DOI: XXX

Recursive Least Squares Semi-blind Beamforming for MIMO Using Decision Directed Adaptation and Constant Modulus Criterion

Xia Hong[1] Sheng Chen[2,3]

[1] Department of Computer Science, School of Mathematical and Physical Sciences, University of Reading, Reading, RG6 6AY, UK (x.hong@reading.ac.uk)

[2] School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK (sqc@ecs.soton.ac.uk)

[3] King Abdulaziz University, Jeddah 21589, Saudi Arabia

Abstract: A new semi-blind adaptive beamforming scheme is proposed for multi-input multi-output (MIMO) induced and space-division multiple-access based wireless systems that employ high order phase shift keying signaling. A minimum number of training symbols, very close to the number of receiver antenna elements, are used to provide a rough initial least squares estimate of the beamformer's weight vector. A novel cost function combining the constant modulus criterion with decision-directed adaptation is adopted to adapt the beamformer weight vector. This cost function can be approximated as a quadratic form with a closed-form solution, based on which we then derive the recursive least squares (RLS) semi-blind adaptive beamforming algorithm. 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. Our proposed semi-blind RLS beamforming algorithm therefore provides an efficient detection scheme for the future generation of MIMO aided mobile communication systems.

Keywords: Multi-input multi-output, space-division multiple-access, beamforming, semi-blind adaptive algorithm, constant modulus criterion, decision-directed adaption

1 Introduction

Mobile communication technology has gone through four generations of development, and currently the industry is actively developing the beyond fourth generation (B4G) or fifth generation (5G) system. A quick review of the evolution of mobile communication systems will serve the motivation for our current study. The first generation (1G) and second generation (2G) mobile networks were based on orthogonal channel access, and they offered limited user capacity because of the limited frequency-time resources. In order to better utilize the available frequency-time resources and to support broadband applications, starting from the third generation (3G) system and especially in the current fourth generation (4G) system, a fundamental paradigm shift was occurred to allow non-orthogonal access. The rapid development of mobile communication technology in turns fuels ever-increasing new applications, including mobile Internet, social networks and social media, demanding higher and higher data rates. According to [1], the global mobile data traffic has reached 3.7 exabytes per month in 2015 and it is expected to reach 30.6 exabytes per month by 2020 (1 exabyte equals to 10^{18} bytes.), Drastic new mobile communication techniques are needed quickly in order to meet this explosively increasing demands, and multi-input multi-output (MIMO) technology [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14] is a promising component for the future 5G system.

The demand for increasing the capacity of mobile communication systems has thus motivated the development of new communication technologies, in particular, the so-called space-division multiple-access (SDMA) technology, in order to further improve the efficiency of spectral utilisation. Equipped with multiple antennas, a base station (BS) becomes capable of serving multiple users with the same frequency-time resource block by exploiting the spatial dimension, and this effectively offers the potentially unlimited communication resources. In the resulting SDMA induced MIMO system, the adaptive beamforming receiver provides the effective means of uplink data detection at the BS [15, 16, 17, 18, 19], while the transmit beamforming or precoding offers the effective way of downlink data transmission [20, 21, 22, 23, 24]. Both uplink detection and downlink precoding require the knowledge of the MIMO channel state information (CSI). Therefore, the performance of a MIMO communication system heavily relies on the accuracy of the MIMO channel matrix estimation [25]. This paper considers the uplink data detection for MIMO systems. The classical beamforming receiver design is the minimum mean square error (MMSE) solution, which can be realized using various training-based adaptive algorithms [25, 26, 27, 28]. However, pure training-based schemes require a high training overhead, thus considerably reducing the achievable system throughput. Pure blind beamforming [29, 30, 31, 32, 33] 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 [34].

An effective means of resolving the estimation and decision ambiguities inherent in pure blind schemes is to employ only a few training symbols to provide a rough initial estimate and then to switch to a blind adaptive scheme or a decision-directed (DD) adaptive algorithm, which leads to

Regular paper
Manuscript received date –/–/–; revised date –/–/–
Recommended by Associate Editor xxxx

various semi-blind schemes [35, 36, 37, 38, 39]. In particular, the work [40] proposed a concurrent constant modulus (CM) algorithm and soft DD scheme to adapt the beamforming receiver, with the beamformer weight vector initialized by the pure training based least squares (LS) estimate based only on a minimum number of training symbols, which is equal to the number of the receiver antenna elements. More specifically, after the initial training, the weight updating formula of the combined blind CM adaptation and soft DD adaptation is based on the combined stochastic constant modulus algorithm (CMA) and a stochastic gradient ascent of the local maximum likelihood of the beamformer output [40]. However, it is well known that stochastic gradient methods are quite sensitive to the selected step size and have a slow convergence rate. In comparison to stochastic gradient methods, the recursive least squares (RLS) algorithm converges much faster. In [41], a blind RLS based CM algorithm, referred to as the RLS+CMA, was proposed for CM signals by approximating the CM cost function as a quadratic form, thus enabling direct application of the well-known RLS algorithm.

