第 29 卷 第 1 期
2003 年 1 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 29, No. 1
Jan., 2003

White Noise Estimation Theory Based on Kalman Filtering[1)]

DENG Zi-Li[1] XU Yan[2]
[1](Institute of Applied Mathematics, Heilongjiang University, Harbin 150080)
[2](Department of Automation, Heilongjiang University, Harbin 150080)
(E-mail: dzl@hlju.edu.cn)

Abstract By using the Kalman filtering method, a unified and general white noise estimation theory is presented for the first time. It can handle the filtering, smoothing and prediction problems in a unified framework for both the input white noise and measurement white noise in linear discrete time-varying and time-invariant stochastic systems. The optimal and steady-state white noise estimators are presented, and white noise innovation filters and Wiener filters are also presented. They can be applied to seismic data processing in oil exploration, and provide a new tool to solve the state and signal estimation problems. Two simulation examples show their effectiveness.

Key words Input white noise estimators, measurement white noise estimators, deconvolution, reflection seismology, Kalman filtering method

1 Introduction

The input white noise estimation problem for stochastic systems is also called deconvolution. It occurs in many fields including reflection seismology, communications, and signal processing. The optimal input white noise estimators with application to oil seismic exploration have been presented based on the Kalman filter by Mendel[1~4], but they do not constitute a unified white noise estimation theory because the measurement white noise estimation problem was not solved. A unified white noise estimation theory has been presented based on the modern time series analysis method by Deng et al.[5], which not only includes the input white noise estimators but also includes the measurement white noise estimators. It has been used to solve the state and signal estimation problems[5~11]. But its limitation is that it does not solve the optimal white noise estimation problems for time-varying systems and only the steady-state white noise estimators are presented for time-invariant systems.

In order to overcome the above drawback and limitation, a new unified and general white noise estimation theory is presented based on the Kalman filter and projection theory in this paper. It can handle the input and measurement white noise filtering, smoothing and prediction problems in a unified framework for time-varying and time-invariant systems. The optimal and steady-state white noise estimators are presented, and the white noise innovation filters and Wiener filters are presented. They provide a new tool for solving the state and signal estimation problems.

Consider the linear discrete time-varying stochastic system

$$x(t+1) = \Phi(t)x(t) + \Gamma(t)w(t) \tag{1}$$

$$y(t) = H(t)x(t) + v(t) \tag{2}$$

where t is the discrete time, the state $x(t) \in R^n$, the measurement $y(t) \in R^m$, $\Phi(t)$, $\Gamma(t)$ and $H(t)$ are known time-varying matrices.

Assumption 1. $w(t) \in R^r$ and $v(t) \in R^m$ are independent white noises with zero mean and variance matrices $Q(t)$ and $R(t)$, respectively.

1) Supported by National Natural Science Foundation of P. R. China (NSFC-69774019), and by Natural Science Foundation of Heilongjiang Province (F01-15)
Received September 12, 2001; in revised form June 14, 2002
收稿日期 2001-09-12; 收修改稿日期 2002-06-14

Assumption 2. The initial state $\boldsymbol{x}(0)$ with mean $\boldsymbol{\mu}_0$ and variance matrix P_0 is uncorrelated with $\boldsymbol{w}(t)$ and $\boldsymbol{v}(t)$.

Assumption 3. The initial measurement time $t_0=1$.

Optimal white noise estimation problems are to find the optimal (linear minimum variance) estimators $\hat{\boldsymbol{v}}(t|t+N)$ and $\hat{\boldsymbol{w}}(t|t+N)$ of $\boldsymbol{v}(t)$ and $\boldsymbol{w}(t)$ based on measurements $(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(1))$. They respectively minimize the mean-square errors

$$J_\theta = E[(\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t\mid t+N))^{\mathrm{T}}(\boldsymbol{\theta}(t)-\hat{\boldsymbol{\theta}}(t\mid t+N))],\ \boldsymbol{\theta}=\boldsymbol{v},\boldsymbol{w} \tag{3}$$

where E denotes the mathematical expectation, T denotes the transpose. For $N=0$, $N>0$ or $N<0$, they are called white noise filters, smoothers or predictors, respectively.

This paper is based on the following Kalman filter[12]

$$\hat{\boldsymbol{x}}(t+1\mid t+1) = \hat{\boldsymbol{x}}(t+1\mid t)+K(t+1)\boldsymbol{\varepsilon}(t+1) \tag{4}$$

$$\hat{\boldsymbol{x}}(t+1\mid t) = \Phi(t)\hat{\boldsymbol{x}}(t\mid t) \tag{5}$$

$$\boldsymbol{\varepsilon}(t+1) = \boldsymbol{y}(t+1)-H(t+1)\hat{\boldsymbol{x}}(t+1\mid t) \tag{6}$$

$$K(t+1) = P(t+1\mid t)H^{\mathrm{T}}(t+1)[H(t+1)P(t+1\mid t)H^{\mathrm{T}}(t+1)+R(t+1)]^{-1} \tag{7}$$

$$P(t+1\mid t) = \Phi(t)P(t\mid t)\Phi^{\mathrm{T}}(t)+\Gamma(t)Q(t)\Gamma^{\mathrm{T}}(t) \tag{8}$$

$$P(t+1\mid t+1) = [I_n-K(t+1)H(t+1)]P(t+1\mid t) \tag{9}$$

where I_n is the $n\times n$ unit matrix, $\hat{\boldsymbol{x}}(0|0)=\boldsymbol{\mu}_0$, $P(0|0)=P_0$. Substituting (9) into (8) yields the Riccati equation

$$P(t+1\mid t)=\Phi(t)[P(t\mid t-1)-P(t\mid t-1)H^{\mathrm{T}}(t)(H(t)P(t\mid t-1)H^{\mathrm{T}}(t)+R(t))^{-1}\times H(t)P(t\mid t-1)]\Phi^{\mathrm{T}}(t)+\Gamma(t)Q(t)\Gamma^{\mathrm{T}}(t) \tag{10}$$

Notice that the innovation $\boldsymbol{\varepsilon}(t)$ is white noise with zero mean and variance matrix

$$Q_\varepsilon(t) = H(t)P(t\mid t-1)H^{\mathrm{T}}(t)+R(t) \tag{11}$$

2 Optimal white noise estimators

2.1 Optimal measurement white noise estimators

Theorem 1. For the time-varying system (1),(2) with Assumptions 1~3, the optimal measurement white noise estimators are given by

$$\hat{\boldsymbol{v}}(t\mid t+N) = \boldsymbol{0}\quad (N<0) \tag{12}$$

$$\hat{\boldsymbol{v}}(t\mid t) = R(t)Q_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{13}$$

$$\hat{\boldsymbol{v}}(t\mid t+N) = \hat{\boldsymbol{v}}(t\mid t+N-1)+M_v(t\mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{14}$$

where $N=1,2,\cdots$, and the smoothing gains are given as

$$M_v(t\mid t+1) =-R(t)K^{\mathrm{T}}(t)\Phi^{\mathrm{T}}(t)H^{\mathrm{T}}(t+1)Q_\varepsilon^{-1}(t+1) \tag{15}$$

$$M_v(t\mid t+N) =-R(t)K^{\mathrm{T}}(t)\Phi^{\mathrm{T}}(t)\Big\{\prod_{i=1}^{N-1}[I_n-K(t+i)H(t+i)]^{\mathrm{T}}\Phi^{\mathrm{T}}(t+i)\Big\}\times H^{\mathrm{T}}(t+N)Q_\varepsilon^{-1}(t+N) \tag{16}$$

