

# Chapter 2

## Adaptively Robust Kalman Filters with Applications in Navigation

Yuanxi Yang

### Contents

2.1	Introduction . . . . .	50
2.2	The Principle of Adaptively Robust Kalman Filtering . . . . .	53
2.3	Properties of the Adaptive Kalman Filter . . . . .	56
2.3.1	Difference of State Estimate . . . . .	56
2.3.2	The Expectation of the State Estimate of the Adaptive Filter . . . . .	57
2.3.3	Posterior Precision Evaluation . . . . .	58
2.4	Three Kinds of Learning Statistics . . . . .	60
2.4.1	Learning Statistic Constructed by State Discrepancy . . . . .	60
2.4.2	Learning Statistic Constructed by Predicted Residual Vector . . . . .	61
2.4.3	Learning Statistic Constructed by the Ratio of Variance Components . . . . .	62
2.4.4	Learning Statistic Constructed by Velocity . . . . .	63
2.5	Four Kinds of Adaptive Factors . . . . .	63
2.5.1	Adaptive Factor by Three-Segment Function . . . . .	63
2.5.2	Adaptive Factor by Two-Segment Function . . . . .	64
2.5.3	Adaptive Factor by Exponential Function . . . . .	64
2.5.4	Adaptive Factor by Zero and One . . . . .	65
2.5.5	Actual Computation and Analysis . . . . .	66
2.6	Comparison of Two Fading Filters and Adaptively Robust Filter . . . . .	68
2.6.1	Principles of Two Kinds of Fading Filters . . . . .	69
2.6.2	Comparison of Fading Filter and Adaptive Filter . . . . .	71
2.6.3	Actual Computation and Analysis . . . . .	72
2.7	Comparison of Sage Adaptive Filter and Adaptively Robust Filter . . . . .	74
2.7.1	IAE Windowing Method . . . . .	74
2.7.2	RAE Windowing Method . . . . .	75
2.7.3	The Problems of the Windowing Estimation for Covariance Matrix of Kinematic Model . . . . .	76

---

Y. Yang (✉)

Xian Research Institute of Surveying and Mapping, Xian 710054, China  
e-mail: yuanxi@pub.xaonline.com

2.8 Some Application Examples . . . . . 77

References . . . . . 80

A new adaptively robust Kalman filtering was developed in 2001. The main achievements of the adaptively robust filter are summarized from the published papers in recent years. These include the establishment of the principle of the adaptively robust filter, the derivation of the corresponding state parameter estimator, the developments of four adaptive factors for balancing the contribution of kinematic model information and measurements, which include three-segment function, two-segment function, exponential function and zero and one function for state component adaptation, and the establishment of four kinds of learning statistics for judging the kinematic model errors, which include state discrepancy statistic, predicted residual statistic, variance component ratio statistic and velocity discrepancy statistic. The relations of the adaptively robust filter with standard Kalman filter, robust filter and some other adaptive Kalman filters as well as some related adjustment methods are depicted by a figure. Other developments of the adaptively robust filter are also presented.

2.1 Introduction

Applications of the Kalman filter in dynamic or kinematic positioning have sometimes encountered difficulties which have been referred to as divergences. These divergences can often be traced to three factors: (1) insufficient accuracy in modelling the dynamics or kinematics (functional model errors of the state equations); (2) insufficient accuracy in modelling the observations (functional model errors of observation equations); and (3) insufficient accuracy in modelling the distributions or the priori covariance matrices of the measurements and the updated parameters (stochastic model errors) (Yang et al. 2001a).

The current basic procedure for the quality control of Kalman filter consists of

- Functional model compensation for model errors by introducing uncertain parameters into the state and/or the observation equations. Any model error term can be introduced into the models arbitrarily. One could then **augment the state** (Jazwinski 1970, p. 308). A similar approach is developed by Schaffrin (1991, pp. 32–34). He partitions the state vector into  $h$  groups, each being affected by a common scale error. Then an  $h \times 1$  vector of scale parameters is introduced into the models. This kind of approach may, of course, lead to a **high-dimensional state vector** which, in turn, greatly **increases the filter computational load** (Jazwinski 1970, p. 305).
- **Stochastic model compensation by introducing a variance–covariance matrix of the model errors.** In taking this approach to prevent divergence, one has to determine what covariance matrix to add. A reasonable covariance matrix may compensate for the model errors. An ineffective covariance matrix, however, adds the model divergence. For instance, when the model is accurate in some dynamic or kinematic periods, an unsuitable increasing of the covariance matrix of model

error will degrade the state estimators. An effective covariance matrix for model errors can only be determined by trial and error.

- **DIA procedure – detection, identification and adaptation** (Teunissen 1990). It uses a recursive testing procedure to eliminate outliers. In the detection step one looks for unspecified model errors. In the identification step one tries to find the cause of the model error and its most likely starting time. After a model error has been detected and identified, the bias in the state estimate caused by the model error has to be eliminated as well. This model recovery from errors is called adaptation (Salzmann 1995). The identification of the model, however, is quite difficult, especially when the measurements are not accurate enough to detect the unspecified model errors.
- **Sequential least squares procedure**. A quite different procedure that has been frequently used for kinematic positioning does not use the dynamic model information at all but determines discrete positions at the measurement epochs (Cannon et al. 1986). In this case, **no assumption on dynamic model is made**; only the measurements at discrete epoch are employed to estimate the state parameters. The model error, therefore, does not affect the estimates of new state parameters. Usually, this method is presented as a sequential least squares algorithm (Schwarz et al. 1989). The current limitation of this approach is that it wastes the good information of the state model when the model accurately describes the dynamic process in cases.
- **Sage adaptive Kalman filtering**. This kind of adaptive filter **evaluates the variance–covariance matrices of the kinematic model error vector and the measurement error vector by windowing method** (Sage and Husa 1969). In the applications, an **innovation-based adaptive Kalman filtering for an integrated INS/GPS is developed by Mohamed and Schwarz** (Mohamed and Schwarz 1999; Wang et al. 2000). The problem is that the algorithm needs to collect the residuals of the measurements or the update series to calculate the underlined variance–covariance matrices; thus it **requires that the measurement dimensions and types at all epochs be the same**.
- **Fading Kalman filtering**. In order to control the influences of prior state errors or kinematic model errors on the present estimated state parameters, the fading filters, using the **fading factors** to restrict the memory length of Kalman filter and to make the most use of present measurements, were developed in the field of statistics as early as the 1960s and 1970s (Fagin 1964; Sorenson and Sacks 1971). We have analysed the basic properties of the fading filter, the abilities in controlling the influences of the kinematic model errors on the state parameter estimates and the problems possibly existing in the practical applications (Yang and Gao 2006c).
- **Robust filter based on min–max robust theory**. The **deviation of observation error distribution from the Gaussian one may also seriously degrade the performance of the Kalman filtering**. Thus, there appears to be considerable motivation for considering filters which are robustised to perform **fairly well in non-Gaussian environments**. Facing this problem, Masreliez and Martin (1977) applied the influence function of min–max **robust theory to replace the score function of the standard Kalman filter**. The basic disadvantages associated with this

kind of robust filter are that the estimator requires the unknown contaminating distribution to be symmetric and it cannot work as well as the standard Kalman filter in Gaussian noise.

- Robust filter based on M estimation theory (Huber 1964) and Bayesian statistics. To resist the bad influences of both state model errors and measurement outliers, a robust M–M filter is developed (Yang 1991, 1993, 1997, 1999; Zhou et al. 1997, p. 299) by which the measurement outliers are controlled by robust equivalent weights of the measurements, and the model errors are resisted by the equivalent weights of the update parameters according to the divergence of the predicted parameters and the estimated ones. Furthermore, a robust filter for rank-deficient observation models was developed by Koch and Yang (1998) by Bayesian statistics and by applying the robust M-estimate.

Different from Sage–Husa adaptive filtering (see Deng 2003, pp. 162–173; Mohamed and Schwarz 1999; Wang and Kubik 1993; Wang et al. 2000) and limited memory filter (Panozzo et al. 2004) as well as the mentioned adaptive filters, a new adaptively robust filter was developed by combining the adaptive Kalman filter and robust estimation (Yang et al. 2001a), which applies a robust estimation principle for measurement vector to resist its outlier effects and introduces an adaptive factor for the model predicted state vector to control its outlying disturbance influences.

After adaptively robust filtering was developed, four learning statistics and four adaptive factors have been set up based on experiences and have been proved effective in practical applications. An accompanying adaptive factor with a three-segment descending function and a learning statistic constructed by using the discrepancy between the predicted state from the kinematic model and the state estimated from the measurements have been established. Three other kinds of adaptive factors have been developed, which are a two-segment descending function (Yang et al. 2001b), an exponential function (Yang and Gao 2005) and a zero and one function for state component adaptation (Ou et al. 2004; Ren et al. 2005). Three additional learning statistics have also been set up, which include a predicted residual statistic (Xu and Yang 2000; Yang and Gao 2006a), a variance component ratio statistic from both the measurements and the predicted states (Yang and Xu 2003) and a velocity discrepancy between the predicted velocity from the kinematic model and the velocity evaluated from the measurements (Cui and Yang 2006).

Two optimal adaptive factors are established, which satisfy the conditions that the theoretical uncertainty of the predicted state outputted from the adaptive filtering equals or nearly equals its actual estimated uncertainty, and/or the theoretical uncertainty of the predicted residual vector equals or nearly equals its actual estimated uncertainty (Yang and Gao 2006b). Furthermore, an adaptively robust filter with classified adaptive factors is developed (Cui and Yang 2006) which is more effective in tracking the disturbances of the vehicle movements. And an adaptively robust filter with multi-adaptive factors is also set up (Yang et al. 2006), which is more general in theory and contains the adaptively robust filters with single adaptive factor and classified adaptive factors.

The new adaptively robust filter is more or less based on the assumption that the measurements at each epoch should be reliable. If it is not true, then the detection and identification and adaptation procedure can be employed (Teunissen 1990) or a robust Kalman filter can be applied (Koch and Yang 1998; Schaffrin 1991, pp. 32–34; Yang 1991, 1997; Zhou et al. 1997).

## 2.2 The Principle of Adaptively Robust Kalman Filtering

Let the linear dynamic system be given by

$$\mathbf{X}_k = \Phi_{k,k-1} \mathbf{X}_{k-1} + \mathbf{W}_k, \quad (1)$$

where  $\mathbf{X}_k$  denotes  $m \times 1$  state vector at epoch  $t_k$ ,  $\Phi_{k,k-1}$  the  $u \times u$  transition matrix and  $\mathbf{W}_k$  the state noise vector. The observational model at epoch  $t_k$  reads

$$\mathbf{L}_k = \mathbf{A}_k \mathbf{X}_k + \mathbf{e}_k, \quad (2)$$

where  $\mathbf{L}_k$  represents  $n_k \times 1$  observation vector,  $\mathbf{A}_k$  the  $n_k \times m$  design matrix and  $\mathbf{e}_k$  the observational noise vector. Let the covariance matrices of  $\mathbf{W}_k$  and  $\mathbf{e}_k$  be taken as  $\Sigma_{\mathbf{W}_k}$  and  $\Sigma_k$ , respectively, and  $\mathbf{W}_k, \mathbf{W}_j, \mathbf{e}_k$  and  $\mathbf{e}_j$  be mutually uncorrelated and meet:

$$E(\mathbf{e}_k) = 0, E(\mathbf{W}_k) = 0, \quad (3)$$

$$E(\mathbf{e}_k \mathbf{e}_k^T) = \Sigma_k, E(\mathbf{W}_k \mathbf{W}_k^T) = \Sigma_{\mathbf{W}_k}. \quad (4)$$

Assume that the residual vector is denoted by  $\mathbf{V}_k$  and the predicted state vector is  $\bar{\mathbf{X}}_k$ ; then the error equation and the predicted state vector are

$$\mathbf{V}_k = \mathbf{L}_k \hat{\mathbf{X}}_k - \mathbf{L}_k, \quad (5)$$

$$\bar{\mathbf{X}}_k = \Phi_{k,k-1} \hat{\mathbf{X}}_{k-1}, \quad (6)$$

with

$$\Sigma_{\bar{\mathbf{X}}_k} = \Phi_{k,k-1} \Sigma_{\hat{\mathbf{X}}_{k-1}} \Phi_{k,k-1}^T + \Sigma_{\mathbf{W}_k}, \quad (7)$$

where  $\hat{\mathbf{X}}_k$  and  $\hat{\mathbf{X}}_{k-1}$  are the estimated state vectors at epochs  $t_k$  and  $t_{k-1}$ .

