# ADAPTIVE MINIMUM BIT ERROR RATE BEAMFORMING ASSISTED RECEIVER FOR WIRELESS COMMUNICATIONS

*S. Chen, L. Hanzo and N.N. Ahmad*

Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

## ABSTRACT

A novel adaptive beamforming technique is proposed for wireless communication application based on the minimum bit error rate (MBER) criterion. It is shown that the MBER approach provides significant performance gain in terms of smaller bit error rate (BER) over the standard minimum mean square error (MMSE) approach. Using the classical Parzen window estimate of probability density function (p.d.f.), both the block-data and sample-by-sample adaptive implementations of the MBER solution are developed.

## 1. INTRODUCTION

Spatial processing with adaptive antenna array has shown real promise for substantial capacity enhancement in mobile communication [1]–[5]. Adaptive beamforming can separate signals transmitted on the same carrier frequency, provided that they are separated in the spatial domain. The beamforming processing, which combines the signals received by the different elements of an antenna array to form a single output, is classically done by minimizing the mean square error between the desired and actual array outputs. However, for a communication system, it is the BER that really matters. We propose a novel beamforming technique based on minimizing the system BER. Adopting Parzen window or kernel density estimation [6]–[8] to approximate the p.d.f. of the beamformer output, a block-data adaptive MBER algorithm is derived. This is then further simplified to develop a stochastic gradient adaptive MBER algorithm called the approximate least bit error rate (ALBER).

## 2. SYSTEM MODEL

The system consists of $M$ users (sources), and each user transmits a binary phase shift keying signal on the same carrier frequency $\omega = 2\pi f$. The baseband signal of user $i$ is

$$m_i(k) = A_i b_i(k), \quad 1 \leq i \leq M, \tag{1}$$

where $b_i(k) \in \{\pm 1\}$ and $A_i^2$ is user $i$ signal power. The source 1 is the desired user and the other sources are interfering users. The linear antenna array consists of $L$ uniformly spaced elements, and signals at the antenna array are

$$x_l(k) = \sum_{i=1}^{M} m_i(k) \exp\left(j\omega t_l(\theta_i)\right) + n_l(k)$$

$$= \bar{x}_l(k) + n_l(k), \quad 1 \leq l \leq L, \tag{2}$$

where $t_l(\theta_i)$ is the relative time delay at element $l$ for source $i$, $\theta_i$ is the direction of arrival for source $i$, and $n_l(k)$ is a complex white Gaussian noise with zero mean and variance $E[|n_l(k)|^2] = 2\sigma_n^2$. The desired signal to noise ratio is SNR$= A_1^2/2\sigma_n^2$ and the desired signal to interferer $i$ ratio is SIR$_i = A_1^2/A_i^2$, for $2 \leq i \leq M$. In vector form, the array input $\mathbf{x}(k) = [x_1(k) \cdots x_L(k)]^T$ can be expressed as

$$\mathbf{x}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k) = \mathbf{P}\mathbf{b}(k) + \mathbf{n}(k) \tag{3}$$

where $E[\mathbf{n}(k)\mathbf{n}^H(k)] = 2\sigma_n^2 \mathbf{I}_L$, the system matrix $\mathbf{P} = [A_1 \mathbf{s}_1 \cdots A_M \mathbf{s}_M]$ with the steering vector for source $i$ $\mathbf{s}_i = [\exp(j\omega t_1(\theta_i)) \cdots \exp(j\omega t_L(\theta_i))]^T$, and the bit vector $\mathbf{b}(k) = [b_1(k) \cdots b_M(k)]^T$. Note that $\bar{\mathbf{x}}(k) \in \mathcal{X} \triangleq \{\bar{\mathbf{x}}_q = \mathbf{P}\mathbf{b}_q, 1 \leq q \leq N_b\}$, where $N_b = 2^M$ and $\mathbf{b}_q, 1 \leq q \leq N_b$, are all the possible sequences of $\mathbf{b}(k)$.