The error variance matrices $P_v(t|t+N)=E[(\boldsymbol{v}(t)-\hat{\boldsymbol{v}}(t|t+N))(\boldsymbol{v}(t)-\hat{\boldsymbol{v}}(t|t+N))^{\mathrm{T}}]$ are given as

$$P_v(t\mid t+N) = R(t)\qquad (N<0) \tag{17}$$

$$P_v(t\mid t) = R(t)-R(t)Q_\varepsilon^{-1}(t)R(t) \tag{18}$$

$$P_v(t\mid t+N) = P_v(t\mid t+N-1)-M_v(t\mid t+N)Q_\varepsilon(t+N)M_v^{\mathrm{T}}(t\mid t+N) \tag{19}$$

Proof. Applying the projection theory[12], $\hat{\boldsymbol{v}}(t|t+N)$ is the projection of $\boldsymbol{v}(t)$ on the linear space $L(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(1))$ generated by measurements $(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(1))$. When $N<0$, Assumptions 1~3 and iteration for (1) and (2) yield that $\boldsymbol{v}(t)$ is uncorrelated (orthogonal) with $L(\boldsymbol{y}(t+N),\cdots,\boldsymbol{y}(1))$, so that formula (12) holds. Applying the projection formula[12] yields

$$\hat{\boldsymbol{v}}(t\mid t) = \hat{\boldsymbol{v}}(t\mid t-1)+E[\boldsymbol{v}(t)\boldsymbol{\varepsilon}^{\mathrm{T}}(t)]Q_\varepsilon^{-1}(t)\boldsymbol{\varepsilon}(t) \tag{20}$$

Equation (2) and $\hat{y}(t|t-1)=H(t)\hat{x}(t|t-1)$ yield

$$\varepsilon(t) = H(t)\tilde{x}(t \mid t-1) + v(t) \tag{21}$$

where $\tilde{x}(t|t-1)=x(t)-\hat{x}(t|t-1)$. It is obvious that $\tilde{x}(t|t-1)\in L(v(t-1),\cdots,v(1), w(t-1),\cdots, w(0),x(0))$. Hence $v(t)$ is uncorrelated with $\tilde{x}(t|t-1)$. Applying (21) yields

$$E[v(t)\varepsilon^{T}(t)] = R(t) \tag{22}$$

Noting that $\hat{v}(t|t-1)=\mathbf{0}$, and substituting (11) and (22) into (20), we obtain formula (13).

Generally, applying the projection formula yields Equation (14), where

$$M_v(t \mid t+N) = E[v(t)\varepsilon^{T}(t+N)]Q_{\varepsilon}^{-1}(t+N) \tag{23}$$

From (1), (5) and (21) we have

$$\varepsilon(t+1) = H(t+1)[\Phi(t)\tilde{x}(t \mid t) + \Gamma(t)w(t)] + v(t+1) \tag{24}$$

with $\tilde{x}(t|t)=x(t)-\hat{x}(t|t)$. From (1), (4) and (5) we have the recursive relation

$$\tilde{x}(t \mid t) = [I_n - K(t)H(t)][\Phi(t-1)\tilde{x}(t-1 \mid t-1) + \Gamma(t-1)w(t-1)] - K(t)v(t) \tag{25}$$

Noting that $v(t)$ is uncorrelated with $\tilde{x}(t-1|t-1)$, from (24) and (25) we obtain

$$E[v(t)\varepsilon^{T}(t+1)] = -R(t)K^{T}(t)\Phi^{T}(t)H^{T}(t+1) \tag{26}$$

Applying (11), (23) and (26) yields formula (15). Finally, from (24) we have

$$\begin{aligned}\varepsilon(t+N) = {} & H(t+N)[\Phi(t+N-1)\tilde{x}(t+N-1 \mid t+N-1) + \\ & \Gamma(t+N-1)w(t+N-1)] + v(t+N)\end{aligned} \tag{27}$$

For $N\geqslant 2$, by iteration for (25), we have the relation

$$\begin{aligned}\tilde{x}(t+N-1 \mid t+N-1) = {} & \Phi(t+N-1,t-1)\tilde{x}(t-1 \mid t-1) + \\ & \sum_{i=t}^{t+N-1}\Phi(t+N-1,i)\{[I_n - K(i)H(i)]\Gamma(i-1)w(i-1) - K(i)v(i)\}\end{aligned} \tag{28}$$

where we define $\Phi(t+N-1,t+N-1)=I_n$, and for $i<t+N-1$, define

$$\begin{aligned}\Phi(t+N-1,i) = {} & [I_n - K(t+N-1)H(t+N-1)]\Phi(t+N-2) \times \cdots \times \\ & [I_n - K(i+1)H(i+1)]\Phi(i)\end{aligned} \tag{29}$$

Noting that $v(t)$ is uncorrelated with $\tilde{x}(t-1|t-1)$, from Assumptions 1～3 and (27)～(29) we have

$$\begin{aligned}E[v(t)\varepsilon^{T}(t+N)] = {} & -R(t)K^{T}(t)\Phi^{T}(t+N-1,t)\Phi^{T}(t+N-1)H^{T}(t+N) = \\ & -R(t)K^{T}(t)\Phi^{T}(t)\left\{\prod_{i=1}^{N-1}[I_n - K(t+i)H(t+i)]^{T}\Phi^{T}(t+i)\right\}H^{T}(t+N)\end{aligned} \tag{30}$$

Substituting (11) and (30) into (23) yields formula (16). It is obvious that Equation (17) holds. From (13) and (22) we directly obtain (18). Since $\hat{v}(t|t+N-1)\in L(\varepsilon(t+N-1),\cdots,\varepsilon(1))$, we see that $\varepsilon(t+N)$ is uncorrelated with $\hat{v}(t|t+N-1)$. Applying (14) and (23), we easily obtain (19). The proof is completed. □

Corollary 1. $M_v(t|t+N)$ can be computed recursively as

$$\begin{aligned}M_v(t \mid t+N) &= D_v(t,N)H^{T}(t+N)Q_{\varepsilon}^{-1}(t+N) \\ D_v(t,j) &= D_v(t,j-1)[I_n - K(t+j-1)H(t+j-1)]^{T}\Phi^{T}(t+j-1) \\ D_v(t,1) &= -R(t)K^{T}(t)\Phi^{T}(t),\ j = 2,3,\cdots,N\end{aligned} \tag{31}$$

Corollary 2. For $N\geqslant 0$, the non-recursive optimal measurement white noise estimators are given by

$$\hat{v}(t \mid t+N) = \sum_{i=0}^{N}M_v(t \mid t+i)\varepsilon(t+i) \tag{32}$$

where $M_v(t|t)=R(t)Q_{\varepsilon}^{-1}(t)$.

