon might occur due to the bandwidth limit of the signal transmission channel, the sensors aging and/or sensor temporal failure. Furthermore, the system may contaminate with the stochastic nonlinearities owing to a variety of reasons such as random failures and repairs of the components, changes in the interconnections of subsystems, and sudden environment changes. For real-time tracking, the system parameters would have to be time-varying. Our objective is, therefore, to design a filter such that, in the simultaneous presence of stochastic nonlinearities and multiple missing measurements, an optimized upper bound for the filtering error covariance is guaranteed.

Motivated by this background, we consider the following discretized maneuvering target tracking system with stochastic nonlinearities and multiple missing measurements:

$$\begin{cases} x_{k+1} = f(x_k) + g(x_k, \eta_k) + D_k \omega_k \\ y_k = \Xi_k h(x_k) + s(x_k, \zeta_k) + \nu_k \end{cases}$$

where

$$f(x_k) = \begin{bmatrix} 0.8x_k^1 + x_k^1 x_k^2 \\ 1.5x_k^2 - x_k^1 x_k^2 \end{bmatrix}, \quad D_k = \begin{bmatrix} 0.01 \\ 0.03 \end{bmatrix},$$
$$h(x_k) = 7.5 \sin(x_k^2),$$

and $x_k = \begin{bmatrix} x_k^1 & x_k^2 \end{bmatrix}^T$ is composed of the position and velocity of the target, $\omega_k \in \mathbb{R}$ and $\nu_k \in \mathbb{R}$ are zero-mean Gaussian white noises with covariances 0.05. Consider the following the case of the probability density function for Ξ_k:

$$p_k^1(s) = \begin{cases} 0.05, & s = 0 \\ 0.10, & s = 0.5 \\ 0.85, & s = 1 \end{cases}.$$

The expectation and variance can be easily calculated as $\mu_k^1 = 0.9$ and $\left(\sigma_k^1\right)^2 = 0.065$.

The stochastic nonlinearities $g(x_k, \eta_k)$ and $s(x_k, \zeta_k)$ are chosen as follows:

$$\begin{aligned} g(x_k, \eta_k) = & \begin{bmatrix} 0.2 \\ 0.3 \end{bmatrix} [0.3\text{sign}\left(x_k^1\right) x_k^1 \eta_k^1 + 0.4\text{sign}\left(x_k^2\right) \\ & \times x_k^2 \eta_k^2] \\ s(x_k, \zeta_k) = & 0.5 \left[0.3\text{sign}\left(x_k^1\right) x_k^1 \zeta_k^1 + 0.4\text{sign}\left(x_k^2\right) x_k^2 \zeta_k^2\right] \end{aligned}$$

where η_k^i and ζ_k^i ($i = 1, 2$) stand for zero-mean uncorrelated Gaussian white noises with unity covariances. It is not difficult to verify that the above stochastic nonlinearities satisfy

$$\mathbb{E}\left\{ \begin{bmatrix} g(x_k, \eta_k) \\ s(x_k, \zeta_k) \end{bmatrix} \middle| x_k \right\} = 0,$$
$$\mathbb{E}\left\{ \begin{bmatrix} g(x_k, \eta_k) \\ s(x_k, \zeta_k) \end{bmatrix} \begin{bmatrix} g(x_k, \eta_k) \\ s(x_k, \zeta_k) \end{bmatrix}^T \middle| x_k \right\}$$
$$= \begin{bmatrix} 0.04 & 0.06 & 0 \\ 0.06 & 0.09 & 0 \\ 0 & 0 & 0.25 \end{bmatrix} x_k^T \begin{bmatrix} 0.09 & 0 \\ 0 & 0.16 \end{bmatrix} x_k.$$

In the simulation, set the initial value of estimation as $\hat{x}_{0|0} = \bar{x}_0 = \begin{bmatrix} 1.8 & 0.2 \end{bmatrix}^T$ and $\Sigma_{0|0} = 20I_2$. The other parameters are chosen as $B_k = \text{diag}\{0.1, 0.2\}$, $E_{k+1} = \begin{bmatrix} 0.1 & 0.15 \end{bmatrix}^T$, $L_k = L_{k+1} = 0.01I_2$, $\gamma_{1,k} =$

