Frontiers of Information Technology & Electronic Engineering
www.zju.edu.cn/jzus; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

A novel approach of noise statistics estimate using $\boldsymbol{H}_\infty$ filter in target tracking*

Xie WANG[1,2], Mei-qin LIU[‡1,2], Zhen FAN[2], Sen-lin ZHANG[2]

([1]*State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China*)

([2]*College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China*)

E-mail: wangxiek@zju.edu.cn; liumeiqin@zju.edu.cn; fanzhen@zju.edu.cn; slzhang@zju.edu.cn

Received Aug. 17, 2015; Revision accepted Feb. 16, 2016; Crosschecked Apr. 11, 2016

Abstract: Noise statistics are essential for estimation performance. In practical situations, however, a priori information of noise statistics is often imperfect. Previous work on noise statistics identification in linear systems still requires initial prior knowledge of the noise. A novel approach is presented in this paper to solve this paradox. First, we apply the H_∞ filter to obtain the system state estimates without the common assumptions about the noise in conventional adaptive filters. Then by applying state estimates obtained from the H_∞ filter, better estimates of the noise mean and covariance can be achieved, which can improve the performance of estimation. The proposed approach makes the best use of the system knowledge without a priori information with modest computation cost, which makes it possible to be applied online. Finally, numerical examples are presented to show the efficiency of this approach.

Key words: Noise estimate, H_∞ filter, Target tracking
http://dx.doi.org/10.1631/FITEE.1500262 **CLC number:** TP274$^+$.2

1 Introduction

Most optimal or suboptimal estimation problems assume a priori information of the process and measurement noise statistics, at least the first-order and the second-order moments. In the last decades, the Kalman filter has demonstrated to be the best linear, minimum variance, unbiased estimator, even if the noise is not satisfied to the white Gaussian assumption. The well known limitation of the Kalman filter to real-world problems is that it requires a priori information of the state process and measurement noise. However, in most practical situations, these noise statistics are inexactly known or unknown. The use of wrong or inexact prior information can lead to poor estimation precision or complete failure.

When confronting the estimation problems lacking noise statistics information, traditional filters such as the Kalman filter may work less efficiently. To address the noise problem, plenty of work has been devoted to noise statistics estimation which is derived from the Kalman filter or the Bayesian filter since 1970s, such as Bayesian estimation (Mehra, 1972; Alspach *et al.*, 1974), maximum likelihood (Bohlin, 1976), correlation methods (Bélanger, 1974; Odelson *et al.*, 2006; Duník *et al.*, 2015), and covariance-matching methods (Carew and Belanger, 1973).

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. 61374021 and 61531015), the Zhejiang Provincial Natural Science Foundation of China (Nos. LZ14F030002, LZ14F030003, and LY15F030007), the Specialized Research Fund for the Doctoral Program of Higher Education of China (Nos. 20130101110109 and 20120101110115), the Open Fund for the Aircraft Marine Measurement and Control Joint Laboratory, China (No. FOM2015OF009), and the Aerospace Science Foundation of China (Nos. 20132076002 and 2015ZC76006)

ORCID: Xie WANG, http://orcid.org/0000-0001-6518-7564

Meanwhile, in modern industrial embedded real-time systems, recursive estimators are in serious need under the computational limits (Li and Bar-Shalom, 1994; Jwo and Huang, 2007; Yadav *et al.*, 2012). Myers and Tapley (1976) presented a simple sequential approach to estimating the noise statistics at little computational expense, which illuminates the basic idea of exploiting the residuals between state estimates and the measurements. The essential framework in Myers and Tapley (1976) was still utilized widely in recent research (Bavdekar *et al.*, 2011; Yadav *et al.*, 2012; Assa and Janabi-Sharifi, 2014; Feng *et al.*, 2014). However, most publications are still constrained by the Kalman filter or the Bayesian method, which requires a priori knowledge about the noise statistics. To release the noise assumption, it is natural for us to consider some other filtering methods.

Without considering the noise statistics, H_∞ filtering has been developed to obtain the system state estimates (Yaesh and Shaked, 1991; Banavar, 1992; Shen and Deng, 1997; Rawicz, 2000). H_∞ filtering is designed to deal with the circumstance with unknown noise statistics. In H_∞ filtering, the noise sources are arbitrary signals with bounded energy or bounded average power. The estimator is such designed that the gain, from the noise signals to the estimation error, is less than a prescribed bound. Without the first- and second-order moments of noise, the H_∞ filter can provide more accurate estimates of the system states (Shen and Deng, 1997; Simon, 2006) than conventional linear filters. Based on the state estimates from the H_∞ filter, we can obtain more precise estimates of the noise statistics.

Furthermore, multiple model (MM) algorithms have been demonstrated to be efficient, unbiased, and optimal (or suboptimal) (Mazor *et al.*, 1998; Li and Jilkov, 2005) when accommodating modern estimation requirements, such as maneuvering target tracking (Li and Jilkov, 2005; Jiang *et al.*, 2014) and speech recognition (Rabiner, 1990; Gales, 2009). In the Markov framework, a lot of filtering methods have been proposed. However, most of these approaches are still based on a priori knowledge of noise statistics. During the process of explaining how to apply the H_∞ filter to the MM algorithms, in which Bayes' rule is employed, we find that the noise statistics cannot be ignored. Li and Jia (2010) announced an interacting multiple model (IMM) algorithm based on the H_∞ filter, and Fu *et al.* (2013) proposed a robust algorithm which combines the diagonal IMM algorithm with the H_∞ filter. Nevertheless, their algorithms treated the noise covariance as a guess and did not solve the noise covariance problem.

We are motivated to find a new way to estimate the noise mean and covariance in the circumstance where a priori knowledge of the noise statistics is unknown. In this paper, a novel approach is established by exploiting the H_∞ filter. First we apply H_∞ filtering to obtain more precise estimates of the system states than conventional linear estimators. Second, based on the better state estimates, we are supposed to receive more exact residual samples about the noise, which can improve the estimation accuracy for the noise statistics.

