BLIND MAXIMUM LIKELIHOOD SEQUENCE DETECTION OVER FAST FADING CHANNELS

*David J. Reader and William G. Cowley**

Communications Division, Defence Science and Technology Organisation, PO Box 1500, Salisbury, SA, 5108
Telephone: +61 8 259 6588 Facsimile: + 61 8 259 6549 E–Mail: david.reader@dsto.defence.gov.au

*The Institute for Telecommunications Research, University of South Australia, The Levels, Pooraka, SA, 5095
Telephone: +61 8 302 3316 Facsimile: + 61 8 302 3873 E–Mail: bill.cowley@unisa.edu.au

ABSTRACT

Maximum *a posteriori* (MAP) sequence detection for channels with intersymbol interference (ISI) has previously required knowledge of the channel sampled impulse response (SIR). Generally the SIR coefficients are determined via least mean square (LMS) or recursive least squares (RLS) estimation algorithms. For many unguided media channels such as mobile radio and high frequency radio which exhibit a time–varying SIR, these estimators must be adaptive. Adaptive estimators often fail to track adequately and are a major source of detector deterioration. A novel, blind maximum likelihood sequence detection (BMLSD) formulation without the need for external channel SIR estimation is proposed. The BMLSD performance is evaluated via simulation over several fast Rayleigh fading channels, which indicates substantial improvement compared to the conventional MLSD.

1 INTRODUCTION

It is well known [1] that the optimum filter for a received signal with no intersymbol interference (ISI) and additive white Gaussian noise is one matched to the signal. Matched filtering of a received signal with ISI will still maximise the signal to noise ratio (SNR) without altering the ISI, but will ‘‘colour” the noise. Any further filtering at this point to reduce the ISI will be at the expense of the SNR. A matched filter followed by an equaliser is the optimum linear system for a received signal with ISI. However for severe multipath channels, a non-linear technique such as decision feedback equalisation or maximum likelihood sequence detection (MLSD) may be required. Although more complex, the optimum MLSD is preferred for such channels, often employing the Viterbi algorithm to reduce computation [2]. The conventional MLSD metric requires knowledge of the channel sampled impulse response (SIR) to form the trellis through which it traces the most likely data sequence. In high frequency radio and mobile communications channels, dispersion causes the SIR coefficients to vary. Conventional least mean square (LMS) and recursive least squares (RLS) algorithms often fail to estimate these time–varying channels accurately enough, even with a large training overhead. Channel estimator performance is further reduced by detection delay and detection errors. For the fast fading two and three path channels this results in an irreducible detector bit error rate (BER) of between 0.01 and 0.0001.

Recently techniques which combine data sequence detection and channel estimation have been investigated to overcome these problems [3]. Note that in all of these methods, channel estimation in some form is included. For example, in [4], the use of per–survivor processing (PSP) eliminated the irreducible BER for the fast fading channel. However, to achieve this, 7% retraining was required for the channel estimator. Here we do not attempt to estimate the channel. Instead we integrate over all possible channels when forming the sequence metric for the MLSD. By removing the channel estimator, no training sequence is required to be transmitted. This is a distinct advantage, expected to outweigh any increases in complexity, in some applications. An earlier version of this approach [5] used a real valued time–varying channel model and numerical integration rather than the complex valued closed form result obtained in this paper. The work in this paper extends that published in [6] by presenting new results for the fast Rayleigh fading channel.

In Section 2 the channel model with symbol–spaced coefficients is described. The new blind maximum likelihood sequence detection (BMLSD) metric is derived in Section 3, and simulation results presented in Section 4.

2 CHANNEL MODEL

The channel is approximated by the discrete time complex baseband model shown in Figure 1.

