GENETIC ALGORITHM ASSISTED MINIMUM BIT ERROR RATE BEAMFORMING

A. Wolfgang, N.N. Ahmad, S. Chen, L. Hanzo

Dept. of ECS Univ. of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-594 508
{aw03r,nna00r,sqc,lh}@ecs.soton.ac.uk
http://www-mobile.ecs.soton.ac.uk

ABSTRACT

In this paper a novel Genetic Algorithm (GA) assisted Minimum Bit Error Rate Rate (MBER) beamforming technique is introduced. The performance of the proposed GAs is characterised by the Probability Density Function (PDF) and the mean value of the achievable Bit Error Rate (BER) at the beamformer's output. The results are also compared to the theoretical bounds. It is shown that GAs are suitable for MBER beamforming and that a trade-off between the complexity of the GA and its robustness against system parameter variations can be found.

1. INTRODUCTION

In recent years an increasing demand for higher wireless teletraffic system capacity was observed. Since bandwidth is a scarce commodity, this increased demand has to be preferably satisfied by exploiting the existing resources more efficiently. Adaptive antenna arrays [1] can be used for enhancing the system capacity by separating users transmitting on the same carrier frequency in the spatial domain. The receiver controls the radiation pattern of the array by adjusting the array weights so that a certain optimisation criterion is met [1].

The conventional beamformer combines the signals received with the aid of each antenna element for the sake of minimising the Mean Square Error (MSE) between a transmitted and a received reference signal. This technique is known as the Minimum Mean Square Error (MMSE) approach. The Sample Matrix Inversion (SMI) algorithm [1], which is based on invoking the so-called temporal reference technique, is an accurate and well established beamforming algorithm for the realization of this approach [1].

The minimisation of the MSE at the output of the beamformer between a transmitted and a received reference signal does not guarantee the minimisation of the Bit Error Rate (BER). Therefore in [2] the BER rather than the MSE was minimised at the beamformer's output. We refer to this novel approach as the Minimum Bit Error Rate (MBER) approach, which has already been successfully used for both Multi User Detection (MUD) [3] and channel equalisation [4]. The associated performance gain however, is achieved by solving a more elaborate cost function optimisation problem than the minimisation of the quadratic MSE cost function. The function describing the BER as a function of the antenna array weights is highly nonlinear and has numerous local minima. For solving this complex optimisation problem, carefully initialised stochastic gradient based algorithms can be employed [2], although the choice of the appropriate algorithmic parameters may turn out to be challenging. For sub-optimum settings, the algorithm may get trapped in a local minimum of the BER surface and thus result in a sub-optimum solution. With the aim of circumventing these difficulties, it seemed natural to employ a random search based heuristic optimisation algorithm such as a Genetic Algorithm (GA) [5].

GA based array weight calculation has been characterised in [6–8]. GAs have been shown to be resistant against local minima problems and can be randomly initialised, while gradient based search techniques are somewhat deficient in this respect. The genetic approach can be interpreted as a guided random search process, which attempts to imitate biological evolution [5]. The algorithm commences its iterations with a set of potential solutions referred to as the initial population, which can be chosen randomly and for each of these potential initial solutions, which are also referred to as GA individuals, the so-called fitness function is evaluated. This function describes the quality or fitness of a potential solution and ensures that the most fit individuals are selected as 'mates' of the GA creating 'offspring', i.e new parameter estimates, which become part of the next generation. By successively repeating the procedure of combining the best parameter estimates for the sake of creating new estimates, the algorithm is likely to converge to the best solution for the fitness function. The motivation of our research is to derive a GA, which is capable of finding the antenna array weights that directly minimise the BER at the beamformer's output, resulting in the MBER solution.

In Section 2 we will first introduce the system model for our investigations and then present an expression for both the true and the estimated BER of the received signal. In Section 3 a GA configuration will be presented, which is used in Section 4.1 for calculating the MBER solution, where the true BER constitutes the basis of the *fitness function*. The GA will be further developed in Section 4.2 and *an estimate of the true BER replaces the true BER in the fitness function*. Our conclusions are offered in Section 5.