2 Problem statement

Shen and Deng (1997) proposed a method to design the H_∞ filter based on game theory. Consider the following time-varying discrete-time system:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{F}_k \boldsymbol{x}_k + \boldsymbol{G}_k \boldsymbol{w}_k, \\ \boldsymbol{y}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_k, \\ \boldsymbol{z}_k = \boldsymbol{L}_k \boldsymbol{x}_k, \end{cases} \tag{1}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the system state, $\boldsymbol{y}_k \in \mathbb{R}^m$ is the observation vector, $\boldsymbol{z}_k$ is the vector to be estimated, and $\boldsymbol{F}_k$, $\boldsymbol{G}_k$, $\boldsymbol{H}_k$, and $\boldsymbol{L}_k$ are known real constant matrices of appropriate dimensions for any k. The process noise $\boldsymbol{w}_k \in \mathbb{R}^l$ and the measurement noise $\boldsymbol{v}_k \in \mathbb{R}^m$ are mutually independent white Gaussian with unknown but bounded covariances.

The initial state $\boldsymbol{x}_0$ is assumed to be Gaussian with mean $\hat{\boldsymbol{x}}_0$ and covariance $\boldsymbol{P}_0$. Without loss of generality, we assume that $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ satisfy the Gaussian distribution, i.e., $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{q}, \boldsymbol{Q})$ and $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{r}, \boldsymbol{R})$, where $\boldsymbol{q}$, $\boldsymbol{Q}$, $\boldsymbol{r}$, and $\boldsymbol{R}$ are process noise mean, process noise covariance, measurement noise mean, and measurement noise covariance, respectively, and are to be estimated. Moreover, $\boldsymbol{Q}$ and $\boldsymbol{R}$ are positive definite symmetric matrices. Assume that $(\boldsymbol{F}_k, \boldsymbol{G}_k)$ is controllable and $(\boldsymbol{H}_k, \boldsymbol{F}_k)$ is observable. Define $\boldsymbol{G}_k^+$ as the Moore-Penrose generalized inverse of $\boldsymbol{G}_k$ and $Y_k = \{\boldsymbol{y}_k, 0 \le k \le N-1\}$. The estimate of state $\hat{\boldsymbol{x}}_k$ at time k is computed based on the measurement history up to $N-1$ (suppose current time $k=N$). We are more interested in $\hat{\boldsymbol{z}}_k$, the estimate of $\boldsymbol{z}_k$, which is the linear combination of state $\boldsymbol{x}_k$.

The filtering problem to be addressed is to obtain an estimate $\hat{\boldsymbol{z}}_k$ of $\boldsymbol{z}_k$ that provides a uniformly small estimation error, i.e., $\boldsymbol{e}_k = \hat{\boldsymbol{z}}_k - \boldsymbol{z}_k$, for any $\boldsymbol{w}_k$, $\boldsymbol{v}_k$, $\boldsymbol{x}_0$, and the estimates of the noise statistics $\hat{\boldsymbol{q}}, \hat{\boldsymbol{Q}}$, $\hat{\boldsymbol{r}}$, and $\hat{\boldsymbol{R}}$.

3 H_∞ filtering

To put the H_∞ filtering problem for system (1) in a stochastic way, we first define

$$\|\boldsymbol{f}\|_{\boldsymbol{P}} = \boldsymbol{f}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{f} \quad \forall \boldsymbol{f} \in \mathbb{R}^n,$$

where $\boldsymbol{P}$ is a given symmetric positive definite matrix.

The recursive H_∞ filtering procedure is briefly presented below. For details, readers can refer to Shen and Deng (1997) and Simon (2006). According to the game theory, the measure of performance is then given by

$$J = \frac{\sum\limits_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{\boldsymbol{S}_k}}{\|\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\|^2_{\boldsymbol{P}_0^{-1}} + \sum\limits_{k=0}^{N-1}\left(\|\boldsymbol{w}_k\|^2_{\boldsymbol{W}_k^{-1}} + \|\boldsymbol{v}_k\|^2_{\boldsymbol{V}_k^{-1}}\right)} < \frac{1}{\theta}, \tag{2}$$

where $((\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0), \boldsymbol{w}_k, \boldsymbol{v}_k) \neq \boldsymbol{0}$, $\hat{\boldsymbol{x}}_0$ is the prior estimate of state $\boldsymbol{x}_0$, $\boldsymbol{S}_k > 0, \boldsymbol{P}_0^{-1} > 0, \boldsymbol{W}_k > 0, \boldsymbol{V}_k > 0$ are the weighting matrices, and $\dfrac{1}{\theta} > 0$ is the upper bound. Moreover, $\boldsymbol{S}_k$, $\boldsymbol{W}_k$, and $\boldsymbol{V}_k$ are designed according to the actual application scenarios by engineers. Without loss of generality, we assume $\hat{\boldsymbol{x}}_0 = \boldsymbol{x}_0 = \boldsymbol{0}$ so that Eq. (2) can be written as

$$J = \frac{\sum\limits_{k=0}^{N-1} \|\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k\|^2_{\boldsymbol{S}_k}}{\sum\limits_{k=0}^{N-1}\left(\|\boldsymbol{w}_k\|^2_{\boldsymbol{W}_k^{-1}} + \|\boldsymbol{v}_k\|^2_{\boldsymbol{V}_k^{-1}}\right)} < \frac{1}{\theta}. \tag{3}$$

Therefore, the recursive H_∞ estimator to system (1) is

$$\begin{cases} \bar{\boldsymbol{S}}_k = \boldsymbol{L}_k^{\mathrm{T}}\boldsymbol{S}_k\boldsymbol{L}_k, \\ \boldsymbol{K}_k = \boldsymbol{P}_k(\boldsymbol{I} - \theta\bar{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{V}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k)^{-1}\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{V}_k^{-1}, \\ \hat{\boldsymbol{x}}_{k+1} = \boldsymbol{F}_k\hat{\boldsymbol{x}}_k + \boldsymbol{F}_k\boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k), \\ \hat{\boldsymbol{z}}_k = \boldsymbol{L}_k\hat{\boldsymbol{x}}_k, \\ \boldsymbol{P}_{k+1} = \boldsymbol{F}_k\boldsymbol{P}_k(\boldsymbol{I} - \theta\bar{\boldsymbol{S}}_k\boldsymbol{P}_k + \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{V}_k^{-1}\boldsymbol{H}_k\boldsymbol{P}_k)^{-1}\boldsymbol{F}_k^{\mathrm{T}} \\ \qquad\quad +\boldsymbol{G}_k\boldsymbol{W}_k\boldsymbol{G}_k^{\mathrm{T}}, \end{cases} \tag{4}$$

where $\boldsymbol{I}$ is the identity matrix.