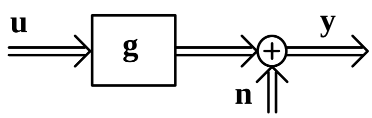

Figure 1 Discrete Time Channel Model

From Figure 1, the received symbol–spaced sequence is

$$\mathbf{y} = [y[0], y[1], ..., y[k]]^{\mathrm{T}} = \mathbf{U}\mathbf{g} + \mathbf{n} \tag{1}$$

where the received sample is

$$y[i] = \sum_{l=0}^{N_g-1} u[i-l]g[l] + n[i] \qquad \text{for } i = 0, ..., k \tag{2}$$

the transmitted symbol at time i is $u[i]$, a data matrix is defined as

$$\mathbf{U} \triangleq \begin{bmatrix} u[0] & u[-1] & \dots & u\left[1-N_g\right] \\ u[1] & u[0] & \dots & u\left[2-N_g\right] \\ \vdots & \vdots & \vdots & \vdots \\ u[k] & u[k-1] & \dots & u\left[k+1-N_g\right] \end{bmatrix} \tag{3}$$

and $\mathbf{g} = \left[g[0], g[1], \dots, g\left[N_g-1\right]\right]^{\mathrm{T}}$, is the channel SIR vector with length N_g. The additive noise component at time i is $n[i]$. Note that the superscripts T and H denote transpose and Hermitian (complex conjugate transpose) respectively.

3 MAP SEQUENCE DETECTION

The maximum *a posteriori* (MAP) transmitted data sequence is

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmax}}\{P_{\mathbf{u}}[\mathbf{u} \mid \mathbf{y}]\} \tag{4}$$

where $P_{\mathbf{u}}[\mathbf{u} \mid \mathbf{y}]$ is the conditional probability that the data sequence $\mathbf{u} = [u[0],...,u[k]]^{\mathrm{T}}$ was transmitted given that the sequence $\mathbf{y}$ has been received. By application of Bayes' Rule

$$P_{\mathbf{u}}[\mathbf{u} \mid \mathbf{y}] = \frac{p_{\mathbf{uy}}(\mathbf{u},\mathbf{y})}{p_{\mathbf{y}}(\mathbf{y})} = \frac{p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u})P_{\mathbf{u}}[\mathbf{u}]}{p_{\mathbf{y}}(\mathbf{y})} \tag{5}$$

where $p_{\mathbf{uy}}(\mathbf{u},\mathbf{y})$ is the joint probability density function (PDF) of the sequences $\mathbf{u}$ and $\mathbf{y}$. As $p_{\mathbf{y}}(\mathbf{y})$ is independent of the selection of $\mathbf{u}$ it can be omitted without affecting the maximisation over $\mathbf{u}$. Equation (4) therefore becomes

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmax}}\left\{p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u})P_{\mathbf{u}}[\mathbf{u}]\right\} \tag{6}$$

where $p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u})$ is the PDF of the received sequence $\mathbf{y}$ conditioned upon the transmitted data sequence, $\mathbf{u}$. Integrating over all $\mathbf{g}$, and as $\mathbf{u}$ and $\mathbf{g}$ are independent, the argument of equation (6) becomes

$$p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u})P_{\mathbf{u}}[\mathbf{u}] = P_{\mathbf{u}}[\mathbf{u}] \int_{\mathbf{g}\in\mathbb{C}_{N_g}} p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u},\mathbf{g})p_{\mathbf{g}}(\mathbf{g})d\mathbf{g} \tag{7}$$

where $p_{\mathbf{g}}(\mathbf{g})$ is the prior PDF of the channel, $\mathbf{g}$. For uncoded uniformly distributed data $P_{\mathbf{u}}[\mathbf{u}]$ is independent of the choice of $\mathbf{u}$ and so may be omitted from equation (7). $p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u},\mathbf{g})$ is simply a function of the noise distribution which is assumed to be complex valued multivariate Gaussian

$$p_{\mathbf{y}}(\mathbf{y} \mid \mathbf{u},\mathbf{g}) = \frac{\exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{Ug})^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}(\mathbf{y} - \mathbf{Ug})\right)}{(2\pi)^{k+1}\det(\mathfrak{R}_{\mathbf{n}})} \tag{8}$$

where $\mathfrak{R}_{\mathbf{n}}^{-1}$ is the inverse of the covariance matrix of $\mathbf{n}$, and Λ is a real valued diagonal weighting matrix defined as

$$\Lambda \stackrel{\Delta}{=} \operatorname{diag}(\lambda[0],\lambda[1],\ldots,\lambda[k]) \tag{9}$$