2. SYSTEM MODEL

The system model of the beamforming process is chosen in analogy to other studies reported in the literature [2]. The user and the interferer are considered as Binary Phase Shift Keying (BPSK) modulated point-sources in the far field of the receiver. Additionally, we consider a one-dimensional L-element linear array of omni-directional antennas having an inter-element spacing of $d = \frac{\lambda}{2}$, where λ is the wave-length of the sources. The receiver noise $\mathbf{n}$ is assumed to be a complex white Gaussian process having a variance of $2\sigma_n$. The antenna array output signal $\mathbf{x}$ at a time instant k can then be written as

$$\mathbf{x}(k) = \mathbf{P}\mathbf{b}(k) + \mathbf{n}(k), \quad (1)$$

The financial support of the EU under the auspices of the Phoenix project is gratefully acknowledged. The authors are also grateful to their colleagues for the enlightenment gained within the Phoenix consortium.

where $\mathbf{b}$ is an M-element vector containing the BPSK modulated users' symbols and $\mathbf{P}$ is the $(L \times M)$-dimensional system matrix, which takes into account the M sources' signal amplitude at the receiver and their angles of incidence θ. The system matrix is defined as

$$\mathbf{P} = \begin{pmatrix} A_1 e^{j\omega t_1(\theta_1)} & \cdots & A_M e^{j\omega t_1(\theta_M)} \\ A_1 e^{j\omega t_2(\theta_1)} & & A_M e^{j\omega t_2(\theta_M)} \\ \vdots & \ddots & \vdots \\ A_1 e^{j\omega t_L(\theta_1)} & \cdots & A_M e^{j\omega t_L(\theta_M)} \end{pmatrix}, \qquad (2)$$

where A_m is the amplitude of the signal received from the m^{th} source and $t_l(\theta_m)$ is the relative time delay of the signal transmitted by the m^{th} source at the l^{th} array element. The beamformer output can now be written as

$$y(k) = \mathbf{w}^H \mathbf{x}(k), \qquad (3)$$

where ^{H} denotes the Hermitian operator, representing the conjugate complex value of a vector.

For M number of BPSK modulated point sources, there exist only 2^M possible transmitted bit sequences $\mathbf{b}_q$ with $1 \leq q \leq N_b$, where $b_{q,1}$ corresponds to the desired user's symbols. Therefore, the beamformer's input signal may assume values from the set $\mathcal{X} = \{\overline{\mathbf{x}}_q = \mathbf{P}\mathbf{b}_q, \quad 1 \leq q \leq N_b\}$. Then it can be shown, that the Bit Error Probability (BEP) for a given weight vector $\mathbf{w}$ may be calculated as follows [2]

$$P_e(\mathbf{w}) = \frac{1}{N_b} \sum_{q=1}^{N_b} Q(g_q(\mathbf{w})) \qquad (4)$$

with

$$g_q(\mathbf{w}) = \frac{\operatorname{sgn}(b_{q,1}) Re\{\mathbf{w}^H \overline{\mathbf{x}}_q\}}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}}. \qquad (5)$$

The true BEP P_e of (4) will be used in the *fitness function* of the GA for an initial BER versus GA complexity study in Section 2. However, in order to aim for a more realistic receiver structure, the true BEP has to be replaced by the estimated $\hat{P}_e$, which can be obtained using *kernel density estimation* [9]. The estimated bit error probability $\hat{P}_e$ may be written as [2]

$$\hat{P}_e(\mathbf{w}) = \frac{1}{K} \sum_{k=1}^{K} Q(\hat{g}_k(\mathbf{w})) \qquad (6)$$

with

$$\hat{g}_q(\mathbf{w}) = \frac{\operatorname{sgn}(b_1(k)) y_R(k)}{\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}}, \qquad (7)$$

where K is the reference sequence length, $y_R(k)$ is the real part of the received reference symbol, $b_1(k)$ is the transmitted reference symbol and ρ_n is the so-called kernel width, also known as the smoothening parameter. Note, that for the kernel density estimation of the Probability Density Function (PDF) of y_R a Gaussian kernel function was used, which has been transformed into the Q-function by integrating it.

