Performance of Minimum-Variance Deconvolution Filter

CHONG-YUNG CHI, MEMBER, IEEE, AND JERRY M. MENDEL, FELLOW, IEEE

Abstract—Recently, we observed zero phase and undershoot patterns in data processed by a minimum-variance deconvolution (MVD) filter. These observations motivated a careful analysis of the MVD filter, which, as we demonstrate in this paper, explains both the zero phase and undershoot patterns. This analysis also connects the MVD filter with the well-known prediction-error filter [6], and Berkhout's two-sided least-squares inverse filter [7]. We show that the performance of the MVD filter depends heavily on the bandwidth of the source wavelet and signal-to-noise ratio, and only slightly on data length.

I. Introduction

MENDEL [1]-[3] and Mendel and Kormylo [4] have developed a minimum variance deconvolution (MVD) filter for the following discrete time convolutional model:

$$z(k) = y(k) + n(k) = \mu(k) * V(k) + n(k)$$

$$= \sum_{j=1}^{k} \mu(j)\, V(k-j) + n(k), \qquad k = 1, 2, \cdots, N \tag{1}$$

in which $y(k)$ is a noise free signal (e.g., output of a communication channel, seismic trace); $n(k)$ is measurement noise which accounts for physical effects not explained by $y(k)$, as well as sensor noise; $V(i)$, $i = 0, 1, 2, \cdots$, is a sequence associated with the signal distorting system (e.g., impulse response of communication channel or seismic source); and $\mu(j)$, $j = 1, 2, \cdots$, is the desired signal sequence (e.g., message, reflectivity sequence). The equation (1) model occurs in many different fields, including astronomy [8], communication systems [9]-[12], and reflection seismology [3], [5], [13].

In MVD, measurements $z(1), z(2), \cdots, z(N)$ are processed linearly so as to obtain a fixed interval smoothed estimate, $\hat{\mu}(k|N)$, of $\mu(k)$ for all $k = 1, 2, \cdots, N$. Recently, we observed zero phase and undershoot patterns in deconvolved data obtained from the MVD filter. These observations motivated a careful analysis of the MVD filter, which, as we demonstrate in this paper, explains both the zero phase and undershoot patterns. This analysis also connects the MVD filter with the well-known prediction-error filter (PEF) [6], and Berkhout's two-sided least-squares inverse filter [7].

The application of the MVD filter requires that $V(k)$ and all statistical parameters for $\mu(k)$ and $n(k)$ be given. In practice, some or all of these quantities may not be known. For example, they are often all unknown in the seismic deconvolution problem, and must somehow be estimated before one can use the MVD filter (see, for example, Mendel [3] and Chi, Mendel, and Hampson [16]). In this paper, we assume that all the information that is needed to implement the MVD filter is known *a priori* and then concentrate on studying the performance of the MVD filter.

In Section II we briefly describe the modeling assumptions associated with convolutional model (1). In Section III we summarize the MVD filter, because its structure is needed in the rest of the paper. In Section IV we compute a number of important filter impulse responses and establish precise relationships between the prediction-error filter and an innovations filter, and between Berkhout's infinite-length two-sided least-squares inverse filter and the steady state MVD filter. In Section V we quantify the zero phase and undershoot patterns observed in deconvolved data obtained from the MVD filter. In Section VI we examine convergence aspects of MVD results and demonstrate the strong effect signal-to-noise ratio has on MVD filter performance. Computer simulations which confirm most of the paper's theoretical results are given in Section VII.

II. Model Assumptions

In convolutional model (1), we assume, as in reflection seismology, that $\mu(k)$ is zero mean and white, with variance

$$E\{\mu^2(k)\} = \sigma_\mu^2. \tag{2}$$

Additionally, noise $n(k)$ is assumed to be zero mean, white, and Gaussian, with variance

$$E\{n^2(k)\} = R \tag{3}$$

and $n(k)$ is assumed to be independent of $\mu(k)$.

Wavelet $V(k)$ (i.e., the impulse response of the signal-distorting system) is assumed to be an nth-order autoregressive moving average (ARMA), whose z-transform is

$$V(z) = \frac{\sum_{i=1}^{n} \beta_i z^{-i+1}}{1 - \sum_{i=1}^{n} \alpha_i z^{-i}}. \tag{4}$$

In MVD we represent (1) and (4) using the following state-variable model:

$$\boldsymbol{x}(k) = \Phi \boldsymbol{x}(k-1) + \gamma \mu(k) \tag{5}$$

and

Manuscript received January 9, 1984; revised April 23, 1984. This work was supported by the National Science Foundation under Grant NSF ECS-8200882, and by the sponsors of the U.S.C. Geo-Signal Processing Program.

C.-Y. Chi was with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089. He is now with the Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109.

J. M. Mendel is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089.

0096-3518/84/1200-1145$01.00

$$z(k) = \boldsymbol{h}'\boldsymbol{x}(k) + n(k), \tag{6}$$

where Φ is an $n \times n$ matrix, $\boldsymbol{\gamma}$ and $\boldsymbol{h}$ are $n \times 1$ vectors, and $\boldsymbol{\Phi}$, $\boldsymbol{\gamma}$, and $\boldsymbol{h}$ are functions of α_i and β_i. Of course, given a transfer function of a linear time-invariant system, there exist many $(\Phi, \boldsymbol{\gamma}, \boldsymbol{h})$'s which generate the same output $z(k)$ (e.g., [3], [17]). For example, Φ, $\boldsymbol{\gamma}$, $\boldsymbol{h}$ for the "controllable canonical form" are

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_n & \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_1 \end{bmatrix},$$

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \boldsymbol{h} = \begin{pmatrix} \beta_n \\ \beta_{n-1} \\ \vdots \\ \beta_1 \end{pmatrix}.$$

One can easily show that the impulse response of the state-variable model, $\boldsymbol{h}'\Phi^k\boldsymbol{\gamma}$ $(k = 0, 1, \cdots)$, equals $V(k)$ associated with the ARMA transfer function (4).

For purposes of later analyses, we express convolutional model (1) in matrix form, as

$$\boldsymbol{z} = V\boldsymbol{\mu} + \boldsymbol{n} \tag{7}$$

where

$$\boldsymbol{z} = \text{col}\,(z(1), z(2), \cdots, z(N)), \tag{8}$$

$$V = \begin{bmatrix} V(0) & 0 & \cdots & 0 \\ V(1) & V(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V(N-1) & V(N-2) & \cdots & V(0) \end{bmatrix}, \tag{9}$$

$$\boldsymbol{\mu} = \text{col}\,[\mu(1), \mu(2), \cdots, \mu(N)], \tag{10}$$

$$\boldsymbol{n} = \text{col}\,[n(1), n(2), \cdots, n(N)], \tag{11}$$

and N is the total number of measurements. Additionally, we assume that V is invertible, i.e., $V(0) = \beta_1 \neq 0$ (see [3] for the case when $V(0) = 0$).

