Performance Analysis of a Blind Channel Identification Method Based on Second-Order Statistics

Wanzhi Qiu and Yingbo Hua *

Department of Electrical and Electronic Engineering
The University of Melbourne
Parkville, Victoria 3052, Australia
fax: +61-3-3446678
email: qiw@ee.mu.oz.au, yhua@ee.mu.oz.au

Abstract

A recent work by Tong et al [1] provides a novel second-order statistics based method for blind channel identification. In this paper, we present a first-order perturbation analysis of the large-symbol-size performance of this method. An explicit expression for the mean square error (MSE) of the channel estimate is derived. Both the theoretical and simulation results are used to analyze the statistical performance of this method. A number of new insights are revealed.

1 Introduction

Channel equalization is necessary to eliminate or mitigate the intersymbol interference (ISI) effect in communications. To equalize a channel, its transfer function or impulse response needs to be identified either implicitly or explicitly. Conventional channel identification relies on non-information-bearing training sequences, which takes a significant amount of channel bandwidth. Blind channel identification achieves the same goal without relying on the training sequences.

Most blind channel identification schemes begin by sampling the channel output at the baud rate to produce a stationary channel output sequence for processing. Consequently, higher than second-order statistics (HOS) is required to identify a possibly nonminimum phase channel. A major problem with the HOS based techniques is their slow convergence rate and local extrema, as shown in [2].

*This work has been supported by the Australian Research Council and the Australian Cooperative Research Centre for Sensor Signal and Information Processing.

Recently, Tong *et al* [1] presented a novel method that exploits the cyclostationarity of an over-sampled channel output to identify a possibly nonminimum phase channel without traning sequences. This method is asymptotically exact. Moreover, since it is based solely on the second-order statistics (SOS), this method offers faster convergence rate than the HOS-based techniques.

However, the performance of this method has not been analyzed in depth before. In this paper, we provide a large-symbol-size analysis of this method. In particular, we derive the MSE of the channel estimate based on the first order perturbation. we then investigate, through both theoretical and simulation ways, how the signal-to-noise ratio (SNR), the symbol size, the sampling rate and observation window length affect the estimation accuracy. The rest of this paper is organized as follows. Section 2 summarizes the blind identification method developed in [1] for easy reference. Section 3 provides a couple of preliminary results on perturbations. Section 4 presents an explicit expression for the MSE of channel estimate. Section 5 includes theoretical analysis, as well as numerical examples which verify the theoretical results and show some insights into the identification method. Section 6 gives the conclusions.

2 The Blind Identification Method

Consider a FIR channel which is driven by a sequence of symbols. The impulse response of the channel is assumed to last for p_0 baud (symbol) intervals. The channel output is sampled T_s times per baud interval. As in [1], the fractionally sampled channel outputs over m baud intervals (the observation window)

1058-6393/95 $4.00 © 1995 IEEE

can be expressed by the vector:

$$x(t) = Hs(t) + n(t) \quad t = 1, 2, \ldots$$

where H is a full column rank $mT_s \times d$ generalized complex Sylvester matrix constructed from the channel impulse response (CIR) vecter $h = [h_{p_0}^T, h_{p_0-1}^T, \cdots, h_1^T]^T$ (with ^{T} denoting transpose and the dimension of each CIR block h_k equal to $T_s \times 1$), $s(t)$ the vector of $d(= m + p_0 - 1)$ adjacent input symbols and $n(t)$ the white noise. Assuming that the input symbols are i. i. d. with zero mean and unit variance (without loss of generality) one can write

$$R_x(0) \triangleq E\{x(t)x^H(t)\} = HH^H + \sigma^2 I_{mT_s} \quad (1)$$

$$R_x(1) \triangleq E\{x(t)x^H(t+1)\} = HJ_1H^H + \sigma^2 J'_{T_s} \quad (2)$$

where $E\{\}$ denotes the expectation, ^{H} the complex conjugate transpose, σ^2 the noise variance, I_{mT_s} the identity matrix, J_k $(0 \leq k < d)$ the $d \times d$ "shifting" matrix with 1's on the kth lower sub-diagonal and 0's else where (also $J_{-k} \triangleq J_k^T$), and J'_k the $mT_s \times mT_s$ "shifting" matrix defined in the same way as J_k.