To guarantee the above estimator to be the solution, the following condition must be satisfied at each time k:

$$\boldsymbol{P}_k^{-1} - \theta\bar{\boldsymbol{S}}_k + \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{V}_k^{-1}\boldsymbol{H}_k > 0. \tag{5}$$

In this section, we have obtained the state estimates without ideal assumptions of the noise statistics. Now we redirect our focus to the estimation of the noise mean and covariance.

4 Estimation of $\boldsymbol{r}$, $\boldsymbol{R}$, $\boldsymbol{q}$, and $\boldsymbol{Q}$

The empirical estimators for the noise statistics are derived in a batch form based on the assumption of a constant value of noise samples. The basic idea behind the covariance-matching technique is to make the residuals consistent with their theoretical covariances. For the measurement noise covariance, consider the linear measurement state relationship at a given observation time k, given $\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k$. The Gaussian distributed noise, which can be proved to be energy bounded, is the necessary condition for the H_∞ filter (see the Appendix). The true state $\boldsymbol{x}_k$ is unknown, so $\boldsymbol{v}_k$ cannot be determined directly through $\boldsymbol{v}_k = \boldsymbol{y}_k - \boldsymbol{H}_k\boldsymbol{x}_k$. However, an intuitive approximation for $\boldsymbol{v}_k$ can be determined by the adoption of state estimates obtained from the formerly described H_∞ filter. Define the following innovation sequence:

$$\begin{aligned} \boldsymbol{r}_k &= \boldsymbol{y}_k - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k \\ &= (\boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k) - \boldsymbol{H}_k\hat{\boldsymbol{x}}_k \\ &= \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k, \end{aligned} \tag{6}$$

where $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k - \hat{\boldsymbol{x}}_k$ is the state estimation error.

It is assumed that $\boldsymbol{v}_k$ ($k = 1, 2, \ldots, N$) are independent, and parameters $\boldsymbol{r}$ and $\boldsymbol{R}$ are constant. If the innovation sequence $\boldsymbol{r}_k$ is assumed to be the representative of $\boldsymbol{v}_k$, it may be considered independent and identically distributed, and a simple parameter estimation problem can be constructed. Assuming $\boldsymbol{r}_k \sim \mathcal{N}(\boldsymbol{r}, \boldsymbol{C}_r)$, $k = 1, 2, \ldots, N$, sequence $\boldsymbol{r}_k$ can be acquired by Eq. (6). Based on the empirical measurements, the unknown distribution $\boldsymbol{r}_k$ is to be estimated first.

An unbiased estimator for $\boldsymbol{r}$ is taken as the sequence mean:

$$\hat{\boldsymbol{r}} = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{r}_k. \tag{7}$$

An unbiased estimator for $\boldsymbol{R}$ is obtained by first constructing an estimator for $\boldsymbol{C}_r$:

$$\hat{\boldsymbol{C}}_r = \frac{1}{N-1}\sum_{k=1}^{N}(\boldsymbol{r}_k - \hat{\boldsymbol{r}})(\boldsymbol{r}_k - \hat{\boldsymbol{r}})^{\mathrm{T}}. \tag{8}$$

Substituting Eq. (6) into Eq. (8), we can obtain the expectation of $\boldsymbol{C}_r$. First, we have

$$\begin{aligned}
&(\boldsymbol{r}_k - \hat{\boldsymbol{r}})(\boldsymbol{r}_k - \hat{\boldsymbol{r}})^{\mathrm{T}}\\
=&\boldsymbol{r}_k\boldsymbol{r}_k^{\mathrm{T}} - \boldsymbol{r}_k\hat{\boldsymbol{r}}^{\mathrm{T}} - \hat{\boldsymbol{r}}\boldsymbol{r}_k^{\mathrm{T}} + \hat{\boldsymbol{r}}\hat{\boldsymbol{r}}^{\mathrm{T}}\\
=&(\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k)(\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k)^{\mathrm{T}} - (\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k)\hat{\boldsymbol{r}}^{\mathrm{T}}\\
&- \hat{\boldsymbol{r}}(\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k + \boldsymbol{v}_k)^{\mathrm{T}} + \hat{\boldsymbol{r}}\hat{\boldsymbol{r}}^{\mathrm{T}}\\
=&\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\boldsymbol{v}_k^{\mathrm{T}} + \boldsymbol{v}_k\tilde{\boldsymbol{x}}_k\boldsymbol{H}_k^{\mathrm{T}} + \boldsymbol{v}_k\boldsymbol{v}_k^{\mathrm{T}}\\
&- \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\hat{\boldsymbol{r}}^{\mathrm{T}} - \boldsymbol{v}_k\hat{\boldsymbol{r}}^{\mathrm{T}} - \hat{\boldsymbol{r}}\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}} - \hat{\boldsymbol{r}}\boldsymbol{v}_k^{\mathrm{T}} + \hat{\boldsymbol{r}}\hat{\boldsymbol{r}}^{\mathrm{T}}\\
=&\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}} + (\boldsymbol{v}_k - \hat{\boldsymbol{r}})(\boldsymbol{v}_k - \hat{\boldsymbol{r}})^{\mathrm{T}} + \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\boldsymbol{v}_k^{\mathrm{T}}\\
&+ \boldsymbol{v}_k\tilde{\boldsymbol{x}}_k\boldsymbol{H}_k^{\mathrm{T}} - \boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\hat{\boldsymbol{r}}^{\mathrm{T}} - \hat{\boldsymbol{r}}\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{H}_k^{\mathrm{T}}.
\end{aligned} \tag{9}$$

Taking the expectation of both sides of Eq. (9), we have

$$\begin{aligned}
&E[(\boldsymbol{r}_k - \hat{\boldsymbol{r}})(\boldsymbol{r}_k - \hat{\boldsymbol{r}})^{\mathrm{T}}]\\
=&\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^{\mathrm{T}} + E[(\boldsymbol{v}_k - \hat{\boldsymbol{r}})(\boldsymbol{v}_k - \hat{\boldsymbol{r}})^{\mathrm{T}}]\\
=&\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^{\mathrm{T}} + \hat{\boldsymbol{R}}.
\end{aligned} \tag{10}$$

Here we apply the assumptions that $\tilde{\boldsymbol{x}}_k$ and $\boldsymbol{v}_k$ are mutually independent, and that $\hat{\boldsymbol{x}}_k$ is an unbiased estimate. Therefore, we have $E[\boldsymbol{H}_k\tilde{\boldsymbol{x}}_k\boldsymbol{v}_k^{\mathrm{T}}] = 0$ and $E[\tilde{\boldsymbol{x}}_k] = 0$. Then we obtain

$$E[\hat{\boldsymbol{C}}_r] = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^{\mathrm{T}} + \hat{\boldsymbol{R}}. \tag{11}$$