III. MVD Filter

The MVD filter estimates the input signal $\mu(k)$ from the measurements, $\boldsymbol{z}$. It is well known (e.g., [3]) that the linear minimum variance estimator of $\boldsymbol{\mu}$, $\hat{\boldsymbol{\mu}}$, is given by

$$\hat{\boldsymbol{\mu}} = E\{\boldsymbol{\mu}\boldsymbol{z}'\}\,[E\{\boldsymbol{z}\boldsymbol{z}'\}]^{-1}\,\boldsymbol{z}. \tag{12}$$

From (7) and our modeling assumptions about μ and n, we determine that

$$E\{\boldsymbol{\mu}\boldsymbol{z}'\} = E\{\boldsymbol{\mu}(V\boldsymbol{\mu} + \boldsymbol{n})'\} = \sigma_\mu^2 V' \tag{13}$$

and

$$E[\boldsymbol{z}\boldsymbol{z}'] \triangleq \Omega = E\,[(V\boldsymbol{\mu} + \boldsymbol{n})(V\boldsymbol{\mu} + \boldsymbol{n})'] = \sigma_\mu^2\, VV' + RI. \tag{14}$$

Substituting (13) and (14) into (12), and then using (7), we determine that

$$\hat{\boldsymbol{\mu}} = \sigma_\mu^2\, V'\Omega^{-1}\, \boldsymbol{z} = \sigma_\mu^2\, V'\Omega^{-1}\,(V\boldsymbol{\mu} + \boldsymbol{n}) = \hat{\boldsymbol{\mu}}_s + \boldsymbol{e}, \tag{15}$$

where

$$\hat{\boldsymbol{\mu}}_s = \sigma_\mu^2\, V'\Omega^{-1}\, V\boldsymbol{\mu} \tag{16}$$

and

$$\boldsymbol{e} = \sigma_\mu^2\, V'\Omega^{-1}\boldsymbol{n}. \tag{17}$$

From (15) and (14), it is straightforward to show that the error covariance matrix for estimator $\hat{\boldsymbol{\mu}}$ is

$$E\{[\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}]\,[\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}]'\} = E\,[\tilde{\boldsymbol{\mu}}\tilde{\boldsymbol{\mu}}'] = \sigma_\mu^2 I - \sigma_\mu^4\, V'\Omega^{-1}\, V. \tag{18}$$

Matrices V and Ω are $N \times N$, and can be quite large; this makes the calculation and inversion of Ω costly. Instead of computing $\hat{\boldsymbol{\mu}}$ via the first term on the right-hand side of (15), it can be computed using Mendel's MVD algorithm, which we summarize next. Because this algorithm uses a Kalman filter, we state the Kalman filter equations first.

- *Predictor:*

$$\hat{\boldsymbol{x}}(k|k-1) = \Phi\hat{\boldsymbol{x}}(k-1|k-1) \tag{19}$$

$$P(k|k-1) = \Phi P(k-1|k-1)\,\Phi' + \sigma_\mu^2\boldsymbol{\gamma}\boldsymbol{\gamma}'. \tag{20}$$

- *Innovations:*

$$\tilde{z}(k|k-1) = z(k) - \boldsymbol{h}'\hat{\boldsymbol{x}}(k|k-1) \tag{21}$$

$$\eta(k) = \boldsymbol{h}'P(k|k-1)\,\boldsymbol{h} + R. \tag{22}$$

- *Corrector:*

$$\boldsymbol{K}(k) = P(k|k-1)\,\boldsymbol{h}\,\eta^{-1}(k) \tag{23}$$

$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}(k)\,\tilde{z}(k|k-1) \tag{24}$$

and

$$P(k|k) = [I - \boldsymbol{K}(k)\,\boldsymbol{h}']\;P(k|k-1). \tag{25}$$

In these equations, $\hat{\boldsymbol{x}}(k|j)$ denotes the state estimate of $\boldsymbol{x}(k)$ given $z(1), z(2), \cdots, z(j)$, and $P(k|j)$ is the corresponding error covariance matrix. Additionally, $\tilde{z}(k|k-1)$ is the innovations process, $\eta(k)$ is its variance, and $\boldsymbol{K}(k)$ is the Kalman gain vector.

The first step of Mendel's MVD algorithm is to pass all of the data, $z(1), z(2), \cdots, z(N)$, through the Kalman filter. Values of $\tilde{z}(k|k-1)$, $\eta(k)$, and $\boldsymbol{K}(k)$ are stored at each iteration of the Kalman filter. The second step of the MVD algorithm is to process the innovations, $\tilde{z}(k|k-1)$, $k = 1, 2, \cdots, N$, by the *anticausal* filter

$$\boldsymbol{r}(k|N) = [I - \boldsymbol{K}(k)\,\boldsymbol{h}']'\,\Phi'\,\boldsymbol{r}(k+1|N) + \boldsymbol{h}\,\eta^{-1}(k)\,\tilde{z}(k|k-1), \tag{26}$$

$k = N, N-1, \cdots, 1$, where $\boldsymbol{r}(N+1|N) = \boldsymbol{0}$. Finally, estimates of $\mu(k)$, $\hat{\mu}(k|N)$ are obtained as

$$\hat{\mu}(k|N) = \sigma_\mu^2\boldsymbol{\gamma}'\,\boldsymbol{r}(k|N). \tag{27}$$

Obviously, the MVD filter is *noncausal* and *time varying*.

IV. Steady-State MVD Filter

For a time-invariant wavelet and stationary noises, the Kalman gain matrix, as well as the error-covariance matrices, will reach steady-state values. When this occurs, both the Kalman

filter and anticausal filter (26) and (27) become time invariant, and we then refer to the MVD filter as a *steady-state MVD filter*. In the sequel we will examine properties of this steady-state filter. To begin, though, we need to establish formulas for the impulse responses (IR's) of a number of subsystems which are associated with the interconnections of the steady-state Kalman and anticausal filters.

In (15) we decomposed $\hat{\boldsymbol{\mu}}$ into its signal and noise components, $\hat{\boldsymbol{\mu}}_s$ and $\boldsymbol{e}$, respectively. Note, from (16) and (17), that $\hat{\boldsymbol{\mu}}_s$ depends only on $\boldsymbol{\mu}$, whereas $\boldsymbol{e}$ depends only on $\boldsymbol{n}$. Fig. 1 shows more clearly how $\hat{\mu}_s(k|N)$ and $e(k)$ [the components of $\hat{\boldsymbol{\mu}}_s$ and $\boldsymbol{e}$, respectively] are constructed from $\mu(k)$ and $n(k)$. Observe that both $z(k)$ and $\tilde{z}(k|k-1)$ have been decomposed into signal and noise components, i.e., $z(k) = y(k) + n(k)$ and $\tilde{z}(k|k-1) = \tilde{z}_s(k|k-1) + \tilde{z}_n(k|k-1)$. Observe, also, that $V_2(k)$ is the IR of the steady-state Kalman filter whose input and output are $z(k)$ and $\tilde{z}(k|k-1)$, respectively, and that $V_3(k)$ is the IR of the anticausal steady-state filter (26) and (27), whose input and output are $\tilde{z}(k|k-1)$ and $\hat{\mu}(k|N)$, respectively. In this section we compute formulas for $V_2(k)$, $V_1(k) = V(k) * V_2(k)$, $V_3(k)$, and $V_{MV}(k) = V_2(k) * V_3(k)$.