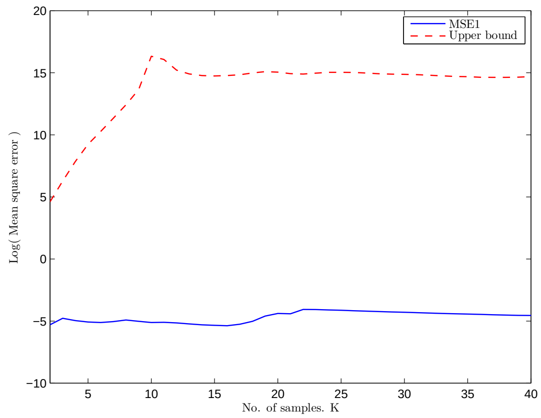

Fig. 2. MSE1 and its upper bound

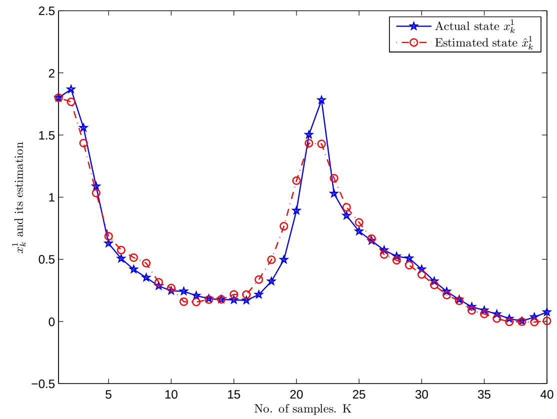

Fig. 4. The actual state x_k^1 and its estimation $\hat{x}_k^1$

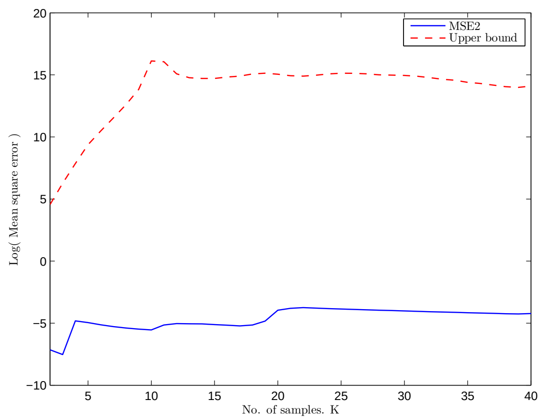

Fig. 3. MSE2 and its upper bound

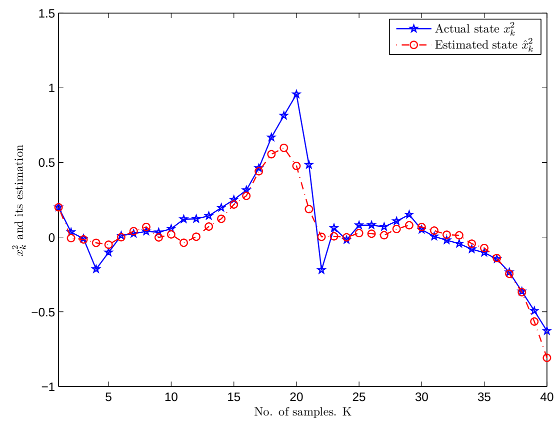

Fig. 5. The actual state x_k^2 and its estimation $\hat{x}_k^2$

0.002, $\gamma_{2,k+1} = 0.002$, $\varepsilon_1 = 0.4$, $\varepsilon_2 = 0.35$, $a_1 = 7.5$ and $a_2 = 0.05$. Let MSE1 denote the mean square error (MSE) for the estimation of the first state, i.e., $(1/K)\sum_{k=1}^{K}\left\{\begin{bmatrix} 1 & 0 \end{bmatrix}\left(x_k - \hat{x}_{k|k}\right)\right\}^2$, where K is the number of the samples. Similarly, MSE2 is the mean square error for the estimation of the second state, i.e., $(1/K)\sum_{k=1}^{K}\left\{\begin{bmatrix} 0 & 1 \end{bmatrix}\left(x_k - \hat{x}_{k|k}\right)\right\}^2$.

According to (28), (29) and (33) in Theorem 3, the upper bound of the filtering error covariance and filter gains at every time step can be recursively calculated. Therefore, the addressed filter design problem can be solved by means of the proposed filter structure (7)-(8). The simulation results are shown in Figs. 2-5. Among them, Figs. 2-3 show the upper bounds $\Sigma_{k|k}^{11}$ and $\Sigma_{k|k}^{22}$ as well as the MSE for the states x_k^1 and x_k^2, which confirm that the MSE stay below their upper bounds. Moreover, the trajectories of the actual states x_k^i and their estimates $\hat{x}_k^i$ $(i = 1, 2)$ are plotted in Figs. 4-5, which illustrate that the presented filter scheme can perform well to estimate the system states. This is due to the fact that we have made specific efforts to compensate the effects of the stochastic nonlinearities and multiple missing measurements.