Hence, we obtain the estimate of $\boldsymbol{R}$ by substituting Eq. (8) into Eq. (11):

$$\begin{aligned}
\hat{\boldsymbol{R}} =& \frac{1}{N-1}\sum_{k=1}^{N}\Bigg[(\boldsymbol{r}_k - \hat{\boldsymbol{r}})(\boldsymbol{r}_k - \hat{\boldsymbol{r}})^{\mathrm{T}}\\
&- \frac{N-1}{N}\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^{\mathrm{T}}\Bigg].
\end{aligned} \tag{12}$$

For the state process noise statistics, we can basically follow the same scheme of the measurement noise estimation but some additional treatments are needed. Consider the linear dynamical state relationship at time k, given by Eq. (1):

$$\boldsymbol{G}_k\boldsymbol{w}_k = \boldsymbol{x}_{k+1} - \boldsymbol{F}_k\boldsymbol{x}_k. \tag{13}$$

Similar to $\boldsymbol{v}_k$, the true states $\boldsymbol{x}_{k+1}$ and $\boldsymbol{x}_k$ are difficult to know, and we can use an intuitive approach for $\boldsymbol{w}_k$ by applying the state estimates obtained from the H_∞ filter. Define the process residual as $\boldsymbol{q}_k$ and then define the following equation from Eq. (13):

$$\boldsymbol{G}_k\boldsymbol{q}_k = \hat{\boldsymbol{x}}_{k+1} - \boldsymbol{F}_k\hat{\boldsymbol{x}}_k. \tag{14}$$

From Eqs. (13) and (14), we obtain

$$\begin{cases}
\boldsymbol{w}_k = \boldsymbol{G}_k^{+}\left(\boldsymbol{x}_{k+1} - \boldsymbol{F}_k\boldsymbol{x}_k\right),\\
\boldsymbol{q}_k = \boldsymbol{G}_k^{+}\left(\hat{\boldsymbol{x}}_{k+1} - \boldsymbol{F}_k\hat{\boldsymbol{x}}_k\right).
\end{cases} \tag{15}$$

It is assumed that the stationary noise $\boldsymbol{w}_k$ ($k = 1, 2, \ldots, N$) are mutually independent, and parameters $\boldsymbol{q}$ and $\boldsymbol{Q}$ are constant. Here $\boldsymbol{q}_k$ is assumed to be the representative of $\boldsymbol{w}_k$, and it can be considered independent and identically distributed. Similarly, by defining a parameter estimation problem, $\boldsymbol{q}_k$ is characterized by Gaussian distribution $\mathcal{N}(\boldsymbol{q}, \boldsymbol{C}_q)$. An unbiased estimate for $\boldsymbol{q}$ is obtained from the sample mean:

$$\hat{\boldsymbol{q}} = \frac{1}{N}\sum_{k=1}^{N}\boldsymbol{q}_k. \tag{16}$$

Now we obtain the estimate of $\boldsymbol{C}_q$ by

$$\hat{\boldsymbol{C}}_q = \frac{1}{N-1}\sum_{k=1}^{N}(\boldsymbol{q}_k - \hat{\boldsymbol{q}})(\boldsymbol{q}_k - \hat{\boldsymbol{q}})^{\mathrm{T}}. \tag{17}$$

Subtracting Eq. (14) by Eq. (13), we obtain

$$\begin{aligned}
\boldsymbol{G}_k\left(\boldsymbol{w}_k - \boldsymbol{q}_k\right) &= (\boldsymbol{x}_{k+1} - \hat{\boldsymbol{x}}_{k+1}) - \boldsymbol{F}_k(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k)\\
&= \tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{F}_k\tilde{\boldsymbol{x}}_k,
\end{aligned} \tag{18}$$

and then

$$\begin{aligned}
&\boldsymbol{G}_k(\boldsymbol{w}_k - \boldsymbol{q}_k)(\boldsymbol{w}_k - \boldsymbol{q}_k)^{\mathrm{T}}\boldsymbol{G}_k^{\mathrm{T}}\\
=&(\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{F}_k\tilde{\boldsymbol{x}}_k)(\tilde{\boldsymbol{x}}_{k+1} - \boldsymbol{F}_k\tilde{\boldsymbol{x}}_k)^{\mathrm{T}}\\
=&\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}} - \boldsymbol{F}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}} - \tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{F}_k^{\mathrm{T}}\\
&+ \boldsymbol{F}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{F}_k^{\mathrm{T}}.
\end{aligned} \tag{19}$$

Taking the expectation of both sides of the equation above, we have

$$\begin{aligned}&\boldsymbol{G}_kE[(\boldsymbol{w}_k-\boldsymbol{q}_k)(\boldsymbol{w}_k-\boldsymbol{q}_k)^{\mathrm{T}}]\boldsymbol{G}_k^{\mathrm{T}}\\ =&E\big[\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}}-\boldsymbol{F}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}}-\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{F}_k^{\mathrm{T}}\\ &+\boldsymbol{F}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{F}_k^{\mathrm{T}}\big]\\ =&E[\tilde{\boldsymbol{x}}_{k+1}\tilde{\boldsymbol{x}}_{k+1}^{\mathrm{T}}]+E[\boldsymbol{F}_k\tilde{\boldsymbol{x}}_k\tilde{\boldsymbol{x}}_k^{\mathrm{T}}\boldsymbol{F}_k^{\mathrm{T}}]\\ =&\boldsymbol{P}_{k+1}+\boldsymbol{F}_k\boldsymbol{P}_k\boldsymbol{F}_k^{\mathrm{T}},\end{aligned}\tag{20}$$

where the estimation errors $\tilde{\boldsymbol{x}}_k$ and $\tilde{\boldsymbol{x}}_{k+1}$ are mutually independent. We still need the relationship between $\boldsymbol{Q}$ and $\boldsymbol{C}_q$. Note that, based on our hypothesis, $\boldsymbol{w}_k$ and $\boldsymbol{q}_k$ have the same mean value. Let us introduce $\hat{\boldsymbol{q}}$ into $(\boldsymbol{w}_k-\boldsymbol{q}_k)(\boldsymbol{w}_k-\boldsymbol{q}_k)^{\mathrm{T}}$ as follows:

$$\begin{aligned}&(\boldsymbol{w}_k-\boldsymbol{q}_k)(\boldsymbol{w}_k-\boldsymbol{q}_k)^{\mathrm{T}}\\ =&[(\boldsymbol{w}_k-\hat{\boldsymbol{q}})-(\boldsymbol{q}_k-\hat{\boldsymbol{q}})][(\boldsymbol{w}_k-\hat{\boldsymbol{q}})-(\boldsymbol{q}_k-\hat{\boldsymbol{q}})]^{\mathrm{T}}\\ =&(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}-(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\\ &-(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}+(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}.\end{aligned}\tag{21}$$