By using the least squares principle,

$$\mathbf{V}_k^T \mathbf{P}_k \mathbf{V}_k + (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \mathbf{P}_{\bar{\mathbf{X}}_k} (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k) = \min, \quad (8)$$

where  $\mathbf{P}_k = \Sigma_k^{-1}$  and  $\mathbf{P}_{\bar{\mathbf{X}}_k} = \Sigma_{\bar{\mathbf{X}}_k}^{-1}$  are the weight matrices of  $\mathbf{L}_k$  and  $\bar{\mathbf{X}}_k$ , respectively; we obtain the estimator of the standard Kalman filter (Koch and Yang 1998):

$$\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k + \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k). \quad (9)$$

Or equivalently

$$\hat{\mathbf{X}}_k = \bar{\mathbf{X}}_k + \mathbf{K}_k (\mathbf{L}_k - \mathbf{A}_k \bar{\mathbf{X}}_k), \quad (10)$$

with

$$\mathbf{K}_k = \Sigma_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T (\mathbf{A}_k \Sigma_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T + \Sigma_k)^{-1}, \quad (11)$$

$$\Sigma_{\hat{\mathbf{X}}_k} = [\mathbf{I} - \mathbf{K}_k \mathbf{A}_k] \Sigma_{\bar{\mathbf{X}}_k}. \quad (12)$$

Changing a little bit of the score function of (8), like

$$\sum_{i=1}^{n_k} p_i \rho(v_i) + \alpha_k (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k)^T \mathbf{P}_{\bar{\mathbf{X}}_k} (\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k) = \min, \quad (13)$$

where  $\rho$  is a convex and continuous function,  $\alpha_k$  is an adaptive factor with values in  $0 < \alpha_k \leq 1$ , we get the estimator of the adaptive filter (Yang et al. 2001a):

$$\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k + \mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{L}_k), \quad (14)$$

or equivalently (Yang et al. 2001b; Xu 2007)

$$\hat{\mathbf{X}}_k = \bar{\mathbf{X}}_k + \bar{\mathbf{K}}_k (\mathbf{L}_k - \mathbf{A}_k \bar{\mathbf{X}}_k), \quad (15)$$

where  $\bar{\mathbf{K}}_k$  is an adaptive gain matrix:

$$\bar{\mathbf{K}}_k = \frac{1}{\alpha_k} \Sigma_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T \left( \frac{1}{\alpha_k} \mathbf{A}_k \Sigma_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T + \bar{\Sigma}_k \right)^{-1}. \quad (16)$$

The posterior covariance matrix of the estimated state vector is

$$\Sigma_{\hat{\mathbf{X}}_k} = (\mathbf{I} - \bar{\mathbf{K}}_k \mathbf{A}_k) \Sigma_{\bar{\mathbf{X}}_k} / \alpha_k. \quad (17)$$

With the variations of adaptive factor  $\alpha_k$  and the equivalent weight matrix, the adaptively robust filter will change into various estimators.

Case 1: if  $\alpha_k = 0$  and  $\Sigma_k = \Sigma_k$  or  $\mathbf{P}_k = \mathbf{P}_k$ , then

$$\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k, \quad (18)$$

which is an LS estimator by using only the new measurements at epoch  $t_k$ . This estimator is suitable to the case that the measurements are not contaminated by outliers and the updated parameters are biased so much that the information of updated parameters should be forgotten completely.

Case 2: if  $\alpha_k = 1$  and  $\hat{\Sigma}_k = \Sigma_k$ , then the standard Kalman filter estimators (8), (9) and (10) are obtained.

Case 3: if  $\alpha_k$  changes between 0 and 1, and  $\hat{P}_k = P_k$ , then

$$\hat{X}_k = (A_k^T P_k A_k + \alpha_k P_{\bar{X}_k})^{-1} (A_k^T P_k L_k + \alpha_k P_{\bar{X}_k} \bar{X}_k), \quad (19)$$

which is an adaptive LS estimator of Kalman filter. It balances the contribution of the updated parameters and the measurements. The only difference between (17) and (8) is the weight matrix of  $L_k$ . The former uses the equivalent weight matrix and the latter uses the original weight matrix of  $L_k$ .

Case 4: if  $\alpha_k = 0$ , then we obtain

$$\hat{X}_k = (A_k^T \bar{P}_k A_k)^{-1} A_k^T \bar{P}_k L_k, \quad (20)$$

which is a robust estimator by using only the new measurements at epoch  $t_k$ .

Case 5: if  $\alpha_k = 1$ , then

$$\hat{X}_k = (A_k^T \bar{P}_k A_k + P_{\bar{X}_k})^{-1} (A_k^T \bar{P}_k L_k + P_{\bar{X}_k} \bar{X}_k), \quad (21)$$

which is an M-LS filter estimator (Yang 1997).

The relations of the adaptively robust filter with other estimators are shown in Fig. 2.1, in which ARF denotes the adaptively robust filter.

If the covariance matrices of the measurement vector  $L_k$  and the predicted state vector  $\bar{X}_k$  are evaluated by Sage windowing method (see Deng 2003), denoted as  $\hat{\Sigma}_k$  and  $\hat{\Sigma}_{\bar{X}_k}$ , respectively, that is,

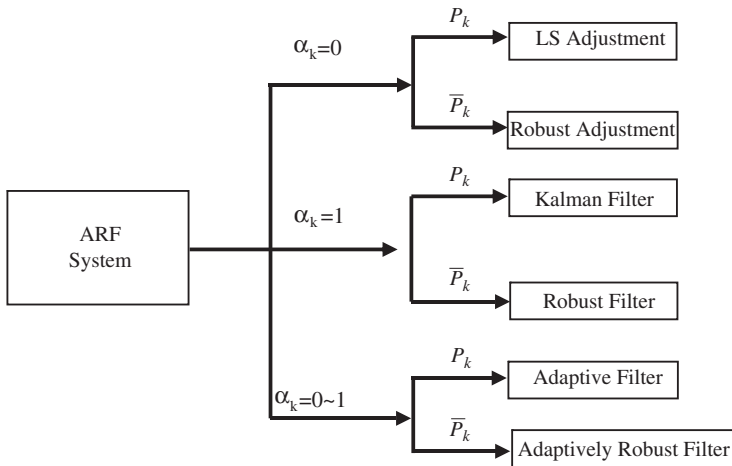


Fig. 2.1 Adaptively robust filter

$$\hat{\Sigma}_k = \frac{1}{m} \sum_{j=0}^m \bar{\mathbf{V}}_{k-j} \bar{\mathbf{V}}_{k-j}^T - \mathbf{A}_k \Sigma_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T, \quad (22)$$

$$\hat{\Sigma}_{\Delta X} = \frac{1}{m} \sum_{j=0}^m \Delta \mathbf{X}_{k-j} \Delta \mathbf{X}_{k-j}^T, \quad (23)$$

where

$$\bar{\mathbf{V}}_k = \mathbf{A}_k \bar{\mathbf{X}}_k - \mathbf{L}_k, \quad (24)$$

$$\Delta \mathbf{X}_k = \hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k, \quad (25)$$

then the adaptively robust filter can include the adaptive Sage filter.

## 2.3 Properties of the Adaptive Kalman Filter

### 2.3.1 Difference of State Estimate

Rewriting (9) and (14) as normal equations, and only considering the least squares situation, that is,  $\bar{\mathbf{P}}_k = \mathbf{P}_k$ , we have

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k}) \hat{\mathbf{X}}_k = \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k + \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k, \quad (26)$$

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \hat{\mathbf{X}}_{\text{ad}} = \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k + \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k, \quad (27)$$

where  $\hat{\mathbf{X}}_k$  and  $\hat{\mathbf{X}}_{\text{ad}}$  denote the state estimates by using standard Kalman filter and adaptive filter, respectively. Let

$$\hat{\mathbf{X}}_k - \hat{\mathbf{X}}_{\text{ad}} = \delta \hat{\mathbf{X}}_k, \quad (28)$$

then (27) is changed into

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \hat{\mathbf{X}}_k - (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \delta \hat{\mathbf{X}}_k = \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k + \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k. \quad (29)$$

Subtracting (29) from (26), we have

$$(1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} \hat{\mathbf{X}}_k + (\mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \delta \hat{\mathbf{X}}_k = (1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} \bar{\mathbf{X}}_k \quad (30)$$

or

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \delta \hat{\mathbf{X}}_k = (1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} (\bar{\mathbf{X}}_k - \hat{\mathbf{X}}_k). \quad (31)$$

Denote

$$\bar{\mathbf{X}}_k - \hat{\mathbf{X}}_{\text{ad}} = \delta \bar{\mathbf{X}}_k. \quad (32)$$



Factually,  $\delta\bar{\mathbf{X}}_k$  can be looked as the bias of the predicted state, and (31) is changed into

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \delta\hat{\mathbf{X}}_k = (1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} (\delta\bar{\mathbf{X}}_k + \hat{\mathbf{X}}_{\text{ad}} - \hat{\mathbf{X}}_k). \quad (33)$$

Finally we have

$$(\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k}) \delta\hat{\mathbf{X}}_k = (1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} \delta\bar{\mathbf{X}}_k, \quad (34)$$

$$\delta\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (1 - \alpha_k) \mathbf{P}_{\bar{\mathbf{X}}_k} \delta\bar{\mathbf{X}}_k. \quad (35)$$

It is easy to see that if  $\alpha_k = 1$ , then  $\delta\hat{\mathbf{X}}_k = 0$ , that is,  $\hat{\mathbf{X}}_{\text{ad}} = \hat{\mathbf{X}}_k$ ; in this case the state estimate of the adaptive filter is equivalent to that of the standard Kalman filter; if  $\alpha_k = 0$ , then

$$\hat{\mathbf{X}}_{\text{ad}} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k)^{-1} \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k. \quad (36)$$

In this case,  $\hat{\mathbf{X}}_{\text{ad}}$  is equivalent to the estimate of the least squares estimation not considering the information of the state equation, and the error of the predicted state  $\bar{\mathbf{X}}_k$  will not affect the updated estimate of the state.

### 2.3.2 The Expectation of the State Estimate of the Adaptive Filter

Considering that the observational vector is unbiased, that is  $\mathbf{E}\mathbf{L}_k = \mathbf{A}_k \mathbf{X}_k$ , then the expectation of  $\hat{\mathbf{X}}_{\text{ad}}$  from (27) is

$$\mathbf{E}\hat{\mathbf{X}}_{\text{ad}} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{E}\bar{\mathbf{X}}_k + \mathbf{A}_k^T \mathbf{P}_k \mathbf{L}_k \mathbf{X}_k). \quad (37)$$

If the predicted state vector  $\bar{\mathbf{X}}_k$  is also unbiased, that is  $\mathbf{E}\bar{\mathbf{X}}_k = \mathbf{X}_k$ , then

$$\mathbf{E}\hat{\mathbf{X}}_{\text{ad}} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k}) \mathbf{X}_k = \mathbf{X}_k. \quad (38)$$

It is obvious that if the observational vector  $\mathbf{L}_k$  and the predicted state vector  $\bar{\mathbf{X}}_k$  are not affected by abnormal biases, then the estimate of the adaptive filter is unbiased.

If the predicted state vector  $\bar{\mathbf{X}}_k$  is biased, and the biased vector is denoted by  $\mathbf{b}_{\bar{\mathbf{X}}_k}$ , that is,

$$\mathbf{E}\bar{\mathbf{X}}_k = \mathbf{X}_k + \mathbf{b}_{\bar{\mathbf{X}}_k} \neq \mathbf{X}_k, \quad (39)$$

then

$$\mathbf{E}\hat{\mathbf{X}}_{\text{ad}} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{X}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k} + \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k \mathbf{X}_k), \quad (40)$$

then we have

$$E\hat{\mathbf{X}}_{\text{ad}} = \mathbf{X}_k + (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k} \quad (41)$$

or

$$\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}} = E\hat{\mathbf{X}}_{\text{ad}} - \mathbf{X}_k = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k}. \quad (42)$$

Equation (42) gives the influence of the bias  $\mathbf{b}_{\bar{\mathbf{X}}_k}$  of  $\bar{\mathbf{X}}_k$  on the bias  $\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}}$  of the estimate expectation of the adaptive filter. Equation (42) also tells us that  $\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}}$  changes with the variations of  $\alpha_k$  and  $\mathbf{b}_{\bar{\mathbf{X}}_k}$ , especially when  $\mathbf{b}_{\bar{\mathbf{X}}_k}$  is beyond a particular region;  $\alpha_k$  tends to zero and, in this case,  $\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}}$  also tends to zero. In other words, the bias of the state estimate of the adaptive filter is controlled by the adaptive factor  $\alpha_k$ . Usually the larger the bias  $\mathbf{b}_{\bar{\mathbf{X}}_k}$  of the predicted state, the smaller the adaptive factor  $\alpha_k$ .