The task of the GA in the system configuration considered is to calculate the complex array weight vector $\mathbf{w}$ in order to minimise (4) and (6), respectively.

3. GENETIC ALGORITHM CONFIGURATION

In this section the GA configuration employed is presented, which is used for calculating the MBER weights of an antenna array. The general structure of a GA entails the following steps [5]:

1. Create initial population;
2. Evaluate the fitness of the individuals in the initial population;
3. Select individuals of the initial population fitness-proportionately, which are subject to cross-over of parents or mates for the sake of creating the new generation of individuals;
4. Perform cross-over for creating the new population having an increased average fitness;
5. Perform mutation for the sake of avoiding premature convergence, which might trap the GA in local minima.

The operation of the GA commences with an initial set of potential array weights, which are chosen randomly from within a given search-space. Each of the potential array weight solutions is represented as a binary string using Binary Coded Decimal (BCD) format with 8 bits assigned for the real and 8 bits for the imaginary part of each array weight w_l. BCD encoded, randomly chosen weight vectors form the initial population, which may also be referred to as the first generation of the GA associated with the generation index $g = 1$. The total number of generations is G.

The fitness f_i of an individual was quantified as

$$f_i = 1 - \frac{1}{1 - \log(P_e)}, \qquad (8)$$

which approached unity, as P_e decreased. The employment of the logarithm of P_e in the objective function has significantly improved the convergence behaviour of the GA. The fitness function of (8) outperforms other functions that rely on a linear relationship between the fitness of an individual and the BER, which is particularly so at low BERs.

Having evaluated the fitness f_i, it is scaled by a multiplicative factor in order to stimulate a certain level of competition among the individuals, thus preventing the algorithm from premature convergence without wider exploration of the entire search space. Numerous fitness scaling methods, such as for example sigma scaling [5] or linear scaling [5] have been investigated, demonstrating that the recently introduced so-called span scaling [10] technique was the most effective. Span scaling modifies the fitness f_i^g of an individual of generation g according to [10]

$$f_{i,new} = \frac{f_i^g - \overline{f}_g}{\sigma_g} + W_g \qquad (9)$$

with

$$W_g = 1 - \frac{1}{1 + \exp\left(-(\frac{10g}{G} - 4)\right)}, \qquad (10)$$

where $\overline{f}_g$ is the mean fitness and σ_g is the standard deviation of the fitness of generation g from the mean $2\sigma_g$. The function W_g in (10) is a modified version of the one used in [10]. In contrast to other scaling methods, adaptive span scaling has the advantage of forcing the GA to converge to a solution after a fixed number of generations by increasing the 'selection pressure', as the generation index increases. In addition to span scaling, elitism [5] was also used in the GA advocated, which guarantees that the most fit individual of a generation is always retained in the forthcoming generation.

The GA features presented in this section have shown in our initial studies, that they improve the GA's performance, although carefully optimised span scaling and elitism have to be used, since they may guide the search towards local minima, potentially preventing the algorithm from converging to the global minimum. In order to ensure that the performance of the GA is not limited by

Bits per array weight	16
Selection type	Roulette Selection
Cross-over type	Single Point
Cross-over Probability	0.9
Scaling	Span Scaling
Mutation type	Bit Inversion
Mutation Probability	0.1
Elitism	performed

Table 1: GA parameters

the GA operators, it is important to carefully monitor the PDF of the BER at the beamformer's output evaluated for numerous differently initialised GA runs.