In the same way, taking the expectation of both sides, we have

$$\begin{aligned}&E[(\boldsymbol{w}_k-\boldsymbol{q}_k)(\boldsymbol{w}_k-\boldsymbol{q}_k)^{\mathrm{T}}]\\ =&E\big[(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}-(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\\ &-(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}+(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\big]\\ =&E\left[(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\right]-E\left[(\boldsymbol{w}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\right]\\ &-E\left[(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{w}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\right]+E\left[(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\right]\\ =&\hat{\boldsymbol{Q}}-\left(E[\boldsymbol{w}_k\boldsymbol{q}_k^{\mathrm{T}}]-E[\boldsymbol{w}_k]\hat{\boldsymbol{q}}^{\mathrm{T}}-\hat{\boldsymbol{q}}E[\boldsymbol{q}_k^{\mathrm{T}}]+\hat{\boldsymbol{q}}\hat{\boldsymbol{q}}^{\mathrm{T}}\right)\\ &-\left(E[\boldsymbol{q}_k\boldsymbol{w}_k^{\mathrm{T}}]-E[\boldsymbol{q}_k]\hat{\boldsymbol{q}}^{\mathrm{T}}-\hat{\boldsymbol{q}}E[\boldsymbol{w}_k^{\mathrm{T}}]+\hat{\boldsymbol{q}}\hat{\boldsymbol{q}}^{\mathrm{T}}\right)+\boldsymbol{C}_q\\ =&\hat{\boldsymbol{Q}}+\boldsymbol{C}_q,\end{aligned}\tag{22}$$

where we apply the hypothesis that $\boldsymbol{w}_k$ and $\boldsymbol{q}_k$ are mutually independent. From Eqs. (20) and (22), we have

$$\boldsymbol{P}_{k+1}+\boldsymbol{F}_k\boldsymbol{P}_k\boldsymbol{F}_k^{\mathrm{T}}=\boldsymbol{G}_k(\hat{\boldsymbol{Q}}+\boldsymbol{C}_q)\boldsymbol{G}_k^{\mathrm{T}}.\tag{23}$$

Following the same steps used for Eq. (11), we have

$$\boldsymbol{G}_kE[\boldsymbol{C}_q]\boldsymbol{G}_k^{\mathrm{T}}=\frac{1}{N}\sum_{k=1}^{N}\left[\boldsymbol{P}_{k+1}+\boldsymbol{F}_k\boldsymbol{P}_k\boldsymbol{F}_k^{\mathrm{T}}\right]-\boldsymbol{G}_k\hat{\boldsymbol{Q}}\boldsymbol{G}_k^{\mathrm{T}}.\tag{24}$$

The estimate of $\boldsymbol{Q}$, after the substitution of Eq. (17), is given by

$$\begin{aligned}\hat{\boldsymbol{Q}}=&\frac{1}{N-1}\sum_{k=1}^{N}\boldsymbol{G}_k^{+}\Big[\frac{N-1}{N}\left(\boldsymbol{F}_k\boldsymbol{P}_k\boldsymbol{F}_k^{\mathrm{T}}+\boldsymbol{P}_{k+1}\right)\\ &-(\boldsymbol{q}_k-\hat{\boldsymbol{q}})(\boldsymbol{q}_k-\hat{\boldsymbol{q}})^{\mathrm{T}}\Big]\left(\boldsymbol{G}_k^{+}\right)^{\mathrm{T}}.\end{aligned}\tag{25}$$

In summary, the estimators for $\boldsymbol{r}, \boldsymbol{R}, \boldsymbol{q}$, and $\boldsymbol{Q}$ are presented in Eqs. (7), (12), (16), and (25), respectively. These estimates are based on the innovation sequence $\boldsymbol{r}_k$ and the residual sequence $\boldsymbol{q}_k$ ($k=1,2,\ldots,N$), which are obtained through the H_∞ filtering estimates and assumed to be statistically independent and identically distributed.

5 Numerical simulation

Example 1 Consider a simple scalar system (Simon, 2006)

$$\begin{cases}x_{k+1}=x_k+u_k+w_k,\\ y_k=x_k+v_k,\end{cases}\tag{26}$$

where $u_k=10$, $w_k\sim\mathcal{N}(0,1)$ and $v_k\sim\mathcal{N}(0,1)$ are both Gaussian white noises.

Set the H_∞ filter parameters $\theta=1/50$, $W_k=V_k=10$. The initial values in the Kalman filter are $Q_0=R_0=1$.

We compare the proposed approach with the sequential estimator in Myers and Tapley (1976), where the two approaches are derived from the same framework. The performance of the estimators is displayed using the root mean square error (RMSE) and the results are obtained from 100 Monte Carlo runs.

The simulation results (Figs. 1 and 2, Table 1) easily confirm that the proposed approach plays much better. Since the H_∞ filter does not rely on the noise statistics, it would naturally provide more precise residual samples, which are used to obtain the noise statistics estimates. On the other hand, once the Kalman filter fails in the estimating procedure, it might lead to complete collapse.