Inclusion of this weighting matrix can be considered as using an altered noise covariance matrix equal to $\Lambda^{-1/2}\mathfrak{R}_{\mathbf{n}}\Lambda^{-1/2}$ and allows the metric contributions to be emphasised based upon the time index, k. For example, if the actual channel, $\mathbf{g}$, or noise covariance matrix, $\mathfrak{R}_{\mathbf{n}}$, is time–varying, less emphasis may be put upon older contributions. The noise correlation may be attributed to the receive filtering.

Note that the choice of the prior PDF, $p_{\mathbf{g}}(\mathbf{g})$, is a very controversial issue [7], and should be based upon some knowledge of the physical system. Here it is assumed that the channel coefficients remain stationary over the observation interval of the detector and that they are from a Gaussian distribution with PDF

$$p_{\mathbf{g}}(\mathbf{g}) = \frac{1}{(2\pi)^{N_g}\det(\Sigma_{\mathbf{g}})}\exp - \frac{1}{2}\left((\mathbf{g} - \mathbf{m}_{\mathbf{g}})^{\mathrm{H}}\Sigma_{\mathbf{g}}^{-1}(\mathbf{g} - \mathbf{m}_{\mathbf{g}})\right) \tag{10}$$

mean value, $\mathbf{m_g}$, and covariance matrix, $\Sigma_{\mathbf{g}}$. This may be viewed as providing the detector with an estimate $\mathbf{m_g}$ of the channel and some information $\Sigma_{\mathbf{g}}$ on how accurate that estimate is. $\Sigma_{\mathbf{g}}$ also allows for correlated channel coefficients, which is often the case in practice. If the mean is close to the actual channel, then the PDF in equation (10) may be termed a "minimally informative" prior [7, pp 55]. If there is limited knowledge of the mean, then $\Sigma_{\mathbf{g}}$ may be increased. For example, if an external channel estimate is used for the mean, then $\Sigma_{\mathbf{g}}$ may be used to inform the detector of its accuracy. Note that the conventional MLSD metric uses a channel estimate which it assumes is correct. For Gaussian data, the joint PDF of a LMS channel estimate is Gaussian [8, pp 335]. Note that this can be used as further justification for the choice of a Gaussian prior, if equation (10) is interpreted as the PDF of $\mathbf{m_g}$.

From equations (6), (7), (8) and (10), setting the appropriate integration limits, the maximum likelihood (ML) data sequence is

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmax}}\left\{\frac{1}{(2\pi)^{k+1}\det(\mathfrak{R}_{\mathbf{n}})(2\pi)^{N_g}\det(\Sigma_{\mathbf{g}})}\int_{\mathbf{g}=-\infty}^{\infty}\exp-(\Gamma + \Omega)d\mathbf{g}\right\} \tag{11}$$

where from equations (8) and (10) the argument of the exponential is

$$\Gamma + \Omega = \frac{1}{2}(\mathbf{y}–\mathbf{Ug})^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}(\mathbf{y}–\mathbf{Ug}) + \frac{1}{2}(\mathbf{g}–\mathbf{m_g})^{\mathrm{H}}\Sigma_{\mathbf{g}}^{-1}(\mathbf{g}–\mathbf{m_g}) \tag{12}$$

which by expanding and then grouping $\mathbf{g}$ terms becomes

$$\begin{aligned} &= \frac{1}{2}\Big(\mathbf{g}^{\mathrm{H}}\big(\mathbf{U}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{U} + \Sigma_{\mathbf{g}}^{-1}\big)\mathbf{g} - \\ &\quad 2\operatorname{Re}\left\{\mathbf{g}^{\mathrm{H}}\big(\mathbf{U}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{y} + \Sigma_{\mathbf{g}}^{-1}\mathbf{m_g}\big)\right\} + \\ &\quad \mathbf{y}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{y} + \mathbf{m}_{\mathbf{g}}^{\mathrm{H}}\Sigma_{\mathbf{g}}^{-1}\mathbf{m_g}\Big) \end{aligned} \tag{13}$$