The beamformer output is

$$y(k) = \mathbf{w}^H \mathbf{x}(k) = \bar{y}(k) + e(k) \tag{4}$$

where $\mathbf{w}$ is the complex beamformer weight vector, and $e(k) = \mathbf{w}^H \mathbf{n}(k)$ is Gaussian with zero mean and $E[|e(k)|^2] = 2\sigma_n^2 \mathbf{w}^H \mathbf{w}$. The estimate of the transmitted bit $b_1(k)$ is

$$\hat{b}_1(k) = \text{sgn}(y_R(k)) \tag{5}$$

where $y_R(k) = \Re[y(k)]$ and sgn$(\cdot)$ the sign function. Note that $\bar{y}(k) \in \mathcal{Y} \triangleq \{\bar{y}_q = \mathbf{w}^H \bar{\mathbf{x}}_q, 1 \leq q \leq N_b\}$. Thus, $\bar{y}_R(k)$ can only take values from the set

$$\mathcal{Y}_R \triangleq \{\bar{y}_{R,q} = \Re[\bar{y}_q], 1 \leq q \leq N_b\} \tag{6}$$

which can be divided into the two subsets

$$\mathcal{Y}_R^{(\pm)} \triangleq \{\bar{y}_{R,q}^{(\pm)} \in \mathcal{Y}_R : b_1(k) = \pm 1\}. \tag{7}$$

## 3. MBER BEAMFORMING SOLUTION

The p.d.f. of $y_R(k)$ is

$$p(y_R) = \frac{1}{N_b\sqrt{2\pi\sigma_n^2\mathbf{w}^H\mathbf{w}}} \sum_{q=1}^{N_b} \exp\left(-\frac{(y_R - \bar{y}_{R,q})^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \tag{8}$$

and it can be shown that the BER is given by

$$P_E(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} Q\left(g_{q,+}(\mathbf{w})\right) \tag{9}$$

where $N_{sb} = N_b/2$ is the number of the points in $\mathcal{Y}_R^{(+)}$,

$$g_{q,+}(\mathbf{w}) = \frac{\text{sgn}(b_{q,1})\bar{y}_{R,q}^{(+)}}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}}, \tag{10}$$

and $b_{q,1}$ is the first element of $\mathbf{b}_q$ related to the desired user. Note that the BER is invariant to a positive scaling of $\mathbf{w}$. Alternatively, the BER can be calculated using $\mathcal{Y}_R^{(-)}$.

The MBER beamforming solution is then defined as

$$\mathbf{w}_{\text{MBER}} = \arg\min_{\mathbf{w}} P_E(\mathbf{w}). \tag{11}$$

The gradient of $P_E(\mathbf{w})$ with respect to $\mathbf{w}$ is

$$\nabla P_E(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{q=1}^{N_{sb}} \exp\left(-\frac{\left(\bar{y}_{R,q}^{(+)}\right)^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right)$$

$$\times \text{sgn}(b_{q,1})\left(\frac{\bar{y}_{R,q}^{(+)}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} - \bar{\mathbf{x}}_q^{(+)}\right). \tag{12}$$

The optimization problem (11) can be solved for iteratively using a simplified conjugated gradient algorithm [9],[10].

## 4. ADAPTIVE MBER BEAMFORMING

Given a block of $K$ training samples $\{\mathbf{x}(k), b_1(k)\}$, a Parzen window estimate of the p.d.f. (8) is given by:

$$\hat{p}(y_R) = \frac{1}{K\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{k=1}^{K} \exp\left(-\frac{(y_R - y_R(k))^2}{2\rho_n^2\mathbf{w}^H\mathbf{w}}\right) \tag{13}$$

where the kernel width $\rho_n$ is related to the noise standard deviation $\sigma_n$. From this estimated p.d.f., the estimated BER is given by

$$\hat{P}_E(\mathbf{w}) = \frac{1}{K} \sum_{k=1}^{K} Q\left(\hat{g}_k(\mathbf{w})\right) \tag{14}$$

with

$$\hat{g}_k(\mathbf{w}) = \frac{\text{sgn}(b_1(k))y_R(k)}{\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}}. \tag{15}$$