It is the structure shown in (1) and (2) that Tong *et al* exploited to estimate H (up to an unknown constant scalar $e^{j\phi}$) from $R_x(0)$ and $R_x(1)$. The impulse response vector h can be easily extracted once H is available. Asuming $M(= N + d - 1)$ symbols are used to carry out the identification, the estimation procedure by Tong *et al* can now be summarized as follows.

step 1. Estimate $R_x(0)$ and $R_x(1)$ from $x(t)$,

$$\hat{R}_x(0) = \frac{1}{N}\sum_{t=1}^{N} x(t)x^H(t)$$

$$\hat{R}_x(1) = \frac{1}{N}\sum_{t=1}^{N} x(t)x^H(t+1)$$

Where ^ denotes the estimate.

step 2. From $\hat{R}_x(0)$, use eigendecomposition to estimate the noise variance σ^2 and the dimension d of the signal subspace [3],

$$\hat{R}_x(0) = \hat{U}_s\hat{\Sigma}_s\hat{U}_s^H + \hat{U}_n\hat{\Sigma}_n\hat{U}_n^H$$

where $\hat{\Sigma}_s = diag[\hat{\lambda}_1, \ldots, \hat{\lambda}_d]$, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_{mT_s}$ are the eigenvalues of $\hat{R}_x(0)$, and $\hat{U}_s = [\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_d]$ contains orthonormal eigenvectors corresponding to $\hat{\Sigma}_s$. Assuming that the dimension d of the signal subspace can be estimated correctly, the noise variance can be estimated as the average of $\hat{\lambda}_i$, $i = d+1, \ldots, mT_s$. Also form $\hat{\Lambda} = \hat{\Sigma}_s - \hat{\sigma}^2 I_d$.

step 3. Compute the SVD of

$$\hat{R} \triangleq \hat{\Lambda}^{-\frac{1}{2}}\hat{U}_s^H\hat{P}\hat{U}_s\hat{\Lambda}^{-\frac{1}{2}}$$

where $\hat{P} \triangleq \hat{R}_x(1) - \hat{\sigma}^2 J'_{T_s}$. Let $\hat{y}_d$ denote the least left singular vector of $\hat{R}$.

step 4. Form an estimate of H (up to a constant scalar $e^{j\phi}$) using the formula

$$\hat{H} = \hat{U}_s\hat{\Lambda}^{\frac{1}{2}}[\hat{y}_d, \hat{R}\hat{y}_d, \cdots, \hat{R}^{(d-1)}\hat{y}_d]$$

and extract the estimate of each CIR block via

$$\hat{h}_n = W\hat{\xi}_n \quad (n = 1, 2, \ldots, p_0) \quad (3)$$

where $\xi_n (\triangleq U_s\Lambda^{\frac{1}{2}}R^{(n-1)}y_d)$ denotes the nth column of H and W the $T_s \times mT_s$ matrix defined as $W = [I_{T_s}\, 0]$.

3 Preliminary Results

In this section, we provide the expressions for the first order perturbations (due to noise and finite symbol size) on the columns of matrix H . We denote a (first-order) perturbation by preceeding the corresponding quantity by Δ. We further assume that the principal eigenvalues of $\hat{R}_x(0)$ ($\lambda_i : i = 1, \ldots, d$) are distinct and the noise $\{n(t)\}$ is circular Gaussian.

Lemma 1 (shown in [4])

$$\begin{aligned} \Delta\xi_n &\triangleq \hat{\xi}_n - \xi_n \\ &= \sum_{i=1}^{2n} A_i\Delta\Lambda U_s^H a_i^{(1)} + \sum_{i=1}^{n-1} B_i\Delta U_s^H a_i^{(2)} + \\ &\quad \sum_{i=1}^{n+1} C_i\Delta U_s a_i^{(3)} + \sum_{i=1}^{n-1} D_i\Delta P a_i^{(4)} + E_1\Delta P^H a_1^{(5)} \end{aligned}$$

where matrices A_i, B_i, C_i, D_i and E_1, vectors $a_i^{(k)}(k = 1, \ldots, 5)$ are all expressed in terms of H, U_s and $H^\dagger$ (the pseudoinverse of H). See *Appendix A* for details.

We further relate $\Delta\xi_n$ to the received data.