Against this background, we propose a novel cost function combining the CM criterion with the DD adaptation for adaptive estimation of the beamformer weight vector for SDMA induced MIMO wireless communication systems that employ phase shift keying (PSK) signaling. More specifically, the optimization cost function is the sum of the CM cost function and the LS error based on the previously detected hard symbols. We adopt the idea of [41] to approximate the CM part of the proposed cost function as a quadratic function. For this composite cost function, we show that there exists a closed-form optimal LS solution. This enables us to derive the proposed RLS semi-blind adaptive beamforming algorithm. In the present semi-blind beamforming application with a minimum number of training pilot symbols, we demonstrate that the proposed RLS+DD+CM algorithm converges very fast and is capable of approaching the performance of the MMSE beamforming solution associated with the perfect MIMO CSI. Moreover, the simulation results show that the proposed semi-blind RLS+DD+CM algorithm significantly outperforms the semi-blind RLS+CMA for high-order PSK signals, in terms of convergence rate and achievable system's symbol error rate (SER).

The remainder of this paper is organized as follows. In Section 2, we briefly introduce the beamforming receiver model for SDMA induced MIMO communication systems, while Section 3 is devoted to our proposed RLS semi-blind algorithm. Our simulation results are presented in Section 4, and our concluding remarks are offered in Section 5.

Throughout this contribution, we adopt the following notational conventions. The complex number field is denoted by $\mathbb{C}$. Boldface capitals and lower case letters stand for matrices and vectors, respectively, while $\boldsymbol{I}_p$ stands for the $p\times p$ identity matrix and $\mathbf{1}_p$ denotes an all-one vector of length p. Additionally, $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ represent the transpose and Hermitian operators, respectively, while $\|\cdot\|$ and $|\cdot|$ denote the norm and magnitude operators, respectively. Furthermore, $(\cdot)^*$ denotes the complex conjugate operation and $\mathrm{j}=\sqrt{-1}$ represents the imaginary axis, while $\mathrm{E}[\cdot]$ denotes the expectation operator and $\mathrm{diag}\{a_1, a_2, \cdots, a_p\}$ represents the diagonal matrix with the diagonal elements of $a_1, a_2, \cdots, a_p$.

2 Beamforming Receiver Model

We consider the coherent MIMO communication system that supports n_T users on the same frequency-time resource block, where each user is equipped with a single antenna and transmits an M-PSK signal on the same angular carrier frequency of ω. In order to achieve user separation in the angular domain, the BS receiver is equipped with a uniformly spaced linear antenna array (ULA) consisting of n_R antenna elements. We further assume that the communication is over flat fading channels. Then, at the symbol index k, the system is described by the following well-known MIMO model

$$\boldsymbol{x}(k) = \boldsymbol{H}\boldsymbol{s}(k) + \boldsymbol{\varepsilon}(k), \tag{1}$$

where $\boldsymbol{x}(k) = \big[x_1(k)\ x_2(k) \cdots x_{n_R}(k)\big]^{\mathrm{T}} \in \mathbb{C}^{n_R\times 1}$ is the received signal vector and $\boldsymbol{\varepsilon}(k) = \big[\varepsilon_1(k)\ \varepsilon_2(k) \cdots \varepsilon_{n_R}(k)\big]^{\mathrm{T}} \in \mathbb{C}^{n_R\times 1}$ is the system's additive Gaussian white noise (AWGN) vector having $\mathrm{E}\big[\boldsymbol{\varepsilon}(k)\boldsymbol{\varepsilon}^{\mathrm{H}}(k)\big] = 2\sigma_\varepsilon^2 \boldsymbol{I}_{n_R}$, while $\boldsymbol{s}(k) = \big[s_1(k)\ s_2(k) \cdots s_{n_T}(k)\big]^{\mathrm{T}} \in \mathbb{C}^{n_T\times 1}$ is the transmitted symbol vector of the n_T users with the symbol energy given by $\mathrm{E}\big[|s_m(k)|^2\big] = \sigma_s^2$ for $1 \le m \le n_T$, and $\boldsymbol{H} \in \mathbb{C}^{n_R\times n_T}$ is the $n_R \times n_T$ MIMO channel matrix.