Similarly, when  $E\bar{\mathbf{X}}_k \neq \mathbf{X}_k$ , the expectation of the state estimate of the standard Kalman filter is

$$E\hat{\mathbf{X}}_k = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{X}_k + \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k} + \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k \mathbf{X}_k) \quad (43)$$

or

$$E\hat{\mathbf{X}}_k = \mathbf{X}_k + (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k}. \quad (44)$$

Let

$$\mathbf{b}_{\hat{\mathbf{X}}_k} = E\hat{\mathbf{X}}_k - \mathbf{X}_k, \quad (\text{bias of } \hat{\mathbf{X}}_k). \quad (45)$$

Then

$$\mathbf{b}_{\hat{\mathbf{X}}_k} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{b}_{\bar{\mathbf{X}}_k}. \quad (46)$$

### 2.3.3 Posterior Precision Evaluation

By applying the variance propagation law we obtain the covariance matrices of the estimated state vectors of the adaptive filter and the standard Kalman filter, respectively, as

$$\Sigma_{\hat{\mathbf{X}}_{\text{ad}}} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} (\alpha_k^2 \mathbf{P}_{\bar{\mathbf{X}}_k} + \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k) (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \hat{\sigma}_{\text{ad}}^2, \quad (47)$$

$$\Sigma_{\hat{\mathbf{X}}_k} = (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k})^{-1} \hat{\sigma}^2, \quad (48)$$

in which

$$\hat{\sigma}_{\text{ad}}^2 = \frac{\alpha_k \delta \hat{\mathbf{X}}_{\text{ad}}^T \mathbf{P}_{\bar{\mathbf{X}}_k} \delta \hat{\mathbf{X}}_{\text{ad}} + \mathbf{V}_{\text{ad}}^T \mathbf{P}_k \mathbf{V}_{\text{ad}}}{r_k}, \quad (49)$$

$$\hat{\sigma}^2 = \frac{\delta \hat{\mathbf{X}}_k^T \mathbf{P}_{\bar{\mathbf{X}}_k} \delta \hat{\mathbf{X}}_k + \mathbf{V}_k^T \mathbf{P}_k \mathbf{V}_k}{r_k}, \quad (50)$$

where  $r_k$  is redundant number of the observations,  $\mathbf{V}_{\text{ad}}$  and  $\mathbf{V}_k$  are residual vectors of the observations with respect to  $\hat{\mathbf{X}}_{\text{ad}}$  and  $\hat{\mathbf{X}}_k$ , respectively.

In practice, the posterior precision is evaluated by the mean-square error (Xu and Rummel 1994), that is,

$$\text{MSE}(\hat{\mathbf{X}}_{\text{ad}}) = \text{E}(\hat{\mathbf{X}}_{\text{ad}} - \mathbf{X}_k)^T (\hat{\mathbf{X}}_{\text{ad}} - \mathbf{X}_k), \quad (51)$$

where  $\mathbf{X}_k$  denotes the true value of the state vector. Changing (51) as

$$\text{MSE}(\hat{\mathbf{X}}_{\text{ad}}) = \text{E}(\hat{\mathbf{X}}_{\text{ad}} - \text{E}\hat{\mathbf{X}}_{\text{ad}} + \text{E}\hat{\mathbf{X}}_{\text{ad}} - \mathbf{X}_k)^T (\hat{\mathbf{X}}_{\text{ad}} - \text{E}\hat{\mathbf{X}}_{\text{ad}} + \text{E}\hat{\mathbf{X}}_{\text{ad}} - \mathbf{X}_k), \quad (52)$$

letting

$$\hat{\mathbf{X}}_{\text{ad}} - \text{E}\hat{\mathbf{X}}_{\text{ad}} = \mathbf{e}_{\hat{\mathbf{X}}_{\text{ad}}} \left( \text{true error vector of } \hat{\mathbf{X}}_{\text{ad}} \right) \quad (53)$$

and considering

$$\text{E}\mathbf{X}_k = \mathbf{X}_k, \text{E}(\hat{\mathbf{X}}_{\text{ad}} - \text{E}\hat{\mathbf{X}}_{\text{ad}}) = 0, \quad (54)$$

we get

$$\text{MSE}(\hat{\mathbf{X}}_{\text{ad}}) = \text{tr}(\mathbf{e}_{\hat{\mathbf{X}}_{\text{ad}}} \cdot \mathbf{e}_{\hat{\mathbf{X}}_{\text{ad}}}^T) + \left\| \mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}} \right\|^2. \quad (55)$$

On the other hand, we have

$$\text{E}(\mathbf{e}_{\hat{\mathbf{X}}_{\text{ad}}} \cdot \mathbf{e}_{\hat{\mathbf{X}}_{\text{ad}}}^T) = \sum_{\hat{\mathbf{X}}_{\text{ad}}} (\text{covariance matrix of } \hat{\mathbf{X}}_{\text{ad}}), \quad (56)$$

which then yields

$$\text{MSE}(\hat{\mathbf{X}}_{\text{ad}}) = \text{tr} \Sigma_{\hat{\mathbf{X}}_{\text{ad}}} + \alpha_k \left\| (\mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \alpha_k P_{\bar{\mathbf{X}}_k})^{-1} \mathbf{P}_{\bar{\mathbf{X}}_k} (\text{E}\bar{\mathbf{X}}_k - \mathbf{X}_k) \right\|^2. \quad (57)$$

It is seen from (57) that

1. If the predicted state vector  $\bar{\mathbf{X}}_k$  is unbiased, that is  $\mathbf{b}_{\mathbf{X}_k} = 0$ , then  $\hat{\mathbf{X}}_{\text{ad}}$  is unbiased, or  $\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}} = 0$ ; in this case, the MSE of  $\hat{\mathbf{X}}_{\text{ad}}$  is the trace of its covariance matrix,

$$\text{MSE}(\hat{\mathbf{X}}_{\text{ad}}) = \text{tr}(\Sigma_{\hat{\mathbf{X}}_{\text{ad}}}). \quad (58)$$

2. If  $\bar{\mathbf{X}}_k$  is biased, but  $\alpha_k = 0$ , (57) is still valid, that is the adaptive factor  $\alpha_k$  controls the bias of the outputs of the adaptive filter.

Therefore, when the predicted state  $\bar{\mathbf{X}}_k$  has any abnormal bias due to some sudden disturbance of the vehicle, the adaptive factor  $\alpha_k$  will be decreased, which leads to the bias  $\mathbf{b}_{\hat{\mathbf{X}}_{\text{ad}}}$  of the state estimate of the adaptive filter to decrease, and the mean-square error of  $\hat{\mathbf{X}}_{\text{ad}}$  tends to the trace of the covariance matrix of  $\hat{\mathbf{X}}_{\text{ad}}$ .

In conclusion, the differences between the adaptive filter and the standard Kalman filter depend on the adaptive factor  $\alpha_k$ . When the predicted states are accurate, then  $\alpha_k$  tends to 1, and the state differences estimated from the adaptive filter and the standard Kalman filter are small.

The unbiasedness of the estimated state vector outputted by adaptively filter is controlled by the adaptive factor  $\alpha_k$ ; if  $\alpha_k$  tends to zero,  $\hat{\mathbf{X}}_{\text{ad}}$  is unbiased.

The MSE of the estimated state vector outputted by the adaptive filter is also controlled by the adaptive factor  $\alpha_k$ ; if  $\alpha_k$  tends to zero, the MSE of  $\hat{\mathbf{X}}_{\text{ad}}$  tends to the trace of its covariance matrix.

The robustness of the adaptive filter outputs has been described in Yang et al. (2001a, b) and Yang and Xu (2003). It has been demonstrated by theory and practical experiments that the adaptive factor plays significant roles in controlling the influences of the outlying disturbances of the dynamical information on the estimated state vector and its MSE.

## 2.4 Three Kinds of Learning Statistics

### 2.4.1 Learning Statistic Constructed by State Discrepancy

In the beginning of the development of the adaptively robust Kalman filter, a learning statistic of the kinematic model errors was constructed by using the difference between the state estimated from measurements and that predicted from the kinematic model at epoch  $t_k$  (Yang et al. 2001a, b). If the number of measurements at epoch  $t_k$  is larger than that of the state components, then the estimated state vector is obtained by using measurement vector  $\mathbf{L}_k$ , based on the robust estimation principle, that is,

$$\tilde{\mathbf{X}}_k = (\mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{A}_k)^{-1} \mathbf{A}_k^T \bar{\mathbf{P}}_k \mathbf{L}_k, \quad (59)$$

where  $\bar{\mathbf{P}}_k$  denotes the equivalent weight matrix of  $\mathbf{L}_k$ , which can be obtained by the Huber function (Huber 1981), three-segment functions (Yang 1994, 1999; Yang et al. 2002a, b; Zhou 1989), etc.

The discrepancy between  $\tilde{\mathbf{X}}_k$  from (59) and  $\bar{\mathbf{X}}_k$  from (6) can be measured by

$$\|\tilde{\mathbf{X}}_k - \bar{\mathbf{X}}_k\| = \left( \Delta \tilde{\mathbf{X}}_{k_1}^2 + \Delta \tilde{\mathbf{X}}_{k_2}^2 + \cdots + \Delta \tilde{\mathbf{X}}_{k_m}^2 \right)^{\frac{1}{2}}. \quad (60)$$

Then the learning statistic expressed by the state discrepancy is set up:

$$|\Delta \tilde{\mathbf{X}}_k| = \frac{\|\tilde{\mathbf{X}}_k - \bar{\mathbf{X}}_k\|}{\sqrt{\text{tr}(\Sigma_{\bar{\mathbf{X}}_k})}}, \quad (61)$$

where “tr” denotes trace.

It is noted that (1) the number of measurements at computation epoch should be larger than the number of state components, otherwise the statistic  $|\Delta \tilde{\mathbf{X}}_k|$  cannot be constructed; (2) the estimated state vector  $\tilde{\mathbf{X}}_k$  should be accurate, otherwise the statistic  $|\Delta \tilde{\mathbf{X}}_k|$  cannot reflect the kinematic model errors; and (3) the learning statistic  $|\Delta \tilde{\mathbf{X}}_k|$  can only reflect the integrated error of the kinematic model; any disturbing of the components of the predicted state vector is treated as the whole state outlier.

#### 2.4.2 Learning Statistic Constructed by Predicted Residual Vector

If the measurement vector  $\mathbf{L}_k$  is reliable, then the predicted residual vector  $\bar{\mathbf{V}}_k$  will reflect the error of predicted state vector  $\bar{\mathbf{X}}_k$ . A learning statistic constructed by the predicted residual vector is (Xu and Yang 2000; Yang and Gao 2006a)

$$\Delta \bar{\mathbf{V}}_k = \left( \frac{\bar{\mathbf{V}}_k^T \bar{\mathbf{V}}_k}{\text{tr}(\Sigma_{\bar{\mathbf{V}}_k})} \right)^{\frac{1}{2}}. \quad (62)$$

If there are  $n_k$  measurements at epoch  $t_k$ , then  $\Delta \bar{\mathbf{V}}_k$  can be expressed as

$$\Delta \bar{\mathbf{V}}_k = \left( \frac{\sum_{i=1}^{n_k} \bar{\mathbf{V}}_{k_i}^2}{\sum_{i=1}^{n_k} \sigma_{\bar{\mathbf{V}}_{k_i}}^2} \right)^{\frac{1}{2}}. \quad (63)$$

It is noted that (1) using the learning statistic constructed by predicted residual  $\Delta \bar{\mathbf{V}}_k$ , we do not need to evaluate the state vector before filtering; (2) it is not necessary that the number of measurements be larger than that of state components; and (3)  $\Delta \bar{\mathbf{V}}_k$  contains more measurement error influence than the state discrepancy statistic  $|\Delta \tilde{\mathbf{X}}_k|$ .

