Sequence Estimation in GSM Using the EM Algorithm *

Linda M. Zeger†and Hisashi Kobayashi
Department of Electrical Engineering
Princeton University

March 17, 1998

Abstract

We apply the EM (Expectation–Maximization) algorithm to sequence estimation for GMSK, and in particular, to GSM. Time varying multipath channel parameters are considered unknown, and are averaged over to obtain the maximum likelihood sequence estimate. A simplified EM algorithm is derived for low noise, and simulation results are presented, showing that this algorithm performs well through a range of practical SNRs. This algorithm could potentially increase the information bit rate and improve sequence estimation in GSM.

I Introduction

There has been much recent interest in improving the performance of reception of signals used in mobile land radio [1]–[2]. A number of equalization and other schemes [3] have been proposed to handle distortion due to multipath fading. Currently, in the GSM mobile cellular phone system the impulse response of the channel is estimated using training data inserted in each time slot. The estimated impulse response function is then used to estimate the transmitted data sequence.

In this paper, we discuss an alternative approach: We consider a multipath channel with an impulse response represented by delta functions for a finite number of paths. Using the EM algorithm, we average the log likelihood of the transmitted data sequence and the multipath channel parameters over the multipath parameters. We then find the data sequence that maximizes this averaged likelihood. However, in order to average the likelihood over the channel parameters, we must obtain their probability density. We estimate the density of the unknown channel parameters based on a previous estimate of the data sequence, and then use the estimated density of the channel parameters in turn to produce an improved estimate of the transmitted data sequence. These steps are repeated until convergence is achieved.

The goal from the point of view of the cellular phone user is to obtain the most probable data that was transmitted, given the received signal; this maximum *a posteriori* probability estimate is precisely what we obtain via the EM algorithm as discussed below. Specific estimates of channel parameters are not necessarily required to obtain this estimate, but rather the probability density of the channel parameters is used here. Presently, sequence estimation in GSM is based on a possibly imperfect estimate of the channel impulse response. In contrast, successive iterations of the EM algorithm allow for improvement of the initially imperfect estimate of the probability density of channel parameters and also of the initial estimate of the transmitted data sequence. In addition, our method uses channel information derived in each time slot in the successive time slot. The information bit rate in GSM could potentially be increased with use of the EM algorithm since the need for training data is reduced. Alternatively, the EM algorithm could possibly be used in conjunction with training data to improve estimation of the transmitted information sequence.

An iterative method for maximizing a likelihood function in the presence of unobserved data, the Expectation–Maximization algorithm [4], has recently been applied to the problem of sequence estimation with unknown channel parameters [5]. Additional recent applications of the EM algorithm to sequence estimation are found in [6]–[7], and to other detection and estimation problems include [8]–[10].

In Section II we describe the multipath channel model and the received GMSK signal. Application of the EM algorithm to sequence estimation, and to GMSK in the time varying channel is discussed in Section III. We derive a simplified version of the EM algorithm, valid for low noise, in Section IV, and present simulation results for this algorithm in Section V. Conclusions are stated in Section VI.

II The Model

We first describe the modulation scheme used in GSM, as well as the multipath model we consider. The trans-

*This work has been supported, in part, by grants from the National Science Foundation, the New Jersey Commission on Science and Technology, The Ogasawara Foundation for the Promotion of Science and Engineering, Asahi Chemical Industry, Co., and Seiko Communications.

†Email: zeger@ee.princeton.edu. Currently with Lucent Technologies.

mitted data sequence is denoted by C_n for $n = 1, 2, 3, ...$ where C_n can take on values of ± 1. The transmitted signal

$$X(t) = A\cos(\omega_c t + \phi(t)) \quad (1)$$

is modulated using GMSK so that

$$\phi(t) = \frac{\pi h_f}{T}\int_{-\infty}^{t}\sum_n C_n g(t' - nT)dt', \quad (2)$$

where the modulation index $h_f = .5$ and the bit period for GSM is $T = 3.69$ μsec. The baseband frequency pulse $g(t)$ is obtained by convolving a NRZ data sequence with a Gaussian low pass filter and is given by

$$g(t) = \left\{\text{erf}\left[\frac{-2\pi B_b}{\sqrt{\ln 2}}t\right] + \text{erf}\left[\frac{2\pi B_b}{\sqrt{\ln 2}}(t+T)\right]\right\}/2, \quad (3)$$

where the bandwidth of the low pass filter B_b satisfies $B_b T = .3$.

