Semi-blind Adaptive Beamforming for High-throughput Quadrature Amplitude Modulation Systems

Sheng Chen   Wang Yao   Lajos Hanzo

School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK

Abstract: A semi-blind adaptive beamforming scheme is proposed for wireless systems that employ high-throughput quadrature amplitude modulation signalling. A minimum number of training symbols, equal to the number of receiver antenna array′s elements, are first utilised to provide a rough initial least squares estimate of the beamformer′s weight vector. A concurrent constant modulus algorithm and soft decision-directed scheme is then applied to adapt the beamformer. This semi-blind adaptive beamforming scheme is capable of converging fast to the minimum mean-square-error beamforming solution, as demonstrated in our simulation study.

Keywords: Beamforming, minimum mean square error, quadrature amplitude modulation, semi-blind adaptive algorithm, constant modulus algorithm, soft decision-directed scheme.

1 Introduction

The ever-increasing demand for wireless communication capacity has motivated the development of antenna array assisted spatial processing techniques in order to further improve the achievable spectral efficiency. A particular technique that has shown real promise in achieving substantial capacity enhancement is the use of adaptive beamforming with antenna arrays[1−3]. Adaptive beamforming is capable of separating signals transmitted on the same carrier frequency, and thus provides a practical means of supporting multiusers in a space division multiple access scenario. For the sake of further improving the achievable bandwidth efficiency, high-throughput quadrature amplitude modulation (QAM) schemes[4] have become popular in numerous wireless network standards. For example, the 16-QAM and 64-QAM schemes were adopted in the WiMax standard[5]. The classical beamforming design is the minimum mean square error (MMSE) solution, which can be realised using the training-based adaptive algorithms[2, 6−9]. Pure training-based schemes, however, require a high proportion of training symbols, which considerably reduces the achievable system throughput. Pure blind beamforming[10−12] does not reduce the achievable system throughput at the expense of high computational complexity and slow convergence. Moreover, blind beamforming results in unavoidable estimation and decision ambiguities[13, 14].

An effective means of resolving the estimation and decision ambiguities inherent in blind schemes is to employ a few training symbols. Our presented work combines this very short training with blind adaptive beamforming which leads to the novel and attractive semi-blind adaptive beamforming algorithm. In particular, we consider an adaptive beamforming assisted receiver for wireless systems that employ high-throughput QAM signalling[4]. Because we will consider the low-complexity MMSE solution as the optimal design for beamforming, the number of users supported by the system is assumed to be no more than the number of receiver antenna array′s elements. The proposed adaptive beamforming method is semi-blind as we employ a minimum number of pilots, which is equal to the number of array

elements, to provide a rough least squares (LS) estimate for the beamformer′s weight vector. In general, this initialisation is not sufficiently accurate to achieve an "opening-eye", and therefore, it is unsafe to carry out decision-directed (DD) adaptation for the beamformer. We then apply a constant modulus algorithm (CMA) assisted soft DD (SDD) blind adaptive algorithm to adapt the beamformer. The concurrent CMA and SDD algorithm was originally derived for blind equalisation of single-input single-output QAM systems[15], and it was extended to single-input multiple-output systems in [16]. This blind adaptive scheme has a very low computational complexity. In the present semi-blind beamforming application, owing to the initial information provided by the training pilots, the algorithm converges much faster than the pure blind adaptation case, and it is capable of approaching the performance of the MMSE beamforming solution based on the perfect channel knowledge, as will be shown in our simulation study.

To the best of our knowledge, this is the first time that a high-performance semi-blind adaptive beamforming scheme with low complexity is proposed for high-order QAM systems. In the related multiple-input multiple-output (MIMO) technologies, many semi-blind schemes have been proposed[17−22]. In these schemes, a few training symbols are first used to provide an initial MIMO channel estimate, and the channel estimator and the maximum likelihood (ML) data detector iteratively exchange their information, where the MIMO channel estimator relies on a DD adaptation. These semi-blind MIMO schemes, however, becomes computationally prohibitive for high-order QAM signalling owing to the high complexity of ML data detection. Moreover, to estimate the MIMO channel matrix, training pilots must include all the users′ data. In the beamforming system, the receiver only has access to the desired user′s data during the training and it does not have access to the interfering users′ data at all. Thus, these semi-blind MIMO schemes cannot be applied to the current beamforming system.

Recently, a spatial-equaliser based semi-blind MIMO scheme[23] is proposed, which can potentially be applied to high-order QAM systems. In this scheme, the spatial equaliser is adapted by minimising the combined cost func-

tion of the training-based sum of the squared errors and a higher-order statistic (HOS) aided criterion using a block-data based gradient algorithm. In [23], however, the authors make an unnecessary assumption of the known MIMO channel matrix[1]. By changing the training-based criterion to include only the desired user's training symbols, this semi-blind scheme can be applied to the beamforming system. In terms of computational requirements, the complexity of the block-data based algorithm in [23] is significantly higher than that of our proposed stochastic gradient algorithm. In terms of the achievable performance, our simpler stochastic gradient scheme actually outperforms the more complex block-data based gradient scheme of [23]. This is because the blind adaptive process in the semi-blind scheme of [23] is based on the HOS (e.g., CMA) criterion, while our blind adaptive process is based on the HOS (CMA) aided SDD criterion. The latter can approach the optimal MMSE solution more accurately and achieve a faster convergence, as a benefit of the fact that SDD adaptation is more like the true training.