Remark 7 *As discussed in [31], the matrices B_k, E_{k+1} and L_k are used to quantitatively characterize the upper bound of the linearization errors obtained from the Taylor series expansion for the nonlinearities. Accordingly, by taking the inequalities (17) and (20) into consideration, the high-order terms in the Taylor series expansions can be approximated. In the simulation, we set the matrix L_k as $\delta_k I$ (δ_k is a positive constant) in order to enhance the feasibility of (30) and (31), and then we can always adjust the values of scaling matrices B_k and E_{k+1} to guarantee the inequalities (17) and (20). Specifically, it is worth mentioning that we can simply set $B_k = 0$ and $E_{k+1} = 0$ if the effects of the linearization errors are negligible for some problems.*

5 Conclusions

In this paper, we have made one of the first few attempts to design the finite-horizon EKF for a class of time-varying systems with stochastic nonlinearities and multiple missing measurements. The stochastic nonlinearities described by statistical means have been taken into account. The phenomenon of multiple missing measurements has been described by any discrete-time distributions with known probability density function. A series of mutually independent random variables has been introduced to characterize the operation behavior of each sensor. By means of the Riccati-like equation approach, we have designed the EKF such that, for both the stochastic nonlinearities and multiple missing measurements, the upper bound of the filtering error covariance exits and is then minimized by properly designing the filter gain at every sampling instant. Finally, the effectiveness and applicability of the developed algorithm has been demonstrated by an illustrative simulation example.

References

[1] A. N. Bishop, A. V. Savkin, and P. N. Pathirana, Vision-based target tracking and surveillance with robust set-valued state estimation, *IEEE Signal Processing Letters*, vol. 17, no. 3, pp. 289–292, Mar. 2010.

[2] M. Basin, P. Shi, and D. Calderon-Alvarez, Central suboptimal H_∞ filter design for nonlinear polynomial systems, *International Journal of Adaptive Control and Signal Processing*, vol. 23, no. 10, pp. 926–939, Oct. 2009.

[3] M. Basin, P. Shi, and D. Calderon-Alvarez, Approximate finite-dimensional filtering for polynomial states over polynomial observations, *International Journal of Control*, vol. 83, no. 4, pp. 724–730, Apr. 2010.

[4] W. Chen and W. Zheng, Exponential stability of nonlinear time-delay systems with delayed impulse effects, *Automatica*, vol. 47, no. 5, pp. 1075–1083, May 2011.

[5] T. M. Cheng, V. Malyavej, and A. V. Savkin, Decentralized robust set-valued state estimation in networked multiple sensor systems, *Computers & Mathematics with Applications*, vol. 59, no. 8, pp. 2636–2646, Apr. 2010.

[6] G. Calafiore, Reliable localization using set-valued nonlinear filters, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 35, no. 2, pp. 189–197, Mar. 2005.

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

[8] F. O. Hounkpevi and E. Yaz, Robust minimum variance linear state estimators for multiple sensors with different failure rates, *Automatica*, vol. 43, no. 7, pp. 1274–1280, Jul. 2007.

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

[10] M. Sahebsara, T. Chen, and S. L. Shah, Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations, *International Journal of Control*, vol. 80, no. 2, pp. 292–301, Feb. 2007.

[11] M. R. James and I. R. Petersen, Nonlinear state estimation for uncertain systems with an integral constraint, *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 2926–2937, Nov. 1998.

[12] A. G. Kallapur, I. R. Petersen, and S. G. Anavatti, A discrete-time robust extended Kalman filter for uncertain systems with sum quadratic constraints, *IEEE Transactions on Automatic Control*, vol. 54, no. 4, pp. 850–854, Apr. 2009.

[13] S. Kluge, K. Reif, and M. Brokate, Stochastic stability of the extended Kalman filter with intermittent observations, *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 514–518, Feb. 2010.

[14] P. Li and J. Lam, Disturbance analysis of nonlinear differential equation models of genetic SUM regulatory networks, *IEEE-ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 1, pp. 253–259, Jan-Feb. 2011.

[15] P. Li, J. Lam, and Z. Shu, H_∞ positive filtering for positive linear discrete-time systems: an augmentation approach, *IEEE Transactions on Automatic Control*, vol. 55, no. 10, pp. 2337–2342, Oct. 2010.

[16] X. Mao. *Stochastic differential equations and applications (2nd ed.)*, Chichester: Horwood Publishing, 2007.

[17] W. NaNacara and E. Yaz, Recursive estimators for linear and nonlinear systems with uncertain observations, *Signal Processing*, vol. 62, no. 2, pp. 215–228, Oct. 1997.

[18] P. N. Pathirana, S. W. Ekanayake, and A. V. Savkin, Fusion based 3D tracking of mobile transmitters via robust set-valued state estimation with RSS measurements, *IEEE Communications Letters*, vol. 15, no. 5, pp. 554–556, May 2011.

