SEMI-BLIND CHANNEL ESTIMATION: A NEW LEAST-SQUARES APPROACH

Michael Joham[1], *Wolfgang Utschick*[1], *Josef A. Nossek*[1], *and Michael D. Zoltowski*[2]

[1]Institute for Network Theory and Circuit Design,
Munich University of Technology
Arcisstr. 21, 80333 Munich, Germany
phone: ++49 89 289 28524
fax: ++49 89 289 28504
e-mail: Wolfgang.Utschick@ei.tum.de

[2]School of Electrical Engineering
Purdue University
West Lafayette, IN 47907-1285
phone: 765-494-3512
fax: 765-494-0880
e-mail: mikedz@ecn.purdue.edu

ABSTRACT

A new semi-blind algorithm for channel estimation is proposed. The intention is to improve the estimation of the channel impulse response based on training data. We deploy the information gained from the properties of the whole received signal in order to reduce the dimension of the optimization problem encountered with the training data based method. The new method is compared with an algorithm based only on training data, and another semi-blind algorithm. As a descriptive measure for evaluation the bit error ratio (BER) is used.

1. INTRODUCTION

The task of *channel estimation* is to find an approximation $\hat{h}(t) \in \mathcal{L}^2$ of the channel impulse response by processing the received signal in baseband in order to remove the distortion caused by the dispersive characteristics of radio channels.
Most of the standards for mobile communications include *training data* which can be utilized for channel estimation. The usage of training data leads to a straight and accurate channel estimation for the price of reduced capacity caused by transmitting training sequences. On the other hand, the *blind channel estimation* approaches do not need any training information. They are based on statistical properties of the received signals [TXK94, MDCM95, KZ98] or exploit the diversity of two or more antennas at the receiver [XLTK95]. Xu et. al presented the *Cross Relation* (CR) method which does not use statistics but deploys the property of commutation of the linear convolution [XLTK95].
Blind channel estimation is inaccurate for the case of noise. Moreover, the algorithms presume that the channel is time-invariant over many symbol times.
In general, *semi-blind* methods combine both approaches. In this work, we present a semi-blind channel estimation method which deploys the CR method. Li and Ding [LD98] presented a similar approach which leads to higher computational costs but does not improve the data transmission. The new algorithm still keeps the idea of channel estimation via a training data based *least-squares* method. However, the solution is constrained to a reduced signal subspace which lies in the manifold of CR solutions.

2. CHANNEL MODEL

We presume that there is more than one receiving antenna but only one transmitting antenna. Moreover, we refer to the narrowband assumption. Hence, the channel impulse responses $h_i(t)$ only differ in the complex factor g_{ip}. Therefore the impulse response of the i-th channel

$$h_i(t) = \sum_{p=1}^{P} g_{ip}\delta(t - \tau_p), \tag{1}$$

where $g_{ip} \in \mathcal{C}$ is the amplitude and $\tau_p \in \mathcal{R}_+$ is the delay time which is modelled as a stochastic process with a uniform distribution over $[0, \tau_{max}]$ *(delay spread)*. Hereby, we approximate the real channel impulse function $h \in \mathcal{L}^2 : \mathcal{R} \to \mathcal{C}$ by its representation in a subspace of a finite set of *dirac* functions. We consider only P dominant paths because paths with small amplitudes are neglected. Without loss of generality we set $\tau_1 = 0$ and $g_{11} = 1$. The received signal at the i-th antenna element in baseband is equal to

$$x_i(t) = h_i(t) * y(t) + n_i(t), \quad i = 1, \dots, M, \tag{2}$$

corrupted by additive, white, gaussian noise $n_i(t)$. The symbol '$*$' denotes the linear convolution. The received signal is sampled l times per symbol duration T_0 over N_s symbols. The data sequence used in the remainder is

$$x_i[n] = x_i(\frac{T_0}{l} n), \quad n = 0, \dots, lN_s - 1. \tag{3}$$

We approximate the unknown channel $h_i(t)$ with

$$\hat{h}_i(t) = \sum_{m=0}^{N_h - 1} \hat{h}_i[m] \delta(t - \frac{T_0}{l} m). \tag{4}$$

This approximation is reasonable since we have to match the original channel only over the bandwidth of the transmitted signal. The number of impulses N_h has influence on the quality of the algorithms presented in this paper. We assume that the delay spread τ_{max} is known, thus, we choose a $N_h = \tau_{max}/T_0$ in order to cover the whole delay spread with the impulse response of the estimation $\hat{h}_i(t)$.