Throughout this contribution, we adopt the following notational conventions. Boldface capitals and lower-case letters stand for matrices and vectors, respectively, while $\boldsymbol{I}_K$ denotes the $K \times K$ identity matrix. Furthermore, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ are the transpose and Hermitian operators, respectively, while $\|\cdot\|$ and $|\cdot|$ denote the norm and magnitude operators, respectively. $\mathrm{E}\left[\cdot\right]$ is the expectation operator, while $(\cdot)^*$ denotes the complex conjugate. Finally, $j = \sqrt{-1}$.

2　Beamforming model

We consider a coherent communication system that supports n_T users, where each user transmits an M-QAM signal on the same angular carrier frequency of ω. In order to achieve user separation in the angular domain[24, 25], the receiver is equipped with a linear antenna array consisting of n_R uniformly spaced elements. Further assume that the communication is over flat fading channels. Then, the system is described by the following model

$$\boldsymbol{x}(k) = \boldsymbol{H}\,\boldsymbol{s}(k) + \boldsymbol{n}(k) \qquad (1)$$

where k denotes the symbol index, $\boldsymbol{x}(k) = [x_1(k)\ \ x_2(k)\ \ \cdots\ \ x_{n_R}(k)]^{\mathrm{T}}$ is the received signal vector, $\boldsymbol{n}(k) = [n_1(k)\ \ n_2(k)\ \ \cdots\ \ n_{n_R}(k)]^{\mathrm{T}}$ is the complex-valued Gaussian white noise vector associated with the system having $\mathrm{E}\left[\boldsymbol{n}(k)\boldsymbol{n}^{\mathrm{H}}(k)\right] = 2\sigma_n^2\boldsymbol{I}_{n_R}$, $\boldsymbol{s}(k) = [s_1(k)\ \ s_2(k)\ \ \cdots\ \ s_{n_T}(k)]^{\mathrm{T}}$ is the transmitted symbols vector of the n_T users with the symbol energy given by $\mathrm{E}\left[|s_m(k)|^2\right] = \sigma_s^2$ for $1 \leqslant m \leqslant n_T$, and $\boldsymbol{H}$ denotes the $n_R \times n_T$ system channel matrix.

More specifically, the system channel matrix $\boldsymbol{H} = [h_{l,m}]$, where $1 \leqslant l \leqslant n_R$ and $1 \leqslant m \leqslant n_T$, is defined by

$$\boldsymbol{H} = [\boldsymbol{h}_1\ \boldsymbol{h}_2\ \cdots\ \boldsymbol{h}_{n_T}] = [A_1\boldsymbol{\eta}_1\ A_2\boldsymbol{\eta}_2\cdots A_{n_T}\boldsymbol{\eta}_{n_T}] \qquad (2)$$

where A_m denotes the non-dispersive channel coefficient for user m, and the steering vector for user m is given by

$$\boldsymbol{\eta}_m = \left[\mathrm{e}^{j\omega t_1(\theta_m)}\ \mathrm{e}^{j\omega t_2(\theta_m)}\ \cdots\ \mathrm{e}^{j\omega t_{n_R}(\theta_m)}\right]^{\mathrm{T}} \qquad (3)$$

---

[1] If the MIMO channel matrix were known, the MMSE spatial equaliser could be designed directly and there would be no need for any semi-blind adaptation.

with θ_m being the angle of arrival for user m and $t_l(\theta_m)$ being the relative time delay at array element l for user m. θ_m is uniformly distributed in $[0,\ 2\pi)$ and the magnitude of A_m is a Rayleigh process. However, the fading is assumed to be sufficiently slow, so that during the time period of a transmission block or frame, all the related entries $h_{l,m}$ in the system channel matrix $\boldsymbol{H}$ is deemed unchanged. From frame to frame, $h_{l,m}$ are assumed to be uncorrelated complex-valued Gaussian processes with zero mean and $\mathrm{E}\left[|h_{l,m}|^2\right] = 1$.

The modulation scheme is the M-QAM and, therefore, the transmitted data symbols $s_m(k)$, $1 \leqslant m \leqslant n_T$, take the values from the M-QAM symbol set defined by

$$\mathcal{S} \triangleq \{s_{i,q} = u_i + ju_q,\ 1 \leqslant i, q \leqslant \sqrt{M}\} \qquad (4)$$

with the real-part symbol $\mathrm{Re}[s_{i,q}] = u_i = 2i - \sqrt{M} - 1$ and the imaginary-part symbol $\mathrm{Im}[s_{i,q}] = u_q = 2q - \sqrt{M} - 1$. The average signal-to-noise ratio (SNR) of the system is defined as $\mathrm{SNR} = n_T \times \sigma_s^2/(2\sigma_n^2)$. Without loss of generality, user one is assumed to be the desired user and the rest of the users are interfering ones. A beamformer

$$y(k) = \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}(k) \qquad (5)$$

is used to detect the transmitted symbols $s_1(k)$ of the desired user, where $\boldsymbol{w}$ denotes the $n_R \times 1$ complex-valued weight vector of the beamformer. With the perfect channel knowledge, the optimal MMSE solution that minimises the mean square error criterion $\mathrm{E}[|s_1(k) - y(k)|^2]$ is given by

$$\boldsymbol{w}_{\mathrm{MMSE}} = \left(\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \frac{2\sigma_n^2}{\sigma_s^2}\boldsymbol{I}_{n_R}\right)^{-1}\boldsymbol{h}_1. \qquad (6)$$