We consider a general multipath model with a total of M paths. The received signal is given by

$$\begin{aligned} Y(t) &= \alpha_1 X(t - \tau_1, \theta_1) + \alpha_2 X(t - \tau_2, \theta_2) \\ &+ \ldots + \alpha_M X(t - \tau_M, \theta_M), \end{aligned}$$

where θ_i, α_i, and τ_i denote the phase, amplitude, and delay of the i^{th} path respectively. We set $\theta_1 = \tau_1 = 0$, and then the remaining phases and time delays are relative to the signal from the first path. Letting

$$q(t) = \int_{-\infty}^{t} g(t')dt', \quad (4)$$

we combine equations (1) through (4) to obtain the complex envelope of the received signal in the presence of multipath fading and additive white Gaussian noise:

$$Y(t) = \sum_{i=1}^{M}\alpha_i \exp\left[j\left(\Phi_i(t, \mathbf{C})\right)\right] + n_I(t) + jn_Q(t), \quad (5)$$

where

$$\Phi_i(t, \mathbf{C}) = \theta_i - \omega_c \tau_i + \frac{\pi h_f}{T}\sum_n C_n q(t - \tau_i - nT), \quad (6)$$

and the inphase and quadrature noise components n_I and n_Q are independent WGN processes with noise density N_o.

III The EM Algorithm

We would like to estimate the transmitted data sequence C_n for $n = 1, 2, 3, ..$, which we represent by $\mathbf{C}$. Using the EM algorithm, we present a method to find the maximum *a posteriori* probability estimate for the sequence of transmitted symbols $\mathbf{C}$, given inphase and quadrature samples of the complex received signal $Y(t)$, denoted collectively by $\mathbf{y}$. Thus we maximize the posterior probability density $f(\mathbf{C} \mid \mathbf{y})$ with respect to $\mathbf{C}$. Equivalently, the ML estimate is found, using a prior density on C_n that restricts its values to ± 1. This prior density is incorporated into the EM algorithm, by using a Viterbi algorithm to find the optimal path through a trellis of allowed states for the M (maximization) step.

The EM algorithm produces maximum likelihood estimates in the presence of unobserved data by repeating two steps until convergence. The E (Expectation) step computes the expected log likelihood of the observed and unobserved data by averaging over the unobserved data. In the M (Maximization) step, the expected log likelihood is maximized. References [4] and [5] provide background on the EM algorithm and its general application to sequence estimation respectively, and equations (7) and (8) below parallel the formulation in [5].

We now discuss the EM algorithm as used in sequence estimation. The channel parameters, denoted collectively by $\beta = \{\alpha_1, \theta_1, \tau_1, ...\alpha_M, \theta_M, \tau_M\}$, are treated as the unobserved data in the EM algorithm. An average over β is taken in the E step in order to compute the expected log likelihood

$$Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p) = E\left[\ln f(\mathbf{y}, \beta \mid \mathbf{C}^{p+1}) \mid \mathbf{y}, \mathbf{C}^p\right] \quad (7)$$

at the $(p+1)^{st}$ iteration, where $f(\mathbf{y}, \beta \mid \mathbf{C}^{p+1})$ is the joint probability density of $\mathbf{y}$ and β conditional on $\mathbf{C}^{p+1}$. The conditional expectation (7) is an average using the conditional density of the channel parameters $f(\beta \mid \mathbf{y}, \mathbf{C}^p)$, given the estimate $\mathbf{C}^p$ of the transmitted symbols from the previous (the p^{th}) iteration. The M step determines the value of $\mathbf{C}^{p+1}$ that maximizes $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p)$ given $\mathbf{C}^p$; hence, terms independent of $\mathbf{C}^{p+1}$ can be eliminated from $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p)$. Since $f(\beta \mid \mathbf{C}^{p+1}) = \rho(\beta)$, where $\rho(\beta)$ denotes the prior density of β which is independent of $\mathbf{C}^{p+1}$, the resulting expected likelihood can be written as

$$\begin{aligned} Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p) &= E\left[\ln f(\mathbf{y} \mid \beta, \mathbf{C}^{p+1}) \mid \mathbf{y}, \mathbf{C}^p\right] \quad (8) \\ &= \int \ln f(\mathbf{y} \mid \beta, \mathbf{C}^{p+1}) f(\beta \mid \mathbf{y}, \mathbf{C}^p) d\beta. \end{aligned}$$