3. TRAINING BASED CHANNEL ESTIMATION

Since the GSM standard includes a $N_t = 26$ bit long training sequence in the midamble of each burst we exemplarily examine the training based channel estimation method for GSM. We have to take into account that the *Gaussian Minimum Shift Keying* (GMSK) signal in GSM has "memory", because all previously sent bits have influence on the current value of the phase. However, after four symbol times the phase portion according to a particular bit has the constant value $\pm\pi/2$ [JKZ99, JU98]. As a consequence, we do not use the samples of the first three symbol times of the training sequence, because we do not know the values of the prior bits, hence our actual training sequence is three bits shorter. So we know the transmitted sequence and the received signal of the remaining $N_t - 3$ bits of the training sequence. The GMSK modulated training sequence reads as

$$y_t[n] = \exp\left(j \frac{\pi}{2} \sum_{m=0}^{N_t - 1} t_m \psi \left(\frac{T_0}{l} n - mT_0 \right) \right), \tag{5}$$

where $t_m \in \{-1, +1\}$ are the known training bits. By collecting the transmitted signals in the convolution matrix

$$\boldsymbol{Y}_t = \begin{bmatrix} y_t[N_h + 3l - 1] & y_t[N_h + 3l - 2] & \cdots & y_t[3l] \\ y_t[N_h + 3l] & y_t[N_h + 3l - 1] & \cdots & y_t[3l + 1] \\ \vdots & \vdots & \ddots & \vdots \\ y_t[N_t l - 1] & y_t[N_t l - 2] & \cdots & y_t[N_t l - N_h] \end{bmatrix}, \tag{6}$$

with $\boldsymbol{Y}_t \in \mathcal{C}^{N_t l - N_h - 3l + 1 \times N_h}$, we end up with the following least squares problem:

$$\min_{\hat{\boldsymbol{h}}} \|\boldsymbol{Y} \hat{\boldsymbol{h}} - \boldsymbol{x}_t\|_2^2, \tag{7}$$

where the transmitted and received signals are arranged in

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_t & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Y}_t \end{bmatrix}, \qquad \text{and} \tag{8}$$

$$\begin{aligned} \boldsymbol{x}_t &= [\boldsymbol{x}_{t1}^T, \boldsymbol{x}_{t2}^T]^T, \\ \boldsymbol{x}_{ti} &= [x_i[N_{trpos} + 3l], \dots, x_i[N_{trpos} + N_t l - N_h]]^T \end{aligned} \tag{9}$$

with $\boldsymbol{x}_t \in \mathcal{C}^{N_t l - N_h - 3l + 1}$. N_{trpos} is the number of the first received sample that belongs to the training sequence. The solution of (7) is the multiplication with the Moore-Penrose-pseudoinverse of $\boldsymbol{Y}$, thus, $\hat{\boldsymbol{h}} = \boldsymbol{Y}^+ \boldsymbol{x}_t = (\boldsymbol{Y}^H \boldsymbol{Y})^{-1} \boldsymbol{Y}^H \boldsymbol{x}_t$.

4. BLIND CHANNEL ESTIMATION

In [XLTK95] Xu, Liu, Tong, and Kailath presented the Cross Relation (CR) method for blind channel estimation. The CR follows from the property of commutation of the linear convolution. For simplicity let us consider two channels $h_1(t)$ and $h_2(t)$ since the CR method can be easily extended for more than two antennas. We excite both of them with the same input signal $y(t)$. Then the output of the ith channel equals $x_i(t) = y(t) * h_i(t)$.