3　The proposed semi-blind algorithm

Let the number of available training symbols be K, and denote the available training data as $\boldsymbol{X}_K = [\boldsymbol{x}(1)\ \boldsymbol{x}(2)\ \cdots\ \boldsymbol{x}(K)]$ and $\boldsymbol{s}_K = [s_1(1)\ s_1(2)\ \cdots\ s_1(K)]^{\mathrm{T}}$. The LS estimate of the beamformer's weight vector based on the training data $\{\boldsymbol{X}_K, \boldsymbol{s}_K\}$ is readily given as

$$\boldsymbol{w}(0) = \left(\boldsymbol{X}_K\boldsymbol{X}_K^{\mathrm{H}}\right)^{-1}\boldsymbol{X}_K\boldsymbol{s}_K^*. \qquad (7)$$

In order to maintain throughput, the number of training pilots should be as small as possible. To ensure that $\boldsymbol{X}_K\boldsymbol{X}_K^{\mathrm{H}}$ has a full rank, we will choose $K = n_R$ as the minimum number of training symbols. Because the training data are insufficient, the initial LS weight vector (7) may not be sufficiently accurate to open the eye. Therefore, DD adaptation is generally unsafe. However, we can apply the concurrent CMA and SDD blind scheme[15, 16] to adapt the beamformer (5) with $\boldsymbol{w}(0)$ of (7) as the initial weight vector. Let the beamformer's weight vector be split into two parts, yielding $\boldsymbol{w} = \boldsymbol{w}_c + \boldsymbol{w}_d$, and denote the beamformer's output at sample k as $y(k) = \boldsymbol{w}^{\mathrm{H}}(k)\boldsymbol{x}(k)$. The initial $\boldsymbol{w}_c$ and $\boldsymbol{w}_d$ are simply set to $\boldsymbol{w}_c(0) = \boldsymbol{w}_d(0) = 0.5\boldsymbol{w}(0)$, with $\boldsymbol{w}(0)$ given in (7).

Specifically, the weight vector $\boldsymbol{w}_c$ is updated using the classical CMA[26, 27]

$$\begin{cases} \varepsilon(k) = y(k)\left(\Delta - |y(k)|^2\right) \\ \boldsymbol{w}_c(k+1) = \boldsymbol{w}_c(k) + \mu_{\mathrm{CMA}}\varepsilon^*(k)\boldsymbol{x}(k) \end{cases} \qquad (8)$$

where $\Delta = \mathrm{E}\left[|s_1(k)|^4\right]/\mathrm{E}\left[|s_1(k)|^2\right]$ and μ_{CMA} is the step size of the CMA. The weight vector $\boldsymbol{w}_d$ by contrast is updated using the SDD scheme[15, 16], which has its root in the blind equalisation scheme of [28]. The complex phasor plane is divided into the $M/4$ rectangular regions, and each region $\mathcal{S}_{i,l}$ contains four symbol points as defined in the following

$$\mathcal{S}_{i,l} = \{s_{p,q},\ p = 2i-1, 2i,\ q = 2l-1,\ 2l\} \qquad (9)$$

where $1 \leqslant i$ and $l \leqslant \sqrt{M}/2$. An illustration of this local decision region is illustrated in Fig. 1. If the beamformer′s output $y(k) \in \mathcal{S}_{i,l}$, a local approximation of the marginal probability density function (PDF) of $y(k)$ is given by [15, 16]

$$\hat{p}(\boldsymbol{w}, y(k)) \approx \sum_{p=2i-1}^{2i} \sum_{q=2l-1}^{2l} \frac{1}{8\pi\rho} \mathrm{e}^{-\frac{|y(k)-s_{p,q}|^2}{2\rho}} \qquad (10)$$

where ρ defines the cluster width associated with the four clusters of each region $\mathcal{S}_{i,l}$. The SDD algorithm is designed to maximise the log of the local marginal PDF criterion $\mathrm{E}[J_{\mathrm{LMAP}}(\boldsymbol{w}, y(k))]$, where $J_{\mathrm{LMAP}}(\boldsymbol{w}, y(k)) = \rho \ln(\hat{p}(\boldsymbol{w}, y(k)))$, via a stochastic gradient optimisation. Specifically, $\boldsymbol{w}_d$ is updated according to [15, 16]

$$\boldsymbol{w}_d(k+1) = \boldsymbol{w}_d(k) + \mu_{\mathrm{SDD}} \frac{\partial J_{\mathrm{LMAP}}(\boldsymbol{w}(k), y(k))}{\partial \boldsymbol{w}_d} \qquad (11)$$