Lemma 2 (shown in [4])

$$\Delta\xi_n = \sum_{j=1}^{I(n)} F'_j(n)\hat{R}_j(n)g_j(n) + \mu'(n) \quad (4)$$

where

$$I(n) = mT_s + nd + n$$

$$\hat{R}_j(n) = \frac{1}{N}\sum_{t=1}^{N} x(t+\alpha_j)x^H(t+\beta_j), \; j = 1, \ldots, I(n)$$

$$(\alpha_j, \beta_j) = \begin{cases} (0,0) & j = 1 \sim mT_s + nd \\ (0,1) & j = mT_s + nd + 1 \sim I(n) - 1 \\ (1,0) & j = I(n) \end{cases}$$

where matrices $F'_j(n)$, and vectors $g_j(n)$, $\mu'(n)$ are independent of received data $\{x(t)\}$, see *Appendix B* for definitions.

4 Main Results

Now we are ready to derive the MSE of the channel estimate. As we mentioned in Section 1, the method by Tong *et al* can identify a channel up to an unknown constant phase (under the assumption that $E\{s_k s_k^*\} = 1$). In order to get rid of the effect of the unknown phase on the error evaluation, we introduce a normalized estimate. Thus the estimate error is given by the difference between the normalized estimate and the true value. Let's adjust the estimate in such a way that the phase of a nonzero element of H, say, $a_0 = H(k_1, k_2)$, is exact. Then, the normalized estimate of the nth column of H is given by

$$\hat{\xi}_n^o \triangleq \hat{\xi}_n \cdot \left(\frac{a_0}{\hat{a_0}} / |\frac{a_0}{\hat{a_0}}|\right)$$

Thus, using (4), we get the perturbation on ξ_n^o as

$$\Delta\xi_n^o \triangleq \hat{\xi}_n^o - \xi_n = \sum_{j=1}^{L} F'_j \hat{R}_j g_j + \mu'(n)$$

where

$$L = I(n) + 2I(k_2)$$

$$F'_j = \begin{cases} F'_j(n) & j = 1 \sim I(n) \\ \frac{-\xi_n}{2a_0} e_{k_1}^H F'_{j-I(n)}(k_2) & j = I(n)+1 \sim I(n)+I(k_2) \\ \frac{\xi_n}{2a_0^*} g_{j-I(n)-I(k_2)}^H (k_2) & j = I(n)+I(k_2)+1 \sim L \end{cases}$$

$$\hat{R}_j = \begin{cases} \hat{R}_j(n) & j = 1 \sim I(n) \\ \hat{R}_{j-I(n)}(k_2) & j = I(n)+1 \sim I(n)+I(k_2) \\ \hat{R}^H_{j-I(n)-I(k_2)}(k_2) & j = I(n)+I(k_2)+1 \sim L \end{cases}$$

$$g_j = \begin{cases} g_j(n) & j = 1 \sim I(n) \\ g_{j-I(n)}(k_2) & j = I(n)+1 \sim I(n)+I(k_2) \\ F'^H_{j-I(n)-2I(k_2)}(k_2) e_{k_1} & j = I(n)+I(k_2)+1 \sim L \end{cases}$$

and e_k is the $mT_s \times 1$ vector with 1 as the ith entry and 0's as others.

Next, observing (3), we get the perturbation on the (normalzed) estimate of CIR blocks h_n as

$$\Delta h_n = \sum_{j=1}^{L} F_j^H \hat{R}_j g_j + \mu \quad (n = 1, 2, \ldots, p_0)$$

where $F_j = F_j'^H W^H$ and $\mu = W\mu'(n)$.

Theorem 1 (shown in [4]): The MSE of the large-symbol-size ($M \approx N \gg 1$) estimate of each CIR block h_n ($n = 1, 2, \ldots, p_0$) can be expressed as

$$\begin{aligned} & E\{\Delta h_n^H \Delta h_n\} \\ = & \frac{1}{N} \sum_{i,j=1}^{L} \sum_{l=l_1}^{l_2} g_i^H J_\beta g_j \mathbf{tr}\left[F_j^H J_\alpha^H F_i\right] + \\ & \frac{|E\{s_k^2\}|^2}{N} \sum_{i,j=1}^{L} \sum_{l=l_3}^{l_4} \varepsilon_i^H J_{l+\alpha_j-\beta_i} Q_j^* Q_i^T J_{l+\beta_j-\alpha_i} \varepsilon_j \\ & + \frac{E\{|s_k|^4\} - |E\{s_k^2\}|^2 - 2}{N} \cdot \sum_{i,j=1}^{L} \sum_{l=l_5}^{l_6} \\ & \left(J_{\beta_i-\tau}\varepsilon_i \odot J_{l+\beta_j-\tau}\varepsilon_j^*\right)^H \cdot \left(J_{\alpha_i-\tau} Q_i \odot J_{l+\alpha_j-\tau} Q_j^*\right) \end{aligned} \tag{5}$$