It can be seen from equation (12) that, with Λ equal to the identity matrix, $\mathbf{I}$, and as time increases ($k \to \infty$), the data PDF terms, Γ, will dominate the prior PDF terms, Ω. The latter then just become initialisation conditions. In this case the data "swamps out" the prior knowledge. However, for weighting matrices with limited memory, the prior knowledge will always have some effect upon equation (12). This is unlike conventional weighted algorithms, for example weighted RLS, where the additional term added to the correlation matrix is also weighted by λ^k [8, pp 484]. This causes the effect of the initialisation condition to decay as k increases. Here the prior information has deliberately been allowed to remain.

Using equation (13) the integral in equation (11) can now be expressed in the form [9]

$$\int_{\mathbf{g}=-\infty}^{\infty}\exp-\left(\mathbf{g}^{\mathrm{H}}\mathbf{Ag} + 2\operatorname{Re}\left\{\mathbf{g}^{\mathrm{H}}\mathbf{b}\right\} + c\right)d\mathbf{g} = \frac{\pi^{N_g}}{\det(\mathbf{A})}\exp\left(\mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{b}–c\right) \tag{14}$$

where from equation (13)

$$\mathbf{A} \stackrel{\Delta}{=} \frac{1}{2}\left(\mathbf{U}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{U} + \Sigma_{\mathbf{g}}^{-1}\right) \tag{15}$$

$$\mathbf{b} \stackrel{\Delta}{=} -\frac{1}{2}\left(\mathbf{U}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{y} + \Sigma_{\mathbf{g}}^{-1}\mathbf{m_g}\right) \tag{16}$$

$$c \stackrel{\Delta}{=} \frac{1}{2}\left(\mathbf{y}^{\mathrm{H}}\Lambda^{1/2}\mathfrak{R}_{\mathbf{n}}^{-1}\Lambda^{1/2}\mathbf{y} + \mathbf{m}_{\mathbf{g}}^{\mathrm{H}}\Sigma_{\mathbf{g}}^{-1}\mathbf{m_g}\right) \tag{17}$$

Substituting equation (14) into (11), the ML data sequence becomes

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmax}}\left\{ \frac{\exp(\mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{b} - c)}{(2\pi)^{k+1} 2^{N_g} \det(\mathcal{R}_{\mathbf{n}}) \det(\boldsymbol{\Sigma}_{\mathbf{g}}) \det(\mathbf{A})} \right\} \tag{18}$$

Omitting the terms in equation (18) which are independent of **u**, taking the logarithm, and changing to a minimisation over **u** the ML detected sequence is given by

$$\hat{\mathbf{u}} = \underset{\mathbf{u}}{\operatorname{argmin}}\{\psi\} \tag{19}$$

where the BMLSD metric is defined as

$$\psi \stackrel{\Delta}{=} \ln(\det(\mathbf{A})) - \mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{b} \tag{20}$$

Note that as ln(.) is a monotonic function, taking the logarithm does not affect the minimisation over **u**.

3.1 Uncorrelated Channel and Noise

The special case of uncorrelated channel coefficients and noise is now considered. Although the channel coefficients are not always uncorrelated, for most time–varying channels this is an acceptable approximation [1]. A whitened matched filter technique [10] can be used to justify the uncorrelated noise assumption. For uncorrelated channel coefficients, **g**, the covariance matrix, $\boldsymbol{\Sigma}_{\mathbf{g}}$, equals $\sigma_g^2\mathbf{I}$ and for uncorrelated noise, **n**, the covariance matrix, $\mathcal{R}_{\mathbf{n}}$, equals $\sigma_n^2\mathbf{I}$.