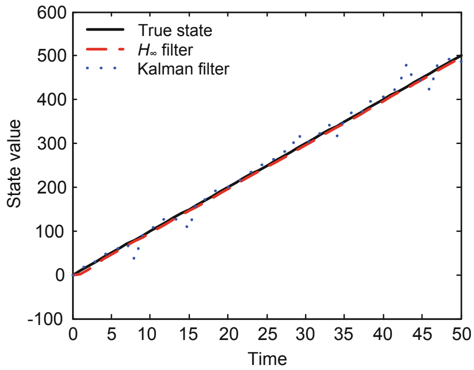

Fig. 1 Comparisons of the true system state x_k and the results under the Kalman filter and H_∞ filter

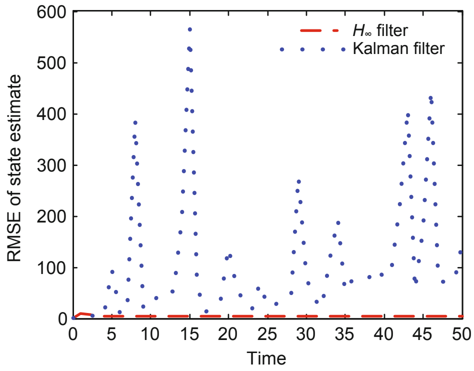

Fig. 2 Root mean square errors (RMSEs) between the true state and the estimates under the Kalman filter and H_∞ filter

Example 2 Consider a maneuvering target similar to Alouani and Blair (1993), which consists of constant velocity (CV) and constant turn (CT) movements. The detailed trajectory is shown in Fig. 3. When $20 < t \le 30$, the target moves in CT mode, and the angular velocity is $w = -0.1$. During the rest of the time, the target moves in CV mode. The CV state function is

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix} \boldsymbol{w}_k, \tag{27}$$

while the CT state function is

$$\boldsymbol{x}_{k+1} = \begin{bmatrix} 1 & \frac{\sin(wT)}{w} & 0 & \frac{-[1-\cos(wT)]}{w} \\ 0 & \cos(wT) & 0 & -\sin(wT) \\ 0 & \frac{1-\cos(wT)}{w} & 1 & \frac{\sin(wT)}{w} \\ 0 & \frac{\sin(wT)}{w} & 0 & \cos(wT) \end{bmatrix} \boldsymbol{x}_k + \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix} \boldsymbol{w}_k. \tag{28}$$

The measurement function for both models is

$$\boldsymbol{y}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \boldsymbol{x}_k + \boldsymbol{v}_k. \tag{29}$$

In this example, we estimate the first and the third elements in $\boldsymbol{x}_k$, which are the X and Y components of the position, respectively. As we can see from Eqs. (27)–(29), both $\boldsymbol{w}_k$ and $\boldsymbol{v}_k$ at each time k consist of two elements such as $[w_1\ w_2]^{\mathrm{T}}$ and $[v_1\ v_2]^{\mathrm{T}}$. In this simulation, we assume that both $\boldsymbol{w}_k$ and v_k are Gaussian white noises and $\boldsymbol{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{1})$, $\boldsymbol{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{30})$.

Table 1 Average root mean square error (RMSE) comparison between the Kalman filter and the H_∞ filter

Method	w_k		v_k	
	Mean	Covariance	Mean	Covariance
Kalman	8.2556	3.7×10^6	56.4969	1.7×10^6
H_∞	0.5839	24.7196	5.8127	11.8625

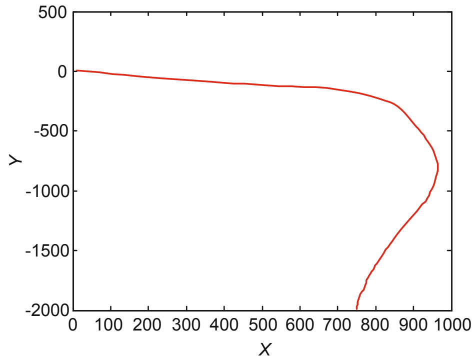

Fig. 3 The noisy trajectory (consisting of constant velocity (CV) and constant turn (CT) movements)

In our approach, the dynamics of the noise is handled as unknown, which is important and distinct from other approaches. What we will do here is to achieve the state estimate $\hat{\boldsymbol{x}}_k$ through H_∞ filtering first. Then by using all the known information, including $\hat{\boldsymbol{x}}_k$ and $\boldsymbol{P}_k$ in Eq. (4), we will try to obtain the estimates of means and covariances of the process noise and measurement noise, respectively.

For comparison, we employ the estimation method in Myers and Tapley (1976). Both methods are derived in a batch form based on the residual samples under the same framework, where the algorithm in Myers and Tapley (1976) uses the Kalman filter, and our algorithm uses the H_∞ filter. Nevertheless, we have to give the former an initial value of the noise statistics, which is a drawback of the algorithm but is also one of the motivations of this study.

From Fig. 4, we can see that the H_∞ filter shows better performance than the Kalman filter. In other words, the algorithm in this study might obtain more precise residual samples (see Eqs. (6) and (14)). In short, the simulation results confirm our

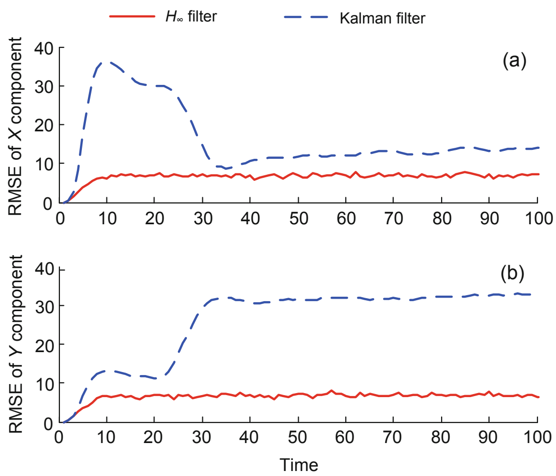

Fig. 4 Root mean square errors (RMSEs) between the system state and the estimated state: (a) $\boldsymbol{X}$ component of the position; (b) $\boldsymbol{Y}$ component of the position

suppositions. The mean and covariance estimates of the process noise and measurement noise are shown in Figs. 5 and 6, respectively. Specifically, in Fig. 6a and 6b, the RMSE of the process noise covariance estimates based on the Kalman filter cannot converge to 0. Although the RMSE of the process noise mean estimate might be a little bit larger (Figs. 5a and 5b), our algorithm can guarantee the convergence as the residual samples accumulate.

Note that θ, $\boldsymbol{P}_0$, $\boldsymbol{S}_k$, $\boldsymbol{W}_k$, $\boldsymbol{V}_k$ in Eq. (2) are free parameters, which should be designed according to special circumstances and instances. The effectiveness of the H_∞ filter and the novel algorithm can be highly sensitive to these weighting parameters, although this algorithm is robust to model uncertainty. The adjustment procedure might be a little complicated (Simon, 2006). However, the cost is worthy of the high performance and simplicity to be embedded into online systems as we can see from the simulation results. After appropriate parameter adjustment, all the results approximate the true values of the means and covariances.