[19] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, Stochastic stability of the discrete-time extended Kalman filter, *IEEE Transactions on Automatic control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.

[20] H. Rotstein, M. Sznaier, and M. Idan, H_2/H_∞ filtering-theory and an aerospace application, *Proceedings of the 1994 American Control Conference*, vol. 2, pp. 1791–1795, 1994.

[21] P. Shi, M. Mahmoud, S. K. Nguang, and A. Ismail, Robust filtering for jumping systems with mode-dependent delays, *Signal Processing*, vol. 86, no. 1, pp. 140–152, Jan. 2006.

[22] 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.

[23] Y. Theodor and U. Shaked, Robust discrete-time minimum-variance filtering, *IEEE Transactions on Signal Processing*, vol. 44, no. 2, pp. 181–189, Feb. 1996.

[24] Z. Wang, D. W. C. Ho, and X. Liu, Variance-constrained filtering for uncertain stochastic systems with missing measurements, *IEEE Transactions on Automatic control*, vol. 48, no. 7, pp. 1254–1258, Jul. 2003.

[25] 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 Transactions on Computational Biology and Bioinformatics*, vol. 6, no. 3, pp. 410–419, Jul-Sep. 2009.

[26] 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.

[27] L. Wu and W. Zheng, Weighted H_∞ model reduction for linear switched systems with time-varying delay, *Automatica*, vol. 45, no. 1, pp. 186–193, Jan. 2009.

[28] L. Xie, Y. C. Soh, and C. E. de Souza, Robust Kalman filtering for uncertain discrete-time systems, *IEEE Transactions on Automatic control*, vol. 39, no. 6, pp. 1310–1314, Jun. 1994.

[29] L. Xie, L. Lu, D. Zhang, and H. Zhang, Improved robust H_2 and H_∞ filtering for uncertain discrete-time systems, *Automatica*, vol. 40, no. 5, pp. 873–880, May 2004.

[30] J. Xiong and J. Lam, Fixed-order robust H_∞ filter design for Markovian jump systems with uncertain switching probabilities, *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1421–1430, Apr. 2006.

[31] K. Xiong, C. Wei, and L. Liu, Robust extended Kalman filtering for nonlinear systems with stochastic uncertainties, *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 40, no. 2, pp. 399–405, Mar. 2010.

[32] K. Xiong, L. Liu, and Y. Liu, Robust extended Kalman filtering for nonlinear systems with multiplicative noises, *Optimal Control Applications and Methods*, vol. 32, no. 1, pp. 47–63, Jan-Feb. 2011.

[33] E. Yaz, On the optimal state estimation of a class of discrete-time nonlinear systems, *IEEE Transactions on Circuits and Systems*, vol. 34, no. 9, pp. 1127–1129, Sep. 1987.

[34] Y. Yaz and E. Yaz, A new formulation of some discrete-time stochastic parameter state estimation problems, *Applied Mathematics Letters*, vol. 10, no. 6, pp. 13–19, Nov. 1997.

[35] E. Yaz and Y. Yaz, State estimation of uncertain nonlinear stochastic systems with general criteria, *Applied Mathematics Letters*, vol. 14, no. 5, pp. 605–610, Aug. 2001.

[36] D. Yue and Q. Han, Network-based robust H_∞ filtering for uncertain linear systems, *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4293–4301, Nov. 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. He is now 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, stochastic systems. He is a very active reviewer for many international journals.

Zidong Wang was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Information Systems and Computing, Brunel University, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 100 papers in refereed international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves as an Associate Editor for 11 international journals, including IEEE Transactions on Automatic Control, IEEE Transactions on Control Systems Technology, IEEE Transactions on Neural Networks, IEEE Transactions on Signal Processing, and IEEE Transactions on Systems, Man, and Cybernetics - Part C. He is a Senior Member of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Huijun Gao 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. Dr Gao's research interests include network-based control, robust control/filter theory, time-delay systems and their engineering applications.

He 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.

Lampros K. Stergioulas received the B.Sc. degree in informatics and physics from the National and Kapodistrian University of Athens, Athens, Greece, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Liverpool, Liverpool, U.K., specializing in information engineering and communications.

He is currently a Reader with the Department of Information Systems and Computing, Brunel University, Uxbridge, U.K. He has published over 150 papers in journals and international conferences. He has been a Principal Investigator in numerous European Union and U.K. projects and a Coordinator of three multisite European projects.

His research interests include signal processing, intelligent information processing, technology-enhanced learning, educational and health information systems, and human-centered computing. Dr. Stergioulas is a Qualified Chartered Engineer.