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

$$\boldsymbol{H} = \big[\boldsymbol{h}_1\ \boldsymbol{h}_2 \cdots \boldsymbol{h}_{n_T}\big] = \big[A_1\boldsymbol{\eta}_1\ A_2\boldsymbol{\eta}_2 \cdots A_{n_T}\boldsymbol{\eta}_{n_T}\big], \tag{2}$$

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

$$\boldsymbol{\eta}_m = \Big[e^{\mathrm{j}\omega t_1(\theta_m)}\ e^{\mathrm{j}\omega t_2(\theta_m)} \cdots e^{\mathrm{j}\omega t_{n_R}(\theta_m)}\Big]^{\mathrm{T}} \in \mathbb{C}^{n_R\times 1}, \tag{3}$$

where θ_m is the angle of arrival for user m, which is assumed to be uniformly distributed in $[0,\ 2\pi)$, and $t_l(\theta_m)$ is

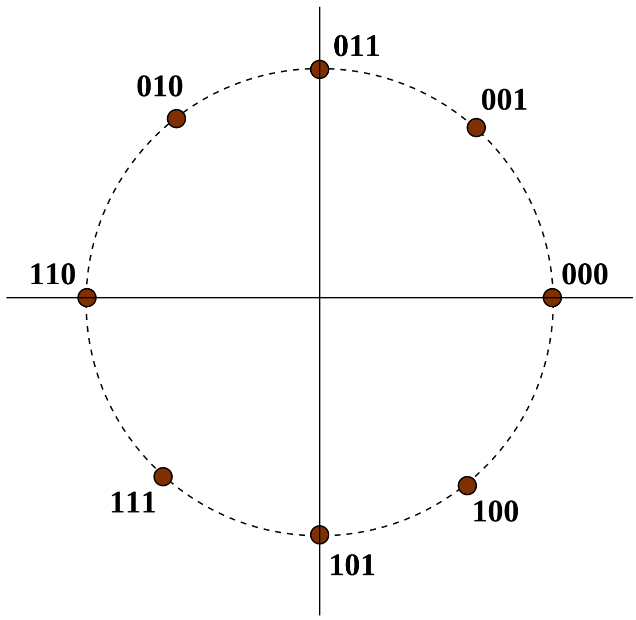

Figure 1: Grey encoded 8PSK constellation diagram with 3 bits per symbol.

the relative time delay at array element l for user m. 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 $\boldsymbol{H}$ are deemed unchanged. From frame to frame, $h_{l,m}$ are assumed to be uncorrelated complex-valued Gaussian processes with zero mean and $\mathrm{E}[|h_{l,m}|^2] = 1$.

The transmitted data symbols $s_m(k)$, $1 \le m \le n_T$, take the values from the M-PSK symbol set defined by

$$\mathcal{S} \triangleq \left\{ s^{(l)} = e^{\mathrm{j}2\pi l/M}, 1 \le l \le M \right\}, \tag{4}$$

where M takes the value of 2, 4, 8, 16 and so on, which correspond to 1 bit per symbol, 2 bits per symbol, 3 bits per symbol, 4 bits per symbol and etc. The system's signal-to-noise ratio (SNR) is defined by $\mathrm{SNR} = n_T\sigma_s^2/\big(2\sigma_\varepsilon^2\big) = n_T/\big(2\sigma_\varepsilon^2\big)$, as $\sigma_s^2 = 1$. The 8PSK symbol set is illustrated in Figure 1.