We begin by deriving the impulse response of the steady-state Kalman (i.e., innovations) filter $V_2(k)$. From (19), (21), and (24) we see that

$$\hat{\boldsymbol{x}}(k|k-1) = \Phi(I - \boldsymbol{K}\boldsymbol{h}')\hat{\boldsymbol{x}}(k-1|k-2) + \Phi\boldsymbol{K}z(k-1) \tag{28}$$

where $\boldsymbol{K}$ is the steady state Kalman gain. From (28) and (21), it is straightforward to show that

$$V_2(k) = \begin{cases} 1, & k = 0 \\ -\boldsymbol{h}'[\Phi(I - \boldsymbol{K}\boldsymbol{h}')]^{k-1}\Phi\boldsymbol{K}, & k > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{29}$$

Fact 1: The negative of $V_2(k)$ in (29) represents a closed-form formula for the coefficients of an infinite-length prediction-error filter.

Proof: Using the fact, from (29), that $V_2(0) = 1$, we can express the innovations as

$$\tilde{z}(k|k-1) = z(k) * V_2(k) = z(k) - \sum_{i=1}^{\infty}[-V_2(i)]\,z(k-i). \tag{30}$$

Note, also, that $\tilde{z}(k|k-1) = z(k) - \hat{z}(k|k-1)$, which means that the second term on the right-hand side of (30) equals $\hat{z}(k|k-1)$, but $\hat{z}(k|k-1)$ is a minimum variance predicted estimate of $z(k)$. Therefore, $\tilde{z}(k|k-1)$ is the output of a PEF [6]. The coefficients of this PEF are $1, -V_2(1), -V_2(2), \cdots$. ∎

We believe that this is the first time a closed-form formula has been given for the PEF coefficients. To obtain these coefficients, one must first obtain a state-variable model from source wavelet data (this establishes Φ, $\boldsymbol{\gamma}$ and $\boldsymbol{h}$), then compute the steady state Kalman gain $\boldsymbol{K}$, and finally compute $V_2(k)$ via (29).

Next, we determine $V_1(k)$ where

$$V_1(k) = V(k) * V_2(k). \tag{31}$$

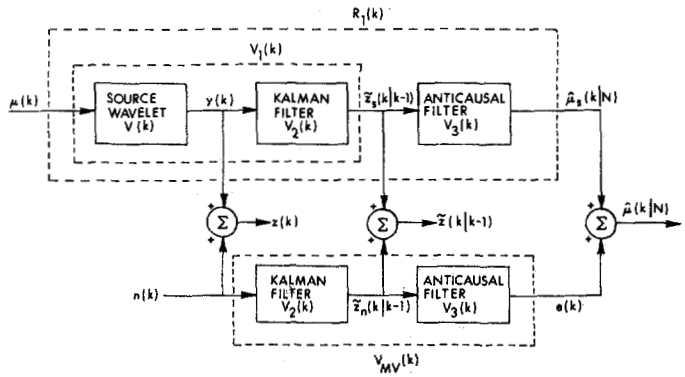

Fig. 1. A block diagram interpretation of MVD filtering.

From Fig. 1, we see that

$$\begin{aligned}\tilde{z}(k|k-1) &= \tilde{z}_s(k|k-1) + \tilde{z}_n(k|k-1) \\ &= \mu(k) * V_1(k) + n(k) * V_2(k).\end{aligned} \tag{32}$$

We determine $V_1(k)$ by letting $n(k) = 0$ in the Kalman filter equations. When $n(k) = 0$

$$z(k) = y(k) = \boldsymbol{h}'\boldsymbol{x}(k) \tag{33}$$

and

$$\tilde{z}(k|k-1) = y(k) - \boldsymbol{h}'\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{h}'\hat{\boldsymbol{x}}(k|k-1) \tag{34}$$

where

$$\tilde{\boldsymbol{x}}(k|k-1) = \boldsymbol{x}(k) - \hat{\boldsymbol{x}}(k|k-1). \tag{35}$$

From (5), (28), (33), and (35), we find that $\tilde{\boldsymbol{x}}(k|k-1)$ satisfies the equation

$$\begin{aligned}\tilde{\boldsymbol{x}}(k|k-1) &= \Phi\boldsymbol{x}(k-1) + \boldsymbol{\gamma}\mu(k) \\ &\quad - \Phi(I - \boldsymbol{K}\boldsymbol{h}')\hat{\boldsymbol{x}}(k-1|k-2) - \Phi\boldsymbol{K}\boldsymbol{h}'\boldsymbol{x}(k-1) \\ &= \Phi(I - \boldsymbol{K}\boldsymbol{h}')\tilde{\boldsymbol{x}}(k-1|k-2) + \boldsymbol{\gamma}\mu(k).\end{aligned} \tag{36}$$

Equations (36) and (34) imply that

$$V_1(k) = \begin{cases} \boldsymbol{h}'[\Phi(I - \boldsymbol{K}\boldsymbol{h}')]^k\boldsymbol{\gamma}, & k \geqslant 0 \\ 0, & \text{otherwise.} \end{cases} \tag{37}$$

Next, we determine the impulse response of the steady-state anticausal filter. Note, from Fig. 1, that $\hat{\mu}(k|N) = V_3(k) * \tilde{z}(k|k-1)$. From (26) and (27) we determine that

$$V_3(k) = \begin{cases} \dfrac{\sigma_\mu^2}{\eta}\boldsymbol{\gamma}'[(I - \boldsymbol{K}\boldsymbol{h}')'\Phi']^{-k}\boldsymbol{h} \\ \quad = \dfrac{\sigma_\mu^2}{\eta}\boldsymbol{h}'[\Phi(I - \boldsymbol{K}\boldsymbol{h}')]^{-k}\boldsymbol{\gamma}, & k \leqslant 0 \\ 0, & \text{otherwise} \end{cases} \tag{38}$$

where η is the steady-state value of $\eta(k)$. Comparing (37) and (38), we see that

$$V_3(k) = \frac{\sigma_\mu^2}{\eta}V_1(-k). \tag{39}$$

Finally, we compute the IR of the steady-state MVD filter, $V_{MV}(k)$, which, from Fig. 1, (39), and (31), is

$$V_{MV}(k) = V_2(k) * V_3(k) = \frac{\sigma_\mu^2}{\eta} V_2(k) * V_1(-k)$$

$$= \frac{\sigma_\mu^2}{\eta} V_2(k) * V(-k) * V_2(-k). \quad (40)$$

Fact 2: The steady-state MVD filter, whose IR is given by $V_{MV}(k)$, is exactly the same as Berkhout's infinite-length two-sided least-squares inverse filter.

Proof: The Fourier transform of $V_{MV}(k)$, $V_{MV}(\omega)$, is

$$V_{MV}(\omega) = \frac{\sigma_\mu^2}{\eta} V^*(\omega) \left| V_2(\omega) \right|^2. \quad (41)$$

From (30) and (1), and the fact that $\tilde{z}(k \mid k-1)$ is white noise, it follows that

$$\eta = \Phi_z(\omega) \left| V_2(\omega) \right|^2 \quad (42)$$

and

$$\Phi_z(\omega) = \sigma_\mu^2 \left| V(\omega) \right|^2 + R. \quad (43)$$

Substitute (42) and (43) into (41) to see that

$$V_{MV}(\omega) = \frac{\sigma_\mu^2 V^*(\omega)}{\Phi_z(\omega)} = \frac{\sigma_\mu^2 V^*(\omega)}{\sigma_\mu^2 \left| V(\omega) \right|^2 + R}, \quad (44)$$

which is exactly the same frequency response as that of Berkhout's infinite-length two-sided least-squares inverse filter [7]. ∎

The steady-state MVD filter is a recursive implementation of Berkhout's infinite-length filter. Additionally, the MVD filter is applicable to time-varying and nonstationary systems, whereas his filter is not.