where μ_{SDD} is the step size of the SDD, and

$$\frac{\partial J_{\mathrm{LMAP}}(\boldsymbol{w}, y(k))}{\partial \boldsymbol{w}_d} = \frac{1}{Z_N} \sum_{p=2i-1}^{2i} \sum_{q=2l-1}^{2l} \mathrm{e}^{-\frac{|y(k)-s_{p,q}|^2}{2\rho}} \times$$
$$(s_{p,q} - y(k))^* \boldsymbol{x}(k) \qquad (12)$$

with the normalisation factor

$$Z_N = \sum_{p=2i-1}^{2i} \sum_{q=2l-1}^{2l} \mathrm{e}^{-\frac{|y(k)-s_{p,q}|^2}{2\rho}}. \qquad (13)$$
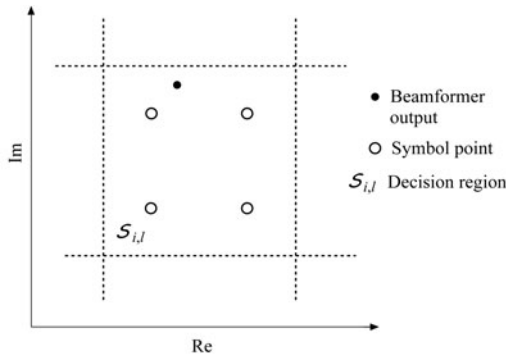

Fig. 1   Illustration of local decision regions for the soft decision-directed adaptation procedure for QAM constellation

The choice of the cluster width ρ, defined in the context of the local PDF (10), should ensure a proper separation of the four clusters of $\mathcal{S}_{i,l}$. Because the minimum distance between the two neighbouring constellation points is 2, ρ is typically chosen to be less than 1. If the value of ρ is too large, a desired degree of separation may not be achieved. On the other hand, if the value of ρ is too small, the algorithm attempts to impose an overly tight control on the

size of clusters and hence may fail to achieve its goal. Apart from these two extreme situations, the performance of the algorithm is not overly sensitive to the value of ρ employed and an appropriate ρ can easily be chosen from a large range of values. More specifically, when the objective of removing interference is accomplished, $y(k) \approx s_1(k) + e(k)$, where $e(k)$ is a Gaussian distributed noise with zero mean. Therefore, the value of ρ is related to the variance of $e(k)$, which is $2\sigma_n^2 \boldsymbol{w}^{\mathrm{H}} \boldsymbol{w}$. Thus, for high SNR situations, a small ρ is desired, while for low SNR cases, large ρ is preferred. Because of the information provided by the training pilots in the form of the initial weight vector (7), a smaller ρ can be used, compared with the case of pure blind adaptation in [15, 16], which leads to better steady-state performance. Soft decision nature can be explicitly seen in (12). Rather than committing to a single hard decision $\mathcal{Q}[y(k)]$ as the hard DD scheme would, where $\mathcal{Q}[\cdot]$ denotes the quantisation operator, alternative decisions are also considered in the local region $\mathcal{S}_{i,l}$ that includes $\mathcal{Q}[y(k)]$. Each tentative decision is weighted by an exponential term $\mathrm{e}^{\{\cdot\}}$, which is a function of the distance between the equaliser′s soft output $y(k)$ and the tentative decision $s_{p,q}$. This soft decision nature substantially reduces the risk of error propagation and achieves faster convergence, compared with the hard DD scheme[15, 16].

It is interesting to point out that a generic partition of the beamformer′s weight vector is

$$\boldsymbol{w} = \alpha \boldsymbol{w}_c + (1-\alpha)\boldsymbol{w}_d \qquad (14)$$

where $0 \leqslant \alpha \leqslant 1$. It is clear that $\alpha = 1$ corresponds to a pure CMA blind spatial equaliser while $\alpha = 0$ is related to a pure SDD blind spatial equaliser. By choosing $\alpha = 0.5$, we arrive at the concurrent CMA and SDD blind beamformer derived in this contribution. Depending on the channel condition, appropriate value of α may be chosen to yield a potentially better beamforming performance. However, this appropriate weighting value is difficult to find. In the absence of any a priori information, the weight vector partition of $\boldsymbol{w}_c = \boldsymbol{w}_d = 0.5\boldsymbol{w}$ can be regarded as an optimal choice.

4   Simulation study

A simulation study was carried out to investigate the proposed semi-blind adaptive beamforming scheme based on the concurrent CMA and SDD algorithm. The achievable performance was assessed in the simulation using the symbol error rate (SER). The analytical SER $P_E(\boldsymbol{w})$ for the beamformer (5) with the weight vector $\boldsymbol{w}$ is given in Appendix.