4. SIMULATIONS

For the sake of arriving at the analytical BER expression of (4), a conjugate gradient algorithm based array weight adaption procedure has been proposed in [2]. This theoretical performance bound will serve as our benchmarker when the performance of the GA is evaluated for the different angular distribution of the users given in Table 2. For a given weight vector calculated with the aid of the GA, the BER was evaluated using (4), so that resorting to Monte Carlo simulation could be avoided. The BER curves presented are averaged over 1000 GA-aided weight-optimisation runs. When we refer to the complexity of a GA, we characterise it by the number of objective function evaluations used. The complexity is then given as the product of the population size P and the number of generations G. If two GAs have an equal complexity, typically the one having a higher population size is preferable, since for a certain generation, the P number of evaluations of the objective function can be processed in parallel.

4.1. GA-Aided Best-Case Performance

For the sake of characterising the achievable upper-bound performance in our initial study, the true bit error probability is assumed to be known at the receiver, so that it can be used for evaluating the fitness of the GA's individuals. The exact configuration of the GA is summarised in Table 1.

The search-space of the GA was limited to the region of $|Re\{w_l\}| \leq 1$ and $|Im\{w_l\}| \leq 1$ with $1 \leq l \leq L$, since the MBER solution found for BPSK modulated users is invariant to linear scaling and thus can be confined to this search-space. The initial population was chosen randomly in conjunction with a uniform probability within the given search-space. The signal of all point sources arranged according the constellations given in Table 2, was assumed to be received with equal power at the antenna array.

Figure 1 (left graph) shows the achievable BER at the beamformer's output for a two-element linear array with the users arranged according to scenario $\mathcal{S}$ of Table 2. The population size has been set to $P = 20$ individuals and the number of generations was chosen to be $G = 20$. It can been seen, that the theoretical MBER solution and the solution calculated with the aid of the GA are identical, even though the GA uses quantised array weights. The convergence behaviour of the GA is characterised with the aid of the PDF of the BER, which is plotted for $g = 1, 5, 10, 20$. It is demonstrated, that the GA converges towards the theoretical best-case performance, as the generation index g increases. After twenty generations the GA reaches the theoretical best-case BER performance with a high probability and hence the PDF of the BER

	User	Interferer				
$\mathcal{S}$	$\theta_0 = 15^\circ$	$\theta_{1,2,3,4} =$	$-30^\circ,$	$60^\circ,$	$-70^\circ,$	80°
$\mathcal{U}$	$\theta_0 = 15^\circ$	$\theta_{1,2,3,4} =$	$4^\circ,$	$26^\circ,$	$-70^\circ,$	80°
$\mathcal{V}$	$\theta_0 = 15^\circ$	$\theta_{1,2,3,4} =$	$9^\circ,$	$21^\circ,$	$-70^\circ,$	80°

Table 2: Angles of incidence relative to the perpendicular of the antenna array for the user constellations $\mathcal{S}, \mathcal{U}$ and $\mathcal{V}$.

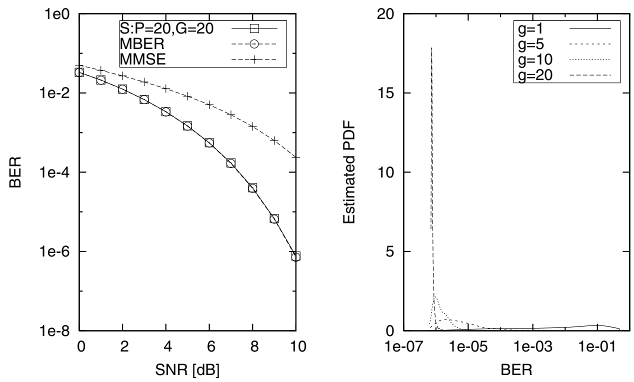

Figure 1: *Left Graph*: BER versus SNR for a two-element linear array receiving five equal-power sources arranged according to constellation $\mathcal{S}$ of Table 2. The GA configuration was as given in Table 1 with a population size of $P = 20$ and $G = 20$ generations. The BER associated with specific values of the fitness function of (8) was assumed to be known at the receiver and was calculated using (4). The GA-based BER results were averaged over 1000 GA runs, when communicating over an AWGN channel. *Right Graph:* PDF of the BER for arrangement $\mathcal{S}$, corresponding to the GA's individual having the highest fitness after $g = 1, 5, 10, 20$ generations. The BER-PDF was estimated using kernel density estimation for SNR=10 dB and 1000 samples over AWGN channels.

achieved in the differently initialised GA-assisted runs becomes a narrow peak, as seen at the right of Figure 1.