V. Properties of the Steady-State MVD Filter

In this section we describe some properties associated with the steady-state MVD filter. These properties were observed first in simulations of the MVD filter, but can be predicted using the results just obtained in Section IV.

Property 1: Let $\hat{\mu}_s(k \mid N)$ denote the "signal" component in $\hat{\mu}(k \mid N)$. Then

$$\hat{\mu}_s(k \mid N) = \mu(k) * R_1(k) \quad (45)$$

where $R_1(k)$ is the autocorrelation function

$$R_1(k) = \frac{\sigma_\mu^2}{\eta} V_1(k) * V_1(-k) \quad (46)$$

and

$$R_1(\omega) = \frac{\sigma_\mu^2 \left| V(\omega) \right|^2}{\sigma_\mu^2 \left| V(\omega) \right|^2 + R}. \quad (47)$$

This means that $\hat{\mu}_s(k \mid N)$ is a zero-phase waveshaped version of $\mu(k)$.

Proof: Expressions for the signal component of $\hat{\mu}(k \mid N)$, $\hat{\mu}_s(k \mid N)$, are easily derived by analyzing the top path in Fig. 1 and using (39). Specifically,

$$\hat{\mu}_s(k \mid N) = \mu(k) * V_1(k) * V_3(k)$$

$$= \mu(k) * \left[\frac{\sigma_\mu^2}{\eta} V_1(k) * V_1(-k) \right] \quad (48)$$

and

$$\hat{\mu}_s(k \mid N) = \mu(k) * R_1(k). \quad (49)$$

Comparing (48) and (49), we determine that $R_1(k)$ is given by (46). Note, also, from Fig. 1 and (44), that

$$R_1(\omega) = V(\omega) V_2(\omega) V_3(\omega)$$

$$= V(\omega) V_{MV}(\omega) = \frac{\sigma_\mu^2 \left| V(\omega) \right|^2}{\sigma_\mu^2 \left| V(\omega) \right|^2 + R}. \quad (50)$$

Observe that $R_1(\omega)$ has zero phase. ∎

Recently, Mendel [15] showed how to accomplish minimum-variance waveshaping by 1) deconvolving $z(k)$ to obtain $\hat{\mu}(k \mid N)$, and 2) convolving $\hat{\mu}(k \mid N)$ with desired waveshaper $d(k)$. Property 1 demonstrates that the final waveshape will not be $d(k)$; instead it will be $R_1(k) * d(k)$. Of course, the more impulsive $R_1(k)$ is, the closer $R_1(k) * d(k)$ will be to $d(k)$.

Property 2: Let σ_μ^2/R be treated as the MVD filter tuning parameter. Then $R_1(k) \to \delta(k)$ as $\sigma_\mu^2/R \to \infty$. In this case the MVD waveshaper has a signal component equal to $\mu(k) * d(k)$.

Proof: Express (47) as

$$R_1(\omega) = \frac{\left| V(\omega) \right|^2 \sigma_\mu^2/R}{1 + \left| V(\omega) \right|^2 \sigma_\mu^2/R} \quad (51)$$

as $\sigma_\mu^2/R \to \infty$ $R_1(\omega) \to 1$, in which case $R_1(k) \to \delta(k)$. ∎

Property 3: Assume that sequence $\mu(k)$ contains only an isolated spike at $k = k_1$. Then $\left| \hat{\mu}_s(k \mid N) \right| \leq \left| \hat{\mu}_s(k_1 \mid N) \right| < \left| \mu(k_1) \right|$ which means that $\hat{\mu}_s(k_1 \mid N)$ undershoots $\mu(k_1)$.

Proof: Sequence $\mu(k)$ can be expressed as

$$\mu(k) = r(k_1)\, \delta(k - k_1) \quad (52)$$

so that

$$\hat{\mu}_s(k \mid N) = r(k_1)\, R_1(k - k_1). \quad (53)$$

A basic property of an autocorrelation function $R_1(j)$ is that $R_1(0) \geq \left| R_1(j) \right|$ for all $j \neq 0$. The maximum value of $R_1(k - k_1)$ occurs at $k = k_1$; thus,

$$\left| \hat{\mu}_s(k \mid N) \right| \leq \left| \hat{\mu}_s(k_1 \mid N) \right|$$

$$= R_1(0) \left| r(k_1) \right|$$

$$= R_1(0) \left| \mu(k_1) \right|. \quad (54)$$

From (47), we see that

$$0 \leq R_1(\omega) < 1 \qquad \forall\, -\pi \leq \omega \leq \pi. \quad (55)$$

From Parseval's relation and (55), we determine that

$$\sum_{k=-\infty}^{\infty} R_1^2(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R_1^2(\omega)\, d\omega < \frac{1}{2\pi} \int_{-\pi}^{\pi} d\omega = 1, \quad (56)$$

from which it follows that

$$R_1(0) < 1; \quad (57)$$

thus,

$$\left| \hat{\mu}_s(k \mid N) \right| \leq \left| \hat{\mu}_s(k_1 \mid N) \right| < \left| \mu(k_1) \right|. \quad (58)$$

∎

If $\mu(k)$ consists of widely separated spikes at $k = k_1, k_2, \cdots, k_M$, then Property 3 explains why $\hat{\mu}_s(k|N)$ will undershoot $\mu(k)$ at $k = k_1, k_2, \cdots, k_M$. On the other hand, if $\mu(k)$ contains closely spaced spikes, for example, at $k = k_4$ and k_5, then it is quite possible for $|\hat{\mu}_s(k_5|N)| > |\mu(k_5)|$. In this case,

$$\begin{aligned}\hat{\mu}_s(k_5|N) &= \sum_{i=1}^{M} r(k_i)R_1(k_5 - k_i) \quad \text{(by linearity)} \\ &= r(k_4)R_1(k_5 - k_4) + r(k_5)R_1(0) \\ &\quad + \sum_{i \neq 4,5} r(k_i)R_1(k_5 - k_i) \\ &\cong r(k_4)R_1(k_5 - k_4) + r(k_5)R_1(0) \qquad (59)\end{aligned}$$

so that if the first term of (59) is in phase with the second term of (59), the two terms will be larger in amplitude than $r(k_5)\ R_1(0)$; thus, it is quite possible that the terms on the right-hand side of (59) will be larger than $\mu(k_5)$ (in amplitude), even though $|r(k_5)R_1(0)| < |\mu(k_5)|$.

Thus far, we have been focusing on the signal component of $\hat{\mu}(k|N), \hat{\mu}_s(k|N)$. From (15) we note that

$$\hat{\mu}(k|N) = \hat{\mu}_s(k|N) + e(k), \qquad (60)$$

and, from Fig. 1, that

$$e(k) = n(k) * V_{MV}(k). \qquad (61)$$

If a lot of noise is present then $\hat{\mu}(k|N)$ may look quite different from $\hat{\mu}_s(k|N)$, and it will be difficult to observe the zero phase and undershoot properties just described. On the other hand, if not much noise is present, then these properties, that were derived for $\hat{\mu}_s(k|N)$, will also be discernable on $\hat{\mu}(k|N)$.