where $\mathbf{tr}$ denotes "the trace of," $\odot$ the elementwise multiplication defined by $(x \odot y)_i = x_i y_i$ (for vectors x and y), $(A \odot B) = \sum_j (a_j \odot b_j)$ (for matrices A and B with a_j and b_j as the respective jth columns),

$$J_\alpha = \sigma^2 J'_{(l+\alpha_j-\alpha_i)T_s} + H J_{l+\alpha_j-\alpha_i} H^H$$
$$J_\beta = \sigma^2 J'_{(l+\beta_j-\beta_i)T_s} + H J_{l+\beta_j-\beta_i} H^H$$
$$\tau = \max\{\beta_i, \alpha_i, l+\alpha_j, l+\beta_j\}$$
$$\varepsilon_i = H^H g_i, \qquad Q_i = H^H F_i$$
$$l_1 = \max\{\beta_i - \beta_j, \alpha_i - \alpha_j\} - \max\{m, d\} + 1$$
$$l_2 = \min\{\beta_i - \beta_j, \alpha_i - \alpha_j\} + \max\{m, d\} - 1$$
$$l_3 = \max\{\beta_i - \alpha_j, \alpha_i - \beta_j\} - d + 1$$
$$l_4 = \min\{\beta_i - \alpha_j, \alpha_i - \beta_j\} + d - 1$$
$$l_5 = \max\{\alpha_i, \beta_i\} - \min\{\alpha_j, \beta_j\} - d + 1$$
$$l_6 = \min\{\alpha_i, \beta_i\} - \max\{\alpha_j, \beta_j\} + d - 1$$

Corollary 1 (shown in [4]): The MSE of the large-symbol-size estimate of CIR h is given by

$$\begin{aligned} E\{\Delta h^H \Delta h\} &= \sum_{n=1}^{p_0} E\{\Delta h_n^H \Delta h_n\} \\ &= \frac{a}{N}\sigma^4 + \frac{b}{N}\sigma^2 + \frac{c}{N} \end{aligned} \tag{6}$$

where definitions of scalars a, b, and c are obvious by comparing (5) and (6).

5 Simulation and Discussions

The MSE of CIR estimate is shown as in (6) to be a quadratic function in terms of the noise variance σ^2, where the coefficients a, b, and c are independent of σ^2 and the symbol size $M(\approx N)$. It is not surprising to note that there is a 'constant' term in (6), which implies that, when N is finite, we can not get exact estimate even when the noise is absent ($\sigma^2 = 0$). This is due to the fact that, even when the noise is absent, infinite number of observation samples are required to obtain the exact covariance matrices $R_x(0)$ and $R_x(1)$, and hence the eigenvalues and eigenvectors to form the channel estimate. This 'constant' term, i. e., c/N, tells us what MSE can be achieved under very high SNR ($\sigma^2 \rightarrow 0$). Next, we can see that the MSE is inversely proportional to the symbol size used to carry out the identification. Under any SNR, we get exact estimate (MSE $\rightarrow 0$) when N tends to infinity. In other words, this method is asymptotically (large symble size) exact. Further, we can see from (6) that, the 'constant' term dominates when SNR is high, and the first term dominates when SNR is low. Thus, the "intersection" of these two terms (i. e., $\sigma^2 = \sqrt{\frac{c}{a}}$), gives us the following information. A sharp rise of MSE is expected when SNR goes down from this intersection. While, when SNR goes up from this intersection, the curve of $MSE \sim SNR$ gets flatter, approching the 'constant' value, and a reasonably good estimate is expected (if, of cause, N is large enough).

Finally, we present a numerical Monte Carlo study which shows more insights into the identification method, and verifies our theoretical work. The channel was chosen as an approximation of a three-ray multipath channel and the CIR was given by

$$\begin{aligned} h(t) &= [0.2c(t+1.25, 0.11) - 0.15c(t+1, 0.11) \\ &\quad +0.5c(t-3.75, 0.11)]G_{6T_0}(t) \end{aligned}$$

where $c(t, \alpha)$ is the raised cosine pulse with roll-off factor α and $G_{6T_0}(t)$ a square window of 6 symbol intervals. The source symbols were drawn from the 16-QAM signal constellation with a uniform distribution and were normalized so that $E\{s_k s_k^*\} = 1$. Define the SNR as (see [1])

$$SNR = 10log_{10}\frac{||h||^2}{\sigma^2 T_s}(dB).$$

100 independent runs were conducted for each simulation scenario. At each run, both input sequence and the noise sequence were independently chosen. Further, define the normalized root-mean-square error (NRMSE) of the CIR estimate as

$$NRMSE \triangleq \frac{1}{||h||}\sqrt{\frac{1}{RUNS}\sum_{i=1}^{RUNS}||\hat{h}_i - h||^2}$$

where $\hat{h}_i$ denotes the ith estimate of h after being adjusted by the phase of the peak value.