Stationary system. A linear antenna array with $n_R = 4$ elements and a half-wavelength element spacing was employed to support $n_T = 4$ 16-QAM users. The angles of arrival for the four users were $10°$, $40°$, $-15°$, and $-45°$, respectively. The simulated stationary channels were $A_m = 1 + j \times 0$, $1 \leqslant m \leqslant 4$. The number of pilot symbols for the semi-blind scheme was $K = 4$. Fig. 2 compares the desired-user′s SERs of the proposed semi-blind beamforming scheme, the training-based beamforming given different numbers of training symbols, and the optimal MMSE beamforming computed with the perfect channel knowledge. In the training-based adaptive beamforming, given K training symbols, the LS estimate of the beamformer′s weight vector

was obtained according to (7), and the resulting SER was calculated. The performance of the training-based beamforming with $K = 64$ training symbols, not shown in Fig. 2, was similar to that of the semi-blind combined CMA and SDD beamforming. It can be seen from Fig. 2 that the semi-blind scheme with four training symbols can closely match the performance of the true MMSE beamforming, while the pure training-based scheme requires at least 64 training symbols to achieve a similar performance.
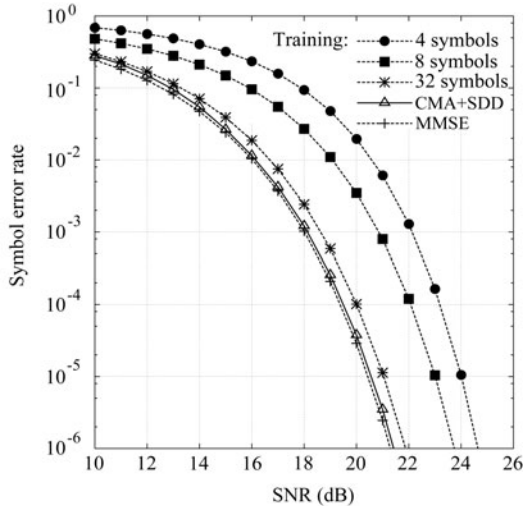

Fig. 2   Desired user-one symbol error rate performance comparison of the proposed semi-blind adaptive beamforming given $K = 4$ training symbols, the training-based adaptive beamforming given different numbers of training symbols, and the true MMSE beamforming given perfect channel knowledge, for the stationary system of four-element array supporting four 16-QAM users

The convergence performance of the proposed semi-blind scheme was investigated. Given the SNR of 18 dB, $K = 4$ training pilots were first used to provide the initial beamforming weight vector according to (7). The appropriate values for the step size of the CMA as well as the step size of the SDD were found empirically, and they were chosen to be $\mu_{\mathrm{CMA}} = 4 \times 10^{-6}$ and $\mu_{\mathrm{SDD}} = 4 \times 10^{-4}$, respectively. A wide range of values were found to be suitable for the cluster width of the SDD. Fig. 3 plots the learning curves of the combined CMA and SDD adaptive algorithm in terms of the SER averaged over ten different runs, given three values of ρ. It is observed from Fig. 3 that, aided by the four training pilots, the convergence rate of the concurrent CMA and SDD algorithm was much faster than the pure blind adaptive counterpart of [15, 16]. Furthermore, the proposed semi-blind scheme is capable of approaching the optimal MMSE solution, as can be seen in Fig. 3. Given the SNR of 18 dB, $K = 4$ training symbols were insufficient for the training-based beamformer, and the eye diagram of the beamformer′s output constellation before the blind adaptation, i.e., with the weight vector $\boldsymbol{w}(0)$ of (7), was completely closed. In contrast, the beamformer′s output constellation after blind adaptation is illustrated in Fig. 4, clearly showing that the eye was opened.
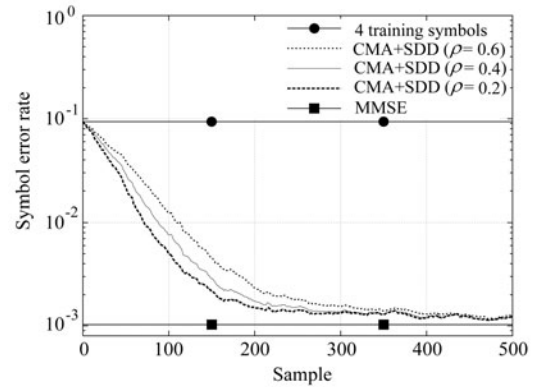

Fig. 3   Learning curves of the concurrent CMA and SDD scheme in terms of the SER averaged over ten different runs for the stationary system of four-element array supporting four 16-QAM users, given the SNR of 18 dB, $\mu_{\mathrm{CMA}} = 4 \times 10^{-6}$, $\mu_{\mathrm{SDD}} = 4 \times 10^{-4}$, and three values of cluster width ρ
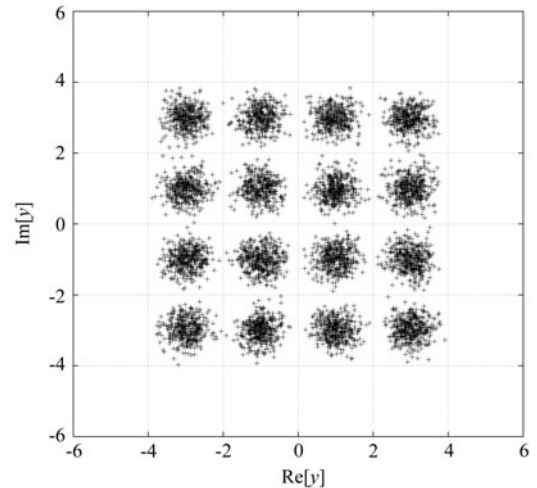