### 2.4.3 Learning Statistic Constructed by the Ratio of Variance Components

If  $\mathbf{L}_k$  and  $\bar{\mathbf{X}}_k$  are treated as two groups of observations at epoch  $t_k$ , their variance components should reflect their accuracies. Thus we can construct a new learning statistic by the ratio of the variance components to judge the kinematic model reliability. The Helmert variance component for  $\mathbf{L}_k$  and  $\bar{\mathbf{X}}_k$  is respectively expressed as (Koch 2000; Koch and Kusche 2002)

$$\hat{\sigma}_{0k}^2 = \frac{\mathbf{V}_k^T \mathbf{P}_k \mathbf{V}_k}{r_k - \text{tr}(\mathbf{N}^{-1} \mathbf{N}_k)}, \quad (64)$$

$$\hat{\sigma}_{0\bar{\mathbf{X}}_k}^2 = \frac{\mathbf{V}_{\bar{\mathbf{X}}_k}^T \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{V}_{\bar{\mathbf{X}}_k}}{m_k - \text{tr}(\mathbf{N}^{-1} \mathbf{P}_{\bar{\mathbf{X}}_k})}, \quad (65)$$

where  $\hat{\sigma}_{0k}^2$  and  $\hat{\sigma}_{0\bar{\mathbf{X}}_k}^2$  denote the estimates of variance components of  $\mathbf{L}_k$  and  $\bar{\mathbf{X}}_k$ , respectively,  $n_k$  is the number of measurements at epoch  $t_k$ ,  $m_k$  is the number of predicted parameters of the state vector,  $\mathbf{V}_k$  is the residual vector of  $\mathbf{L}_k$  expressed by (5) and  $\mathbf{V}_{\bar{\mathbf{X}}_k}$  is the residual vector of  $\bar{\mathbf{X}}_k$ , that is,

$$\mathbf{V}_{\bar{\mathbf{X}}_k} = \hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k = \hat{\mathbf{X}}_k - \Phi_{k,k-1} \hat{\mathbf{X}}_{k-1} \quad (66)$$

and

$$\mathbf{N}_k = \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k, \mathbf{N} = \mathbf{N}_k + \mathbf{P}_{\bar{\mathbf{X}}_k} = \mathbf{A}_k^T \mathbf{P}_k \mathbf{A}_k + \mathbf{P}_{\bar{\mathbf{X}}_k}. \quad (67)$$

The approximate estimates of the Helmert variance components  $\hat{\sigma}_{0k}^2$  and  $\hat{\sigma}_{0\bar{\mathbf{X}}_k}^2$  are

$$\sigma_{0k}^2 \approx \frac{\mathbf{V}_k^T \mathbf{P}_k \mathbf{V}_k}{n_k} \quad (68)$$

and

$$\sigma_{0\bar{\mathbf{X}}_k}^2 \approx \frac{\mathbf{V}_{\bar{\mathbf{X}}_k}^T \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{V}_{\bar{\mathbf{X}}_k}}{m_k}. \quad (69)$$

The ratio of  $\hat{\sigma}_{0k}^2$  and  $\hat{\sigma}_{0\bar{\mathbf{X}}_k}^2$  is defined as the learning statistic

$$S_k = \frac{\sigma_{0\bar{\mathbf{X}}_k}^2}{\sigma_{0k}^2} \approx \frac{\mathbf{V}_{\bar{\mathbf{X}}_k}^T \mathbf{P}_{\bar{\mathbf{X}}_k} \mathbf{V}_{\bar{\mathbf{X}}_k}}{m_k \sigma_{0k}^2}. \quad (70)$$

It is noted that (1) the computation of the learning statistic  $S_k$  should have redundant observations, otherwise the learning statistic will not reliably reflect the model

errors; (2)  $\mathbf{V}_k$  and  $\mathbf{V}_{\tilde{\mathbf{X}}_k}$  should correspond to the same estimated state vector  $\hat{\mathbf{X}}_k$ ; and (3) the computation burden is heavier than the other two learning statistics,  $|\Delta\tilde{\mathbf{X}}_k|$  and  $\Delta\tilde{\mathbf{V}}_k$ , if the iterative computation procedures are applied.

#### 2.4.4 Learning Statistic Constructed by Velocity

Based on the robust estimate  $\tilde{\mathbf{X}}_k$  of the position parameter vector from the measurements, the estimate  $\hat{\mathbf{X}}_{k-1}$  of the state estimate at epoch  $t_{k-1}$  and the sample interval  $t_k - t_{k-1}$ , we obtain the predicted velocity vector (Cui and Yang 2006)

$$\tilde{\dot{\mathbf{X}}}_k = \frac{\tilde{\mathbf{X}}_k - \hat{\mathbf{X}}_{k-1}}{t_k - t_{k-1}}. \quad (71)$$

Then the learning statistic for judging the kinematic model disturbing corresponding to the predicted velocity information is constructed as

$$\Delta\tilde{\dot{\mathbf{X}}}_k = \left\| \tilde{\dot{\mathbf{X}}}_k - \bar{\dot{\mathbf{X}}}_k \right\| / \sqrt{\text{tr}(\Sigma_{\tilde{\dot{\mathbf{X}}}_k})}, \quad (72)$$

where  $\bar{\dot{\mathbf{X}}}_k$  denotes the predicted velocity vector from the kinematic model and  $\Sigma_{\tilde{\dot{\mathbf{X}}}_k}$  is its covariance matrix.

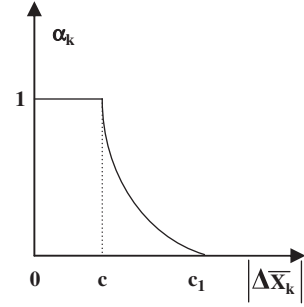
It is noted that (1) if  $\Delta\tilde{\dot{\mathbf{X}}}_k$  is significant outlying, then it indicates that the predicted velocity is outlying or that the kinematic model has large errors and (2) the computation of the learning statistic  $\Delta\tilde{\dot{\mathbf{X}}}_k$  should also have redundant observations, otherwise  $\tilde{\dot{\mathbf{X}}}_k$  cannot be obtained.

## 2.5 Four Kinds of Adaptive Factors

### 2.5.1 Adaptive Factor by Three-Segment Function

An adaptive factor  $\alpha_k$  of a three-segment function is combined by three parts, that is, if a learning statistic is smaller than a particular criterion, then the adaptive factor  $\alpha_k$  is equal to 1, if the learning statistic is significantly outlying, then the adaptive factor  $\alpha_k$  is equal to 0, otherwise  $\alpha_k$  decreases with the statistic growing. We employ the learning statistic  $|\Delta\tilde{\dot{\mathbf{X}}}_k|$  as an example to express the adaptive factor (Yang et al. 2001a):

**Fig. 2.2** Adaptive factor of three-segment function



$$\alpha_k = \begin{cases} 1, & |\Delta \tilde{X}_k| \leq c_0, \\ \frac{c_0}{|\Delta \tilde{X}_k|} \left( \frac{c_1 - |\Delta \tilde{X}_k|}{c_1 - c_0} \right)^2, & c_0 < |\Delta \tilde{X}_k| \leq c_1, \\ 0, & |\Delta \tilde{X}_k| > c_1, \end{cases} \quad (73)$$

where  $c_0$  and  $c_1$  are two criterion constants, with practical values of  $c_0 = 1.0 - 1.5$  and  $c_1 = 3.0 - 4.5$ .

Obviously, if the value of  $|\Delta \tilde{X}_k|$  increases,  $\alpha_k$  decreases.  $\alpha_k$  changes between  $[0, 1]$  (see Fig. 2.2). This kind of adaptive factor is a redescending function, because  $\alpha_k$  descends to zero when the statistic  $|\Delta \tilde{X}_k|$  is larger than the rejection boundary  $c_1$ .

### 2.5.2 Adaptive Factor by Two-Segment Function

We still employ the learning statistic  $|\Delta \tilde{X}_k|$  as an example to express the two-segment adaptive factor (Yang et al. 2001b):

$$\alpha_k = \begin{cases} 1, & |\Delta \tilde{X}_k| \leq c, \\ \frac{c}{|\Delta \tilde{X}_k|}, & |\Delta \tilde{X}_k| > c, \end{cases} \quad (74)$$

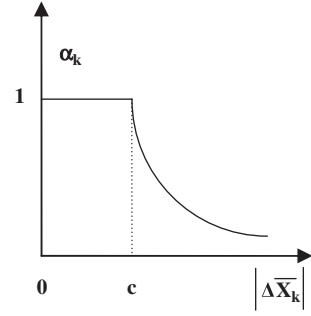
where  $c$  is a constant, the optimal value being 1.0 (Yang and Gao 2006a). It is a descending function; see Fig. 2.3.

### 2.5.3 Adaptive Factor by Exponential Function

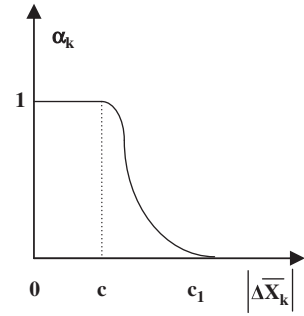
An adaptive factor of exponential function is (Yang and Gao 2005)



**Fig. 2.3** Adaptive factor of two-segment function



**Fig. 2.4** Adaptive factor of exponential function



$$\alpha_k = \begin{cases} 1, & \left| \Delta \tilde{\mathbf{X}}_k \right| \leq c, \\ e^{-(|\Delta \tilde{\mathbf{X}}_k| - c)^2}, & \left| \Delta \tilde{\mathbf{X}}_k \right| > c, \end{cases} \quad (75)$$

where  $c$  is the same as in (74). It is also a descending function; see Fig. 2.4.

### 2.5.4 Adaptive Factor by Zero and One

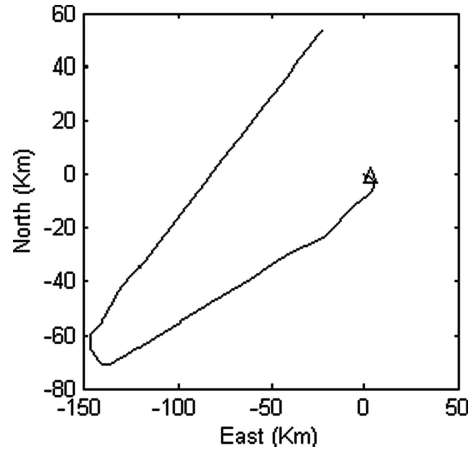
If a state parameter is normal then the adaptive factor equals 1, otherwise it equals 0 (Ou et al. 2004; Ren et al. 2005):

$$\alpha_k = \begin{cases} 1, & \left| \Delta \tilde{\mathbf{X}}_{ki} \right| \leq c, \\ 0, & \left| \Delta \tilde{\mathbf{X}}_{ki} \right| > c, \end{cases} \quad (76)$$

where  $\Delta \tilde{\mathbf{X}}_{ki}$  is the  $i$ th component of the discrepancy state vector.

We can use another two learning statistics,  $\Delta \tilde{\mathbf{V}}_k$  from the predicted residual vector and  $S_k$  from the ratio of variance components of the predicted state and measurements, to construct the same kinds of adaptive factors.

**Fig. 2.5** Position relative to fixed receiver



### 2.5.5 Actual Computation and Analysis

The computation example is the same as that in Yang et al. (2001a). A data set was collected by using a Trimble 4000SSE receiver mounted on an aircraft with the reference receiver fixed at a site about 1 km from the initial aircraft location. After about 10 min of static tracking, the aircraft took off for a flight time of about 90 min; see Fig. 2.5.