VI. Analysis of Estimation Error

In this section we study the behavior of the variance of the error between $\hat{\mu}(k|N)$ and $\mu(k)$, in order to learn if any definitive statements can be made about the convergence of $\hat{\mu}(k|N)$ to $\mu(k)$. This error variance is denoted $\sigma^2(k|N)$.

To begin, we define signal-to-noise ratio (SNR) as (note that the definition of SNR seems to vary within the signal processing community)

$$\text{SNR} = \frac{E\{y^2(k)\}}{E\{n^2(k)\}}. \qquad (61)$$

It is straightforward to show that

$$\text{SNR} = \frac{\sigma_\mu^2 E_v}{R}, \qquad (62)$$

where E_v is the energy of the wavelet $V(k)$, i.e.,

$$E_v = \sum_{k=0}^{\infty} V^2(k). \qquad (63)$$

The following results are proven in Chi [14] for *all* k and N.

1) The error variance $\sigma^2(k|N)$ is bounded from below, i.e.,

$$\sigma^2(k|N) \geq \rho_{k,N}^2 > 0 \qquad (64)$$

where

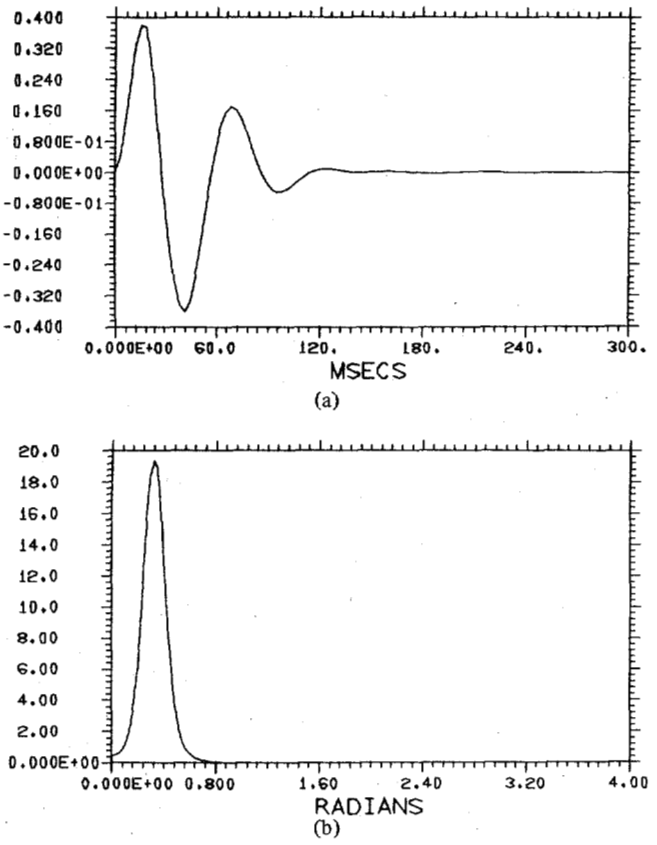

Fig. 2. (a) Fourth-order wavelet and (b) its squared amplitude spectrum.

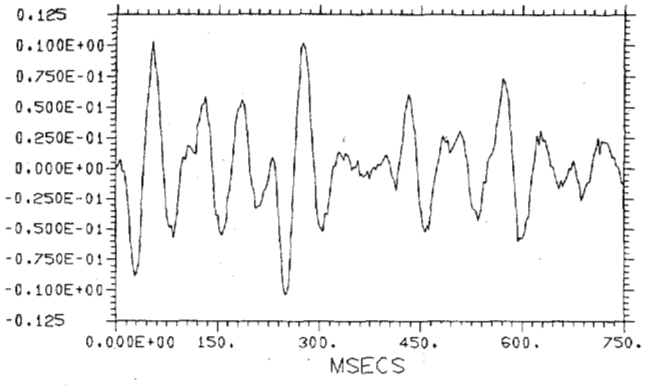

Fig. 3. Synthetic noisy data (SNR = 100).

$$\rho_{k,N} = \left(\frac{\sigma_\mu^2}{1 + \dfrac{\sigma_\mu^2}{R} E_{k,N}}\right)^{1/2} \qquad (65)$$

and

$$E_{k,N} = \sum_{i=0}^{N-k} V^2(i). \qquad (66)$$

For finite SNR, the limiting value of $\sigma^2(k|N)$, as $N \to \infty$, is bounded by ρ^2, where