Without loss of generality, user one is assumed to be the desired user and the rest of the users are interfering ones. A beamforming receiver [18], specified by

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

is used to detect the transmitted symbols $s_1(k)$, where $\boldsymbol{w} \in \mathbb{C}^{n_R\times 1}$ is the $n_R \times 1$ complex-valued weight vector of the beamformer. With the perfect MIMO CSI, the MMSE solution that minimizes the mean squares error criterion $\mathrm{E}\big[|s_1(k) - y(k)|^2\big]$ is given by

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

3 Proposed Recursive Least Squares Semi-blind Algorithm

Given the training data $\boldsymbol{X}_K = [\boldsymbol{x}(1)\ \boldsymbol{x}(2)\cdots\boldsymbol{x}(K)] \in \mathbb{C}^{n_R\times K}$ and $\boldsymbol{s}_K = \big[s_1(1)\ s_1(2)\cdots s_1(K)\big]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$, the LS estimate of the beamformer's weight vector is readily given as

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

In order to maintain the achievable system throughput, the number of training symbols K should be as small as possible. But to ensure that $\boldsymbol{X}_K\boldsymbol{X}_K^{\mathrm{H}}$ has a full rank of n_R, it must be $K \ge n_R$. We will choose K close to the minimum number of training symbols, i.e., $K = n_R$ or $n_R + 1$. Given the initial weight vector $\boldsymbol{w}(0)$, 'blind' adaptation then takes place online, i.e., the beamformer weight vector is computed recursively in time, so that $\boldsymbol{w}(k)$ at time k is given as a modification of $\boldsymbol{w}(k-1)$, upon the arrival of the new received data $\{\boldsymbol{x}(k)\}$.

The widely adopted DD adaption minimizes the cost function given by

$$J_{\mathrm{DD}} = \mathrm{E}\big[\big|\widehat{s}_1(n) - y(n)\big|^2\big], \tag{8}$$

where $\widehat{s}_1(n)$ is the hard decision for $s_1(n)$ based on the current beamformer output $\widehat{y}(n|n-1) = \boldsymbol{w}^{\mathrm{H}}(n-1)\boldsymbol{x}(n)$, denoted by $\widehat{s}_1(n) = \mathrm{dec}\big(\widehat{y}(n|n-1)\big)$. Because the training data are insufficient, the initial LS weight vector (7) with $K \approx n_R$ may not be sufficiently accurate to open the eye. Therefore, the DD adaptation is generally unsafe. Alternatively, the well-known CMA penalizes the deviation of the beamformer output from a constant modulus, and the CM cost function is defined by

$$J_{\mathrm{CM}} = \mathrm{E}\big[\big(|y(n)|^2 - 1\big)^2\big]. \tag{9}$$

To utilize both the advantages of the DD adaptation and the CMA, we consider the combined cost function

$$J = \mathrm{E}\big[\big(|y(n)|^2 - 1\big)^2\big] + \mathrm{E}\big[\big|\widehat{s}_1(n) - y(n)\big|^2\big]. \tag{10}$$

Replacing the statistical expectation operator in (10) with an exponentially weighted time average sum yields

$$J \approx \sum_{n=1}^{k}\lambda^{k-n}\Big(\big(|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}(n)|^2 - 1\big)^2 + |\widehat{s}_1(n) - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}(n)|^2\Big), \tag{11}$$

where λ is a forgetting factor that is slightly less than 1, e.g., 0.99 to 0.95. Note that the first term in J is not quadratic, and we use the idea given in [41] to approximate it as

$$\big|\boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}(n)\big|^2 \approx \boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}(n), \tag{12}$$

where

$$\boldsymbol{z}(n) = \boldsymbol{x}(n)\boldsymbol{x}^{\mathrm{H}}(n)\boldsymbol{w}(k-1) \in \mathbb{C}^{n_R\times 1}. \tag{13}$$

With this approximation, we have

$$J \approx \sum_{n=1}^{k}\lambda^{k-n}\Big(\big(\boldsymbol{w}^{\mathrm{H}}\boldsymbol{z}(n) - 1\big)^2 + |\widehat{s}_1(n) - \boldsymbol{w}^{\mathrm{H}}\boldsymbol{x}(n)|^2\Big). \tag{14}$$

Algorithm 1 RLS+DD+CM semi-blind algorithm.