Let us assume that we filter the signal $x_1(t)$ with $h_2(t)$ and $x_2(t)$ with $h_1(t)$. Fig. 1 depicts this configuration. The output signals $x_{12}(t)$ and $x_{21}(t)$ are the same due to the property of commutation $h_1(t) * h_2(t) = h_2(t) * h_1(t)$. Since $x_1(t)$ and $x_2(t)$ are known we can substitute $y(t) * h_i(t)$ by $x_i(t)$ and $h_i(t)$ by $\hat{h}_i(t)$. The basic result of the CR method is

$$x_2(t) * \hat{h}_1(t) - x_1(t) * \hat{h}_2(t) = 0. \tag{10}$$

After receiving the transmitted signal, it is sampled l times per symbol time, and since $\hat{h}_i(t)$ is a weighted sum of Dirac delta functions we can use the sampled version of the CR. We again collect the samples $x_i[n]$ in a convolution matrix

$$\boldsymbol{X}_i = \begin{bmatrix} x_i[N_h - 1] & x_i[N_h - 2] & \cdots & x_i[0] \\ x_i[N_h] & x_i[N_h - 1] & \cdots & x_i[1] \\ \vdots & \vdots & \ddots & \vdots \\ x_i[lN_s - 1] & x_i[lN_s - 2] & \dots & x_i[lN_s - N_h] \end{bmatrix}, \tag{11}$$

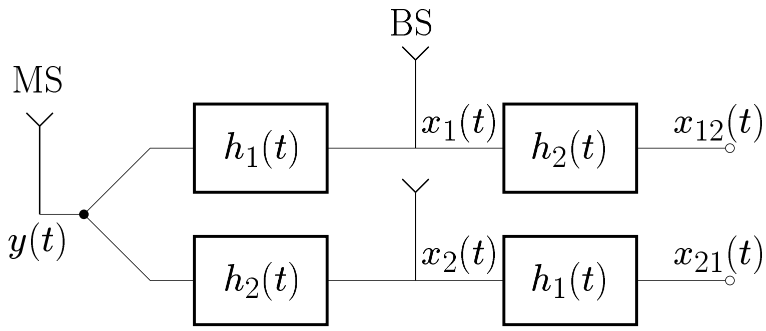

Figure 1: Principle of the cross relation (CR) method.

with $\boldsymbol{X}_i \in \mathcal{C}^{lN_s - N_h + 1 \times N_h}$, and put the taps of $\hat{h}_i(t)$ in the vector

$$\hat{\boldsymbol{h}} = \left[\hat{h}_1[0], \ldots, \hat{h}_1[N_h - 1], \hat{h}_2[0], \ldots, \hat{h}_2[N_h - 1]\right]^T \quad (12)$$

$\in \mathcal{C}^{2N_h \times 1}$. This leads to following system of equations:

$$[\boldsymbol{X}_2, -\boldsymbol{X}_1]\hat{\boldsymbol{h}} = \boldsymbol{X}\hat{\boldsymbol{h}} = \boldsymbol{0}. \quad (13)$$

To find the solution we apply the *Singular Value Decomposition* (SVD) to the $lN_s - N_h + 1 \times 2N_h$ matrix $\boldsymbol{X} = [\boldsymbol{X}_2, -\boldsymbol{X}_1]$, thus, we get

$$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^H. \quad (14)$$

If the column space of $\boldsymbol{X}$ is not of full rank the $2N_h - r$ smallest singular values $\sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_{2N_h}$ (where $r = \text{rank}(\boldsymbol{X})$) are equal to zero. In this noisefree case the column vectors $\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \ldots, \boldsymbol{v}_{2N_h}$ span the nullspace of matrix $\boldsymbol{X}$ and the manifold of nontrivial solutions of (13). In the following we refer to the rank r as the rank of signal subspace and to $\boldsymbol{X}_s$ resp. $\boldsymbol{X}_n$ as the noisefree (s)ignal and (n)oise part of $\boldsymbol{X}$.
In practical applications, the samples $x_i[n]$ and the matrices $\boldsymbol{X}_i$ are noisy and the column space of matrix $\boldsymbol{X}$ is numerically of full rank. If we know or estimate (cf. [WK85]) the reduced rank r of the noisefree signal subspace, represented by matrix "$\boldsymbol{X}_s$", the SVD of $\boldsymbol{X}$ can be denoted by