Fig. 4   The beamformer′s output constellation after blind adaptation given SNR of 18 dB for the stationary system of four-element array supporting four 16-QAM users

Flat fading system. A beamforming system with $n_T = 4$, $n_R = 5$, and the 16-QAM modulation scheme was simulated. The system′s channel impulse response taps $h_{l,m}$, $1 \leqslant l \leqslant 5$ and $1 \leqslant m \leqslant 4$, were uncorrelated complex-valued Gaussian processes with zero mean and $\mathrm{E}\left[|h_{l,m}|^2\right] = 1$, and the performance was averaged over 100 system realisations. The number of pilot symbols used for the semi-blind scheme was $K = 5$. The average SER performance for the purely training based scheme with 5, 10, and 40 training symbols, respectively, as well as the proposed semi-blind beamforming scheme with the aid of 5 training symbols are shown in Fig. 5, in comparison with the achievable performance of the MMSE beamforming given the perfect channel knowledge. The step size of the CMA as well as the step size and cluster width of the SDD were empirically set to $\mu_{\mathrm{CMA}} = 1 \times 10^{-7}$, $\mu_{\mathrm{SDD}} = 2 \times 10^{-4}$, and $\rho = 0.4$. The blind adaptive process was observed to achieve convergence typically within 200 to 400 samples. It can be seen from Fig. 5 that to achieve a similar performance as the semi-blind CMA-SDD scheme, the training based scheme required 40 training symbols.
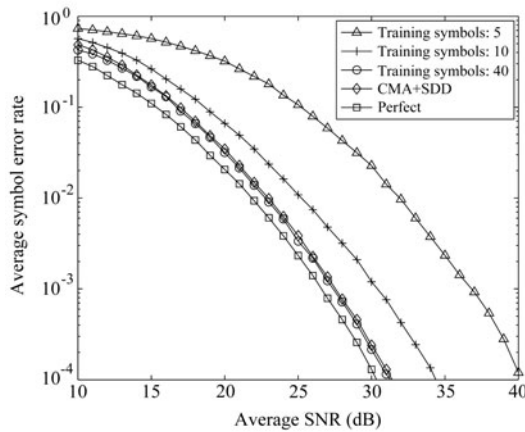
Fig. 5 Average symbol error rate performance of the proposed semi-blind beamforming scheme with five training symbols, in comparison with the cases of training only based on different numbers of training symbols and the MMSE beamforming with the perfect channel knowledge, averaged over 100 realisations of the flat-fading 5×4 16-QAM beamforming system

5 Conclusions

A semi-blind adaptive beamforming scheme of low complexity has been proposed for wireless systems that employ high throughput QAM signalling. A minimum number of training symbols, equal to the number of receiver antenna-array's elements, is used to provide a rough LS estimate of the beamformer's weight vector for initialisation. The CMA aided SDD blind adaptive scheme is then adopted to adapt the beamformer. Our simulation study has confirmed that this semi-blind concurrent CMA and SDD algorithm converges much faster than its pure blind counterpart and the proposed semi-blind adaptive beamformer is capable of approaching the performance of the optimal MMSE beamforming solution.

Appendix

The analytical SER for the beamformer (5) with the weight vector $\boldsymbol{w}$ is given in [29]. Define the combined system response as $\boldsymbol{w}^{\mathrm{H}} \boldsymbol{H} = [c_1 \ c_2 \cdots c_{n_T}]$, and assume that $c_1 = c_{R_1} + j c_{I_1}$ satisfies $c_{R_1} > 0$ and $c_{I_1} = 0$. The MMSE beamforming solution (6) meets this condition. For our proposed semi-blind beamformer, this condition is generally met, as it approaches the MMSE solution. If this condition is not satisfied, a rotation operation can always be performed on the weight vector to guarantee this condition[29]. The SER is expressed as

$$P_E(\boldsymbol{w}) = P_{E_R}(\boldsymbol{w}) + P_{E_I}(\boldsymbol{w}) - P_{E_R}(\boldsymbol{w})P_{E_I}(\boldsymbol{w}) \qquad (\text{A1})$$

where $P_{E_R}(\boldsymbol{w})$ and $P_{E_I}(\boldsymbol{w})$ are the real-part and imaginary-part SERs, respectively. Note $\boldsymbol{x}(k) = \bar{\boldsymbol{x}}(k) + \boldsymbol{n}(k)$ and $y(k) = \bar{y}(k) + e(k)$, where $e(k)$ is Gaussian distributed with zero mean and $\mathrm{E}[|e(k)|^2] = 2\sigma_n^2 \boldsymbol{w}^{\mathrm{H}}\boldsymbol{w}$. The noise-free part $\bar{y}(k)$ takes values from the set $\mathcal{Y}$ that contains $N_s = M^{n_T}$ points. $\mathcal{Y}$ can be divided into the M subsets conditioned on $s_1(k)$ as