1: Initialize $\boldsymbol{w}(0)$ according to (7) and set $\boldsymbol{P}_0 = \left(\boldsymbol{X}_K\boldsymbol{X}_K^{\mathrm{H}}\right)^{-1}$ based on the training data set $\{\boldsymbol{X}_K, \boldsymbol{s}_K\}$.
2: **for** time step $k = 1, 2, \cdots,$ **do**
3: Calculate

$$\left\{\begin{array}{l}
\widehat{y}(k|k-1) = \boldsymbol{w}^{\mathrm{H}}(k-1)\boldsymbol{x}(k) \\
\widehat{s}_1(k) = \left\{\begin{array}{ll} s_1(k) & \text{if } k \le K \\ \mathrm{dec}\big(\widehat{y}(k|k-1)\big) & \text{if } k > K \end{array}\right. \\
\alpha_k = \big|\widehat{y}(k|k-1)\big|^2 + 1 \\
\beta_k = \widehat{s}_1^{*}(k) - \big|\widehat{y}(k|k-1)\big|^2\widehat{y}^{*}(k|k-1) \\
\boldsymbol{P}_k = \dfrac{1}{\lambda}\left(\boldsymbol{P}_{k-1} - \dfrac{\alpha_k\boldsymbol{P}_{k-1}\boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{H}}(k)\boldsymbol{P}_{k-1}}{\lambda + \alpha_k\boldsymbol{x}^{\mathrm{H}}(k)\boldsymbol{P}_{k-1}\boldsymbol{x}(k)}\right) \\
\boldsymbol{w}(k) = \boldsymbol{w}(k-1) + \beta_k\boldsymbol{P}_k\boldsymbol{x}(k)
\end{array}\right.$$

4: **end for**

The approximation (12) is reasonable since the difference between $\boldsymbol{w}^{\mathrm{H}}(k)\boldsymbol{x}(n)$ and $\boldsymbol{w}^{\mathrm{H}}(k-1)\boldsymbol{x}(n)$ is usually small when n is close to k, whereas when n is much less than k the difference will be 'forgotten' or attenuated by the factor λ^{k-n}.

Let $\widehat{\boldsymbol{s}}_1(k) = \left[\widehat{s}_1(1)\ \widehat{s}_1(2)\cdots\widehat{s}_1(k)\right]^{\mathrm{T}} \in \mathbb{C}^{k\times 1}$ and $\boldsymbol{\Lambda}_k = \mathrm{diag}\{\lambda^{k-1},\cdots,\lambda,1\} \in \mathbb{C}^{k\times k}$. Further denote $\boldsymbol{X}_k = \left[\boldsymbol{X}_{k-1}\ \boldsymbol{x}(k)\right] \in \mathbb{C}^{n_R\times k}$ and $\boldsymbol{Z}_k = \left[\boldsymbol{Z}_{k-1}\ \boldsymbol{z}(k)\right] \in \mathbb{C}^{n_R\times k}$ with with $\boldsymbol{Z}_1 = \boldsymbol{z}(1)$ and $\boldsymbol{X}_1 = \boldsymbol{x}(1)$.

The cost function (11) can be equivalently expressed as

$$\begin{aligned} J \approx & \left(\mathbf{1}_k - \boldsymbol{Z}_k^{\mathrm{H}}\boldsymbol{w}\right)^{\mathrm{H}}\boldsymbol{\Lambda}_k\left(\mathbf{1}_k - \boldsymbol{Z}_k^{\mathrm{H}}\boldsymbol{w}\right) \\ & + \left(\widehat{\boldsymbol{s}}_1^*(k) - \boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{w}\right)^{\mathrm{H}}\boldsymbol{\Lambda}_k\left(\widehat{\boldsymbol{s}}_1^*(k) - \boldsymbol{X}_k^{\mathrm{H}}\boldsymbol{w}\right). \end{aligned} \tag{15}$$

The minimizer of J is given as

$$\boldsymbol{w}(k) = \boldsymbol{P}_k\left(\boldsymbol{Z}_k\boldsymbol{\Lambda}_k\mathbf{1}_k + \boldsymbol{X}_k\boldsymbol{\Lambda}_k\widehat{\boldsymbol{s}}_1^*(k)\right), \tag{16}$$

where $\boldsymbol{P}_k = \left(\boldsymbol{Z}_k\boldsymbol{\Lambda}_k\boldsymbol{Z}_k^{\mathrm{H}} + \boldsymbol{X}_k\boldsymbol{\Lambda}_k\boldsymbol{X}_k^{\mathrm{H}}\right)^{-1} \in \mathbb{C}^{n_R\times n_R}$.