If the number of array elements is increased, the complexity of the GA also has to be increased, when aiming for a narrow BER-PDF, since the search-space expands significantly. For the optimisation of the elements in a four-element array characterised in Figure 2 (left graph) the population size was increased to $P = 40$ and the number of generations was set to $G = 40$. In the case, when the point sources have been arranged according to scenario $\mathcal{S}$ of Table 2, a good agreement is observed between the theoretical best-case BER performance and the results obtained with the aid of the GA. However, if the two interferers arriving at angles of $\theta = -30^\circ$ and $\theta = 60^\circ$ are moved towards the desired user, as in constellation $\mathcal{U}$ of Table 2, there is a performance gap between both solutions, as seen in Figure 2.

Furthermore, a lower probability of convergence is observed in the right hand side graph of Figure 2, where the PDF of the BER corresponding to the GA's individual having the maximum fitness in the generation is plotted after $g = 1, 10, 20, 40$ generations for SNR=10 dB and for the user constellation $\mathcal{S}$ of Table 2. The GA is still capable of converging to the theoretical best-case performance, albeit with a lower probability than for the two-element array characterised in Figure 1. However, it is important to note that even if the GA does not converge to the ideal solution, it remains capable of reaching a reasonable BER.

For the eight-element array a similar behaviour to that recorded for the four-element scheme can be observed in Figure 3. In the scenario, when the users are arranged according to constellation $\mathcal{S}$ of Table 2, a good agreement was found between the theoretical

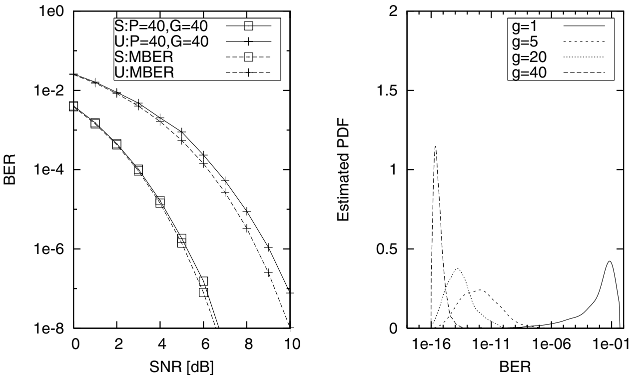

Figure 2: *Left Graph*: BER versus SNR for a four-element linear array receiving five equal-power sources arranged according to constellations $\mathcal{S}$ and $\mathcal{U}$ of Table 2. The GA configuration was as given in Table 1 with a population size of $P = 40$ and $G = 40$ number of generations. The BER associated with specific values of the fitness function of (8) was assumed to be known at the receiver and was calculated using (4). The GA-based BER results were averaged over 1000 GA runs, when communicating over an AWGN channel. *Right Graph:* PDF of the BER for arrangement $\mathcal{S}$, corresponding to the GA's individual having the highest fitness after $g = 1, 5, 20, 40$ generations. The BER-PDF was estimated using kernel density estimation for SNR=10 dB and 1000 samples over AWGN channels.