The EM algorithm is now applied to GMSK modulation and the channel model described in Section II. The conditional density of $\mathbf{y}$ can be obtained from equation (5) and is:

$$\begin{aligned} f(\mathbf{y} \mid \beta, \mathbf{C}) = \left(\frac{1}{2\pi N_o}\right)^K \exp -\frac{1}{2N_o}\sum_{l=1}^{K}\Bigg\{ \\ \left[y_I(t_l) - \sum_{i=1}^{M}\alpha_i \cos\left(\Phi_i(t_l, \mathbf{C})\right)\right]^2 \\ + \left[y_Q(t_l) - \sum_{i=1}^{M}\alpha_i \sin\left(\Phi_i(t_l, \mathbf{C})\right)\right]^2\Bigg\}, \quad (9) \end{aligned}$$

where K is the number of samples of the received signal, and the inphase and quadrature components of the samples taken at time t_l are denoted by $y_I(t_l)$ and $y_Q(t_l)$ respectively. Substituting equation (9) evaluated at $\mathbf{C}^{p+1}$

into (8) and eliminating terms and overall factors independent of $\mathbf{C}^{p+1}$ yields

$$\begin{aligned}-Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p) = & \\ & \sum_{l=1}^{K} \int \Bigg\{ \left[y_I(t_l) - \sum_{i=1}^{M} \alpha_i \cos\left(\Phi_i(t_l, \mathbf{C}^{p+1})\right) \right]^2 \\ & + \left[y_Q(t_l) - \sum_{i=1}^{M} \alpha_i \sin\left(\Phi_i(t_l, \mathbf{C}^{p+1})\right) \right]^2 \Bigg\} \\ & \times f(\beta \mid \mathbf{y}, \mathbf{C}^p) \quad d\beta. \qquad (10)\end{aligned}$$

Furthermore,

$$f_s(\beta \mid \mathbf{y}, \mathbf{C}^p) = \kappa f_s(\mathbf{y} \mid \beta, \mathbf{C}^p) \rho_s(\beta), \qquad (11)$$

where κ is a constant independent of $\mathbf{C}^{p+1}$, and $\rho_s(\beta)$ is the prior probability density of β. For later use, we have added the subscript s, which refers to time slot number s. The density $f(\mathbf{y} \mid \beta, \mathbf{C}^p)$ as determined from equation (9) evaluated at $\mathbf{C}^p$ can be used in (11), which in turn can be substituted into (10).

Minimization of equation (10) with respect to $\mathbf{C}^{p+1}$ for the M step is performed at iteration $p+1$, where the values of $\mathbf{C}^p$ are the estimates obtained from the M step of the previous iteration. The E and M steps are repeated until $\mathbf{C}^{p+1} = \mathbf{C}^p$, and convergence is achieved.

For each time slot, we use the EM algorithm to estimate the symbols transmitted in that time slot. Because the channel will not change too much during one time slot, the approximation

$$f_s(\beta \mid \mathbf{y}, \mathbf{C}^1) \approx f_{s-1}(\beta \mid \mathbf{y}, \mathbf{C}^P), \qquad (12)$$

can be used for the first iteration of time slot s, where $f_{s-1}(\beta \mid \mathbf{y}, \mathbf{C}^P)$ denotes the density obtained from the final (P^{th}) EM iteration of the previous time slot. The density used in subsequent EM iterations of time slot s can be obtained from (11) and equation (9) used to calculate $f(\mathbf{y} \mid \beta, \mathbf{C}^p)$, where the prior density $\rho_s(\beta)$ can be constructed using $f_{s-1}(\beta \mid \mathbf{y}, \mathbf{C}^P)$. Training data could be used to form the prior density $\rho_1(\beta)$ for the first time slot.

IV Low Noise Algorithm

The EM algorithm could be simply implemented if the multiple integral in (8) could be performed analytically. However, for GMSK modulation in the presence of the fading multipath channel, computation of $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p)$ in (10) would require numerical integration for each trial sequence $\mathbf{C}$. The EM algorithm can be implemented more efficiently when the noise is not too high. We now discuss an algorithm for estimating $\mathbf{C}$, which the EM algorithm reduces to for the case of low noise.