$$\rho^2 = \frac{\sigma_\mu^2}{1 + \text{SNR}}. \qquad (67)$$

2) $\sigma^2(k|N) \to 0$ if and only if SNR $\to \infty$. This means that,

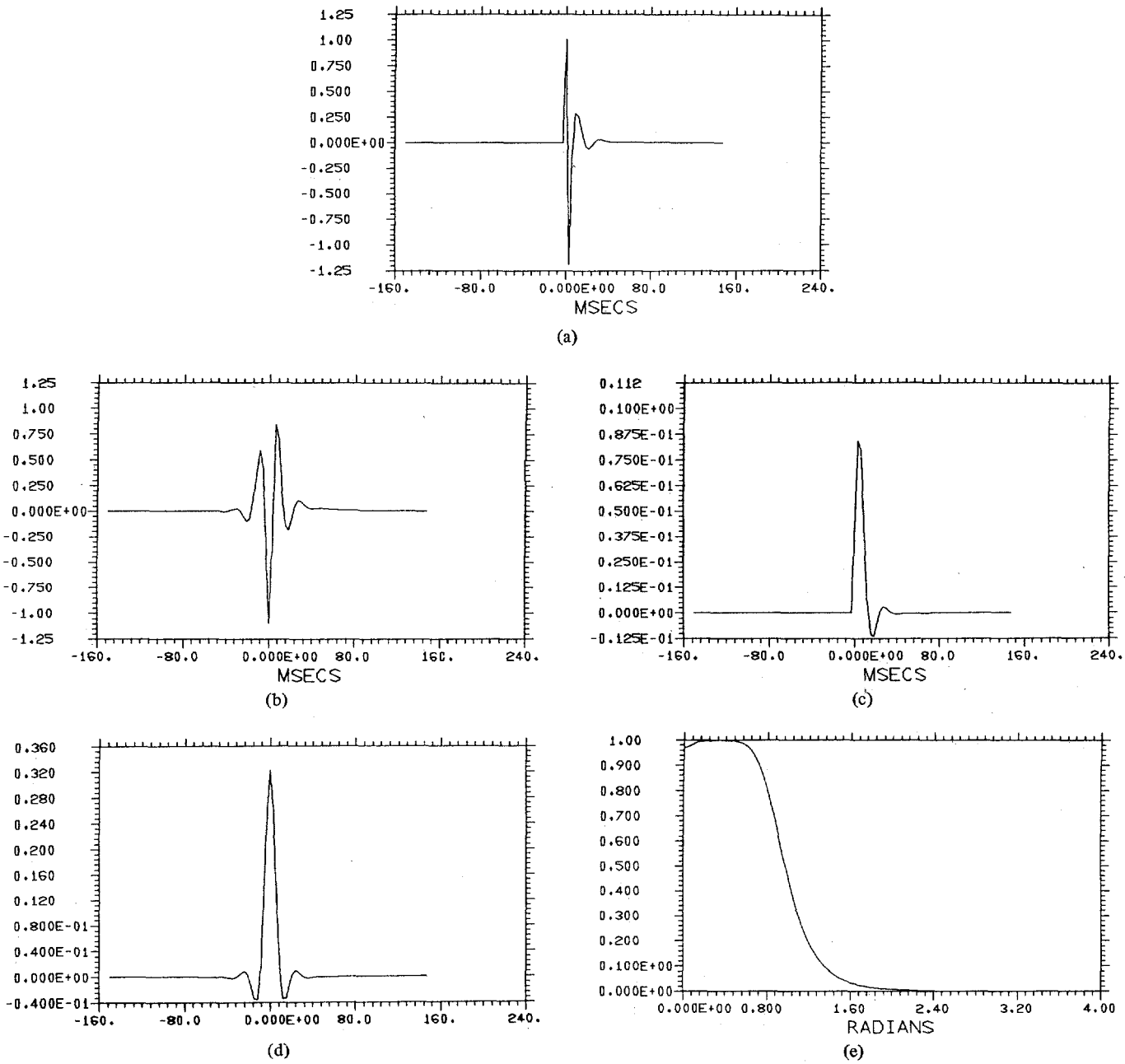

Fig. 4. (a) $V_2(k)$, (b) $V_{MV}(k)$, (c) $V_1(k)$, (d) $R_1(k)$, and (e) $R_1(\omega)$ for SNR = 100.

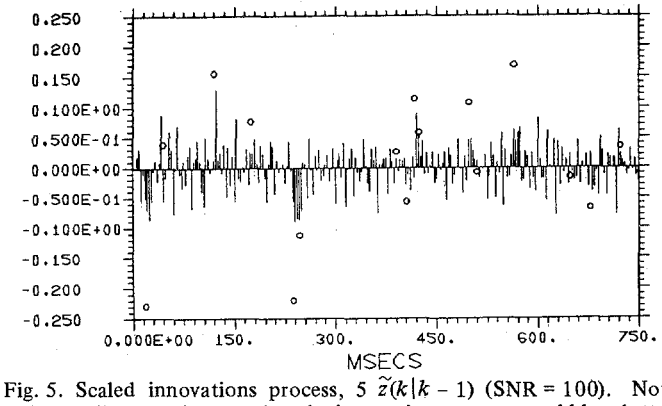

Fig. 5. Scaled innovations process, 5 $\tilde{z}(k|k-1)$ (SNR = 100). Note that scaling was done so that the innovations process could be plotted on the same plot as $\mu(k)$, and that the circles depict true nonzero values of $\mu(k)$.

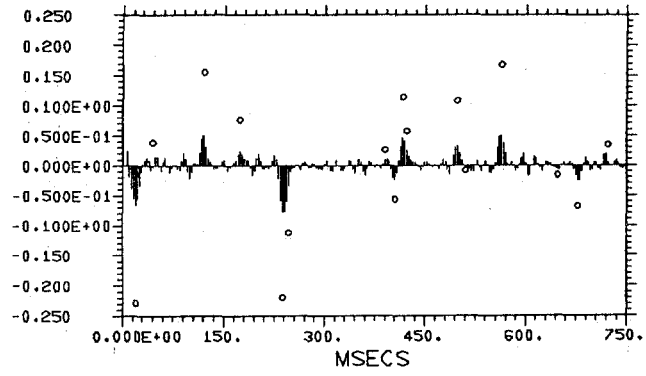

Fig. 6. $\hat{\mu}(k|N)$ (SNR = 100). Circles mark true values of $\mu(k)$ and bars mark the corresponding estimates.

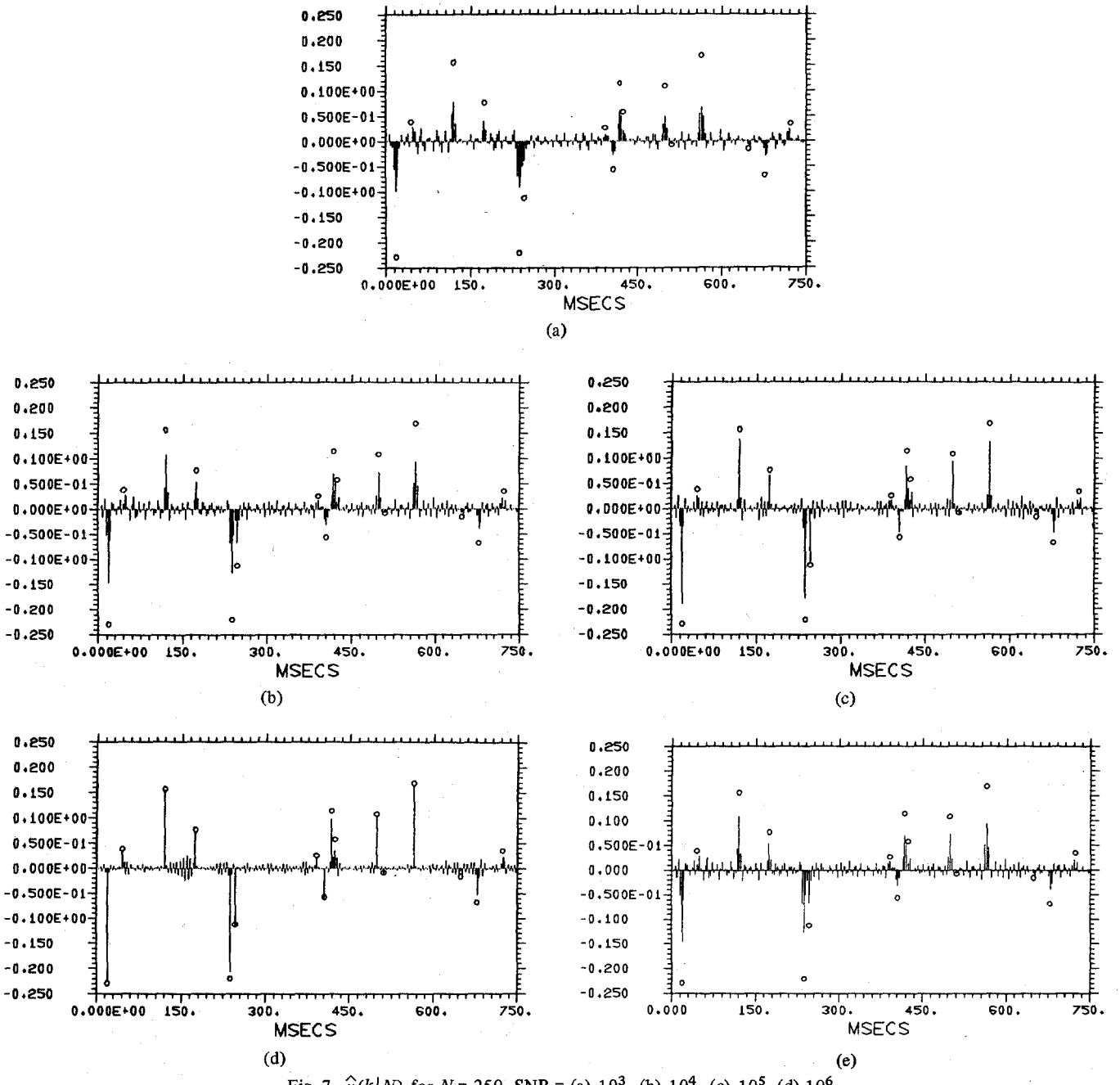

Fig. 7. $\hat{\mu}(k|N)$ for $N = 250$, SNR = (a) 10^3, (b) 10^4, (c) 10^5, (d) 10^6, and (e) for $N = 2500$ and SNR = 10^4. Circles mark true values of $\mu(k)$ and bars mark the corresponding estimates.

in general, $\hat{\mu}(k|N)$ does not converge to $\mu(k)$ in a mean-square sense.