2.2 Optimal input white noise estimators

Theorem 2. For the time-varying system (1)～(2) with Assumptions 1～3, the optimal input white noise estimators are given as

$$\hat{w}(t \mid t+N) = \mathbf{0} \qquad (N \leqslant 0) \tag{33}$$

$$\hat{w}(t \mid t+1) = Q(t)\Gamma^{\mathrm{T}}(t)H^{\mathrm{T}}(t+1)Q_{\varepsilon}^{-1}(t+1)\boldsymbol{\varepsilon}(t+1) \tag{34}$$

$$\hat{w}(t \mid t+N) = \hat{w}(t \mid t+N-1) + M_w(t \mid t+N)\boldsymbol{\varepsilon}(t+N) \tag{35}$$

where $N=1,2,\cdots$, and the smoothing gains are given as

$$M_w(t \mid t+1) = Q(t)\Gamma^{\mathrm{T}}(t)H^{\mathrm{T}}(t+1)Q_{\varepsilon}^{-1}(t+1) \tag{36}$$

$$M_w(t \mid t+N) = Q(t)\Gamma^{\mathrm{T}}(t)\left\{\prod_{i=1}^{N-1}[I_n - K(t+i)H(t+i)]^{\mathrm{T}}\Phi^{\mathrm{T}}(t+i)\right\}\times H^{\mathrm{T}}(t+N)Q_{\varepsilon}^{-1}(t+N) \tag{37}$$

The error variance matrices are given by

$$P_w(t \mid t+N) = P_w(t \mid t+N-1) - M_w(t \mid t+N)Q_{\varepsilon}(t+N)M_w^{\mathrm{T}}(t \mid t+N) \tag{38}$$

with initial value $P_w(t|t)=Q(t)$.

Proof. The proof is very similar to that for Theorem 1, and is omitted.

Corollary 3. $M_w(t|t+N)$ can be computed recursively as

$$\begin{aligned} M_w(t \mid t+N) &= D_w(t,N)H^{\mathrm{T}}(t+N)Q_{\varepsilon}^{-1}(t+N) \\ D_w(t,j) &= D_w(t,j-1)[I_n - K(t+j-1)H(t+j-1)]^{\mathrm{T}}\Phi^{\mathrm{T}}(t+j-1) \\ D_w(t,1) &= Q(t)\Gamma^{\mathrm{T}}(t),\ j = 2,3,\cdots,N \end{aligned} \tag{39}$$

Corollary 4. For $N \geqslant 0$, the non-recursive optimal input white noise estimators are given by

$$\hat{w}(t \mid t+N) = \sum_{i=1}^{N} M_w(t \mid t+i)\boldsymbol{\varepsilon}(t+i) \tag{40}$$

Remark 1. Theorem 1 is presented for the first time in the literature. Theorem 2 was given by Mendel[1,2] under the assumption that $\Gamma(t)$ is of full column rank. Here the method of the proof for Theorem 2 is different to Mendel's as in [1] and [2], and the assumption that $\Gamma(t)$ is of full column rank is avoided.

Remark 2. Equations (12) and (33) show that the white noise predictors are equal to zero because the white noises are orthogonal to linear space generated by the past measurement data.

3 Steady-state white noise estimators

Assumption 4[12]. System (1), (2) is the time-invariant system with constant matrices $\Phi(t)=\Phi$, $\Gamma(t)=\Gamma$, $H(t)=H$, $Q(t)=Q$ and $R(t)=R$, and satisfies the condition that Φ is a stable matrix, or (Φ, H) is a completely observable pair and (Φ, Γ) is a completely controllable pair, or (Φ, H) is a completely detectable pair, and $(\Phi, \Gamma\bar{Q})$ is a completely stabilizable pair for any $\bar{Q}$ with $\bar{Q}\bar{Q}^{\mathrm{T}}=Q$.

For the time-invariant system (1), (2) with Assumptions 1~4, there exists the steady-state Kalman filter[6,12], i. e., $P(t+1|t)\to\Sigma$, $K(t)\to K$, as $t\to\infty$, and from (4)~(7) and (10) the steady-state Kalman predictor is given by

$$\hat{x}(t+1 \mid t) = \Phi\hat{x}(t \mid t-1) + K_p\boldsymbol{\varepsilon}(t) \tag{41}$$

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{y}(t) - H\hat{x}(t \mid t-1) \tag{42}$$

$$K_p = \Phi K \tag{43}$$

$$K = \Sigma H^{\mathrm{T}}[H\Sigma H^{\mathrm{T}} + R]^{-1} \tag{44}$$

where the innovation $\boldsymbol{\varepsilon}(t)$ is computed recursively via (41) and (42) with arbitrary initial value $\hat{x}(1|0)$, and Σ is a unique positive definite solution of the steady-state Riccati equation

$$\Sigma = \Phi[\Sigma - \Sigma H^{\mathrm{T}}(H\Sigma H^{\mathrm{T}} + R)^{-1}H\Sigma]\Phi^{\mathrm{T}} + \Gamma Q\Gamma^{\mathrm{T}} \tag{45}$$

Notice that from (11) the variance matrix Q_{ε} of the steady-state innovation is given as

$$Q_{\varepsilon} = H\Sigma H^{\mathrm{T}} + R \tag{46}$$

In Theorem 1 and Theorem 2, as $t\to\infty$, we obtain the following Theorem 3 and Theorem 4.

Theorem 3. For the time-invariant system (1),(2) with Assumptions 1~4, the steady-state measurement white noise estimators are given as

$$\hat{\boldsymbol{v}}(t \mid t+N)=\mathbf{0} \qquad (N<0) \tag{47}$$

$$\hat{\boldsymbol{v}}(t \mid t)=RQ_{\varepsilon}^{-1}\boldsymbol{\varepsilon}(t) \tag{48}$$

$$\hat{\boldsymbol{v}}(t \mid t+N)=\hat{\boldsymbol{v}}(t \mid t+N-1)+M_{v}(N)\boldsymbol{\varepsilon}(t+N) \tag{49}$$

$$M_{v}(N)=-RK^{\mathrm{T}}\Phi^{\mathrm{T}}[(I_{n}-KH)^{\mathrm{T}}\Phi^{\mathrm{T}}]^{N-1}H^{\mathrm{T}}Q_{\varepsilon}^{-1} \tag{50}$$

and the steady-state error variance matrices are given by

$$P_{v}(t \mid t+N)=R(N<0) \tag{51}$$

$$P_{v}(t \mid t)=R-RQ_{\varepsilon}^{-1}R \tag{52}$$

$$P_{v}(t \mid t+N)=P_{v}(t \mid t+N-1)-M_{v}(N)Q_{\varepsilon}M_{v}^{\mathrm{T}}(N) \tag{53}$$

Theorem 4. For the time-invariant system (1),(2) with Assumptions 1~4, the steady-state input white noise estimators are given as

$$\hat{\boldsymbol{w}}(t \mid t+N)=\mathbf{0} \quad (N<0) \tag{54}$$

$$\hat{\boldsymbol{w}}(t \mid t)=\mathbf{0} \tag{55}$$

$$\hat{\boldsymbol{w}}(t \mid t+N)=\hat{\boldsymbol{w}}(t \mid t+N-1)+M_{w}(N)\boldsymbol{\varepsilon}(t+N) \tag{56}$$

$$M_{w}(N)=Q\Gamma^{\mathrm{T}}[(I_{n}-KH)^{\mathrm{T}}\Phi^{\mathrm{T}}]^{N-1}H^{\mathrm{T}}Q_{\varepsilon}^{-1} \tag{57}$$

and the steady-state error variance matrices are given by

$$P_{w}(t \mid t+N)=Q \qquad (N \leqslant 0)$$

$$P_{w}(t \mid t+N)=P_{w}(t \mid t+N-1)-M_{w}(N)Q_{\varepsilon}M_{w}^{\mathrm{T}}(N) \tag{58}$$

with initial value $P_{w}(t|t)=Q$.