We simplify the integral in (10) by noting that equations (9–11) indicate that for a relatively uniform prior $\rho_s(\beta)$, most of the contribution to the integral comes from values of β that make the likelihood (9) large. We focus on this range of β and consider the estimate $\tilde{\beta}_s^p$ of the true β at iteration p of time slot s that maximizes $f(\mathbf{y} \mid \beta, \mathbf{C}^p)$, and is thus defined by

$$\tilde{\beta}_s^p = \tilde{\beta}_s^p(\mathbf{y}, \mathbf{C}^p) = \arg\min_{\beta} \lambda(\beta, \mathbf{y}, \mathbf{C}^p), \qquad (13)$$

where the negative log likelihood $\lambda(\beta, \mathbf{y}, \mathbf{C}^p)$ is

$$\begin{aligned}\lambda(\beta, \mathbf{y}, \mathbf{C}^p) &= \sum_{l=1}^{K} \left[y_I(t_l) - \sum_{i=1}^{M} \alpha_i \cos\left(\Phi_i(t_l, \mathbf{C}^p)\right) \right]^2 \\ &\quad + \left[y_Q(t_l) - \sum_{i=1}^{M} \alpha_i \sin\left(\Phi_i(t_l, \mathbf{C}^p)\right) \right]^2 \qquad (14)\end{aligned}$$

Equation (10) can now be written as

$$Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p) = -\int \lambda(\beta, \mathbf{y}, \mathbf{C}^{p+1}) f(\beta \mid \mathbf{y}, \mathbf{C}^p) d\beta. \qquad (15)$$

When β is close to $\tilde{\beta}_s^p$,

$$\begin{aligned} f(\mathbf{y} \mid \beta, \mathbf{C}^p) &\approx \left(\frac{1}{2\pi N_o} \right)^K \exp - \left[\frac{1}{2N_o} \lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^p) \right] \\ & \times \exp - \left[\frac{1}{2N_o} (\beta - \tilde{\beta}_s^p)^T A (\beta - \tilde{\beta}_s^p)/2 \right], \qquad (16)\end{aligned}$$

where the elements of the inverse covariance matrix A are given by

$$A_{ij} = \frac{\partial^2 \lambda(\beta, \mathbf{y}, \mathbf{C}^p)}{\partial \beta_i \partial \beta_j} \mid_{\tilde{\beta}_s^p} .$$

We consider a uniform prior density $\rho_s(\beta)$ that restricts β sufficiently close to $\tilde{\beta}_s^p$ so that the approximation (16) holds. Most of the contribution to $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p)$ then comes from β in a region $R(A/N_o)$, which grows with N_o, around $\tilde{\beta}_s^p$.

In the maximization of $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p)$ in the M step, we consider the result of minimizing $\lambda(\beta, \mathbf{y}, \mathbf{C}^{p+1})$ for various values of β, and note that a range of β will produce the same optimal sequence $\mathbf{C}^{p+1}$. We define S to be the set of all β such that the same sequence $\mathbf{C}^{p+1}$ minimizes $\lambda(\beta, \mathbf{y}, \mathbf{C}^{p+1})$ as minimizes $\lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^{p+1})$. If

$$R(A/N_o) \subset S, \qquad (17)$$

then the values of β that contribute significantly to the integral in (15) all produce the same sequence $\mathbf{C}^{p+1}$ upon minimizing $\lambda(\beta, \mathbf{y}, \mathbf{C}^{p+1})$. Therefore, for the purpose of finding the optimal sequence $\mathbf{C}^{p+1}$, only $\lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^{p+1})$ needs to be minimized and the integral in (15) need not be evaluated. Hence, the density

$$f(\beta \mid \mathbf{y}, \mathbf{C}^p) = \delta(\beta - \tilde{\beta}_s^p) \qquad (18)$$

can be used in (15).

The E step is then reduced to computing $\tilde{\beta}_s^p = \tilde{\beta}_s^p(\mathbf{y}, \mathbf{C}^p)$ from (13). The M step in turn becomes finding $\mathbf{C}^{p+1}$ to maximize $Q(\mathbf{C}^{p+1} \mid \mathbf{C}^p) = -\lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^{p+1})$, and the Viterbi Algorithm can be used to perform this maximization. Therefore, when the noise is low enough so that (17) is satisfied, the EM algorithm reduces to iteratively estimating β and $\mathbf{C}$. These estimations at iteration p are accomplished by minimizing $\lambda(\beta, \mathbf{y}, \mathbf{C}^p)$ with respect to β for the E step, and then $\lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^{p+1})$ with respect to $\mathbf{C}^{p+1}$ for the M step. One EM iteration of this low noise algorithm is thus analogous to the current equalization method used in GSM in the sense that there the channel impulse response function is estimated, and then the transmitted sequence is estimated from that impulse response function.