The gradient of $\hat{P}_E(\mathbf{w})$ is

$$\nabla\hat{P}_E(\mathbf{w}) = \frac{1}{2K\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{k=1}^{K} \exp\left(-\frac{y_R^2(k)}{2\rho_n^2\mathbf{w}^H\mathbf{w}}\right)$$

$$\times \text{sgn}(b_1(k))\left(\frac{y_R(k)\mathbf{w}}{\mathbf{w}^H\mathbf{w}} - \mathbf{x}(k)\right). \tag{16}$$

By substituting $\nabla P_E(\mathbf{w})$ with $\nabla\hat{P}_E(\mathbf{w})$ in the conjugate gradient updating mechanism, a block-data adaptive MBER algorithm is readily obtained. The step size $\mu$ and the kernel width $\rho_n$ are the two algorithm parameters.

An alternative Parzen window estimate is given by

$$\tilde{p}(y_R) = \frac{1}{K\sqrt{2\pi}\rho_n} \sum_{k=1}^{K} \exp\left(-\frac{(y_R - y_R(k))^2}{2\rho_n^2}\right) \tag{17}$$

and an approximation of the BER is

$$\tilde{P}_E(\mathbf{w}) = \frac{1}{K} \sum_{k=1}^{K} Q\left(\tilde{g}_k(\mathbf{w})\right) \tag{18}$$

with

$$\tilde{g}_k(\mathbf{w}) = \frac{\text{sgn}(b_1(k))y_R(k)}{\rho_n}. \tag{19}$$

This approximation is valid provided that the width $\rho_n$ is chosen appropriately. Adopting this approach and considering sample-by-sample adaptation leads to the stochastic gradient adaptive MBER algorithm called ALBER:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\frac{\text{sgn}(b_1(k))}{2\sqrt{2\pi}\rho_n} \exp\left(-\frac{y_R^2(k)}{2\rho_n^2}\right) \mathbf{x}(k). \tag{20}$$

This ALBER algorithm has a similar computational complexity to the very simple least mean square algorithm.
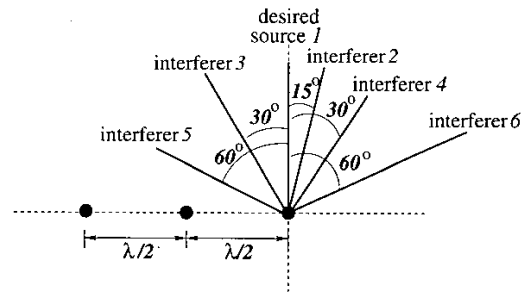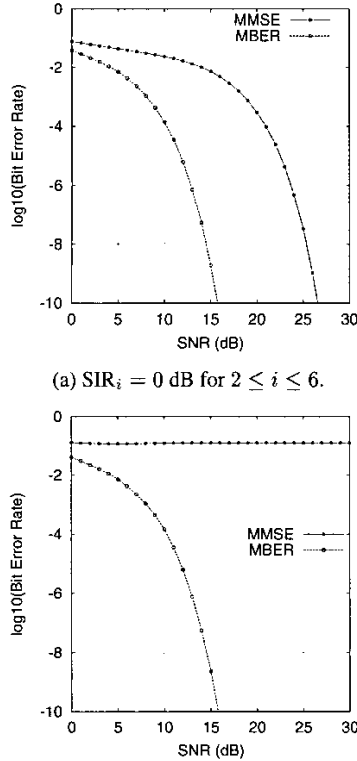


Figure 1: Locations of the desired source and the interfering sources with respect to the three-element linear array with $\lambda/2$ element spacing, $\lambda$ being the wavelength.

(a) $SIR_i = 0$ dB for $2 \le i \le 6$.



(b) $SIR_i = 0$ dB for $i = 2, 4, 5, 6$, and $SIR_3 = -6$ dB.

Figure 2: Comparison of bit error rate performance.