3) $\hat{\mu}(k|N)$ does not even converge to $\mu(k)$ in probability, unless SNR $\to \infty$. Additionally, one can show that

4) In the region of time where both the Kalman and anticausal filters (i.e., the MVD filter) are in steady state, the steady-state value of $\sigma^2(k|N)$, σ^2, is given by

$$\sigma^2 = \sigma_\mu^2 [1 - R_1(0)]. \tag{68}$$

Note, from (54), that $1 - R_1(0)$ equals the undershoot between an isolated spike and its MVD estimate; thus, the larger this undershoot is, the larger σ^2 is.

5) $\sigma^2 \to 0$ if and only if SNR $\to \infty$.

These results demonstrate the strong effect SNR has on MVD filter performance.

VII. Computer Simulations

In this section we provide two examples which support and demonstrate our preceding theoretical results. For each example we generated noisy data by convolving a white Bernoulli-Gaussian input sequence, $\mu(k)$, with a wavelet $V(k)$ and adding white Gaussian noise to those results. We then deconvolved the noisy data using the MVD filter, to obtain $\hat{\mu}(k|N)$.

A Bernoulli-Gaussian sequence [3] can be represented as

$$\mu(k) = r(k)\, q(k) \tag{69}$$

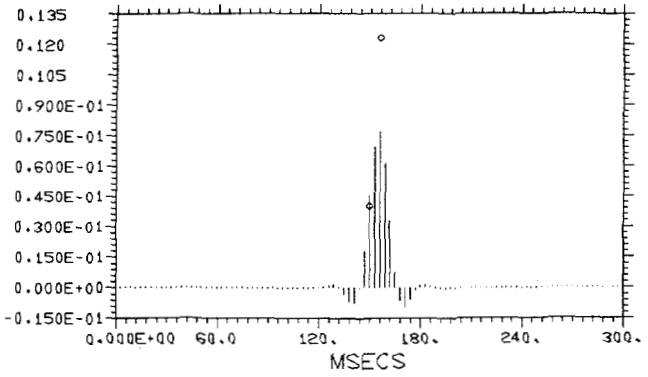

Fig. 8. $\hat{\mu}_s(k)$ (SNR = 50). Circles mark true values of $\mu(k)$ and bars mark the corresponding estimates. The larger true spike is scaled by a factor 0.5.

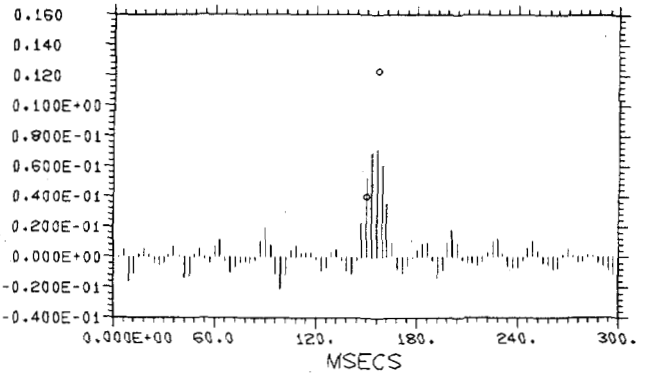

Fig. 9. $\hat{\mu}(k|N)$ (SNR = 50). Circles mark true values of $\mu(k)$ and bars mark the corresponding estimates. The larger true spike is scaled by a factor 0.5.

where $r(k)$ is a white Gaussian amplitude sequence with variance C, and $q(k)$ is a Bernoulli sequence with parameter λ.

Example 1: In this example, we assumed that $\lambda = 0.05$, $C = 0.0225$, $R = 0.1536 \times 10^{-4}$, and $N = 250$, and, we used the narrow-band fourth-order wavelet depicted in Fig. 2(a). The squared amplitude spectrum of $V(k)$ is depicted in Fig. 2(b), and the transfer function of $V(k)$ is

$$V(z) = \frac{0.0378417 - 0.0306517z^{-2}}{1 - 3.4016497z^{-1} + 4.5113732z^{-2} - 2.7553363z^{-3} + 0.6561z^{-4}}. \tag{70}$$

The resulting noisy seismic data, for SNR = 100, are depicted in Fig. 3.

Fig. 4(a)-(e) depicts $V_2(k)$, $V_{MV}(k)$, $V_1(k)$, $R_1(k)$, and $R_1(\omega)$. Observe that $V_2(k)$ is causal with $V_2(0) = 1$; $V_{MV}(k)$ is noncausal, but is not zero phase; $V_1(k)$ is causal and has a much shorter length than $V(k)$; $R_1(k)$ is a zero-phase signal with much shorter length than $V(k)$; and $R_1(\omega)$ [computed from (51)] has a low-pass character. Note that $R_1(0) = 0.3213$, which means that for sparse spikes $\hat{\mu}_s(k)$ will undershoot $\mu(k)$ quite a lot.

Fig. 5, which depicts a scaled innovations process, [i.e., $5\tilde{z}(k|k-1)$], demonstrates that the deconvolved data using a prediction-error filter is very noisy and quite unsatisfactory. Fig. 6, which depicts $\hat{\mu}(k|N)$, demonstrates a more predictable result. Observe that $\hat{\mu}(k|N)$ undershoots $\mu(k)$ by about 70 percent and is comprised of a series of noncausal signals.

Deconvolved results for larger SNR values equal to 10^3, 10^4, 10^5, and 10^6 are depicted in Fig. 7(a)-(d). Observe that, as SNR increases, $\hat{\mu}(k|N)$ becomes more spike-like, and that as SNR $\to \infty$, $\hat{\mu}(k|N) \to \mu(k)$. For most values of SNR $\hat{\mu}(k|N)$ almost always undershoots $\mu(k)$.

Fig. 7(e) and (b) should be compared. They are virtually indistinguishable. Both are for the same SNR = 10^4, but the former is for a very long data record (i.e., $N = 2500$), whereas the latter is for a shorter data record (i.e., $N = 250$).

Fig. 7(a)-(e) supports our theoretical conclusions that MVD filter performance depends very much on SNR but not on data length N.

Chi [14] has repeated this experiment for a broad-band (in frequency) wavelet. Space does not permit us to include his results. Instead, we mention that a broad-band source wavelet can, as predicted by (51), overcome the shortcomings of low SNR. For example, results from a prediction-error filter are quite good; $\hat{\mu}(k|N)$ undershoots $\mu(k)$ much less than for a narrow-band wavelet; and $R_1(k)$ is very spiky so that a noncausal pattern is not observed in $\hat{\mu}_s(k|N)$ or $\hat{\mu}(k|N)$.