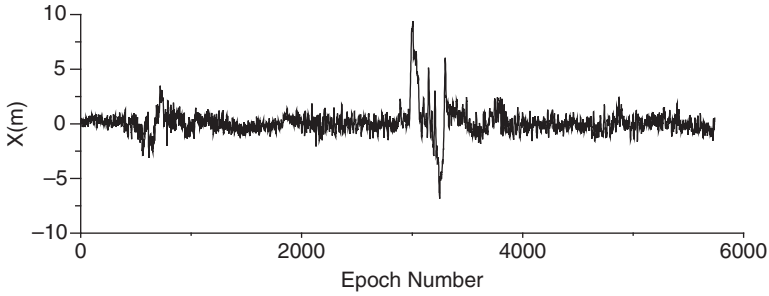
In order to analyse the roles of the adaptive factors in adaptive filtering, the highly precise results from double-differenced carrier measurements were used as “true values” for comparing with the results from the code measurements. The constant-velocity model of the Kalman filtering was employed. The initial variances for positions, velocities and P2-code measurements were selected separately as  $0.2\text{m}^2$ ,  $9 \times 10^{-5}\text{m}^2/\text{s}^2$  and  $1.0\text{m}^2$ . The spectral density for velocities was chosen to be  $0.01\text{m}^2/\text{s}^2$ . The selected dynamic model covariance matrix was the same as that used in Schwarz et al. (1989), Yang et al. (2001a), Yang and Xu (2003) and Yang and Wen (2003).

The following four schemes were carried out:

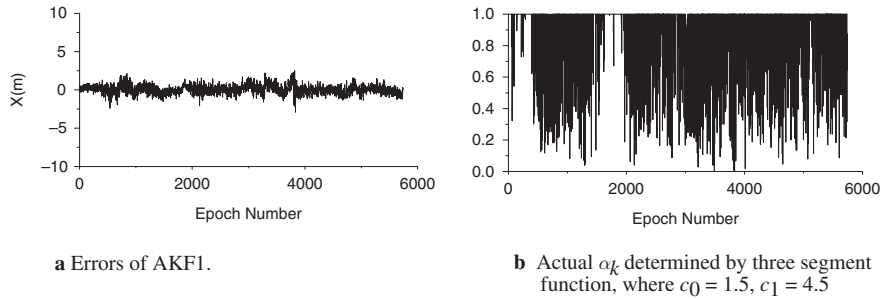
- Scheme 1: Classical Kalman filtering (KF)
- Scheme 2: Adaptive Kalman filtering based on the three-segment function of the state discrepancy (AKF1)
- Scheme 3: Adaptive Kalman filtering based on the two-segment function of the state discrepancy (AKF2)
- Scheme 4: Adaptive Kalman filtering based on the exponential function of the state discrepancy (AKF3)

The computation results are shown in Figs. 2.6, 2.7, 2.8 and 2.9 and Table 2.1.

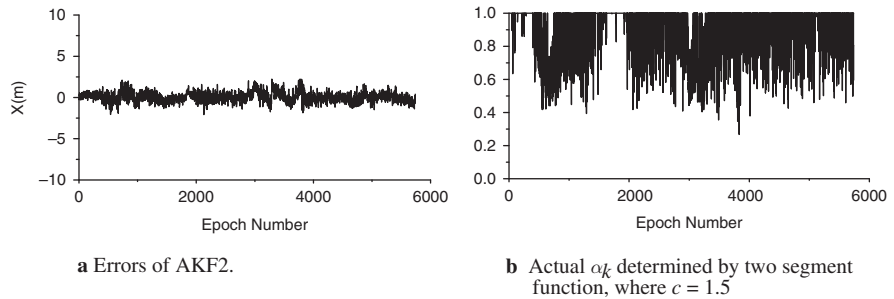
From the calculation results above, we find that



**Fig. 2.6** Errors of classical Kalman filtering



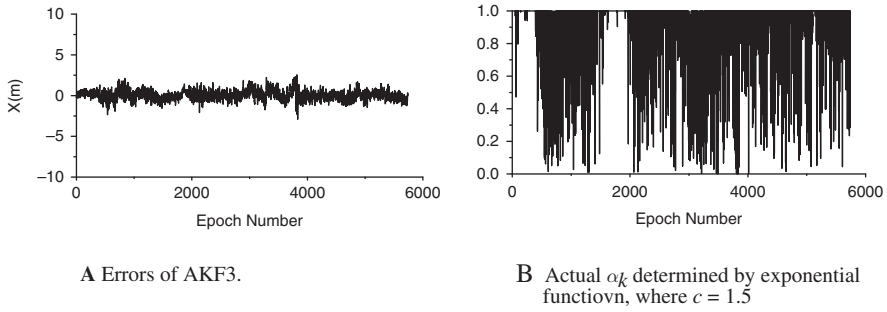
**Fig. 2.7** (a) Errors of AKF1. (b) Actual  $\alpha_k$  determined by three-segment function, where  $c_0 = 1.5$  and  $c_1 = 4.5$



**Fig. 2.8** (a) Errors of AKF2. (b) Actual  $\alpha_k$  determined by two-segment function, where  $c = 1.5$

**Table 2.1** Comparison of RMS (unit: m)

	KF	AKF1	AKF2	AKF3
X	1.1630	0.5648	0.5948	0.5839
Y	1.5070	0.4438	0.5119	0.4766
Z	1.5455	0.7804	0.8201	0.8028



**Fig. 2.9** (a) Errors of AKF3. (b) Actual  $\alpha_k$  determined by exponential function, where  $c = 1.5$

- (1) The influences of the disturbances of the fly around the epoch 1,000 and during the epoch 3,000 through 4,000 on the Kalman filtering are very significant; see Fig. 2.6. The RMS of the position are 1.1630, 1.5070 and 1.5455 m, respectively.
- (2) The adaptive filtering based on the three kinds of adaptive factors gives reasonable results, and the influences of the disturbances of the fly are controlled.
- (3) It is shown, by the theoretical and practical charting curves of the adaptive factors, that the result corresponding to the three-segment function is superior to those corresponding to the two-segment function and the exponential function. The reason is that the three-segment function decreases the adaptive factor quickly when the errors of the predicted state increase and gives the significant outlying predicted state zero factor; this kind of outlying predicted state from the kinematic model does not have any effect on the filtering results; see the second, third and fourth columns of Table 2.1 (AKF1, AKF2 and AKF3), respectively.

## 2.6 Comparison of Two Fading Filters and Adaptively Robust Filter

In order to control the influences of prior state errors or kinematic model errors on the present estimated state parameters, the fading filters, using the fading factors to restrict the memory length of Kalman filter and to make the most use of present measurements, were developed in the field of statistics as early as the 1960s and 1970s (Fagin 1964; Sorenson and Sacks 1971). The basic properties of the fading filter have been analysed, and the abilities in controlling the influences of the kinematic model errors on the state parameter estimates and the possibly existing problems in the practical applications have been discussed, respectively (Yang and Gao 2006c).

### 2.6.1 Principles of Two Kinds of Fading Filters

Assume that the estimated state vector and the residual vector are, respectively,  $\hat{\mathbf{X}}_{k-1}$  and  $\mathbf{V}_{\hat{\mathbf{X}}_{k-1}}$  at epoch  $t_{k-1}$ ; the corresponding re-estimated state vector of  $\mathbf{X}_{k-1}$  by using new measurement vector  $\mathbf{L}_k$  at epoch  $t_k$  is denoted by  $\hat{\hat{\mathbf{X}}}_{k-1}$ ; then the corresponding error equations are, respectively,

$$\mathbf{V}_{\hat{\mathbf{X}}_{k-1}} = \hat{\hat{\mathbf{X}}}_{k-1} - \hat{\mathbf{X}}_{k-1}, \quad (77)$$

$$\hat{\mathbf{w}}_{k-1} = \hat{\mathbf{X}}_k - \Phi_{k,k-1} \hat{\hat{\mathbf{X}}}_{k-1}. \quad (78)$$

The fading filtering result is the same as (10).

If the covariance matrix of the estimated state vector  $\hat{\mathbf{X}}_{k-1}$  at epoch  $t_{k-1}$  is inflated, and the  $\hat{\mathbf{X}}_{k-1}$  is treated as a stochastic vector which is like the observational vector, then the corresponding risk function is like (Yang and Gao 2006c)

$$\Omega(k) = \mathbf{V}_k^T \Sigma_k^{-1} \mathbf{V}_k + \frac{1}{\lambda_k} \mathbf{V}_{\hat{\mathbf{X}}_{k-1}}^T \Sigma_{\hat{\mathbf{X}}_{k-1}}^{-1} \mathbf{V}_{\hat{\mathbf{X}}_{k-1}} + \hat{\mathbf{w}}_k^T \Sigma_{\mathbf{w}_k}^{-1} \hat{\mathbf{w}}_k = \min, \quad (79)$$

where  $\lambda_k$  is the fading factor which satisfies  $\lambda_k \geq 1$ . The corresponding covariance matrix  $\Sigma_{\hat{\mathbf{X}}_{k-1}}$  follows:

$$\Sigma_{\bar{\mathbf{X}}_k} = \lambda_k \Phi_{k,k-1} \Sigma_{\hat{\mathbf{X}}_{k-1}} \Phi_{k,k-1}^T + \Sigma_{\mathbf{w}_k}, \quad (80)$$

where  $\Sigma_{\hat{\mathbf{X}}_{k-1}}$  is the covariance matrix of the state estimated at epoch  $t_{k-1}$ , which results in fading filtering.

The obvious difference between the fading filter and the standard Kalman filter is that the prior state covariance matrix in the fading filter is inflated by  $\lambda_k$  times in order to reduce the contribution of the prior state and strengthen the contribution of the present measurements on the last state estimate.

Another fading filter is based on the following risk function:

$$\Omega(k) = \mathbf{V}_k^T \Sigma_k^{-1} \mathbf{V}_k + \mathbf{V}_{\hat{\mathbf{X}}_{k-1}}^T \Sigma_{\hat{\mathbf{X}}_{k-1}}^{-1} \mathbf{V}_{\hat{\mathbf{X}}_{k-1}} + \frac{1}{\lambda_k} \hat{\mathbf{w}}_k^T \Sigma_{\mathbf{w}_k}^{-1} \hat{\mathbf{w}}_k = \min. \quad (81)$$

Then

$$\Sigma_{\bar{\mathbf{X}}_k} = \Phi_{k,k-1} \Sigma_{\hat{\mathbf{X}}_{k-1}} \Phi_{k,k-1}^T + \lambda_k \Sigma_{\mathbf{w}_k}. \quad (82)$$

The two fading filters above are very similar, and both of them are different from the standard Kalman filter with the risk function (8) or

$$\Omega(k) = \mathbf{V}_k^T \Sigma_k^{-1} \mathbf{V}_k + \mathbf{V}_{\hat{\mathbf{X}}_{k-1}}^T \Sigma_{\hat{\mathbf{X}}_{k-1}}^{-1} \mathbf{V}_{\hat{\mathbf{X}}_{k-1}} + \hat{\mathbf{w}}_k^T \Sigma_{\mathbf{w}_k}^{-1} \hat{\mathbf{w}}_k = \min. \quad (83)$$

The key problem of fading filter is to construct a reasonable fading factor. Two kinds of fading factors, based on the optimization theory, are established (Fagin 1964; Fang 1998), one of which is expressed as

$$\lambda_k = \max\{1, \frac{1}{n} \text{tr}(\mathbf{N}_k \mathbf{M}_k^{-1})\}, \quad (84)$$

where  $\text{tr}[\cdot]$  denotes the trace of matrix.  $\mathbf{M}_k$  and  $\mathbf{N}_k$  are defined as

$$\mathbf{M}_k = \mathbf{A}_k \Phi_{k,k-1} \Sigma_{\hat{\mathbf{x}}_{k-1}} \Phi_{k,k-1}^T \mathbf{A}_k^T, \quad (85)$$

$$\mathbf{N}_k = \Sigma_{\tilde{\mathbf{v}}_k} - \mathbf{A}_k \Sigma \mathbf{w}_k \mathbf{A}_k^T - \Sigma_k, \quad (86)$$

where  $\Sigma_{\tilde{\mathbf{v}}_k}$  is the covariance matrix of the predicted residual vector  $\tilde{\mathbf{v}}_k$  (Yang et al. 2001; Yang and Xu 2003):

$$\Sigma_{\tilde{\mathbf{v}}_k} = E(\tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T). \quad (87)$$

Usually  $\Sigma_{\tilde{\mathbf{v}}_k}$  is calculated by windowing estimation method (Xia et al. 1990), similar to the Sage filtering (Yang and Xu 2003), that is,

$$\hat{\Sigma}_{\tilde{\mathbf{v}}_k} = \frac{1}{k-1} \sum_{i=1}^k \tilde{\mathbf{v}}_i \tilde{\mathbf{v}}_i^T. \quad (88)$$

One simpler expression of (84) is defined as follows (Sorensen and Sacks 1971):

$$\lambda_k = \max\{1, \text{tr}(\mathbf{N}_k / \mathbf{M}_k)\}. \quad (89)$$

Theoretically, the fading factor  $\lambda_k$  above is optimal. Increasing the predicted residual vector  $\tilde{\mathbf{v}}_{\tilde{\mathbf{x}}_k}$  will increase the covariance matrix  $\Sigma_{\tilde{\mathbf{v}}_k}$  based on (88) and results in an optimal fading factor  $\lambda_k$ .