At the time index $(k-1)$, the weight vector (16) is in the form of

$$\boldsymbol{w}(k-1) = \boldsymbol{P}_{k-1}\Big(\boldsymbol{Z}_{k-1}\boldsymbol{\Lambda}_{k-1}\mathbf{1}_{k-1} + \boldsymbol{X}_{k-1}\boldsymbol{\Lambda}_{k-1}\widehat{\boldsymbol{s}}_1^*(k-1)\Big). \tag{17}$$

Using (13), it is easy to verify that

$$\begin{aligned} \boldsymbol{P}_k^{-1} =& \lambda\boldsymbol{P}_{k-1}^{-1} + \boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{H}}(k) + \boldsymbol{z}(k)\boldsymbol{z}^{\mathrm{H}}(k) \\ =& = \lambda\boldsymbol{P}_{k-1}^{-1} + \alpha_k\boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{H}}(k) \end{aligned} \tag{18}$$

with $\alpha_k = \left|\widehat{y}(k|k-1)\right|^2 + 1$, where $\widehat{y}(k|k-1) = \boldsymbol{w}^{\mathrm{H}}(k-1)\boldsymbol{x}(k)$. Using the famous matrix inversion lemma, we can recursively calculate $\boldsymbol{P}_k$ according to

$$\boldsymbol{P}_k = \frac{1}{\lambda}\left(\boldsymbol{P}_{k-1} - \frac{\alpha_k\boldsymbol{P}_{k-1}\boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{H}}(k)\boldsymbol{P}_{k-1}}{\lambda + \alpha_k\boldsymbol{x}^{\mathrm{H}}(k)\boldsymbol{P}_{k-1}\boldsymbol{x}(k)}\right). \tag{19}$$

Similarly, using (13) and (18), we have

$$\begin{aligned} \boldsymbol{Z}_k\boldsymbol{\Lambda}_k\mathbf{1}_k + \boldsymbol{X}_k\boldsymbol{\Lambda}_k\widehat{\boldsymbol{s}}_1^*(k) =& \lambda\big(\boldsymbol{Z}_{k-1}\boldsymbol{\Lambda}_{k-1}\mathbf{1}_{k-1} + \\ & \boldsymbol{X}_{k-1}\boldsymbol{\Lambda}_{k-1}\widehat{\boldsymbol{s}}_1^*(k-1)\big) + \boldsymbol{z}(k) + \boldsymbol{x}(k)\widehat{s}_1^*(k) \\ =& \lambda\boldsymbol{P}_{k-1}^{-1}\boldsymbol{w}(k-1) + \boldsymbol{z}(k) + \boldsymbol{x}(k)\widehat{s}_1^*(k) \\ =& \boldsymbol{P}_k^{-1}\boldsymbol{w}(k-1) + \beta_k\boldsymbol{x}(k), \end{aligned} \tag{20}$$

where $\beta_k = \widehat{s}_1^*(k) - \left|\widehat{y}(k|k-1)\right|^2\widehat{y}^*(k|k-1)$.

Substituting (20) into (16) leads to the recursive formula for updating the beamformer weight vector as

$$\boldsymbol{w}(k) = \boldsymbol{w}(k-1) + \beta_k\boldsymbol{P}_k\boldsymbol{x}(k). \tag{21}$$

We summarize the proposed RLS semi-blind beamforming algorithm using the combined DD adaptation and CM criterion, referred to as the RLS+DD+CM, in Algorithm 1. Like standard RLS, the proposed algorithm has a computational complexity on the order of $\mathrm{O}(n_R^2)$, mainly due to calculating $\boldsymbol{P}_k$.

Remarks:

The difference between the proposed approach with [40]: The proposed approach is very different from [40]. Firstly, the previous work [40] is based on stochastic gradient type algorithms, which have a much slower convergence rate than RLS (it takes tens of thousands of samples to converge). Secondly the previous work [40] is based on the combined blind CM adaptation and soft DD adaptation by maximizing a local approximation of marginal probability density function, while the proposed approach is based on the combined blind CM adaptation and hard DD adaptation using a composite least squares error cost function. Soft DD is probabilistic model based and it fits well stochastic gradient type algorithms but there is no obvious way of implementing the soft DD adaption with RLS framework. Hence the