$$\mathcal{Y}^{(l,q)} \triangleq \{\bar{y}_i^{(l,q)} \in \mathcal{Y}, 1 \leqslant i \leqslant N_{sb} : s_1(k) = s_{l,q}\} \qquad (\text{A2})$$

for $1 \leqslant l, q \leqslant \sqrt{M}$, where the size of $\mathcal{Y}^{(l,q)}$ is $N_{sb} = N_s/M$. The subset $\mathcal{Y}^{(l,q)}$ is completely specified by the system channel matrix $\boldsymbol{H}$. The SER $P_E(\boldsymbol{w})$ can be calculated based on a single subset $\mathcal{Y}^{(l,q)}$[29]. Expressing $\bar{y}_i^{(l,q)} = \bar{y}_{R_i}^{(l,q)} + j\bar{y}_{I_i}^{(l,q)}$, it can be shown that[29]

$$P_{E_R}(\boldsymbol{w}) = \gamma \frac{1}{N_{sb}} \sum_{i=1}^{N_{sb}} Q(g_{R_i}^{(l,q)}(\boldsymbol{w})) \qquad (\text{A3})$$

$$P_{E_I}(\boldsymbol{w}) = \gamma \frac{1}{N_{sb}} \sum_{i=1}^{N_{sb}} Q(g_{I_i}^{(l,q)}(\boldsymbol{w})) \qquad (\text{A4})$$

where

$$\gamma = \frac{2\sqrt{M} - 2}{\sqrt{M}}$$

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{z^2}{2}} \mathrm{d}z \qquad (\text{A5})$$

$$g_{R_i}^{(l,q)}(\boldsymbol{w}) = \frac{\bar{y}_{R_i}^{(l,q)} - c_{R_1}(u_l - 1)}{\sigma_n \sqrt{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{w}}} \qquad (\text{A6})$$

$$g_{I_i}^{(l,q)}(\boldsymbol{w}) = \frac{\bar{y}_{I_i}^{(l,q)} - c_{R_1}(u_q - 1)}{\sigma_n \sqrt{\boldsymbol{w}^{\mathrm{H}}\boldsymbol{w}}}. \qquad (\text{A7})$$

References

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

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

[3] A. J. Paulraj, D. A. Gore, R. U. Nabar, H. Bölcskei. An overview of MIMO communications — A key to gigabit wireless. *Proceedings of the IEEE*, vol. 92, no. 2, pp. 198–218, 2004.

[4] L. Hanzo, S. X. Ng, T. Keller, W. Webb. *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-coded, Turbo-equalised and Space-time Coded OFDM, CDMA and MC-CDMA Systems*, 2nd Edition, Chichester, UK: John Wiley, 2004.

[5] Air Interface for Fixed Broadband Wireless Access System, IEEE Standard 802.16, Section 8, 2004.

[6] B. Widrow, P. E. Mantey, L. J. Griffiths, B. B. Goode. Adaptive antenna systems. *Proceedings of the IEEE*, vol. 55, no. 12, pp. 2143–2159, 1967.

[7] L. J. Griffiths. A simple adaptive algorithm for real-time processing in antenna arrays. *Proceedings of the IEEE*, vol. 57, no. 10, pp. 1696–1704, 1969.

[8] S. Haykin. *Adaptive Filter Theory*, 3rd Edition, Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[9] S. Chen. Adaptive linear filtering design with minimum symbol error probability criterion. *International Journal of Automation and Computing*, vol. 3, no. 3, pp. 291–303, 2006.

[10] J. J. Shynk, R. P. Gooch. The constant modulus array for cochannel signal copy and direction finding. *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 652–660, 1996.

[11] J. Sheinvald. On blind beamforming for multiple non-Gaussian signals and the constant-modulus algorithm. *IEEE Transactions on Signal Processing*, vol. 46, no. 7, pp. 1878–1885, 1998.

[12] K. Yang, T. Ohira, Y. Zhang, C. Y. Chi. Super-exponential blind adaptive beamforming. *IEEE Transactions on Signal Processing*, vol. 52, no. 6, pp. 1549–1563, 2004.

[13] L. Tong, R. W. Liu, V. C. Soon, Y. F. Huang. Indeterminacy and identifiability of blind identification. *IEEE Transactions on Circuits and Systems*, vol. 38, no. 5, pp. 499–509, 1991.

[14] Y. Inouye, R. W. Liu. A system-theoretic foundation for blind equalization of an FIR MIMO channel system. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, vol. 49, no. 4, pp. 425–436, 2002.

[15] S. Chen, E. S. Chng. Concurrent constant modulus algorithm and soft decision directed scheme for fractionally-spaced blind equalization. In *Proceedings of IEEE International Conference on Communications*, IEEE, Paris, France, vol. 4, pp. 2342–2346, 2004.

[16] S. Chen, A. Wolfgang, L. Hanzo. Constant modulus algorithm aided soft decision directed scheme for blind space-time equalisation of SIMO channels. *Signal Processing*, vol. 87, no. 11, pp. 2587–2599, 2007.

[17] A. Medles, D. T. M. Slock. Semiblind channel estimation for MIMO spatial multiplexing systems. In *Proceedings of the 54th IEEE Vehicular Technology Conference*, IEEE, pp. 1240–1244, 2001.