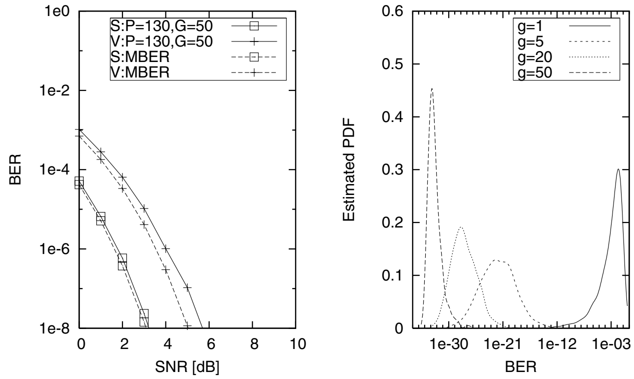

Figure 3: *Left Graph*: BER versus SNR for an eight-element linear array receiving five equal-power sources arranged according to constellations $\mathcal{S}$ and $\mathcal{V}$ of Table 2. The GA configuration was as given in Table 1 with a population size of $P = 130$, $G = 50$ number of generations and a reduced mutation probability $p_m = 0.01$. The BER associated with specific values of the fitness function of (8) was assumed to be known at the receiver and calculated using (4). The GA-based results were averaged over 1000 GA runs, when communicating over an AWGN channel. *Right Graph:* PDF of the BER for arrangement $\mathcal{S}$, corresponding to the GA's individual with highest fitness after $g = 1, 5, 20, 50$ generations. The BER-PDF was estimated using kernel density estimation for SNR=10 dB and 1000 samples over AWGN channels.

best-case performance and the GA-based solution. If the two interferers at angles of $\theta = -30^\circ$ and $\theta = 60^\circ$ are moved closer to the desired source, as seen in scenario $\mathcal{V}$ of Table 2, there is again a performance gap between the two solutions. Informal investigations suggest that the performance gap is due to the coarsely quantised weights of the GA. To elaborate a little further, assume that the array weights of user constellation S have to be calculated, when the BER versus array weights surface is characterised by a wide valley, which represents the MBER solution, although this BER valley not explicitly shown here owing to lack of space. In the case of this wide valley even coarsely quantised weights are very likely to identify this minimum sufficiently accurately. However, if the BER surface is rendered more complex by moving the interferer closer to the desired source, the valley in the BER surface becomes narrower. This implies even for finely quantised weights, which deviate only little from the optimum solution that an increased BER may be recorded.

As a result of the two-, four- and eight-element array based study, we surmise the rule of thumb, that *the complexity of the GA increases quadratically with the array size*, when aiming for a similar BER-PDF spread.

4.2. GA-Aided Performance Using Kernel Estimates

The motivation for the employment of a GA in the context of MBER beamforming was to circumvent the algorithmic difficulties of stochastic gradient based MBER algorithms, such as for example the Least Bit Error Rate (LBER) algorithm [2, 3], which are sensitive to BER gradient estimation errors.

4.2.1. Static GA Configuration

In Figure 4 the BER versus SNR performance is illustrated for the users arranged according to constellation $\mathcal{S}$ of Table 2 and for INR=SNR, where INR is the Interference to Noise to Noise Ratio. The GA parameters are given in Table 1. It can be seen that for a reference length of $K = 512$ (dashed lines) and for $\rho_n = 0.1$ as well as $\rho_n = 1$, respectively, the GA associated with $P = 20$ and $G = 20$ occasionally fails to converge towards the MBER solution, except perhaps for $\rho_n = 0.01$. Further investigations have shown that if the complexity of the GA is increased by opting for $P = 40$ and $G = 40$ (solid lines), the algorithm is capable of compensating for the BER estimation errors, even if a shorter reference sequence length of $K = 256$ is used. As expected, for an inappropriately chosen value of ρ_n the minimum of the BER surface is shifted away from the theoretical solution [3]. In this case the GA becomes incapable of converging to the MBER solution. From the results presented in Figure 4 it is inferred, that a trade-off between the complexity of the GA and its sensitivity against BER estimation errors can be found.