Corollary 5. For the time-invariant system (1),(2) with Assumptions 1~4, the non-recursive optimal white noise estimators are given as

$$\hat{\boldsymbol{v}}(t \mid t+N)=\sum_{i=0}^{N}M_{v}(i)\boldsymbol{\varepsilon}(t+i) \tag{59}$$

$$\hat{\boldsymbol{w}}(t \mid t+N)=\sum_{i=0}^{N}M_{w}(i)\boldsymbol{\varepsilon}(t+i) \tag{60}$$

with definition of $M_{v}(0)=RQ_{\varepsilon}^{-1}$, $M_{w}(0)=0$.

4 White noise innovation filters and Wiener filters

Theorem 5. For the time-invariant system (1),(2) with Assumptions 1~4, the white noise innovation filters are given by

$$\hat{\boldsymbol{v}}(t \mid t+N)=L_{N}^{v}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{61}$$

$$\hat{\boldsymbol{w}}(t \mid t+N)=L_{N}^{w}(q^{-1})\boldsymbol{\varepsilon}(t+N) \tag{62}$$

where q^{-1} is the backward shift operator, $q^{-1}\boldsymbol{\varepsilon}(t)=\boldsymbol{\varepsilon}(t-1)$, and the polynomial matrices are defined as

$$L_{N}^{\theta}(q^{-1})=0 \quad (N<0)$$

$$L_{N}^{\theta}(q^{-1})=\sum_{i=0}^{N}M_{\theta}(i)q^{i-N},\ \theta=v,w,N\geqslant 0 \tag{63}$$

Proof. Introducing the forward shift operator q, $q\boldsymbol{\varepsilon}(t)=\boldsymbol{\varepsilon}(t+1)$, from (59) we have

$$\hat{\boldsymbol{v}}(t \mid t+N)=\sum_{i=0}^{N}M_{v}(i)q^{i-N}\varepsilon(t+N) \tag{64}$$

which yields (61) and (63) with $\theta=v$. Similarly, Equations (62) and (63) with $\theta=w$ hold. □

Remark 3. The innovation filters (61) and (62) can be considered as filters with the transfer functions $L_{N}^{v}(q^{-1})$ and $L_{N}^{w}(q^{-1})$ and input $\boldsymbol{\varepsilon}(t+N)$.

It is well known that a filter is called Wiener filter if it can be expressed as a transfer function form with the measurement signal as input. In order to obtain the white noise

Wiener filters, we shall discover the relation between the innovation $\boldsymbol{\varepsilon}(t)$ and measurement $\boldsymbol{y}(t)$. This relation will be described by the autoregressive moving average (ARMA) innovation model. Substituting (42) into (41) yields

$$\hat{\boldsymbol{x}}(t+1 \mid t) = \Psi_p \hat{\boldsymbol{x}}(t \mid t-1) + K_p \boldsymbol{y}(t) \tag{65}$$

$$\Psi_p = \Phi - K_p H \tag{66}$$

where Ψ_p is a stable matrix[12]. Equation (65) can be rewritten in the transfer function form as

$$\hat{\boldsymbol{x}}(t+1 \mid t) = (I_n - q^{-1}\Psi_p)^{-1} K_p \boldsymbol{y}(t) \tag{67}$$

Substituing (67) into (42) yields

$$\boldsymbol{y}(t) = H(I_n - q^{-1}\Psi_p)^{-1} K_p \boldsymbol{y}(t-1) + \boldsymbol{\varepsilon}(t) \tag{68}$$

Applying the formula of matrix inverse, $(I_n - q^{-1}\Psi_p)^{-1} = \mathrm{adj}(I_n - q^{-1}\Psi_p)/\det(I_n - q^{-1}\Psi_p)$. From (68) we obtain the ARMA innovation model

$$A(q^{-1})\boldsymbol{y}(t) = \psi(q^{-1})\boldsymbol{\varepsilon}(t) \tag{69}$$

where we define

$$\psi(q^{-1}) = \det(I_n - q^{-1}\Psi_p)$$

$$A(q^{-1}) = \psi(q^{-1}) I_m - H\mathrm{adj}(I_n - q^{-1}\Psi_p) K_p q^{-1} \tag{70}$$

where $\psi(q^{-1})$ is a stable polynomial, i. e. , all zeros of $\psi(q)$ lie outside the unit circle because Ψ_p is a stable matrix.

Theorem 6. For the time-invariant system (1), (2) with Assumptions 1~4, the asymptotically stable white noise Wiener filters are given by

$$\hat{\boldsymbol{v}}(t \mid t+N) = \psi^{-1}(q^{-1}) L_N^v(q^{-1}) A(q^{-1}) \boldsymbol{y}(t+N) \tag{71}$$

$$\hat{\boldsymbol{w}}(t \mid t+N) = \psi^{-1}(q^{-1}) L_N^w(q^{-1}) A(q^{-1}) \boldsymbol{y}(t+N) \tag{72}$$

They can be expressed as

$$\psi(q^{-1})\, \hat{\boldsymbol{v}}(t \mid t+N) = K_N^v(q^{-1}) \boldsymbol{y}(t+N) \tag{73}$$

$$\psi(q^{-1})\, \hat{\boldsymbol{w}}(t \mid t+N) = K_N^w(q^{-1}) \boldsymbol{y}(t+N) \tag{74}$$

where we define that

$$K_N^v(q^{-1}) = L_N^v(q^{-1}) A(q^{-1}), \quad K_N^w(q^{-1}) = L_N^w(q^{-1}) A(q^{-1}) \tag{75}$$

Proof. From (69) we have $\boldsymbol{\varepsilon}(t+N) = \psi^{-1}(q^{-1}) A(q^{-1}) \boldsymbol{y}(t+N)$; substituting it into (61) and (62) yields (71) and (72). Since $\psi(q^{-1})$ is a stable polynomial, the Wiener filters (71)~(74) are asymptotically stable. Therefore, Theorem 6 holds. □

5 Simulation examples—Bernoulli-Gaussian white noise estimators

The Bernoulli-Gaussian white noise can be used to describe the reflectivity sequence for seismic data processing in oil exploration[1,2]. The Bernoulli-Gaussian white noise $\theta(t)$ is defined as

$$\theta(t) = b_\theta(t) g_\theta(t) \tag{76}$$

where $b_\theta(t)$ is a Bernoulli white noise taking values 1 and 0 with probabilities

$$P(b_\theta(t) = 1) = \lambda_\theta, \; P(b_\theta(t) = 0) = 1 - \lambda_\theta \tag{77}$$

and $g_\theta(t)$ is a Gaussian white noise with zero mean and variance $\sigma_{g_\theta}^2$, and is independent of $b_\theta(t)$. Hence the mean and variance σ_θ^2 of $\theta(t)$ are given as

$$\mathrm{E}\theta(t) = 0, \; \sigma_\theta^2 = \lambda_\theta \sigma_{g_\theta}^2 \tag{78}$$