When the noise is high enough such that (17) is no longer satisfied, our algorithm of iteratively estimating β and $\mathbf{C}$ is no longer an EM algorithm. In this case, implementation of the EM algorithm requires evaluation of the integrals resulting from substitution of (9) into (11) and (10).

In order to start the low noise EM algorithm, equations (12) and (18) indicate that the initial estimate of β in time slot s is given by the final (P^{th}) estimate from the previous slot:

$$\beta_s^1 = \beta_{s-1}^P. \tag{19}$$

V Simulations

Simulations of the low noise EM algorithm were performed for 300 consecutive time slots of 136 bits each. The values of the true channel parameters were updated after each slot, incorporating fast fading for a two path model. The maximum fade is taken as 30dB. There are 30 fades per second of the primary path, and the mobile speed is 50 kilometers per hour.

We use a Viterbi algorithm with a memory of L symbols to minimize $\lambda(\tilde{\beta}_s^p, \mathbf{y}, \mathbf{C}^{p+1})$ with respect to $\mathbf{C}^{p+1}$ for the M step. The memory $L = 3 + \tau_{\max}/T$ comes from three symbols from GMSK spreading plus multipath spreading which can be as large as the maximum multipath delay $\tau_{\max}$.

We consider two different channel models, which we refer to as Channel I and Channel II, which have different average channel parameters. Figures 1 and 2 display BER vs. SNR for Channel I and Channel II respectively. The SNR is defined as

$$SNR = \frac{\sum_{i=1}^{M} \alpha_i^2}{N_o}, \tag{20}$$

where equation (20) is averaged over all the time slots used in the simulation. Convergence in most time slots is accomplished within 3 EM iterations.

In each figure the BER is plotted when the channel parameters are unknown and sequence estimation is done with the low noise EM algorithm. For comparison, the BER is also displayed for the case in which the channel parameters are known exactly and need not be estimated. The same time varying channel was used for the cases of known and unknown channels. The ML estimate of the sequence for the known channel is computed by minimizing the negative log likelihood (14) with β taking on its known true values, with one pass of the Viterbi Algorithm. For the case of unknown channel parameters, the sequence is estimated using the low noise EM algorithm discussed in Section IV.

From Figures 1 and 2, it is seen that throughout the whole range of SNR plotted, the low noise EM algorithm performs quite well; the BER is almost as low as for the ML estimate of the sequence when the channel parameters are known. It is seen that the low noise EM algorithm's performance, relative to the case of the known channel, improves as the SNR increases. As the SNR increases, the condition (17) is satisfied for a larger fraction of the time during a mobile's travels.

We note that the BER here can be improved, since our simulations include no interleaving or coding, and the received signal was sampled only once per bit. Furthermore, since there was no interleaving or coding, the absolute values of the BER are strongly dependent on the particular fades encountered in the 300 time slots. However, comparison of the BER for the low noise EM algorithm used for the unknown channel to that of the ML estimate for the same channel when it is known is noteworthy.

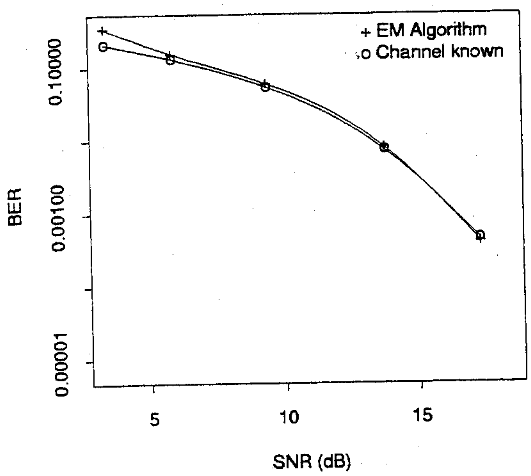
While the low noise algorithm is the EM algorithm whenever (17) holds and it therefore produces the ML sequence estimate, in the presence of noise, the final (P^{th}) channel parameter estimate $\tilde{\beta}_{s-1}^P$ in slot $s-1$ may differ from the true β in that slot. This difference along with the slowly varying true channel parameters could potentially make (19) far enough from the true channel parameters, so that the correct local minimum cannot be found. We assume that (19) is close enough to the true channel parameters in each time slot that the ML sequence estimate and the approximate local minimum corresponding to the true channel parameters can be found. If not, there are two potential methods to correct for this deviation. First, a smaller number of bits could be chosen over which the EM algorithm is run. Another option is to use training data every N slots to realign the channel parameters, where N is chosen to be less than the number of slots over which (19) begins to deviate too much from the true β.