For $m = 5$ and $T_s = 4$, Fig. 1 shows the NRMSE versus SNR for different symbol sizes M. Note that the theoretical results are more consistent with the simulation ones for larger M. This is because the analytical expression for the MSE is based on the assumption of large-symbol-size. It can be seen that the transition from 'steep' region to 'flat' region is at approximately 21 (dB) (corresponding to $\sigma^2 = \sqrt{\frac{c}{a}} \approx 0.002$).

For $M = 500$ and $SNR = 25dB$, Fig. 2(a) shows NRMSE versus m (for $T_s = 4$) and Fig. 2(b) shows NRMSE versus T_s (for $m = 7$). We see that, for a fixed sampling rate, the MSE decreases when we increase the observation window length. However, for a channel with CIR lasting for p_0 baud intervals, there is virtually no improvment by increasing m beyond p_0. Also, the MSE decreases while T_s increases for a fixed observation window length. However, when the sampling rate is higher than roughly 4 times the baud rate, the MSE decreases very slowly. Note, increasing m or T_s (when other system parameters are fixed) leads to larger dimensions of data covariance matrices $R_x(0)$ and $R_x(1)$, which means a heavier computational burden in carrying out the identification.

6 Concluding Remarks

In this paper, we have analyzed the statistical performance of the blind channel identification method proposed by Tong *et al* [1]. The explicit formula for the MSE and the insights revealed can be a helpful tool for the identification system designers.

Appendix A

$$A_i = \begin{cases} -\frac{1}{2}HJ_{i-1}H^{\dagger}U_s & i = 1 \sim n-1 \\ \frac{1}{2}HJ_nJ_{-1}H^{\dagger}U_s & i = n \\ -\frac{1}{2}HJ_{i-n}H^{\dagger}U_s & i = n+1 \sim 2n-1 \\ \frac{1}{2}U_s & i = 2n \end{cases}$$

$$a_i^{(1)} = \begin{cases} U_s^H\eta_{n-i+1} & i = 1 \sim n \\ U_s^H\eta_{2n-i} & i = n+1 \sim 2n-1 \\ U_s^H\eta_n & i = 2n \end{cases}$$

$$C_i = \begin{cases} HJ_iH^H & i = 1 \sim n-1 \\ -HJ_nJ_{-1}H^H & i = n \\ I_{mT_s} & i = n+1 \end{cases}$$

$$a_i^{(3)} = \begin{cases} U_s^H \eta_{n-i} & i = 1 \sim n-1 \\ U_s^H \eta_1 & i = n \\ U_s^H \xi_n & i = n+1 \end{cases}$$

$$\begin{aligned} B_i &= HJ_{i-1}H^\dagger U_s & i &= 1 \sim n-1 \\ a_i^{(2)} &= \xi_{n-i+1} & i &= 1 \sim n-1 \\ D_i &= HJ_{i-1}H^\dagger & i &= 1 \sim n-1 \\ a_i^{(4)} &= \eta_{n-i} & i &= 1 \sim n-1 \end{aligned}$$

$$E_1 = -HJ_nH^\dagger, \qquad a_1^{(5)} = \eta_1$$

where η_i denotes the ith column of $(H^\dagger)^H$.

Appendix B

$$F_j'(n) =$$

$$\begin{cases} A_j' + \frac{r_1+r_2+r_3}{mT_s-d} u_j^H & j = 1 \sim d \\ -\frac{r_1+r_2+r_3}{mT_s-d} e_{j-d}^H & j = d+1 \sim d+mT_s \\ B'_{j-(d+mT_s)} & j = d+mT_s+1 \sim mT_s+nd \\ D_{j-(mT_s+nd)} & j = mT_s+nd+1 \sim I(n)-1 \\ E_1 & j = I(n) \end{cases}$$