The Bernoulli-Gaussian white noise with a time-varying coefficient $c(t)$ is defined as

$$\theta(t) = c(t) b_\theta(t) g_\theta(t) \tag{79}$$

Example 1. Consider the time-varying system

$$\boldsymbol{x}(t+1) = \begin{bmatrix} 1 & 0 \\ 0.5 + \sin\dfrac{2\pi t}{300} & 0 \end{bmatrix} \boldsymbol{x}(t) + \begin{bmatrix} 2\cos\dfrac{2\pi t}{300} \\ 1 \end{bmatrix} \boldsymbol{w}(t) \tag{80}$$

$$y(t) = \left[0,\ 1+0.2\cos\frac{2\pi t}{300}\right]x(t) + v(t) \tag{81}$$

$$w(t) = \left(1+0.1\sin\frac{2\pi t}{60}\right)b_w(t)g_w(t),\quad v(t) = \left(1+0.1\cos\frac{2\pi t}{60}\right)b_v(t)g_v(t) \tag{82}$$

where $w(t)$ and $v(t)$ are independent Bernoulli-Gaussian white noises with the time-varying coefficients. Taking $N=0,3$, $\sigma^2_{g_v}=1$, $\sigma^2_{g_w}=0.03$, $\lambda_v=0.19$, $\lambda_w=0.2$, $\hat{x}(0|0)=\mathbf{0}$, $P(0|0)=0.1I_2$, and applying Theorem 1, we obtain the optimal measurement white noise estimators $\hat{v}(t|t)$ and $\hat{v}(t|t+3)$. The simulation results are shown in Fig. 1 and Fig. 2, where the points denote $\hat{v}(t|t)$ or $\hat{v}(t|t+3)$, and the lines denote $v(t)$. The error variances are shown in Fig. 3, where the dashed line denotes $P_v(t|t)$ and the solid line $P_v(t|t+3)$. From Fig. 1 and Fig. 2 we see that the accuracy of $\hat{v}(t|t+3)$ is higher than that of $\hat{v}(t|t)$, and from Fig. 3 we see that the curve of $P_v(t|t+3)$ lies under that of $P_v(t|t)$ because Equation (19) yields $P_v(t|t+3)\leqslant P_v(t|t)$.

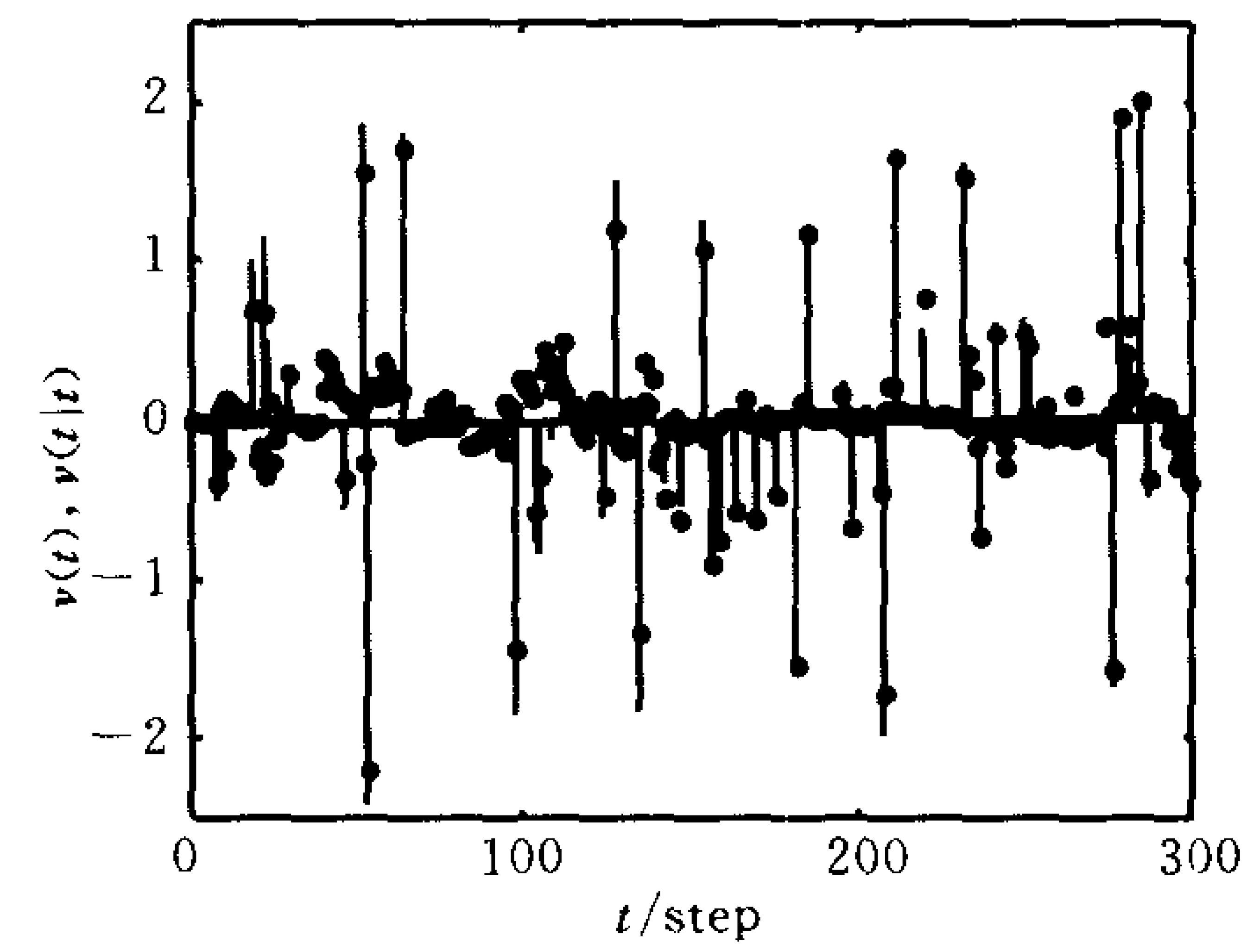

Fig. 1 Bernoulli-Gaussian measurement white noise $v(t)$ and filter $\hat{v}(t|t)$