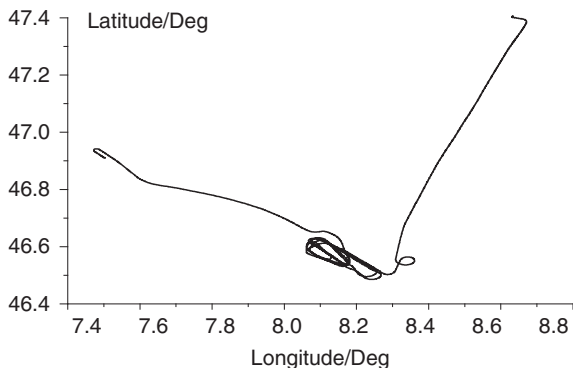
Formula (88) can be improved as (Fang 1998)

$$\Sigma_{\tilde{\mathbf{v}}_k} = \begin{cases} \frac{\lambda_{k-1} \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T}{1 + \lambda_{k-1}}, & k > 1, \\ \frac{1}{2} \tilde{\mathbf{v}}_0 \tilde{\mathbf{v}}_0^T, & k = 1, \end{cases} \quad (90)$$

where  $\tilde{\mathbf{v}}_0$  is the predicted residual vector when  $k = 0$ .

The improved  $\Sigma_{\tilde{\mathbf{v}}_k}$  expressed by (90) is more sensitive than (88) in reflecting the kinematic model errors at present epoch, since  $\Sigma_{\tilde{\mathbf{v}}_k}$  in (90) does not average the historical information, which uses the present information directly.

**Fig. 2.10** The orbit of the airplane movement



### 2.6.2 Comparison of Fading Filter and Adaptive Filter

The adaptively robust filter mentioned above is intermediate between fading filter and standard Kalman filter. It does not distinguish the errors of  $\hat{\mathbf{X}}_{k-1}$  from the kinematic model errors. It treats the predicted state vector  $\bar{\mathbf{X}}_k$  as a whole and adopts the principle that  $\alpha_k \mathbf{V}_{\bar{\mathbf{X}}_k}^T \Sigma_{\bar{\mathbf{X}}_k}^{-1} \mathbf{V}_{\bar{\mathbf{X}}_k}$  is a minimum, in which the adaptive factor  $\alpha_k$  changes between  $[0, 1]$ . When the kinematic model error increases or the vehicle movement is in an unstable state (Fig. 2.10),  $\alpha_k$  is smaller than 1 or even equals 0. When the kinematic model error is small enough,  $\alpha_k$  equals 1, and the adaptive filtering changes into the standard Kalman filtering.

Analysing the adaptively robust filtering and fading filtering, we find that the primary differences are as follows:

1. The adaptively robust filtering adopts the principle of robust estimation, and it can control the influences of the measurement outliers on the estimated state vector.
2. The adaptive factor  $\alpha_k$  acts on the covariance matrix of the predicted state vector  $\bar{\mathbf{X}}_k$ , while the fading factor acts on the covariance matrix of the previous estimated state vector  $\hat{\mathbf{X}}_{k-1}$ .
3. In the fading factor, the matrix  $\mathbf{N}_k$  expressed by (86) may be indefinite, which usually leads to the failure of the filter. If the adaptive filter is expressed by (14), the adaptive factor can be changed in  $[0, 1]$ . If the adaptive filter is expressed by (15) and (16), then the adaptive factor can be changed in  $[0, 1]$ . The adaptive factor is determined by observational information and state predicted information, which is capable of adapting in real time. Because the adaptive factor is constructed by the discrepancy between the predicted state vector and the estimated state vector by measurements, it has strong adaptation ability and real time flexibility.

### 2.6.3 Actual Computation and Analysis

The actual data was collected by using two Trimble 4000SSI receivers mounted on an aircraft. To analyse the roles of the adaptive factors in the adaptive filtering, the highly precise results from double-differenced carrier measurements were used as “true values” to compare with the results from the code measurements. The constant-velocity model of the Kalman filtering was employed. The initial variances for positions, velocities and C/A-code were selected separately as  $0.2 \text{ m}^2$ ,  $9 \times 10^{-5} \text{ m}^2/\text{s}^2$  and  $1.0 \text{ m}^2$ , respectively. The spectral density for velocities was chosen to be  $0.01 \text{ m}^2/\text{s}^2$ . The selected kinematic model covariance matrix was the same as that used in Jazwinski (1970) and the following three schemes were carried out:

- Scheme 1: Standard Kalman filtering (SKF)
- Scheme 2: Fading Kalman filtering (FKF)
- Scheme 3: Adaptively robust Kalman filtering (ARKF)

The errors of the  $X$  axis of the three schemes relative to the “true values” are plotted in Figs. 2.11, 2.12 and 2.13. Because the errors in the  $X$ ,  $Y$  and  $Z$  axes are similar, only the errors of the  $X$  axis are given. The comparison of RMS is shown in Table 2.2.

From the calculation results above, we find that

1. The influences of the disturbances during the flight on the standard Kalman filtering are very significant; see Fig. 2.11 and Table 2.2.
2. From Figs. 2.11 and 2.12 and Table 2.2, we find that the fading filtering can control the influences of the kinematic model disturbances on the navigation results to a certain extent, and the results are obviously superior to the standard Kalman filtering.
3. The results of adaptively robust Kalman filtering are slightly superior to the fading filtering. The adaptively robust Kalman filtering can not only control the

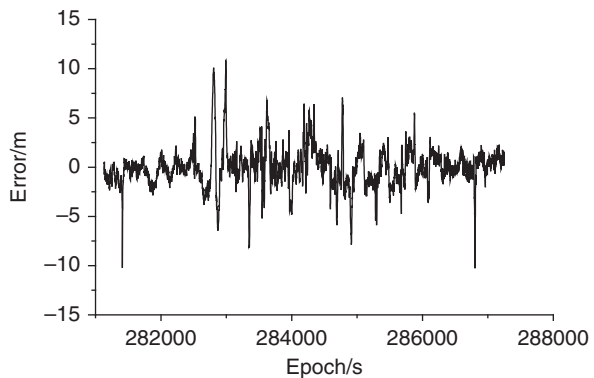
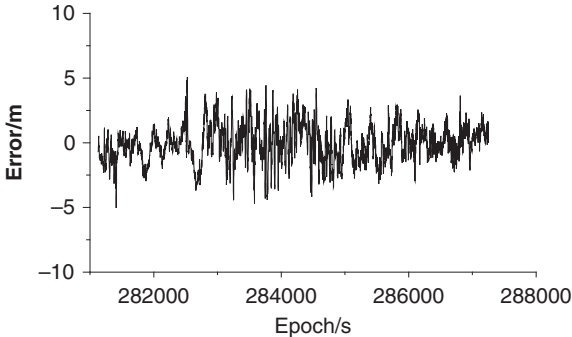


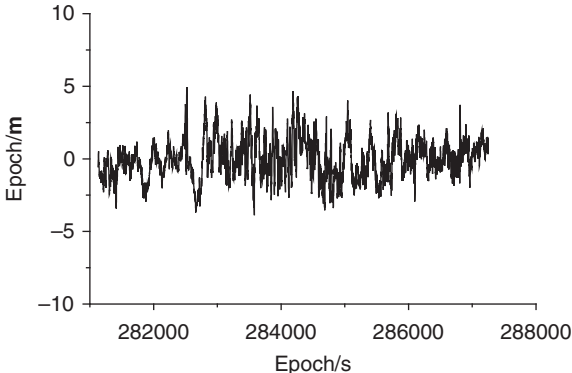
Fig. 2.11 Errors of scheme 1



**Fig. 2.12**    Errors of scheme 2



**Fig. 2.13**    Errors of scheme 3



**Table 2.2**    Comparison of RMS (unit: m)

		CKF	FKF	ARKF
RMS	X	2.006	1.415	1.331
	Y	1.338	0.759	0.675
	Z	1.936	1.539	1.452

influences of the kinematic model disturbances but also control the influences of the measurement outliers on the navigation results; see Fig. 2.13 and Table 2.2. The results of the adaptively robust filtering are very stable and robust, and the calculation is very flexible.

In conclusion, the fading filtering can control the influences of the kinematic model disturbances on the navigation results to some extent, and its results are obviously superior to standard Kalman filtering. The fading filtering uses the fading factor to inflate the covariance matrix of the estimated state vector at the former

epoch in order to reduce the influences of the state model errors on the new estimated results, but it is difficult to distinguish the model errors from the errors of the former estimated state vector. When the disturbances of the kinematic model are large enough, it is difficult to control their influences by the fading factor  $\lambda_k$ .

## 2.7 Comparison of Sage Adaptive Filter and Adaptively Robust Filter

In the adaptive Kalman filtering algorithms, use of the Sage–Husa filter (Sage and Husa 1969) is very popular for approaching the variance–covariance matrices by the windowing method (see Jazwinski 1970) and keeping a good consistency between the predicted residuals and the corresponding statistics. A windowing approach of innovation-based adaptive estimation has been studied by Mehra (1970). It makes the covariance matrices of the observation equation and the state errors adapt to the observation information (Mohamed and Schwarz 1999).

The principle of the Sage adaptive filter is

$$\Omega(k) = \mathbf{V}_k^T \hat{\Sigma}_k^{-1} \mathbf{V}_k + \mathbf{V}_{\bar{\mathbf{X}}_k}^T \hat{\Sigma}_{\bar{\mathbf{X}}_k}^{-1} \mathbf{V}_{\bar{\mathbf{X}}_k} = \min, \quad (91)$$

where  $\hat{\Sigma}_k$  and  $\hat{\Sigma}_{\bar{\mathbf{X}}_k}$  are estimated by window method.

If the covariance matrix of the current observational errors is computed by the innovation vectors from the previous  $m$  epochs, then the adaptive filter is called IAE (innovation-based adaptive estimation) filter. If the covariance matrix is computed by the residual vectors, then the adaptive filter is called RAE (residual-based adaptive estimation) filter. These two adaptively windowing estimations have appeared many times in the literature (Mohamed and Schwarz 1999; Wang et al. 1999; Hu and Ou 1999).

### 2.7.1 IAE Windowing Method

Suppose that the observational errors are normally distributed. If the width of moving windows is chosen as  $m$ , the estimators  $\hat{\Sigma}_{\bar{\mathbf{V}}_k}$  of the covariance matrix  $\Sigma_{\bar{\mathbf{V}}_k}$  can be given by

$$\hat{\Sigma}_{\bar{\mathbf{V}}_k} = \frac{1}{m} \sum_{j=0}^m \bar{\mathbf{V}}_{k-j} \bar{\mathbf{V}}_{k-j}^T. \quad (92)$$

Considering (24) we have the relation of the covariance matrices of the measurement vector, predicted residual vector and predicted state vector as

$$\Sigma_k = \Sigma_{\bar{\mathbf{V}}_k} - \mathbf{A}_k \Sigma_{\bar{\mathbf{A}}_k} \mathbf{A}_k^T. \quad (93)$$

Substituting (92) into (93), we obtain the covariance matrix  $\hat{\Sigma}_k$  of observation information at epoch  $t_k$  as

$$\hat{\Sigma}_k = \hat{\Sigma}_{\bar{\mathbf{V}}_k} - \mathbf{A}_k \hat{\Sigma}_{\bar{\mathbf{X}}_k} \mathbf{A}_k^T. \quad (94)$$

### 2.7.2 RAE Windowing Method

Similar to (92), the covariance matrix of the observational residual vector  $\mathbf{V}_k$  can be expressed as

$$\hat{\Sigma}_{\mathbf{V}_k} = \frac{1}{m} \sum_{j=0}^m \mathbf{V}_{k-j} \mathbf{V}_{k-j}^T. \quad (95)$$

From (95), we can estimate the covariance matrix  $\Sigma_k$  of the observational vector at epoch  $t_k$  as

$$\hat{\Sigma}_k = \hat{\Sigma}_{\mathbf{V}_k} + \mathbf{A}_k \hat{\Sigma}_{\hat{\mathbf{X}}_k} \mathbf{A}_k^T. \quad (96)$$

In order to estimate adaptively  $\hat{\Sigma}_k$  from (96),  $\hat{\Sigma}_{\hat{\mathbf{X}}_k}$  and residual vector  $\mathbf{V}_k$  at epoch  $t_k$  are required, while to estimate  $\hat{\Sigma}_{\bar{\mathbf{X}}_k}$  we must first have  $\hat{\Sigma}_k$ . Therefore, the covariance matrix of the observational vector at epoch  $t_k$  can only be computed using the measurement residuals from the previous  $m$  epochs before  $t_{k-1}$ :

$$\hat{\Sigma}_{\mathbf{V}_{k-1}} = \frac{1}{m} \sum_{j=1}^{m+1} \mathbf{V}_{k-j} \mathbf{V}_{k-j}^T. \quad (97)$$

Then (96) can be changed into

$$\hat{\Sigma}_k = \hat{\Sigma}_{\mathbf{V}_{k-1}} + \mathbf{A}_{k-1} \Sigma_{\hat{\mathbf{X}}_{k-1}} \mathbf{A}_{k-1}^T. \quad (98)$$

After having  $\hat{\Sigma}_k$ , the weight matrix,  $\mathbf{P}_k$ , of the observational vector at epoch  $t_k$  is computed.