## 5. SIMULATION STUDY

The example consisted of six sources and a three-element antenna array. Fig. 1 shows the locations of the desired source and the interfering sources graphically. Fig. 2 compares the BER performance of the MBER solution with that of the MMSE solution under two different conditions: (a) the desired user and all the five interfering sources had equal power, and (b) the desired user and the interfering sources $2, 4, 5, 6$ had equal power, but the interfering source 3 had 6 dB more power than the desired user. Under the condition given in Fig. 2 (b), the MMSE beamformer had a very high error rate floor of above $\frac{1}{8}$. The reason for this was investigated. Given SNR= 14 dB, $SIR_i = 0$ dB for $i = 2, 4, 5, 6$ and $SIR_3 = -6$ dB, Fig. 3 compares the conditional p.d.f. given $b_1(k) = +1$ of the MMSE beamformer with that of the MBER beamformer, where the beamformer weight vector had been normalized to a unit length. It can be seen that under the given condition the resulting $y_R^{(-)}$ and $y_R^{(+)}$ for the MMSE beamforming were linearly inseparable. There were $N_{sb} = 32$ points in $y_R^{(+)}$, and a cluster of four points was on the wrong side of the decision boundary $y_R = 0$ for
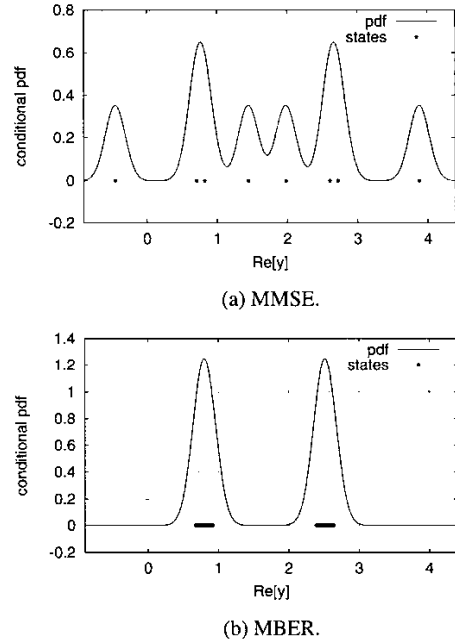


(a) MMSE.



(b) MBER.

Figure 3: Conditional probability density function of beamformer given $b_1(k) = +1$. SNR= 14 dB, $SIR_i = 0$ dB for $i = 2, 4, 5, 6$, and $SIR_3 = -6$ dB.

the MMSE beamforming.

Performance of the block-data gradient adaptive MBER algorithm was next studied. Fig. 4 illustrates the convergence rates of the algorithm given SNR= 14 dB, $SIR_i = 0$ dB for $2 \le i \le 6$ and the two different initial weight vectors. It can be seen that this block-data adaptive MBER algorithm generally converged rapidly. As the BER surface is highly complicated, the initial condition will influence convergence rate. We have found out that the MMSE solution $w_{MMSE}$ is typically not a good initial choice in terms of convergence rate. Performance of the stochastic gradient adaptive MBER algorithm was also investigated. Fig. 5 shows the learning curves of the ALBER algorithm under the same conditions of Fig. 4, where DD denotes decision-directed adaptation with $\hat{b}_1(k)$ substituting $b_1(k)$. It can be seen that the ALBER algorithm had a reasonable convergence speed. Note that the steady-state BER misadjustment was higher when the initial weight vector was set to $w_{MMSE}$, compared with the other initial condition.

## 6. CONCLUSIONS

An adaptive MBER beamforming technique has been proposed. It has been demonstrated that the MBER beam-
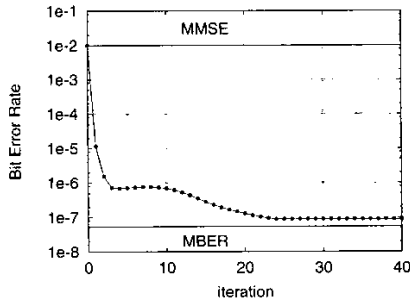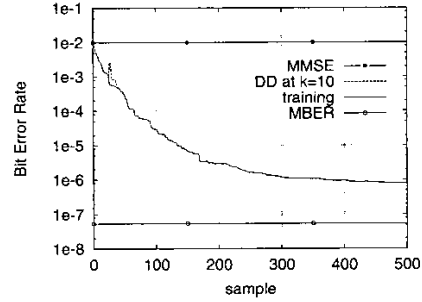
(a) Initial $w_0 = w_{MMSE}$.
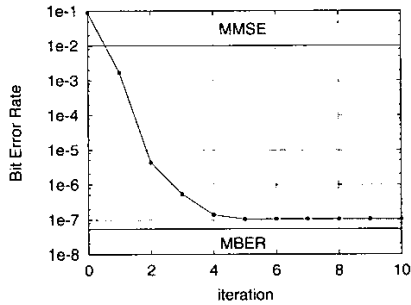