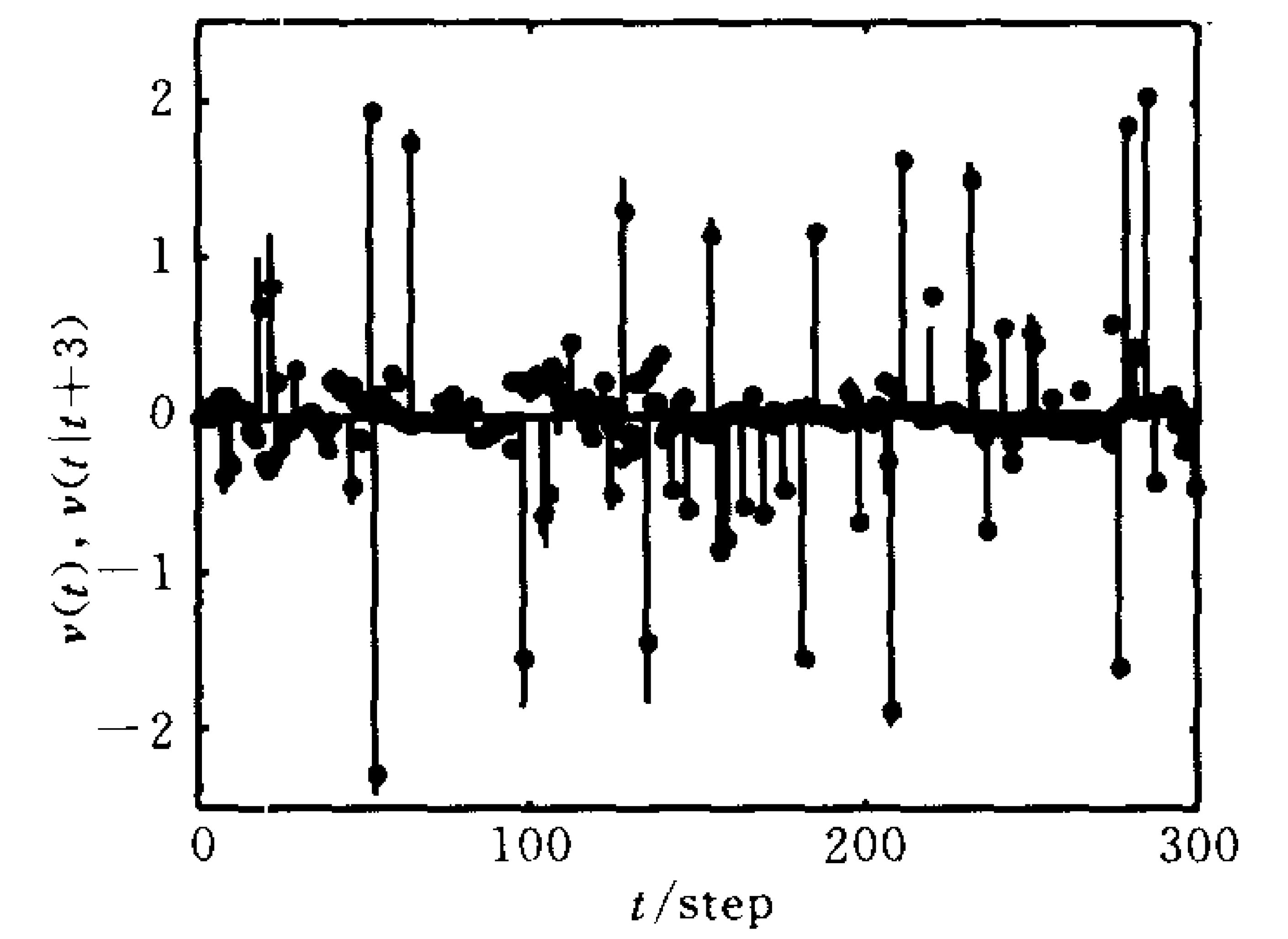

Fig. 2 Bernoulli-Gaussian measurement white noise $v(t)$ and smoother $\hat{v}(t|t+3)$