Comparing the IAE estimators (92) and (93) with the RAE estimators (97) and (98), we can make following inferences:

1. The covariance matrix  $\hat{\Sigma}_k$  estimated by IAE includes the errors of predicted state vector  $\bar{\mathbf{X}}_k$ . The larger the error of  $\bar{\mathbf{X}}_k$ , the larger the error of  $\bar{\mathbf{V}}_k$ , which leads to a poor reliability of the covariance matrices  $\hat{\Sigma}_{\bar{\mathbf{V}}_k}$  and  $\hat{\Sigma}_k$ .
2. The covariance matrix  $\hat{\Sigma}_k$  estimated by RAE is indeed the covariance matrix  $\hat{\Sigma}_{k-1}$  at epoch  $t_{k-1}$ . In order to compute  $\hat{\Sigma}_{\bar{\mathbf{V}}_k}$  and  $\hat{\Sigma}_k$  by using the residual vector  $\mathbf{V}_k$ , we have to compute  $\hat{\mathbf{X}}_k$  at first, while to compute  $\hat{\mathbf{X}}_k$  we must have  $\hat{\Sigma}_k$ .

Thus we have to use the residuals before the epoch  $t_{k-1}$ . The reliability of this prediction depends on the consistency between the measurement accuracy of the current epoch and those of the previous epochs. Otherwise, the representation and reliability of the covariance matrix from this prediction can hardly be ensured.

3. The covariance matrix  $\hat{\Sigma}_k$  estimated by IAE is likely to be negative-definite, that is, the covariance matrix  $\hat{\Sigma}_{\tilde{\mathbf{v}}_k}$  is possibly smaller than  $\mathbf{A}_k \Sigma_{\bar{\mathbf{x}}_k} \mathbf{A}_k^T$ .
4. In general, the covariance matrix  $\hat{\Sigma}_{\tilde{\mathbf{v}}_k}$  computed from (92) is far smaller than  $\hat{\Sigma}_{\tilde{\mathbf{v}}_k}$  computed from (90).
5. The computation of  $\hat{\Sigma}_k$  based on IAE or RAE needs the measurement residuals or the innovation vectors from the previous  $m$  epochs, which increases the storage load of previous information. In addition, the width of moving window  $m$  is difficult to determine.
6. The covariance matrix  $\hat{\Sigma}_k$  computed from (93) or (96) is an average of previous accuracy information, which is almost impossible to use to describe the undulations of the observations at the present epoch. So this kind of adaptive estimation is difficult to use for realizing a real self-adaptation.
7. To estimate the covariance matrix  $\hat{\Sigma}_k$  whether using IAE or RAE requires that observational vectors not only have the same dimension at all epochs but also be the same observation type. Otherwise it is impossible to compute the covariance matrix  $\hat{\Sigma}_k$  using (93) or (96). It is a fatal weakness of IAE or RAE to estimate  $\hat{\Sigma}_k$ . In turn, it makes IAE or RAE hard to apply in kinematic positioning or navigation.

### 2.7.3 The Problems of the Windowing Estimation for Covariance Matrix of Kinematic Model

Let the correction vector of the predicted state vector be given by  $\Delta \mathbf{X}_k$ ; then

$$\Delta \mathbf{X}_k = \hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k. \quad (99)$$

It is easy to deduce that

$$\Sigma \mathbf{w}_k = \Sigma \Delta \mathbf{X}_k + \Sigma \hat{\mathbf{x}}_k - \Phi_{k,k-1} \Sigma \hat{\mathbf{x}}_{k-1} \Phi_{k,k-1}^T. \quad (100)$$

It should be noted that  $E(\Delta \mathbf{X}_k) = 0$ ; then the covariance matrix of  $\Delta \mathbf{X}_k$  can be expressed as

$$\hat{\Sigma}_{\Delta \mathbf{X}} = \frac{1}{m} \sum_{j=0}^m \Delta \mathbf{X}_{k-j} \Delta \mathbf{X}_{k-j}^T. \quad (101)$$

The estimate  $\hat{\Sigma}_{\mathbf{w}_k}$  of  $\Sigma \mathbf{w}_k$  can be obtained as

$$\hat{\Sigma}_{\mathbf{W}_k} = \hat{\Sigma}_{\Delta \mathbf{X}_k} + \Sigma_{\hat{\mathbf{X}}_k} - \Phi_{k,k-1} \Sigma_{\hat{\mathbf{X}}_{k-1}} \Phi_{k,k-1}^T. \quad (102)$$

The estimation of  $\hat{\Sigma}_{\mathbf{W}_k}$  with (102) encounters the following problems: (1) the expression of  $\hat{\Sigma}_{\mathbf{W}_k}$  includes the covariance matrix  $\Sigma_{\hat{\mathbf{X}}_k}$  of the state parameter estimates at epoch  $t_k$ ; however, the computation of  $\Sigma_{\hat{\mathbf{X}}_k}$  needs  $\hat{\Sigma}_{\mathbf{W}_k}$ ; (2) the expression of  $\hat{\Sigma}_{\mathbf{W}_k}$  includes  $\hat{\Sigma}_{\Delta \mathbf{X}_k}$ , which is computed by using the  $\Delta \mathbf{X}_j = \hat{\mathbf{X}}_j - \bar{\mathbf{X}}_j$  from  $m$  epochs and involves  $\hat{\mathbf{X}}_k - \bar{\mathbf{X}}_k$  at epoch  $t_k$  which also needs  $\hat{\Sigma}_{\mathbf{W}_k}$ ; (3) even if we can estimate  $\hat{\Sigma}_{\Delta \mathbf{X}_{k-1}}$  and  $\hat{\Sigma}_{\mathbf{W}_{k-1}}$  using the previous  $\Delta \mathbf{X}_j$  of  $m$  epochs from epoch  $t_{k-1}$  and take the latter as an approximation of  $\hat{\Sigma}_{\mathbf{W}_k}$ , it is hard to make  $\hat{\Sigma}_{\mathbf{W}_k}$  adapt to the real kinematic noise level of the motion of the vehicle because the state disturbance at epoch  $t_k$  cannot be reliably reflected by the disturbances from the previous  $m$  epochs. It is the same case that  $\hat{\Sigma}_{\Delta \mathbf{X}_{k-1}}$  estimated with the average of  $\Delta \mathbf{X}_j \Delta \mathbf{X}_j^T$  of  $m$  epochs cannot reflect the state noise level at epoch  $t_k$ , especially when there is a notable sudden change in state.

In order to avoid the problems mentioned above, we directly estimate  $\hat{\Sigma}_{\mathbf{W}_k}$ . Considering (4), (8) and (20), we express  $\Delta \mathbf{X}_k$  as

$$\Delta \mathbf{X}_k = -\mathbf{K}_k \bar{\mathbf{V}}_k. \quad (103)$$

Then

$$\Sigma_{\Delta \mathbf{X}_k} = \mathbf{K}_k \Sigma_{\bar{\mathbf{V}}_k} \mathbf{K}_k^T. \quad (104)$$

Once the estimate of  $\hat{\Sigma}_{\bar{\mathbf{V}}_k}$  is obtained with (92), we can estimate  $\Sigma_{\Delta \mathbf{X}_k}$  as

$$\hat{\Sigma}_{\Delta \mathbf{X}_k} = \mathbf{K}_k \hat{\Sigma}_{\bar{\mathbf{V}}_k} \mathbf{K}_k^T. \quad (105)$$

In the stable state,  $\hat{\Sigma}_{\mathbf{W}_k}$  can be approximately substituted by  $\hat{\Sigma}_{\Delta \mathbf{X}_k}$ , that is,

$$\hat{\Sigma}_{\mathbf{W}_k} = \mathbf{K}_k \hat{\Sigma}_{\bar{\mathbf{V}}_k} \mathbf{K}_k^T. \quad (106)$$

It should be mentioned that there exists another pair of contradictions in the above adaptive filtering process. Increasing the covariance matrix of state noise using the adaptive estimation is equal to decreasing the covariance matrix of observation noise, and vice versa. So if  $\Sigma_{\mathbf{W}_k}$  and  $\Sigma_k$  are increased or decreased at the same time, a contradiction arises, which sometimes makes an infinite loop and leads to divergence.

## 2.8 Some Application Examples

*Example 1.* A flight experiment is chosen as an example. The data sets were collected by using Trimble 4000SSE on a flight (Yang et al. 2001a, b). The available

measurements are C/A-code, P2-code pseudoranges, L1 and L2 carrier phases and Doppler measurements with 1s data rate. The rover receiver was mounted on an aircraft, and the reference receiver was fixed at a site about 1 km from the initial aircraft location. After about 10 min static tracking, the aircraft took off, and the flight time was about 90 min. The flight states have two notable sudden changes, one is close to epoch 1,000 since the plane takes off and the other is between epoch 3,000 and 4,000 since the flight turns round.

The double-differenced C/A-code and P2-code measurements are employed in the test performance. The constant-velocity model of Kalman filter is employed. An initial variance of  $0.2 \text{ m}^2$  for positions, of  $9 \times 10^{-6} \text{ m}^2 \text{ s}^{-2}$  for velocities, of  $1 \text{ m}^2$  for code measurements, and with spectral density of  $0.2 \text{ m}^2 \text{ s}^{-3}$  for velocities are selected. The dynamic model covariance matrix is chosen as (Schwarz et al. 1989)

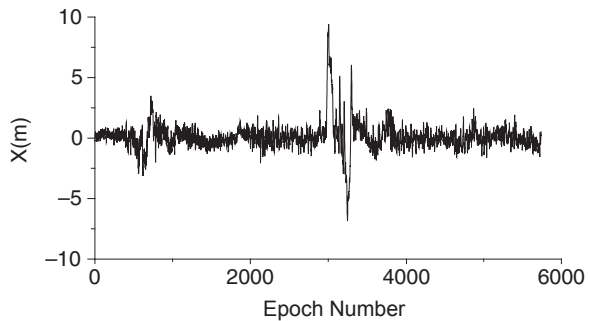
$$\Gamma_t \Sigma_{W_t} \Gamma_t^T = \begin{bmatrix} \frac{1}{3} Q_2 \Delta t^3 & \frac{1}{2} Q_2 \Delta t^2 \\ \frac{1}{2} Q_2 \Delta t^2 & Q_2 \Delta t \end{bmatrix},$$

where  $Q_2$  denotes the spectral density for velocities and  $\Delta t$  denotes a sampling time interval.