VI Conclusions

The EM algorithm can be used to average over unknown multipath channel parameters to estimate a data sequence transmitted with GMSK. Training data would then only need to be used occasionally to realign the channel parameter estimates of the fast fading channel.

When the noise is low, the EM algorithm reduces to iteratively estimating the channel parameters and the transmitted sequence. Simulation results indicate that this low noise EM algorithm is successful for most of the range of practical SNR. Further simulations need to be performed to consider multiple reflected paths and a broader range of channel parameters.

This work has focused on GSM; future investiga-



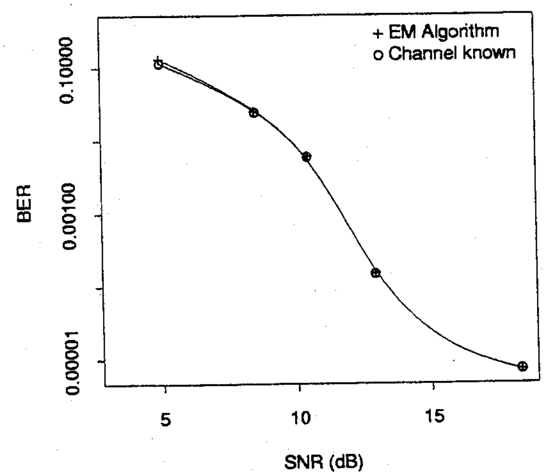


Figure 1: BER vs. SNR of low noise version of EM Algorithm for the unknown channel, and of ML sequence estimate for the corresponding known channel. Channel I: average parameter values: $\alpha_1 = 1.0, \alpha_2 = .7, \tau_2 = 2.5\mu$sec.

Figure 2: BER vs. SNR of low noise version of EM Algorithm for the unknown channel, and of ML sequence estimate for the corresponding known channel. Channel II: average parameter values: $\alpha_1 = 1.0, \alpha_2 = .5, \tau_2 = 6.2\mu$sec.

tion could determine its possible applicability to other wireless technologies using GMSK, such as CDPD and DECT.

References

[1] Y. Iwanami, "Performance of Sequence Estimation Scheme of Narrowband Digital FM Signals with Limiter–Discriminator Detection", *IEEE J. Selected Areas on Comm.*, vol. 12, Feb. 1995.

[2] A. Trachtman, I. Kalet, S. Shamai, "Limiter–Discriminator Detection of CPM Tomlinson Filtering", *IEEE Trans. Commun.*, vol. 42, pp. 819–825, March 1994.

[3] G. D'Aria, F. Muratore, and V. Palestini, "Simulation and Performance of the Pan–European Land Mobile Radio System", *IEEE Trans. Vehic. Techn.*, vol. 41, pp. 177-189, May 1992.

[4] A.P. Dempster, N.M. Laird, and D.B. Rubin, "Maximum likelihood from incomplete data via the EM Algorithm", *J. Royal Statistical Soc., Ser. B*, vol. 39, no. 1, pp.1-38, 1977.

[5] C. N. Georghiades, "Sequence Estimation in the Presence of Random Parameters Via the EM Algorithm", *IEEE Trans. Commun.*, vol. 45, pp. 300-308, March 1997.

[6] C.N. Georghiades and D.L. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection," *IEEE Trans. Comm.*, vol. 39, pp. 54-61, Jan. 1991.

[7] J.W. Modestino, "Reduced-complexity iterative maximum-likelihood sequence estimation on channels with memory," *Proc. Int. Symp. on Information Theory*, San Antonio, TX, Jan. 1993.

[8] M. Feder and E. Weinstein, "Parameter Estimation of Superimposed Signals Using the EM Algorithm", *IEEE Trans. Acoustics, Speech, Signal Proces.*, vol. 36, pp. 477-489, April 1988.

[9] A. Ansari and R. Viswanathan, "Application of Expectation-Maximization Algorithm to the Detection of a Direct-Sequence Signal in Pulsed Noise Jamming", *IEEE Trans. Commun.*, vol. 41, pp. 1151-1154, Aug. 1993.

[10] S. Zabin and H.V. Poor, "Efficient Estimation of Class A Noise Parameters Via the EM Algorithm", *IEEE Trans. Info. Theory*, vol. 37, pp.60-72, Jan. 1991.