Example 2: Our second example demonstrates a case in which overshoot occurs between $\hat{\mu}(k|N)$ and $\mu(k)$. The same λ, C, and $V(k)$ are used in this example as in Example 1; however, SNR = 50. By design, there are only two close spikes in $\mu(k)$. Fig. 8 depicts $\hat{\mu}_s(k|N)$; it undershoots $\mu(k)$ for the larger spike and overshoots $\mu(k)$ for the smaller spike. The same situation occurs for $\hat{\mu}(k|N)$, depicted in Fig. 9. The overshoot behavior depicted in these two figures is due to the close spacing of the two spikes and the broad (in time domain) nature of $R_1(k)$.

VIII. Discussion and Conclusions

Many filters have been developed for seismic deconvolution, channel equalization, etc., which attempt to remove effects of the signal distorting system and noise. The prediction-error, two-sided least-squares, and MVD filters are three such filters. Although their derivations are quite different, we have shown that the former two filters are indeed related to the MVD filter. More specifically, we have shown the following.

1) The steady-state innovations filter is an infinite-length prediction-error filter, whose coefficients are given in closed form by (29).

2) The steady-state MVD filter is equivalent to Berkhout's infinite-length least-squares inverse filter, the coefficients of which can be computed from (40).

3) The signal part of $\hat{\mu}(k|N)$, $\hat{\mu}_s(k|N)$, is obtained by convolving a zero-phase wavelet, $R_1(k)$, with $\mu(k)$, where $R_1(k)$ is an autocorrelation function.

4) $R_1(\omega)$ depends on $\sigma_\mu^2|V(\omega)|^2$ and R [see (51)]; it does not depend on the phase of the source wavelet. When $\sigma_\mu^2|V(\omega)|^2 >> R$ for some band of frequencies, then $R_1(k)$ behaves like a bandpass or low-pass filter. When $\sigma_\mu^2|V(\omega)|^2 >>$

R for a broad range of frequencies, then $R_1(k)$ is quite impulsive so that $\hat{\mu}_s(k|N) \simeq \mu(k)$.

5) When the input $\mu(k)$ is a sparse spike train, $\hat{\mu}(k|N)$ usually undershoots $\mu(k)$.

6) Convergence of the MVD filter depends strongly on SNR instead of N.

From these conclusions we note, finally, that the performance of the MVD filter depends heavily on the wavelet and SNR and only slightly on N.

References

[1] J. M. Mendel, "Minimum-variance deconvolution," *IEEE Trans. Geosci., Remote Sensing*, vol. GE-19, pp. 161–171, 1981.

[2] —, "White-noise estimators for seismic data processing in oil exploration," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 694–706, 1977.

[3] —, *Optimal Seismic Deconvolution: An Estimation-Based Approach*. New York: Academic, 1983.

[4] J. M. Mendel and J. Kormylo, "New fast optimal white-noise estimators for deconvolution," *IEEE Trans. Geosci. Electron.*, vol. GE-5, pp. 32–41, 1977.

[5] E. A. Robinson, "Predictive decomposition of seismic traces," *Geophysics*, vol. 22, pp. 767–778, Oct. 1957.

[6] M. T. Silvia and E. A. Robinson, *Deconvolution of Geophysical Time Series in the Exploration for Oil and Natural Gas*. New York: Elsevier, 1979.

[7] A. B. Berkhout, "Least-squares inverse filtering and wavelet deconvolution," *Geophysics*, vol. 42, pp. 1369–1383, Dec. 1977.

[8] J. D. Scargle, "Studies in astronomical time series analysis: Modeling random processes in the time domain," NASA TM81148, 1979.

[9] C. A. Belfiore and J. H. Park, Jr., "Decision feedback equalization," *Proc. IEEE*, pp. 1143–1156, 1979.

[10] R. W. Lucky, "A survey of the communication theory literature, 1968–1973," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 725–739, 1973.

[11] —, "Techniques of adaptive equalization of digital communication systems," *Bell Syst. Tech. J.*, vol. 45, pp. 255–286, 1966.

[12] —, "Automatic equalization of digital communication systems," *Bell Syst. Tech. J.*, vol. 44, pp. 547–588, 1965.

[13] G. W. Webster, *Deconvolution*, vols. I and II (Geophysics reprint series). Tulsa, OK: Soc. Explor. Geophys., 1978.

[14] C-Y. Chi, "Single-channel and multichannel deconvolution," Ph.D. dissertation, Univ. Southern California, Los Angeles, Apr. 1983.

[15] J. M. Mendel, "Minimum-variance and maximum-likelihood recursive waveshaping," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 599–604, June 1983.

[16] C.-Y. Chi, J. M. Mendel, and D. Hampson, "A computationally-fast approach to maximum-likelihood deconvolution," *Geophysics*, May 1984.

[17] S. A. Tretter, *Introduction to Discrete-Time Signal Processing*. New York: Wiley, 1976.

Chong-Yung Chi (S'83–M'83) was born in Taiwan, Republic of China, on August 7, 1952. He received the B.S. degree from the Tatung Institute of Technology, Taipei, Taiwan, in 1975, the M.S. degree from the National Taiwan University, Taipei, in 1977, and the Ph.D. degree from the University of Southern California, Los Angeles, in 1983, all in electrical engineering.

In September 1979 he was appointed a Teaching/Research Assistant, and later became a Research Assistant in the Department of Electrical Engineering–Systems at the University of Southern California. Since July 1983 he has been a member of the technical staff in the Radar Science and Engineering Section at the Jet Propulsion Laboratory at Pasadena, CA. Currently, he is engaged in the studies of spaceborne radar scatterometer systems and radar image processing. His research interests include digital signal processing, system identification, and estimation theory.

Jerry M. Mendel (S'59–M'61–SM'72–F'78) received the B.S. degree in mechanical engineering and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1959, 1960, and 1963, respectively.

His experience has included teaching courses in electrical engineering at the Polytechnic Institute of Brooklyn, from 1960 to 1963, and has also included various consulting positions. From July 1963 to January 1974 he was with McDonnell Douglas Astronautics Company on a full-time basis. Currently he is Professor and Chairman of Electrical Engineering–Systems at the University of Southern California, Los Angeles. He was Associate Chairman from 1982 to 1984. He teaches courses in estimation theory and seismic data processing for oil exploration, and was Director of the U.S.C. Geo-Signal Processing Program (1980–1983). He has published over 150 technical papers and is author of the monograph *Optimal Seismic Deconvolution: An Estimation-Based Approach* (New York: Academic, 1983), the text *Discrete Techniques of Parameter Estimation: The Equation Error Formulation* (New York: Dekker, 1973), and co-editor (with K. S. Fu of Purdue University) of *Adaptive, Learning and Pattern Recognition Systems* (Academic, 1970).

Dr. Mendel served as Editor of the IEEE Control Systems Society's IEEE TRANSACTIONS ON AUTOMATIC CONTROL, is Consulting Editor of the Control and Systems Theory Series for Marcel Dekker, Inc., and is Associate Editor of *Automatica* and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. He received the SEG 1976 Outstanding Presentation Award for a paper on the application of Kalman Filtering to deconvolution. He is a Distinguished Member of the IEEE Control Systems Society, an elected member of the Administrative Committee and Vice-President for Technical Activities of the IEEE Control Systems Society, a member of the IEEE Geoscience and Remote Sensing Society, the Society of Exploration Geophysicists, the European Association for Exploration Geophysicists, Tau Beta Pi, and Pi Tau Sigma, and a registered Professional Control Systems Engineer in California.