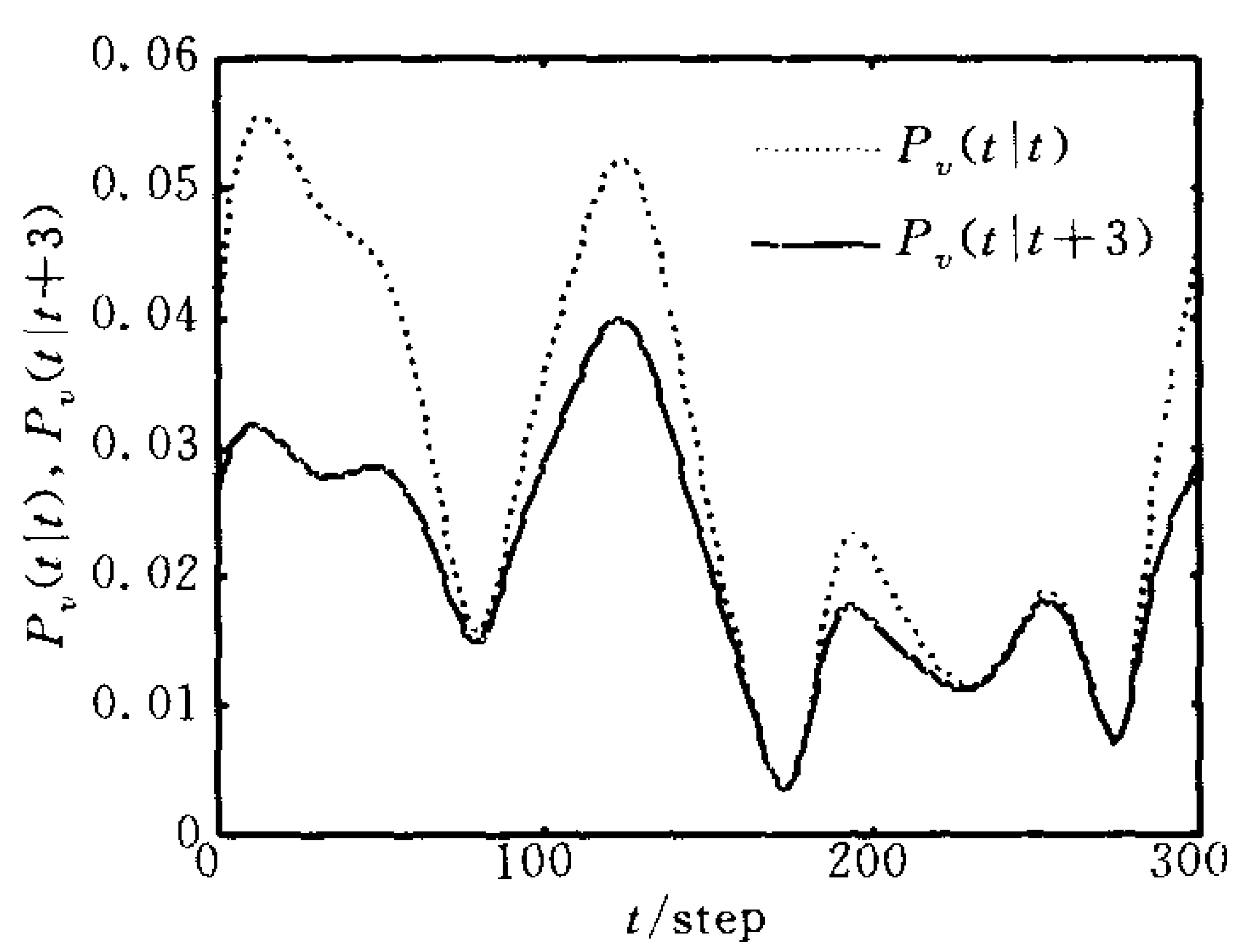

Fig. 3 The error variances $P_v(t|t)$ and $P_v(t|t-3)$

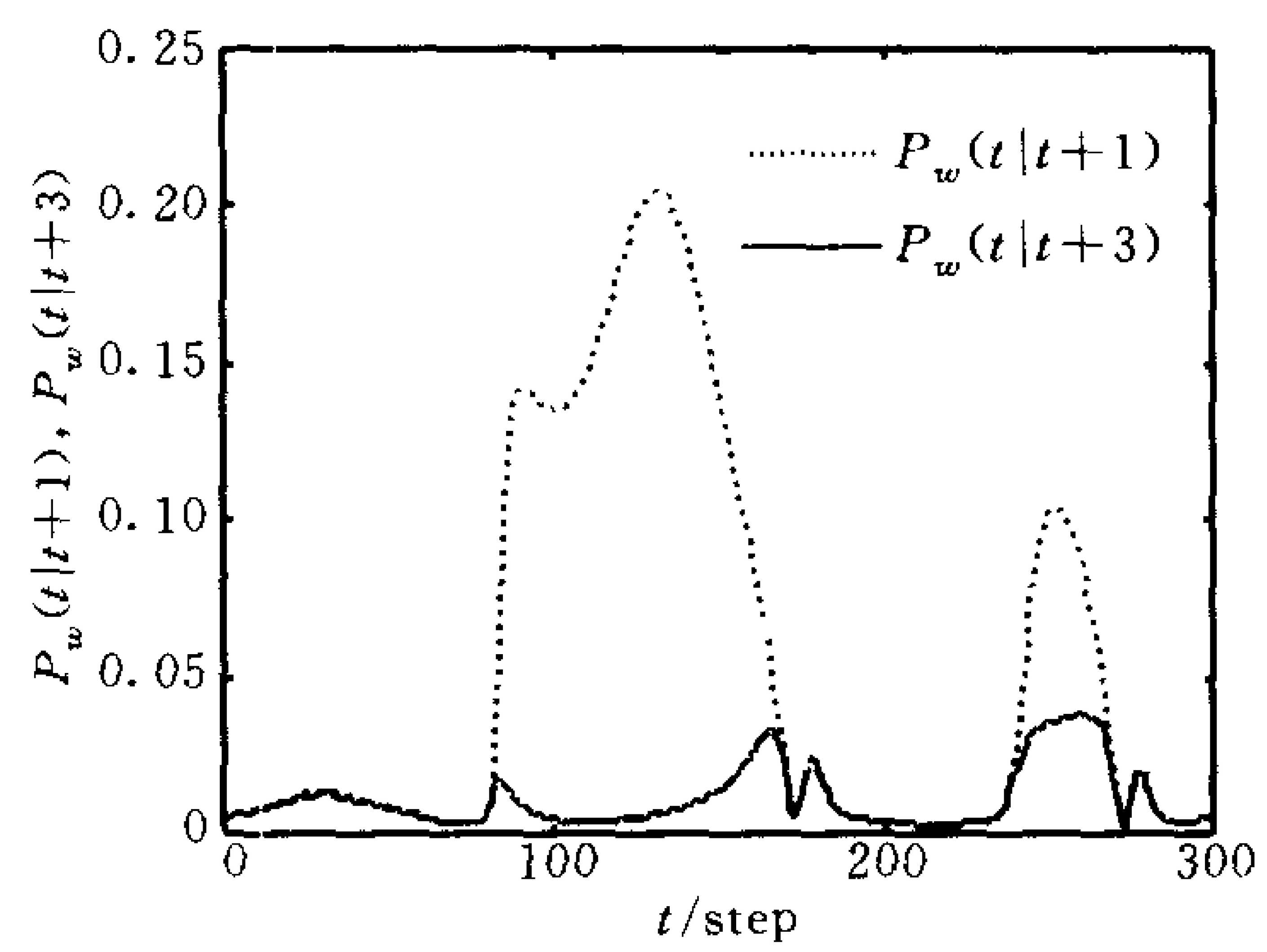

Fig. 4 The error variances $P_w(t|t+1)$ and $P_2(t|t+3)$

Example 2. Consider the time-varying system (80)~(82). Taking $N=1,3$, $\sigma^2_{g_w}=1$, $\sigma^2_{g_v}=0.01$, $\lambda_w=0.2$, $\lambda_v=0.3$, $\hat{x}(0|0)=\mathbf{0}$, $P(0|0)=0.1I_2$, and applying Theorem 2, we obtain the optimal input white noise smoothers $\hat{w}(t|t-1)$ and $\hat{w}(t|t+3)$. The simulation results are shown in Fig. 5 and Fig. 6, where the points denote the estimates $\hat{w}(t|t+1)$ or $\hat{w}(t|t+3)$, and the lines denote the true values $w(t)$. We see that the accuracy of $\hat{w}(t|t+3)$ is higher than that of $\hat{w}(t|t+1)$. The error variances are shown in Fig. 4, where we see that $P_w(t|t+3)\leqslant P_w(t|t+1)$.