$$g_j(n) =$$

$$\begin{cases} u_j & j = 1 \sim d \\ e_{j-d} & j = d+1 \sim d+mT_s \\ \tilde{a}^{(2)}_{j-(d+mT_s)} & j = d+mT_s+1 \sim mT_s+nd \\ a^{(4)}_{mT_s+(d+1)n-j} & j = mT_s+nd+1 \sim I(n) \\ a_1^{(5)} & j = I(n) \end{cases}$$

$$\mu'(n) = -\frac{1}{2}\xi_n$$

where

$$\begin{aligned} A_j' &= \sum_{i=1}^{2n} a_{ij}^{(1)} a_j(i) u_j^H + \sum_{i=1}^{n} a_{ij}^{(3)} C_i U_s \Sigma_j' U_s^H + \\ &\quad a_{(n+1)j}^{(3)} \left(U_s \Sigma_j' U_s^H + \frac{I_{mT_s} - U_s U_s^H}{\lambda_j - \sigma^2} \right) \end{aligned}$$

$$B_j' = b_{(j-1)_d+1} \left(\frac{j-1-(j-1)_d}{d} + 1 \right) u^H_{(j-1)_d+1}$$

$$\tilde{a}_j^{(2)} = U_s \Sigma'_{(j-1)_d+1} U_s^H a^{(2)}_{\frac{j-1-(j-1)_d}{d}+1}$$

$$r_1 = \sum_{i=1}^{2n} A_i a_i^{(1)}, \; r_2 = \sum_{i=1}^{n-1} D_i J'_{T_s} a_i^{(4)}, \; r_3 = E_1 J'_{-T_s} a_1^{(5)}$$

$a_j(i)$ denotes the jth column of A_i, $b_j(i)$ the jth column of B_i , $a_{ij}^{(1)}$ the jth entry of $a_i^{(1)}$, $a_{ij}^{(3)}$ the jth entry of $a_i^{(3)}$ and $\Sigma_j' = diag[(\lambda_j - \lambda_1)^{-1}, \ldots, (\lambda_j - \lambda_{j-1})^{-1}, 0, (\lambda_j - \lambda_{j+1})^{-1}, \ldots, (\lambda_j - \lambda_d)^{-1}]$.

References

[1] L. Tong, G. Xu and T. Kailath, "Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach, " *IEEE Trans. Information Theory*, Vol. 40, No. 2, pp. 340-349, March 1994.

[2] C.R. Johnson, "Admissibility in Blind Adaptive Channel Equalization," *IEEE Control Systems Magazine*, pp. 3-15, January 1991.

[3] G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, 1973.

[4] W. Qiu and Y. Hua, "Performance Analysis of a Blind Channel Identification Method Based on Second-Order Statistics," to be submitted.

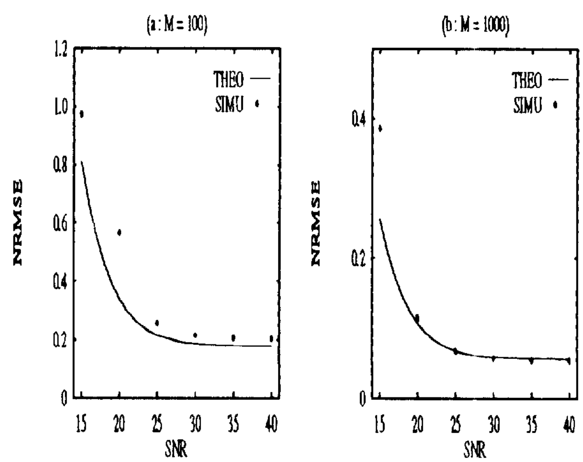

Figure 1

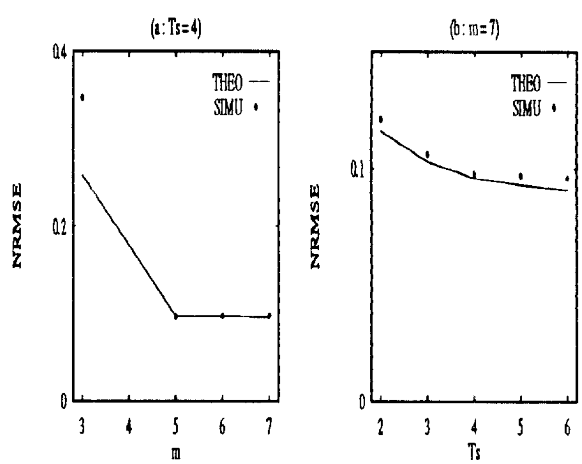

Figure 2