3.2 Metric Interpretations

The metric in equation (20) is now interpreted as a modified Euclidean distance metric based upon per–survivor Bayesian minimum mean square error (BMMSE) channel estimates. For a weighting matrix, Λ, equal to the identity matrix, **I**, it may be shown that the BMMSE estimate of the channel, as defined in [11, pp 364], is

$$\hat{\mathbf{g}}_{\mathbf{MSE}} = \mathbf{A}^{-1}\mathbf{b} \tag{21}$$

where **A** and **b** are defined in equations (15) and (16). Note that this estimate is a function of **u**. Now substituting equation (21) into equation (18), the argument of the exponential in equation (18) becomes

$$\mathbf{b}^{\mathrm{H}}\mathbf{A}^{-1}\mathbf{b} - c = \mathbf{b}^{\mathrm{H}}\, \hat{\mathbf{g}}_{\mathbf{MSE}} - c = -\hat{\mathbf{g}}_{\mathbf{MSE}}^{\mathrm{H}}\mathbf{A}\hat{\mathbf{g}}_{\mathbf{MSE}} + 2\operatorname{Re}\left\{\hat{\mathbf{g}}_{\mathbf{MSE}}^{\mathrm{H}}\mathbf{b}\right\} - c \tag{22}$$

which using equations (15) to (17), simplifies to

$$= -\frac{1}{2}\left(\mathbf{y} - \mathbf{U}\hat{\mathbf{g}}_{\mathbf{MSE}}\right)^{\mathrm{H}} \mathcal{R}_{\mathbf{n}}^{-1}\left(\mathbf{y} - \mathbf{U}\hat{\mathbf{g}}_{\mathbf{MSE}}\right) - \frac{1}{2}\left(\hat{\mathbf{g}}_{\mathbf{MSE}} - \mathbf{m}_{\mathbf{g}}\right)^{\mathrm{H}} \boldsymbol{\Sigma}_{\mathbf{g}}^{-1}\left(\hat{\mathbf{g}}_{\mathbf{MSE}} - \mathbf{m}_{\mathbf{g}}\right) \tag{23}$$

Minimising the metric in equation (20) is equivalent to maximising equation (23). The first term in equation (23) is the conventional Euclidean distance metric with a BMMSE estimate of the channel based upon each stored sequence, **u**. The second term is the Euclidean distance between the BMMSE channel estimate and the channel mean, and is due to the prior PDF of **g**.

In [6] we have shown that as the variance of the channel coefficients becomes small, ie: $\sigma_g^2 \to 0$, the metric in equation (18) approaches the conventional Euclidean distance metric [10] with the channel coefficient mean, $\mathbf{m_g}$, as the channel estimate. We have also shown in [6] that as the variance of the channel coefficients becomes large, ie: $\sigma_g^2 \to \infty$, using the metric in equation (18) is equivalent to a ''per–survivor'' processing (PSP) detector with least square channel estimates. Therefore the new BMLSD metric is a generalisation of both the conventional MLSD and PSP techniques.

4 PERFORMANCE

To determine the performance of the detection metrics derived in Section 3, differentially encoded binary phase shift keyed (DBPSK) and quaternary phase shift keyed (DQPSK) data symbols were transmitted over fast Rayleigh fading two and three path channel models. A multistage, multirate filter structure was used to generate the channel coefficients with a lowpass Gaussian Doppler power spectrum shape. The 3dB Doppler spread of each coefficient, normalised to the symbol rate, was 0.005. The channel noise was additive white Gaussian. The variance of the noise was adjusted to obtain the required average ratio of bit energy to noise power spectral density, E_b/N_0. The bit error rate (BER) at each E_b/N_0 value was measured for three different noise seeds and the average value plotted. The error bars in Figures 2 and 3 indicate the maximum and minimum BER measured. A BER measurement ended when at least 325,000 symbols (approximately 1000 channel coefficient fades) had been transmitted and at least 220 symbol errors had been recorded, or when a maximum of 2×10^7 samples had been transmitted. The value of 220 was chosen to obtain measurements with an uncertainty of 0.1 with 95% confidence. The simulation limit shown in Figures 2 and 3 is equal to $220/2 \times 10^7 = 1.1 \times 10^{-5}$.