(b) Initial $w_0 = [0.1 + j0.0\ \ 0.1 + j0.0\ \ 0.1 + j0.0]^T$.

Figure 4: Convergence rates of block-data adaptive gradient MBER algorithm. SNR= 14 dB and SIR$_i$ = 0 dB for $2 \leq i \leq 6$. Block size $K = 200$, $\mu = 0.6$ and $\rho_n^2 = 4\sigma_n^2 = 0.08$.

former utilizes the system resource more intelligently than the standard MMSE beamformer and, consequently, achieves a better performance in terms of a smaller BER. The results also suggest that the MBER solution is robust to the near-far effect. Adaptive implementation of the MBER beam-forming solution has been developed based on the classical approach of Parzen window estimate for the p.d.f. of the beamformer output. A block-data conjugate gradient adaptive MBER algorithm has been shown to converge rapidly and requires a reasonably small block size to accurately approximate the theoretical MBER solution. A LMS-style stochastic gradient adaptive MBER algorithm called the AL-BER has been shown to perform well. Current work is investigating the extension of the proposed adaptive MBER beamforming to other modulation schemes.

## 7. REFERENCES

[1] J.H. Winters, J. Salz and R.D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Communications*, Vol.42, No.2, pp.1740–1751, February/March/April 1994.

[2] J. Litva and T. K.Y. Lo, *Digital Beamforming in Wireless Communications*. London: Artech House, 1996.

(a) Initial $w_0 = w_{MMSE}$, $\mu = 0.03$ and $\rho_n^2 = 0.04$.
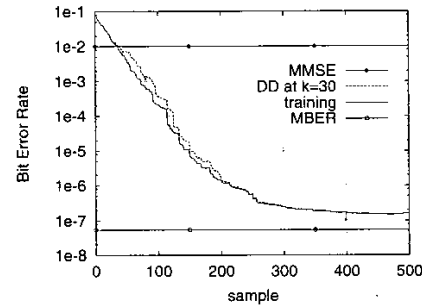


(b) Initial $w_0 = [0.1 + j0.0\ \ 0.1 + j0.0\ \ 0.1 + j0.0]^T$, $\mu = 0.02$ and $\rho_n^2 = 0.08$.

Figure 5: Learning curves of stochastic gradient adaptive MBER algorithm averaged over 30 runs. SNR= 14 dB and SIR$_i$ = 0 dB for $2 \leq i \leq 6$.

[3] L. C. Godara, "Applications of antenna arrays to mobile communications, Part I: Performance improvement, feasibility, and system considerations," *Proc. IEEE*, Vol.85, No.7, pp.1031–1060, 1997.

[4] J.H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, Vol.5, No.1, pp.23–27, 1998.

[5] J.S. Blogh and L. Hanzo, *Third Generation Systems and Intelligent Wireless Networking – Smart Antenna and Adaptive Modulation*. Chichester: John Wiley, 2002.

[6] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, Vol.33, pp.1066–1076, 1962.

[7] B.W. Silverman, *Density Estimation*. London: Chapman Hall, 1996.

[8] A.W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. Oxford: Oxford University Press, 1997.

[9] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley, 1993.

[10] S. Chen, A.K. Samingan, B. Mulgrew and L. Hanzo, "Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels," *IEEE Trans. Signal Processing*, Vol.49, No.6, pp.1240–1247, 2001.