6 Conclusions

In this paper, we have presented a novel approach to estimating the process noise and measurement noise statistics in a linear discrete system. In the target tracking circumstance where a priori information of the noise statistics is unavailable but necessary, we naturally introduced the H_∞ filter to

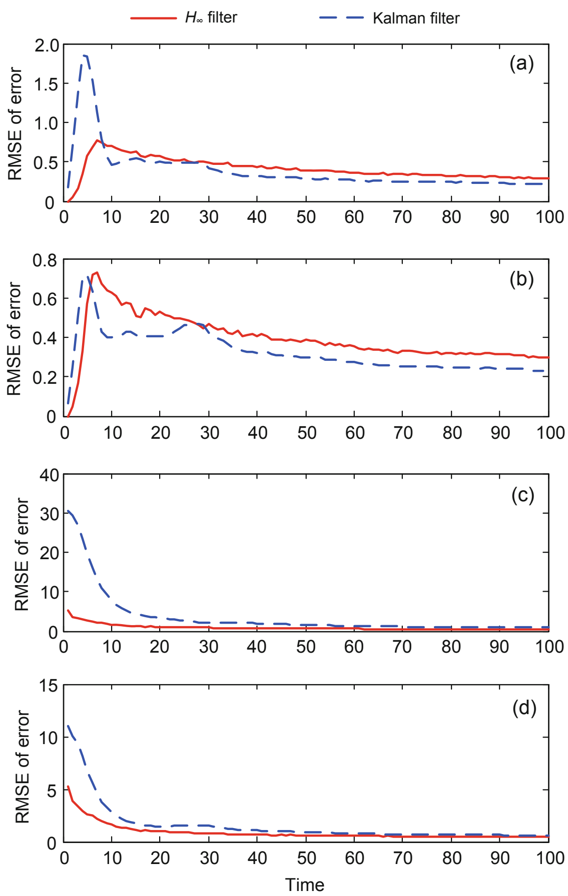

Fig. 5 Root mean square errors (RMSEs) between the process noise mean estimates and the actual values for w_1 (a) and w_2 (b), and the RMSEs between the measurement noise mean estimates and the actual values for v_1 (c) and v_2 (d)

the procedure, which can achieve better system state estimation. The basic idea of our algorithm is to take the best advantage of the residual samples between the system state estimates and the measurements, and the residual samples between the state estimates at time $k+1$ and the ones at time k. According to statistical theory, we can obtain the rational noise statistics estimates. To acquire the noise statistics, the key point is to apply the H_∞ filter to bypass the noise details.

The readers might ask why we would estimate the mean and covariance of the noise while the H_∞ filter is suited efficiently. Since in some situations, the information of the noise is still important and cannot be ignored. Yet in previous work about noise statistics identification, a priori information is still needed. What we have done here is to make the best

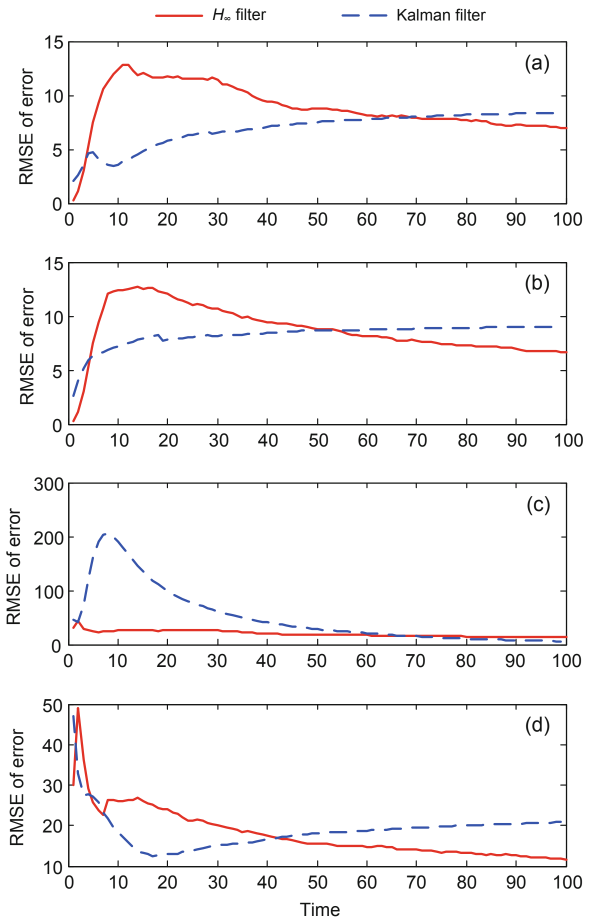

Fig. 6 Root mean square errors (RMSEs) between the process noise covariance estimates and the actual values for Q_{11} (a) and Q_{22} (b), and the RMSEs between the measurement noise covariance estimates and the actual values for R_{11} (c) and R_{22} (d)

use of system knowledge completely without a priori information. Furthermore, the algorithm can work at a modest additional computational cost, which makes it possible to be employed in an online framework. In our future work, this procedure may be applied in the MM algorithm which requires noise statistics.

References

Alouani, A.T., Blair, W.D., 1993. Use of a kinematic constraint in tracking constant speed, maneuvering targets. *IEEE Trans. Autom. Contr.*, **38**(7):1107-1111. http://dx.doi.org/10.1109/9.231465

Alspach, D.L., Scharf, L.L., Abiri, A., 1974. A Bayesian solution to the problem of state estimation in an unknown noise environment. *Int. J. Contr.*, **19**(2):265-287. http://dx.doi.org/10.1080/00207177408932628

Assa, A., Janabi-Sharifi, F., 2014. A robust vision-based sensor fusion approach for real-time pose estimation. *IEEE Trans. Cybern.*, **44**(2):217-227. http://dx.doi.org/10.1109/TCYB.2013.2252339

Banavar, R.N., 1992. A Game Theoretic Approach to Linear Dynamic Estimation. PhD Thesis, Texas University, Austin, USA.