4.2.2. Adaptive GA Configuration

In the previous section, a fixed kernel width of ρ_n was used for BER estimation. It is known however that the perfect kernel width has to be adjusted according to the variance of the received training sequence and is thus dependent on both the SNR and the INR. Silverman has provided a simple rule of thumb for the estimation of ρ_n given as [9]

$$\rho_n = \left(\frac{\hat{\sigma}^5}{3K}\right)^{\frac{1}{5}} \approx 1.06\hat{\sigma}K^{-\frac{1}{5}}, \tag{11}$$

where K is the reference sequence length and $\hat{\sigma}$ is the standard deviation of the received reference sequence. Equation (11) tends to over-smoothen the BER estimate, but the results presented in Figure 4 suggest that the GA is capable of compensating for this inaccuracy.

For the sake of investigating a GA design, which invokes an adaptive kernel width, the GA given in Table 1 was enhanced by incorporating two additional features [5]:

1. Incest Prevention: Prevents that two identical individuals are earmarked for the cross-over operation.

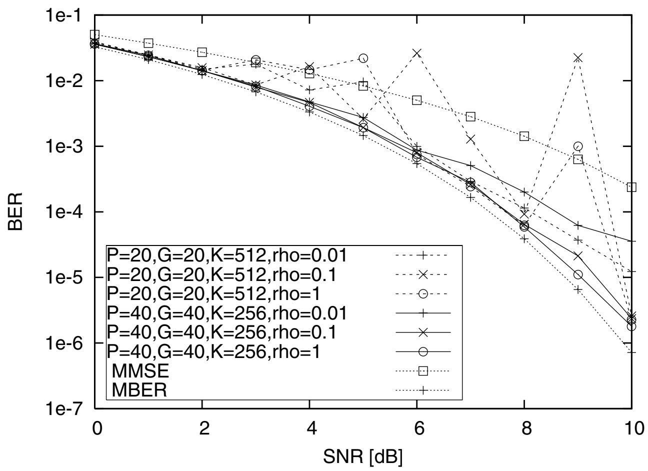

Figure 4: BER versus SNR for a two-element linear array receiving five equal-power sources arranged according to constellation $\mathcal{S}$ of Table 2. The GA configuration was as given in Table 1 with a population size P and G number of generations as given in the legend. The BER associated with specific values of the fitness function of (8) was estimated based on a 256 and a 512 symbol training sequence using (6). The GA results were averaged over 1000 GA runs, when communicating over an AWGN channel.

2. Weighted Mutation: Reduces the mutation probability p_m of the more significant bits of each individual, as the generation index g approaches the maximum number of generations G.

Figure 5 (left graph) shows the achievable BER versus SNR performance in conjunction with the enhanced GA configuration for $P = 50$ and $G = 30$, when detecting equal-power users (SNR=INR$_{1,2,3,4}$ labelled as *equ*) as well as for a scenario studying the near-far effect using SNR=INR$_{3,4}$ and SIR$_{1,2}$ =-6 dB (labelled as *diff*). The users were arranged according to constellation $\mathcal{S}$ of Table 2. The graphs illustrate the superiority of the MBER approach over the MMSE approach in terms of counteracting the near-far effects, while demonstrating that the GA adopts automatically to the new BER surface without being reconfigured. Reducing the training sequence length from $K = 256$ to $K = 64$ does not significantly affect the mean BER, but when considering the BER-PDF after $g = G = 30$ generations, it can be seen that the algorithm's convergence behaviour is noticeably improved. For a longer training sequence the GA is more likely to converge to the MBER solution.

5. CONCLUSIONS AND CURRENT RESEARCH

A novel GA assisted MBER beamformer has been developed. It has been shown that a GA can be used for approximating the MBER solution, which directly minimises the BER at the beamformer's output. A rule of thumb, which relates the complexity of a GA to the array size was provided. The proposed GA may be used for approximating the theoretical best-case MBER solution. A further improved GA associated with an adaptively chosen kernel width was used in Section 4.2.2 for investigating, how the algorithm reacts to time-variant interference conditions, which was capable of accommodating, when the BER surface changed shape owing to changing the sources' power. In this respect the GA outperforms the gradient based algorithms, such as the LBER [3] algorithm, which has to be reconfigured for each BER surface variation.