[18] C. Cozzo, B. L. Hughes. Joint channel estimation and data detection in space-time communications. *IEEE Transactions on Communications*, vol. 51, no. 8, pp. 1266–1270, 2003.

[19] S. Buzzi, M. Lops, S. Sardellitti. Performance of iterative data detection and channel estimation for single-antenna and multiple-antennas wireless communications. *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1085–1104, 2004.

[20] T. Wo, P. A. Hoeher, A. Scherb, K. D. Kammeyer. Performance analysis of maximum-likelihood semiblind estimation of MIMO channels. In *Proceedings of the 63rd IEEE Vehicular Technology Conference*, IEEE, Melbourne, Australia, pp. 1738–1742, 2006.

[21] A. K. Jagannatham, B. D. Rao. Whitening-rotation-based semi-blind MIMO channel estimation. *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 861–869, 2006.

[22] M. Abuthinien, S. Chen, A. Wolfgang, L. Hanzo. Joint maximum likelihood channel estimation and data detection for MIMO systems. In *Proceedings of IEEE International Conference on Communications*, IEEE, Glasgow, Scotland, UK, pp. 5354–5358, 2007.

[23] Z. Ding, T. Ratnarajah, C. F. N. Cowan. HOS-based semi-blind spatial equalization for MIMO Rayleigh fading channels. *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 248–255, 2008.

[24] A. Paulraj, R. Nabar, D. Gore. *Introduction to Space-time Wireless Communications*, Cambridge, UK: Cambridge University Press, 2003.

[25] D. Tse, P. Viswanath. *Fundamentals of Wireless Communication*, Cambridge, UK: Cambridge University Press, 2005.

[26] D. N. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications*, vol. 28, no. 11, pp. 1867–1875, 1980.

[27] J. R. Treichler, B. G. Agee. A new approach to multipath correction of constant modulus signals. *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, no. 2, pp. 459–472, 1983.

[28] S. Chen, S. McLaughlin, P. M. Grant, B. Mulgrew. Multi-stage blind clustering equaliser. *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 701–705, 1995.

[29] S. Chen, H. Q. Du, L. Hanzo. Adaptive minimum symbol error rate beamforming assisted receiver for quadrature amplitude modulation systems. In *Proceedings of the 63rd IEEE Vehicular Technology Conference*, IEEE, Melbourne, Australia, pp. 2236–2240, 2006.

Sheng Chen received the B. Eng. degree from Huadong Petroleum Institute, Dongying, PRC in January 1982, and the Ph. D. degree from the City University, London, UK in September 1986, both in control engineering. He was awarded the D. Sc. degree by the University of Southampton, Southampton, UK in 2005.

From October 1986 to August 1999, he held research and academic appointments at the University of Sheffield, the University of Edinburgh, and the University of Portsmouth, all in UK. Since September 1999, he has been with the School of Electronics and Computer Science, University of Southampton, Southampton, UK. He is a fellow of IET and a fellow of IEEE. In the database of the world's most highly cited researchers, compiled by Institute for Scientific Information (ISI) of the USA, he is on the list of the highly cited researchers in the engineering category.

His research interests include wireless communications, adaptive signal processing for communications, machine learning and neural networks, finite-precision digital controller design, networked control systems, and evolutionary computation methods.

E-mail: sqc@ecs.soton.ac.uk (Corresponding author)
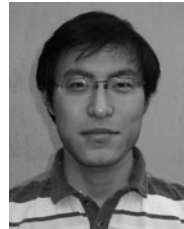
Wang Yao received the B. Sc. degree in information and computing science from Beijing University of Posts and Telecommunications (BUPT), Beijing, PRC in 2006 and the M. Sc. degree with distinction in radio frequency communication systems from the University of Southampton, Southampton, UK in 2007. He is currently working toward the Ph. D. degree with the Communications Research Group, School of Electronics and Computer Science, University of Southampton, UK.

His research interests include multiuser transmission and multiuser detection in multiple-input and multiple-output (MIMO) systems, and signal processing for communications.

E-mail: wy07r@ecs.soton.ac.uk

Lajos Hanzo received the master degree in electronics in 1976 and his doctorate in 1983. In 2004, he was awarded the D. Sc. degree by the University Southampton, Southampton, UK.

During his 30-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the UK. Since 1986, he has been with the School of Electronics and Computer Science, the University of Southampton, UK, where he holds the chair in telecommunications.

He has co-authored 20 John Wiley/IEEE Press books totalling about 20000 pages on mobile radio communications, published in excess of 800 research papers, organised and chaired conference sessions, presented overview lectures and has been awarded a number of distinctions. He is an enthusiastic supporter of industrial-academic liaison. He also offers a range of industrial research overview courses. He is a fellow of the Royal Academy of Engineering (FREng), UK. He is an IEEE Distinguished Lecturer of both the Communications Society and the Vehicular Technology Society as well as a fellow of both IEEE and IEE.

His research interests include multimedia communications, wireless communications and mobile communication.

E-mail: lh@ecs.soton.ac.uk