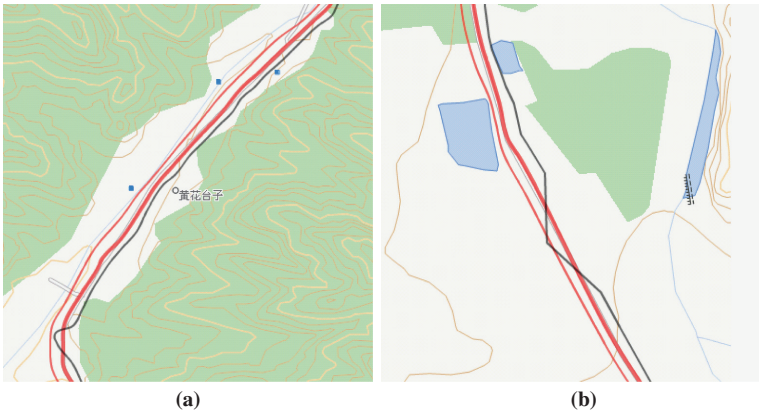
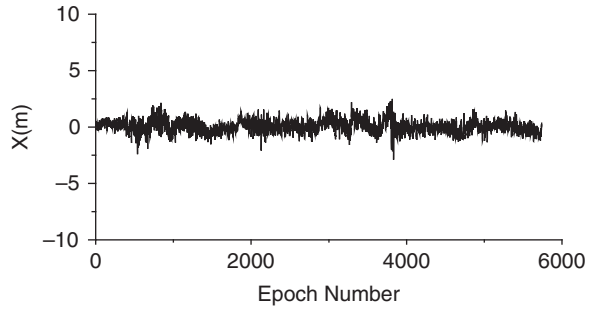
The highly precise results from double-differenced carrier measurements are used only as “true values” for comparing with the results from the code measurements, in which the ambiguities are resolved on the fly using LAMBDA method (Teunissen et al. 1997). The following two schemes are performed:

- Scheme 1: Standard Kalman filtering, i.e.  $\alpha_k = 1$  and  $\bar{\mathbf{P}}_k = \mathbf{P}_k$ ; see Fig. 2.14
- Scheme 2: Adaptive Kalman filtering, in which the adaptive factor  $\alpha_k$  is determined by (73) and  $\bar{\mathbf{P}}_k = \mathbf{P}_k$ ; see Fig. 2.15

Figures 2.14 and 2.15 show that the two unstable states of the flight are obviously reflected in the results of the standard Kalman filtering (Fig. 2.14) and the adaptively robust filter does resist the influences of the dynamic model errors (Fig. 2.15).



**Fig. 2.14** Standard Kalman filtering

**Fig. 2.15** Adaptively robust filtering**Fig. 2.16** Trajectories by differential positioning (*right*), adaptively robust filtering (*middle*) and receiver navigation (*left*)

*Example 2.* Road information updating by GPS: A GPS receiver is mounted on a vehicle, and a referenced GPS is fixed on a known station (Yang et al. 2004). The pseudorange measurements are employed in the test. Two schemes are performed:

- Scheme 1: Differential GPS positioning
- Scheme 2: Adaptively robust filtering

The two kinds of results and the navigation trajectory results are shown in an image map with the scale 1/50,000; see Fig. 2.16.

Figure 2.16a, b explicitly shows that the receiver navigation results and the differential positioning results have significant systematic errors. If the differential measurements number is less than the number of states, the differential positioning method will not give any position result or give an outlying result. The adaptively robust filtering usually gives reasonable results.

Theoretical analyses and many applications have illustrated that the adaptively robust filter with the corresponding adaptive factors and learning statistics can not

only resist the influences of the measurement outliers but also have a strong ability to control the influences of the state disturbances. It is flexible in computation because the adaptively robust filter is very similar to the standard Kalman filter.

The adaptively robust filter can be applied in some other fields, for example, in crustal deformation, in which the adaptive factor can be used with the geophysical deformation model information, and the robust equivalent weights can be employed with the repeated measurements (e.g. GPS).

## References

- Cannon, M.E., Schwarz, K.P. and Wong, R.V.C. (1986) Kinematic positioning with GPS: an analysis of road test. Proceedings of the 4th International Geodetic Symposium on Satellite Positioning, Austin, TX, 28 April–2 May
- Cui, X. and Yang, Y. (2006) Adaptively robust filtering with classified adaptive factor. *Prog. Nat. Sci.*, 16(8), 846–851
- Deng, Z. (2003) Self-Tuning Filtering Theory with Applications—Modern Time Series Analysis Method. Press of Harbin Institute of Technology, Harbin
- Fagin, S.L. (1964) Recursive linear regression theory, optimal filter theory and error analysis of optimal system. *IEEE Conv. Rec.*, 12, 216–240
- Fang, J. (1998) Study and Application of the Optional Estimation Theories and Method in Integrated Navigation System. Dongnan University, Nanjing (in Chinese)
- Hu, G. and Ou, J. (1999) The improved method of adaptive Kalman filtering for GPS high kinematic positioning. *Acta. Geodaetica. Et. Cartographica. Sinica.*, 28(4), 290–294
- Huber, P.J. (1964) Robust estimation of a location parameter. *Ann. Math. Statist.*, 35, 73–101
- Huber, P.J. (1981) *Robust Statistics*. Wiley, New York
- Jazwinski, A.H. (1970) *Stochastic Processes and Filtering Theory*. Mathematics in Science and Engineering. Academic Press, New York
- Koch, K.R. (2000) *Einführung in die Bayes-Statistik*. Springer, Heidelberg
- Koch, K.R. and Kusche, J. (2002) Regularization of geopotential determination from satellite data by variance components. *J. Geod.*, 76(5), 259–268
- Koch, K.R. and Yang, Y. (1998) Robust Kalman filter for rank deficient observation model. *J. Geod.*, 72(8), 436–441
- Masreliez, C.J. and Martin, R.D. (1977) Robust Bayesian estimation for the linear model and robustifying the Kalman filter. *IEEE Trans. Autom. Contr.*, AC-22, 361–371
- Mehra, R.K. (1970) On the identification of variances and adaptive Kalman filtering [J]. *IEEE Trans. Automat. Contr.*, 15(2), 175–184
- Mohamed, A.H. and Schwarz, K.P. (1999) Adaptive Kalman filtering for INS/GPS. *J. Geod.*, 73, 193–203
- Ou, J., Chai, Y. and Yuan, Y. (2004) Adaptive filtering for kinematic positioning by selection of the parameter weights. In: Zhu, Y. and Sun, H. (eds) *Progress in Geodesy and Geodynamics*. Hubei Science & Technology Press, Hubei, 816–823 (in Chinese)
- Panozzo, T., Born, G.H. and Bernelli-Zazzera, F. (2004) Kalman filtering of accelerometric data for aerobraking navigation. Presented in the 6th Conference of Dynamic and Control of Systems and Structures in Space, Riomaggiore, July 2004
- Ren, C., Ou, J. and Yuan, Y. (2005) Application of adaptive filtering by selecting the parameter weight factor in precise kinematic GPS positioning. *Prog. Nat. Sci.*, 15(1), 41–46
- Sage, A.P. and Husa, G.W. (1969) Adaptive filtering with unknown prior statistics. In *Proceeding of Joint Automatic Control Conference*, Boulder Colorado, 1969, pp. 760–769
- Salzmann, M. (1995) Real-time adaptation for model errors in dynamic systems. *Bull. Geod.*, 69, 81–91



- Schaffrin, B. (1991) Generating robustified Kalman filters for the integration of GPS and INS. Technical Report, No. 15, Institute of Geodesy, University of Stuttgart, Stuttgart
- Schwarz, K.P., Cannon, M.E. and Wong, R.V.C. (1989) A comparison of GPS kinematic models for determination of position and velocity along a trajectory. *Manuscripta Geodaeica*, 14(6), 345–353
- Sorenson, H.W. and Sacks, J.E. (1971) Recursive fading memory filtering. *Inform. Sci.*, 3, 101–119
- Teunissen, P.J.G. (1990) Quality control in integrated navigation systems. *IEEE Aerosp. Electron. Syst. Mag.*, 5(7), 35–41
- Teunissen, P.J.G., de Jonge, P.J. and Tiberius, C.C.J.M. (1997) Performance of the LAMBDA method for fast GPS ambiguity resolution. *Navigation, J. Inst. Nav.*, 44(3), 373–383
- Wang, J., Stewart, M.P. and Tsakiri, M. (2000) Adaptive Kalman filtering for integration of GPS with GLONASS and INS. In: Schwarz, K.P. (ed) *Geodesy Beyond 2000*. Springer, Heidelberg, 325–330
- Wang, Y.J. and Kubik, K.K. (1993) Robust Kalman filter and its geodetic applications [J]. *Manuscripta Geodaeica*, 18(6), 349–354
- Xia, Q., Sun, Y. and Zhou, C. (1990) An optimal adaptive algorithm for fading Kalman filter and its application. *Acta Automat. Sin.*, 16(3), 210–216 (in Chinese)
- Xu, G. (2007) *GPS Theory, Algorithms and Applications*, 2nd Ed. Springer, Berlin
- Xu, P. and Rummel, R. (1994) A simulation study of smoothness methods in recovery of regional gravity fields. *Geophys. J. Int.*, 117, 472–486
- Xu, T. and Yang, Y. (2000) Modified sage adaptive filtering. *Sci. Surv. Mapp.*, 25, 22–24
- Yang, Y. (1991) Robust Bayesian estimation. *Bull. Geodesique*, 65(3), 145–150
- Yang, Y. (1993) *The Theory and Application of Robust Estimation*. Bayi Publication House, Beijing (in Chinese)
- Yang, Y. (1994) Robust estimation for dependent observations. *Mans. Geod.*, 19, 10–17
- Yang, Y. (1997) Robust Kalman filter for dynamical system. *J. Zhengzhou Inst. Surv. Mapp.*, 14(2), 79–84 (in Chinese)
- Yang, Y. (1999) Theory and Comparisons of Various Robust Filter Models. *Collection Papers of Geodesy. Surveying and Mapping Press, Beijing*, pp. 51–57 (in Chinese)
- Yang, Y., Cheng, M.K., Shum, C.K. and Tapley, B.D. (1999) Robust estimation of systematic errors of satellite laser range. *J. Geod.*, 73, 345–349
- Yang, Y. and Cui, X. (2008) Adaptively robust filter with multi adaptive factors. *Surv. Rev.*, 40(309), 260–270
- Yang, Y. and Gao, W. (2005) Influence comparison of adaptive factors on navigation results. *J. Nav.*, 58, 471–478
- Yang, Y. and Gao, W. (2006a) An optimal adaptive Kalman filter. *J. Geod.*, 80, 177–183
- Yang, Y. and Gao, W. (2006b) A new learning statistic for adaptive filter based on predicted residuals. *Prog. Nat. Sci.*, 16(8), 833–837
- Yang, Y. and Gao, W. (2006c) Comparison of Two Fading Filters and Adaptively Robust Filter. *Geomatics and Inf. Sci. Wuhan Univ.*, 31(11), 980–982
- Yang, Y., He, H. and Xu, G. (2001a) Adaptively robust filtering for kinematic geodetic positioning. *J. Geod.*, 75(2/3), 109–116
- Yang, Y., Song, L. and Xu, T. (2002a) Robust estimator for correlated observations based on bifactor equivalent weights. *J. Geod.*, 76(6–7), 353–358
- Yang, Y., Song, L. and Xu, T. (2002b) Robust parameter estimation for correlated geodetic observations. *Selected Papers for English Edition, Acta Geodeatica Cartography Sinica*, pp. 18–24
- Yang, Y., Tang, Y., Li, Q. and Zou, Y. (2004) Experiments of adaptive filters for kinematic GPS positioning applied in road information updating in GIS. *International Symposium on Inertial Navigation Technology and Intelligent Traffic*, Nanjing, October 15–17.
- Yang, Y. and Wen, Y. (2003) Synthetically adaptive robust filtering for satellite orbit determination. *Sci. China (series D)*, 32(11), 1112–1119
- Yang, Y. and Xu, T. (2003) An adaptive Kalman filter based on Sage windowing weights and variance components. *J. Nav.*, 56(2), 231–240

- Yang, Y., Xu, T. and He, H. (2001b) On adaptively kinematic filtering. Selected Papers for English of *Acta Geodetica et Cartographica Sinica*, pp. 25–32
- Zhou, J. (1989) Classical theory of errors and robust estimation. *Acta Geodaetica et Cartographica Sinica*, 18(2), 115–120 (in Chinese)
- Zhou, J., Huang, Y., Yang, Y. and Ou, J. (1997) *Robust Least Squares Method*. Press of Huazhong Technical University, Wuhan

Sciences of Geodesy - I

Advances and Future Directions

Xu, G. (Ed.)

2010, XXVI, 488 p., Hardcover

ISBN: 978-3-642-11740-4