Bavdekar, V.A., Deshpande, A.P., Patwardhan, S.C., 2011. Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *J. Process Contr.*, **21**(4):585-601. http://dx.doi.org/10.1016/j.jprocont.2011.01.001

Bélanger, P.R., 1974. Estimation of noise covariance matrices for a linear time-varying stochastic process. *Automatica*, **10**(3):267-275. http://dx.doi.org/10.1016/0005-1098(74)90037-5

Bohlin, T., 1976. Four cases of identification of changing systems. *Math. Sci. Eng.*, **126**:441-518. http://dx.doi.org/10.1016/S0076-5392(08)60878-4

Carew, B., Belanger, P., 1973. Identification of optimum filter steady-state gain for systems with unknown noise covariances. *IEEE Trans. Autom. Contr.*, **18**(6):582-587. http://dx.doi.org/10.1109/TAC.1973.1100420

Duník, J., Straka, O., Šimandl, M., 2015. Estimation of noise covariance matrices for linear systems with nonlinear measurements. Proc. 17th Symp. on System Identification, p.1130-1135. http://dx.doi.org/10.1016/j.ifacol.2015.12.283

Feng, B., Fu, M., Ma, H., *et al.*, 2014. Kalman filter with recursive covariance estimation—sequentially estimating process noise covariance. *IEEE Trans. Ind. Electron.*, **61**(11):6253-6263. http://dx.doi.org/10.1109/TIE.2014.2301756

Fu, X., Jia, Y., Du, J., *et al.*, 2013. H_∞ filtering with diagonal interacting multiple model algorithm for maneuvering target tracking. Proc. American Control Conf., p.6187-6192.

Gales, M.J.F., 2009. Acoustic modelling for speech recognition: hidden Markov models and beyond? Proc. IEEE Workshop on Automatic Speech Recognition & Understanding, p.44. http://dx.doi.org/10.1109/ASRU.2009.5372953

Jiang, T.Y., Liu, M.Q., Wang, X., *et al.*, 2014. An efficient measurement-driven sequential Monte Carlo multi-Bernoulli filter for multi-target filtering. *J. Zhejiang Univ.-Sci. C (Comput. & Electron.)*, **15**(6):445-457. http://dx.doi.org/10.1631/jzus.C1400025

Jwo, D.J., Huang, C.M., 2007. An adaptive fuzzy strong tracking Kalman filter for GPS/INS navigation. Proc. 33rd Annual Conf. of the IEEE Industrial Electronics Society, p.2266-2271. http://dx.doi.org/10.1109/IECON.2007.4460302

Li, W., Jia, Y., 2010. Distributed interacting multiple model H_∞ filtering fusion for multiplatform maneuvering target tracking in clutter. *Signal Process.*, **90**(5):1655-1668. http://dx.doi.org/10.1016/j.sigpro.2009.11.016

Li, X.R., Bar-Shalom, Y., 1994. A recursive multiple model approach to noise identification. *IEEE Trans. Aerosp. Electron. Syst.*, **30**(3):671-684. http://dx.doi.org/10.1109/7.303738

Li, X.R., Jilkov, V.P., 2005. Survey of maneuvering target tracking. Part V: multiple-model methods. *IEEE Trans. Aerosp. Electron. Syst.*, **41**(4):1255-1321. http://dx.doi.org/10.1109/TAES.2005.1561886

Mazor, E., Averbuch, A., Bar-Shalom, Y., *et al.*, 1998. Interacting multiple model methods in target tracking: a survey. *IEEE Trans. Aerosp. Electron. Syst.*, **34**(1):103-123. http://dx.doi.org/10.1109/7.640267

Mehra, R., 1972. Approaches to adaptive filtering. *IEEE Trans. Autom. Contr.*, **17**(5):693-698. http://dx.doi.org/10.1109/TAC.1972.1100100

Myers, K., Tapley, B., 1976. Adaptive sequential estimation with unknown noise statistics. *IEEE Trans. Autom. Contr.*, **21**(4):520-523. http://dx.doi.org/10.1109/TAC.1976.1101260

Odelson, B.J., Rajamani, M.R., Rawlings, J.B., 2006. A new autocovariance least-squares method for estimating noise covariances. *Automatica*, **42**(2):303-308. http://dx.doi.org/10.1016/j.automatica.2005.09.006

Rabiner, L.R., 1990. A tutorial on hidden Markov models and selected applications in speech recognition. *In*: Waibel, A., Lee, K.F. (Eds.), Readings in Speech Recognition. Morgan Kaufmann Publishers Inc., USA, p.267-296.

Rawicz, P.L., 2000. H_∞/H_2/Kalman Filtering of Linear Dynamical Systems via Variational Techniques with Applications to Target Tracking. PhD Thesis, Drexel University, Philadelphia, USA.

Shen, X., Deng, L., 1997. Game theory approach to discrete H_∞ filter design. *IEEE Trans. Signal Process.*, **45**(4):1092-1095. http://dx.doi.org/10.1109/78.564201

Simon, D., 2006. Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches. Wiley, New York, USA.

Yadav, A., Naik, N., Ananthasayanam, M.R., *et al.*, 2012. A constant gain Kalman filter approach to target tracking in wireless sensor networks. Proc. 7th IEEE Int. Conf. on Industrial and Information Systems, p.1-7. http://dx.doi.org/10.1109/ICIInfS.2012.6304803

Yaesh, I., Shaked, U., 1991. A transfer function approach to the problems of discrete-time systems: H_∞-optimal linear control and filtering. *IEEE Trans. Autom. Contr.*, **36**(11):1264-1271. http://dx.doi.org/10.1109/9.100935

Appendix: Proof of the energy bounded property of Gaussian distributed noise

The proof needs the following Parseval's theorem:

Theorem 1 (Parseval's theorem) If $f(x)$ is integrable and square integrable, then we have

$$\int_{-\infty}^{+\infty} f^2(x)\mathrm{d}x = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |F(w)|^2\mathrm{d}w,$$

where $F(w)$ represents the continuous Fourier transform of $f(x)$.

Consider $f(x) \sim \mathcal{N}(\mu, \delta^2)$. According to the Gaussian distribution property, $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$, and $f(x)$ is integrable and square integrable. The energy spectral density function is defined as

$$\Phi(w) = \left|\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} f^2(t)\mathrm{e}^{-\mathrm{i}wt}\mathrm{d}t\right|^2 = \frac{F(w)F^*(w)}{2\pi}, \quad \text{(A1)}$$

where '*' represents the conjugate transpose.

From the above theorem, we can obtain

$$\int_{-\infty}^{+\infty} \Phi(w)\mathrm{d}w = \int_{-\infty}^{+\infty} |f(x)|^2\mathrm{d}x. \quad \text{(A2)}$$

Combined with the condition $\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1$, Eq. (A2) is limited. Therefore, the Gaussian distributed noise is energy bounded.