$$\boldsymbol{X} = \underbrace{\boldsymbol{U}_s\boldsymbol{\Sigma}_{ss}\boldsymbol{V}_s^H}_{\boldsymbol{X}_s} + \underbrace{\boldsymbol{U}_s\boldsymbol{\Sigma}_{sn}\boldsymbol{V}_s^H + \boldsymbol{U}_n\boldsymbol{\Sigma}_{nn}\boldsymbol{V}_n^H}_{\boldsymbol{X}_n}, \quad (15)$$

where

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{ss} + \boldsymbol{\Sigma}_{sn} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_{nn} \end{bmatrix}. \quad (16)$$

The matrix $\boldsymbol{\Sigma}_{nn} = \text{diag}(\sigma_{r+1}, \sigma_{r+2}, \ldots, \sigma_{2N_h})$ represents the diagonal matrix of the $2N_h - r$ smallest singular values of Σ. Then it can be easily shown that the corresponding vectors of $\boldsymbol{V}_n = [\boldsymbol{v}_{r+1}, \boldsymbol{v}_{r+2}, \ldots, \boldsymbol{v}_{2N_h}]$ span the linear manifold of nontrivial solutions of minimum residuum $\|\boldsymbol{U}_n\boldsymbol{\Sigma}_{nn}\boldsymbol{\beta}\|_2^2$ where $\boldsymbol{\beta} \in \mathcal{R}^{2N_h - r}$.

The CR method inherits a serious disadvantage: the Cross Relation is unintentionally fulfilled if the estimated channels have highpass characteristics. The received signal is zero outside the bandwidth of the transmitted signal. If the $\hat{h}_i(t)$ are highpass, then the convolution will also be zero inside the bandwidth, and the difference is zero over the whole band. To avoid these bad estimates Zoltowski and Tseng [ZT98] proposed the Maximum in Band Energy method (MIE). The MIE uses a solution out of the nullspace of $\boldsymbol{X}$, which maximizes the energy inside the bandwidth of the transmitted signal to avert trivial highpass estimates.

5. SEMI-BLIND ESTIMATION

Obviously, a combined approach of channel estimation both utilizing training sequences and a blind method is supposed to outperform the accuracy of a single strategy solution. To accomplish that, Li et. al. [LD98] proposed the combination of the training method and the Cross Relation (CR) method (13) in form of the least squares problem:

$$\min_{\hat{\boldsymbol{h}}} \|\tilde{\boldsymbol{X}}\hat{\boldsymbol{h}} - \tilde{\boldsymbol{x}}\|_2^2, \quad (17)$$

where

$$\tilde{\boldsymbol{X}} = \begin{bmatrix} \boldsymbol{Y}_t & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Y}_t \\ \boldsymbol{X}_2 & -\boldsymbol{X}_1 \end{bmatrix} \quad \text{and} \quad \tilde{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{x}_t \\ \boldsymbol{0} \end{bmatrix}. \quad (18)$$

Since this approach leads to a degradation (see 6. Simulation Results) of the channel estimation performance for low *signal to noise ratio* (SNR) we reformulate the problem: If training sequences are available the channel estimation task basically remains the least squares problem in (7) and blind approaches are merely considered as a method to include extra knowledge. Therefore, the plain combination of both

problems (17) can be outperformed by a more problem adequate approach.
First, in noisefree cases the CR method is correct and the nullspace of $\boldsymbol{X}$ exists. Consequently, the nontrivial solution of (13) lies in the nullspace.
In practical applications which imply noise, the sampled matrix reads as

$$\boldsymbol{X} = \boldsymbol{X}_s + \boldsymbol{X}_n \tag{19}$$

and (13) is replaced by

$$\min_{\hat{\boldsymbol{h}}} \|\boldsymbol{X}\hat{\boldsymbol{h}}\|_2^2 \quad \Leftrightarrow \quad \min_{\hat{\boldsymbol{h}}} \|\boldsymbol{X}_s\hat{\boldsymbol{h}} + \boldsymbol{X}_n\hat{\boldsymbol{h}}\|_2^2. \tag{20}$$