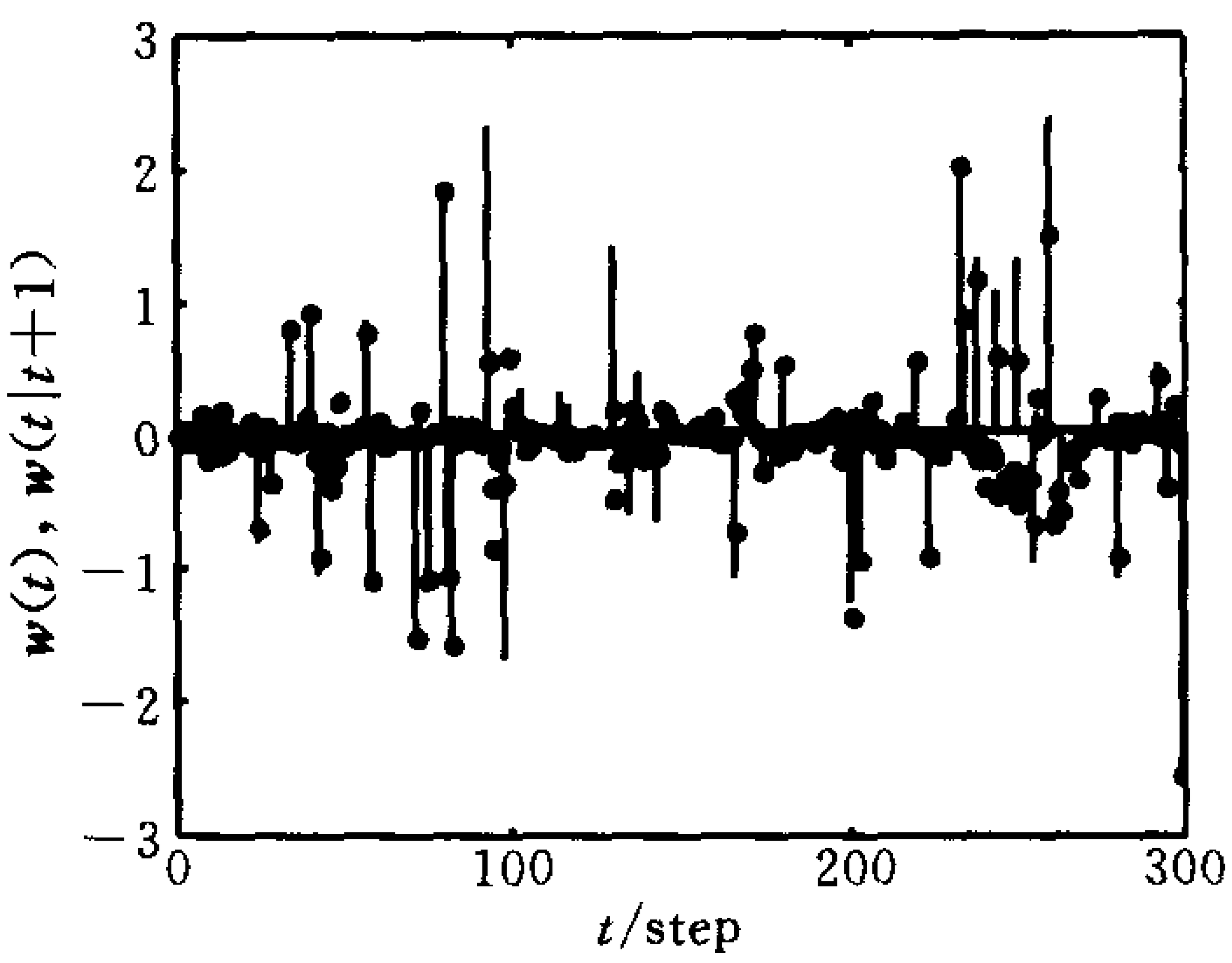

Fig. 5 Bernoulli-Gaussian input white noise $w(t)$ and smoother $\hat{w}(t|t+1)$

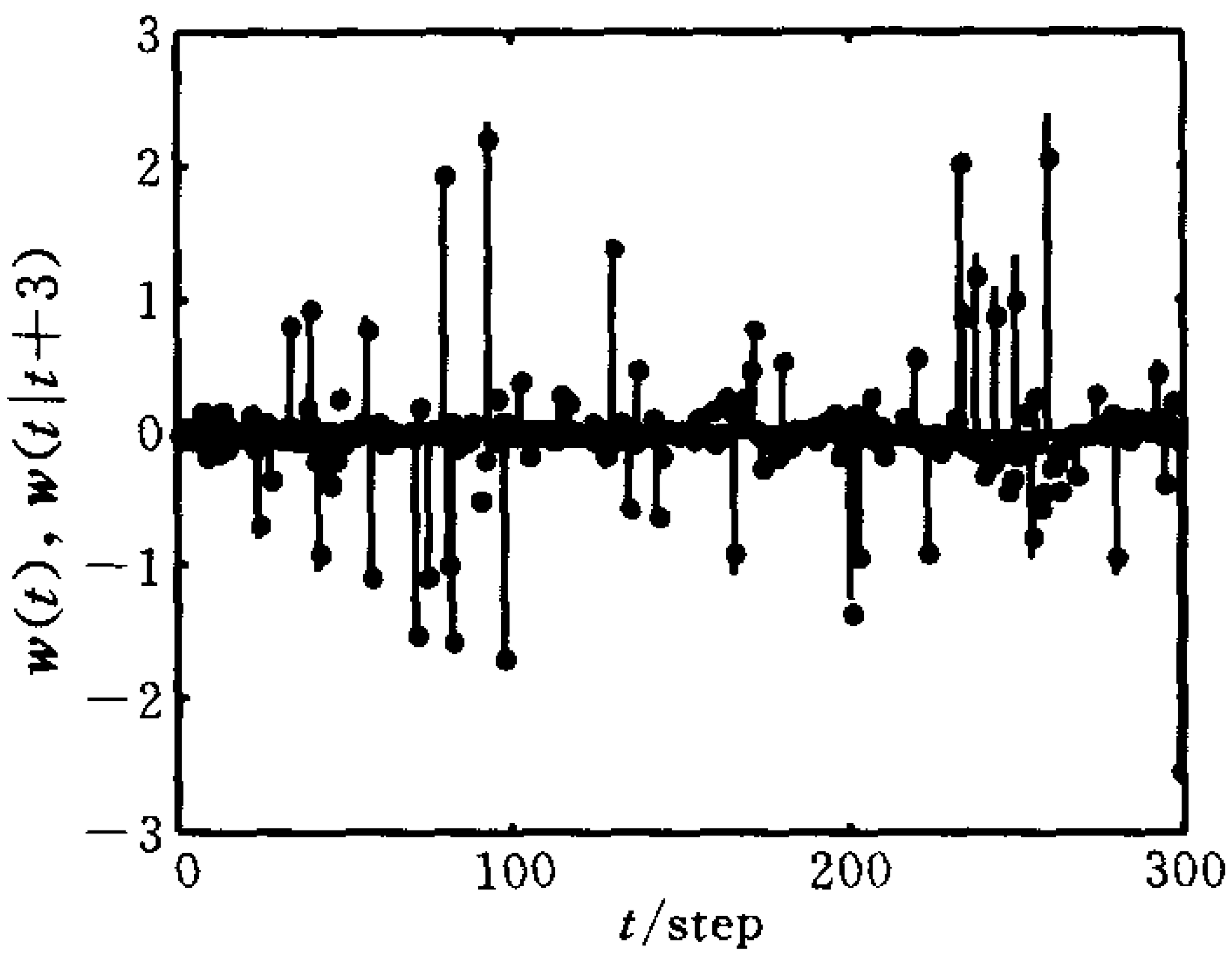

Fig. 6 Bernoulli-Gaussian input white noise $w(t)$ and smoother $\hat{w}(t|t+1)$

6 Conclusion

A unified and general white noise estimation theory has been presented based on the Kalman filter. It overcomes the drawback and limination of Mendel's optimal input white noise estimators and Deng's steady-state white noise estimation theory. It has solved the following five problems: input and measurement white noise estimation; white noise estimation for both time-varying and time-invariant systems; optimal and steady-state white noise estimatiors; white noise innovation filters and Wiener filters; white noise filtering, smoothing and prediction. It can be applied to oil seismic exploration, communications, signal processing and state estimation, and provides a new tool to solve the state and signal estimation problems.

References

1 Mendel J M. White-noise estimators for seismic date processing in oil exploration. *IEEE Transactions on Automatic Control*, 1977, **22**(5): 694~706

2 Mendel J M. Optimal Seismic Deconvolution: An Estimation-Based Approach. New York: Academic Press, 1983. 1~93

3 Mendel J M, Kormylo J. New fast optimal white-noise estimators for deconvolution. *IEEE Transactions on Geoscience Electronics*, 1977, **15**(1): 32~41

4 Mendel J M. Minimum variance deconvolution. *IEEE Transactions on Geoscience and Remote Sensing*, 1981, **19**(3): 161~171

5 Deng Zi-Li, Zhang Huan-Shui, Liu Shu-Jun, Zhou Liu. Optimal and self-tuning white noise estimators with applications to deconvolution and filtering problems. *Automatica*, 1996, **32**(2): 199~216

6 Deng Zi-Li. Optimal Filtering Theory and Applications——Modern Time Series Analysis Method. Harbin: Harbin Institute of Technology Press, 2000. 1~378 (in Chinese)

7 Deng Zi-Li, Xu Yan. Descriptor Wiener state estimators. *Automatica*, 2000, **36**(9): 1761~1766

8 Deng Zi-Li, Liu Yu-Mei. Descriptor Kalman estimators. *International Journal of Systems Science*, 1999, **30**(11): 1205~1212

9 Deng Zi-Li. Time-domain approaches to multichannel optimal deconvolution. *International Journal of Systems Science*, 2000, **31**(6): 787~796

10 Deng Zi-Li, Guo Jin-Zhu, Xu Yan. ARMA optimal recursive state estimators for descriptor systems. *Acta Automatica Sinica*, 2000, **26**(2): 240~244(in Chinese)

11 Deng Zi-Li, Guo Jin-Zhu, Sun Shu-Li. New approach to Wiener deconvolution filters design. *Acta Automatica Sinica*, 2001, **27**(2): 241~246(in Chinese)

12 Anderson B D O, Moore J B. Optimal Filtering. Englewood Cliffs, New Jersey: Prentice-Hall, 1979. 1~305

DENG Zi-Li Professor of Heilongjiang University. His research interests include state estimation, signal processing, and time series analysis.

XU Yan Received her Master's degree from Heilongjiang University in 1998. She is currently an associate professor. Her research interests include optimal filtering and deconvolution.

基于 Kalman 滤波的白噪声估计理论

邓自立 许 燕
（黑龙江大学应用数学研究所 哈尔滨 150080）
（黑龙江大学自动化系 哈尔滨 150080）
（E-mail：dzl@hlju.edu.cn）

摘 要 应用 Kalman 滤波方法，首次提出了一种统一的和通用的白噪声估计理论. 它可统一处理线性离散时变和定常随机系统的输入白噪声和观测白噪声的滤波、平滑和预报问题. 提出了最优和稳态白噪声估值器，且提出了白噪声新息滤波器和 Wiener 滤波器. 它们可应用于石油勘探地震数据处理，且为解决状态和信号估计问题提供一种新工具. 两个仿真例子说明了其有效性.

关键词 输入白噪声估值器，观测白噪声估值器，反卷积，反射地震学，Kalman 滤波方法
中图分类号 O211.64