4.1 Implementation

The Viterbi algorithm was used for ML detection of the input sequence **u**. The detection delay was set to $5N_g$ [1]. The BMLSD metric mean, $\mathbf{m_g}$, equalled (0,0), the variance, σ_g^2, equalled 0.1. The exact value of σ_g^2 chosen was not critical to the performance of the BMLSD. The weighting matrix was

$$\lambda[i] = \lambda^{k-i} \qquad 0 < \lambda < 1 \quad \text{and} \quad 0 \le i \le k \tag{24}$$

which gave an exponentially decaying metric memory with forgetting factor, λ. The best value for λ was chosen via simulation to be 0.7. For a detection metric with a memory length of L, the detector should have M^L states. However only a M^{Ng} state detector was used here. The reduced state implementation did not cause any noticeable performance degradation. The BMLSD metric in equation (20) was implemented both directly, and using LU decomposition [8]. The required computation was of the order of N_g^3. For length two and three channels, this resulted in a increase in execution time of approximately ten to twenty with respect to the conventional MLSD. Recursive algorithms have been developed to reduce computation.

For comparison a conventional MLSD and LMS channel estimator were also implemented. The MLSD was implemented using the Viterbi algorithm.

4.2 Simulation Results

For conventional MLSD over fast fading channels an irreducible bit error rate exists. This is due to two types of error. The first is termed ''phase flipping''. With both the channel and data unknown the detector is able to choose two equally valid solutions, out of phase with each other by a multiple of $2\pi/M$ radians, where M is the size of the data symbol alphabet. The detector then may ''flip'' between these solutions, causing detection errors. Methods to prevent phase flipping have been derived and tested successfully in [12]. Modifying the traceback part of the Viterbi algorithm gave the best results and was used here. The second type of error is termed ''delay ambiguity''. This occurs when the detector chooses a time shifted version of the correct transmitted data sequence. A method which successfully prevents delay ambiguity is presented in [6]. It involves periodically rotating the phase of the transmitted data symbols and was also used here.

In Figure 2, the performance of the new BMLSD metric over the

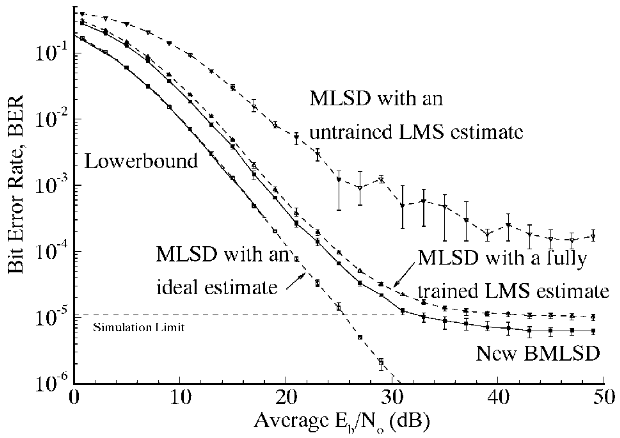

Figure 2 BER Performance of BMLSD Over a Fast Fading Two Path Channel with DBPSK Data

fast fading two path channel is shown. The lowerbound for detection of a differentially encoded isolated pulse transmitted over a Rayleigh fading two path channel is also shown. Note the substantial performance improvement obtained by the new BMLSD over the untrained LMS/MLSD system. The irreducible BER is reduced by between one and two orders of magnitude. Due to the inherent PSP of the BMLSD it even outperforms the fully trained LMS/MLSD system. Similar performance improvements were also obtained over three path channels and with DQPSK data as shown in Figure 3.