The prementioned benefit from the blind channel estimation for the training bases estimation is that the solution of (7) must still suffice

$$\boldsymbol{X}_s\hat{\boldsymbol{h}} = \boldsymbol{0}. \tag{21}$$

Unfortunately, the noisefree matrix $\boldsymbol{X}_s$ is not known. However, assuming that we know rank r of the signal subspace, we yield the nullspace estimate of $\boldsymbol{X}_s$ as: null$(\boldsymbol{X}_s) \approx$ span$(\boldsymbol{V}_n)$. Then the reformulated least squares problem reads as

$$\min_{\hat{\boldsymbol{h}}} \|\boldsymbol{Y}\hat{\boldsymbol{h}} - \boldsymbol{x}_t\|_2^2 \quad \text{s.t.} \quad \hat{\boldsymbol{h}} \in \text{span}(\boldsymbol{V}_n) \quad \Leftrightarrow \tag{22}$$

$$\min_{\hat{\boldsymbol{h}}} \|\boldsymbol{Y}\boldsymbol{V}_n\boldsymbol{\beta} - \boldsymbol{x}_t\|_2^2. \tag{23}$$

With the Moore-Penrose-pseudoinverse of $\boldsymbol{Y}\boldsymbol{V}_n$ the estimation
can be denoted as $\hat{\boldsymbol{h}} = \boldsymbol{V}_n(\boldsymbol{V}_n^H\boldsymbol{Y}^H\boldsymbol{Y}\boldsymbol{V}_n)^{-1}\boldsymbol{V}_n^H\boldsymbol{Y}^H\boldsymbol{x}_t$.
The new semi-blind channel estimation algorithm is summarized as follows:

---

1. *Nullspace estimation:* estimate the signal subspace dimension:
$$\text{rank}([\boldsymbol{X}_2, -\boldsymbol{X}_1]),$$
based on the blind method (CR), and compute V_n as the $2N_h - r$ least dominant right singular vectors of $\boldsymbol{X} = [\boldsymbol{X}_2, -\boldsymbol{X}_1]$.

2. *Training sequence matrices:* collect the transmitted and received signals in
$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_t & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Y}_t \end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{x}_t = [\boldsymbol{x}_{t1}^T, \boldsymbol{x}_{t2}^T]^T.$$

3. *Channel estimation:* calculate the least squares solution of the channel response function $\hat{\boldsymbol{h}} = \boldsymbol{V}_n\boldsymbol{\beta}^*$ with
$$\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}} \|\boldsymbol{Y}\boldsymbol{V}_n\boldsymbol{\beta} - \boldsymbol{x}_t\|_2^2.$$

---

6. SIMULATION RESULTS

In our simulations, we used a multipath channel with $P = 10$. The bit sequence and all channel parameters were chosen randomly for each Monte Carlo run except $g_{11} = 1$ and $\tau_1 = 0$. We assumed that the delay spread τ_{max} is about $10\mu s$, thus, we set $\tau_{max} = 3T_0$, since $T_0 = 3.69\mu s$. The received signal $x(t)$ was sampled $l = 3$ times per T_0. We exploited the whole GSM burst for the semi-blind methods, therefore $N_s = 142$. The length N_h of the channel estimation $\hat{\mathbf{h}}_i$ was set to $3l$ to cover the whole delay spread. In the new semi-blind channel estimation algorithm, we set $2N_h - r = N_h$.
Figure 2 depicts the *bit error rate* (BER) of the different methods compared to the BER of an *additive white gaussion noise* (AWGN) channel. The plots are the mean of 10000 Monte Carlo runs. We show the results for the AWGN channel to demonstrate the influence of the channel estimation and equalization on the performance of the demodulation with a *maximum likelihood sequence estimation* (MLSE) demodulator which is a *Viterbi Algorithm* (VA) implementation.
First, it can be seen that the semi-blind algorithm of Li et. al. [LD98] does not improve the channel estimation of the training sequence method. For low *signal to noise ratio* (SNR) the semi-blind approach increases the BER, thus, the result is worse. For greater SNR Li's method improves the demodulation only slightly.
Our new semi-blind algorithm outperforms the other methods. The improvement in respect to the training sequence method depends on the SNR, because the CR which we exploit is sensitive to noise. For low SNR the CR is very unprecise and the plot of the new method moves very close to the plot of the training sequence method for $SNR < 2dB$. On the other hand, the semi-blind algorithm has the same BER at $9dB$ as the training sequence method at $10dB$. Hence we gain about $1dB$ at a SNR of $10dB$.