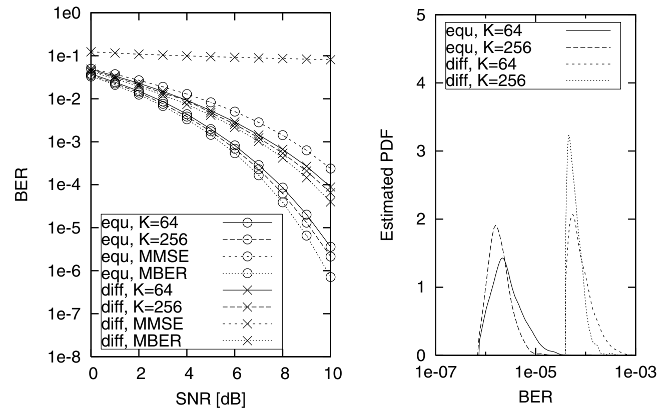

Figure 5: *Left Graph*: BER versus SNR for a two-element linear array receiving five sources arranged according to constellation $\mathcal{S}$ of Table 2. The simulations were carried out for equal-power users with SNR=INR$_{1,2,3,4}$ (labelled as *equ*) and for a constellation where the two interferer of incident angles $\theta = -30°$ and $\theta = 60°$ had 6 dB higher power than the desired source (SNR=INR$_{3,4}$ and SIR$_{1,2}$ =-6 dB labelled as *diff*). The GA configuration was as given in Table 1 with a populations size of $P = 50$, $G = 30$ number of generations and the additional features incest prevention and weighted mutation. The BER associated with specific values of the fitness function of (8) was estimated using (6) and (11). The GA results were averaged over 1000 GA runs, when communicating over an AWGN channel. *Right Graph:* PDF of the BER corresponding to the GA's individual with highest fitness after $g = G = 30$ number of generations corresponding to the SNR-BER curves in the left graph. The BER-PDF was estimated using kernel density estimation for a SNR=10 dB and 1000 samples over an AWGN channel.

6. REFERENCES

[1] J.S. Blogh, L. Hanzo: Third-Generation Systems and Intelligent Wireless Networking - Smart Antennas and Adaptive Modulation, John Wiley, 2002.

[2] S. Chen, N.N. Ahmad and L. Hanzo, "Adaptive Minimum Bit Error Rate Beamforming," *to appear in IEEE Transactions on Wireless Communications*, April 2002.

[3] S. Chen, A.K. Samingan, B. Mulgrew, L. Hanzo, "Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels", *IEEE Transactions on Signal Processing*", no. 6, vol. 49, pp.1240-1247, 2001.

[4] C-C. Yeh, J.R Barry, "Adaptive minimum bit-error rate equalization for binary signaling", *IEEE Transactions on Communications*", no. 7, vol. 48, pp.1226-1235, July 2000.

[5] D. E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning, " Addison-Wesley, 1989.

[6] A. Tennant, M.M. Dawoud, and A.P. Anderson, "Array Pattern Nulling by Element Position Perturbations using GA " *Electronics Letters*, vol.30, pp.174-176, 1994.

[7] R.J. Mitchell, B. Chambers, and A.P. Anderson, "Array Pattern Synthesys in the Complex Plane Optimized by a GA, " *Electronics Letters*, vol.32, pp. 1843-1845, 1996.

[8] R.L. Haupt, "Phase-only adaptive nulling with a GA, " *IEEE Transactions on Antennas and Propagation*, vol. 45, pp.1009-1015, 1997.

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

[10] F. Leclerc, J.-Y. Potvin, "A fitness scaling method based on a span measure," *IEEE International Conference on Evolutionary Computation*, vol. 2, pp.561-565, 1995.