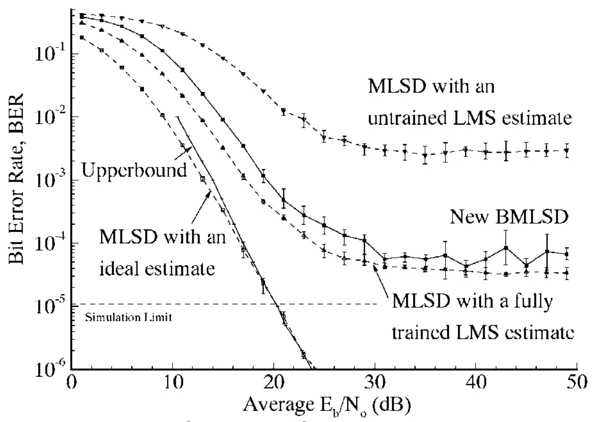

Figure 3 BER Performance of BMLSD over a Fast Fading Three Path Channel with DQPSK Data

The upperbound on the probability of error is for MLSD of data transmitted over a Rayleigh fading three path channel.

5 CONCLUSIONS

The performance shown by the new BMLSD metric over a fast Rayleigh fading channel is superior to that obtained with the conventional untrained LMS/MLSD system. The improved performance is at the expense of a moderate increase in computation. Fast recursive algorithms are expected to reduce this increase. The new detection algorithm has also allowed information about the accuracy of the channel estimate to be given to the detector, in the form of $\mathbf{\Sigma_g}$. For example the technique could be used with a channel estimator providing both $\mathbf{m_g}$ and σ_g^2 to improve overall performance. The metric is a generalisation of both the conventional MLSD and PSP techniques. A fractionally–spaced version of the new BMLSD has now been developed [12] [13].

REFERENCES

[1] Proakis J.G., *Digital Communications*, 2nd Ed., McGraw–Hill, 1989

[2] Forney G.D.Jr., "The Viterbi Algorithm", *Proc. of the IEEE*, Vol. 61, No. 3, pp 268–278, March 1973

[3] Haykin S., Ed., *Blind Deconvolution*, Prentice–Hall, 1994

[4] Polydoros A., Raheli R., Tzou C–K., "Applications Of Per–Survivor Processing To Mobile Digital Communications", *Proc. of the IEEE Global Telecommunications Conf.*, GLOBECOM '93, Houston, U.S.A., pp 77–81, November 1993

[5] Reader D.J., Scholz J.B., "Blind Maximum Likelihood Sequence Estimation For Fading Channels", *Proc. of the IEEE Int. Conf. on Acoustics, Speech & Signal Processing*, ICASSP '94, Adelaide, South Australia, Vol. IV, No. 3, pp 401–404, April 1994

[6] Reader D.J., Cowley W.C., "Blind Maximum Likelihood Sequence Estimation", *Proc. of the IEEE Int. Conf. on Acoustics, Speech & Signal Processing*, ICASSP '96, Atlanta, U.S.A., 1996

[7] Quinn A.P., "Bayesian Point Inference in Signal Processing", *Doctoral Thesis*, University of Cambridge, UK, 1992

[8] Haykin S., *Adaptive Filter Theory*, 2nd Edition, Prentice–Hall, 1991

[9] Farrell R.H., *Techniques of Multivariate Calculation*, Lecture Notes in Mathematics, 520, Springer–Verlag, 1976

[10] Forney G.D. Jr., "Maximum–Likelihood Sequence Estimation Of Digital Sequences In The Presence Of Intersymbol Interference", *IEEE Trans. on Information Technology*, Vol. 18, No. 3, pp 363–378, 1972

[11] Kay M. S., *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice–Hall, Englewood Cliffs, NJ, 1993

[12] Reader D.J., "Blind Maximum Likelihood Sequence Detection Over Fast Fading Communication Channels", *Doctoral Thesis*, University of South Australia, Submitted March 1996

[13] Reader D.J., Cowley W.C., "Fractionally–Spaced Blind Maximum Likelihood Sequence Detection Over Fast Fading Fractionally–Spaced Channels", Submitted to the *1996 IEEE Digital Signal Processing Workshop*, Loen, Norway, September 1996