7. REFERENCES

[JKZ99] M. Joham, T.P. Krauss, and M.D. Zoltowski. Channel Estimation and Equalization for GSM with Multiple Antennas. In *Proceedings of the 49th Annual International Vehicular Technology Conference*, 1999. to appear.

[JU98] Michael Joham and Wolfgang Utschick. A novel semi-blind channel estimation algorithm. Technical Report TUM-LNS-TR-98-4, Munich University of Technology, 1998.

[KZ98] T.P. Krauss and M.D. Zoltowski. Multiuser Second-Order Statistics Based Blind Identification Using a Linear Parameterization of the

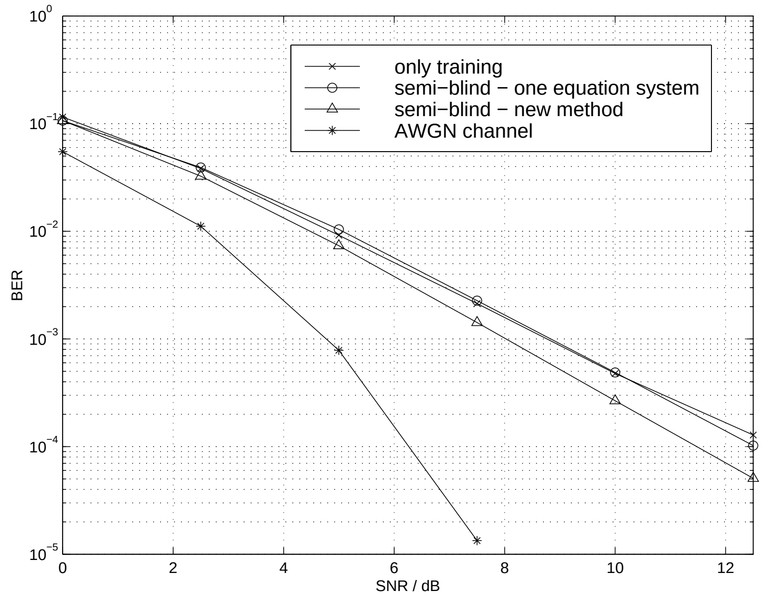

Figure 2: New semi-blind algorithm in comparison with other training sequence based methods

Channel Matrix. *IEEE Transactions on Signal Processing*, 1998.

[LD98] G. Li and Z. Ding. A Semi-Blind Channel Identification Method for GSM Receivers. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, Seattle, 1998.

[MDCM95] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue. Subspace Methods for the Blind Identification of Multichannel FIR Filters. *IEEE Transactions on Signal Processing*, 43:516–525, 1995.

[TXK94] L. Tong, G. Xu, and T. Kailath. Blind Identification and Equalization Based on Second-Order Statistics: A Time Domain Approach. *IEEE Transactions on Information Theory*, 40:340–349, 1994.

[WK85] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(2):387–392, 1985.

[XLTK95] G. Xu, H. Liu, L. Tong, and T. Kailath. A Least-Squares Approach to Blind Channel Identification. *IEEE Journal on Signal Processing*, 43:2982–2993, 1995.

[ZT98] M.D. Zoltowski and D.-F. Tseng. A Weighted Energy Concentration Criterion for Improving the Performance of Deterministic Least Squares Blind Channel Identification at Low SNR. In *Proceedings of Thirty-Second Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 1998.