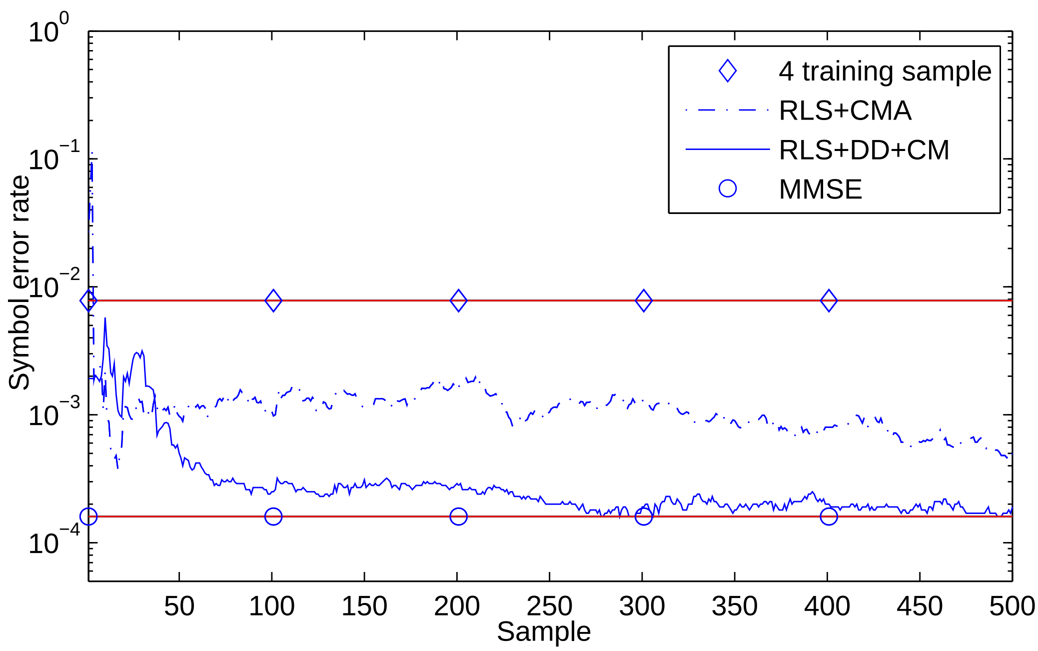

Figure 2: Learning curves, in terms of desired user-one SER, averaged over 200 runs for the stationary system of the four-element ULA supporting four users under the senario of SNR = 12 dB and QPSK modulation scheme.

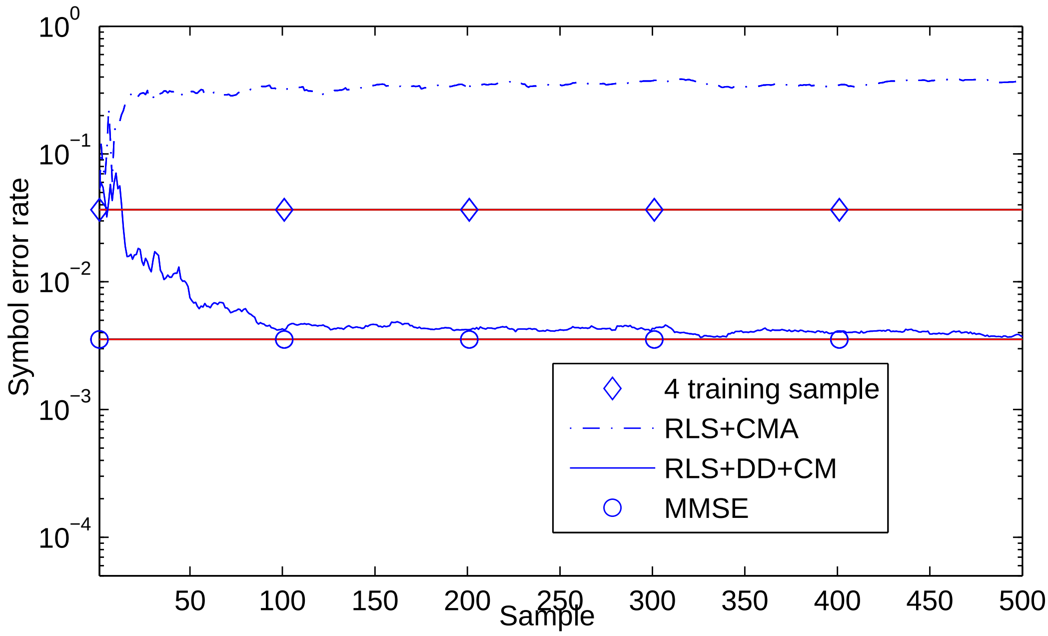

Figure 3: Learning curves, in terms of desired user-one SER, averaged over 200 runs for the stationary system of the four-element ULA supporting four users under the senario of SNR = 15 dB and 8PSK modulation scheme.

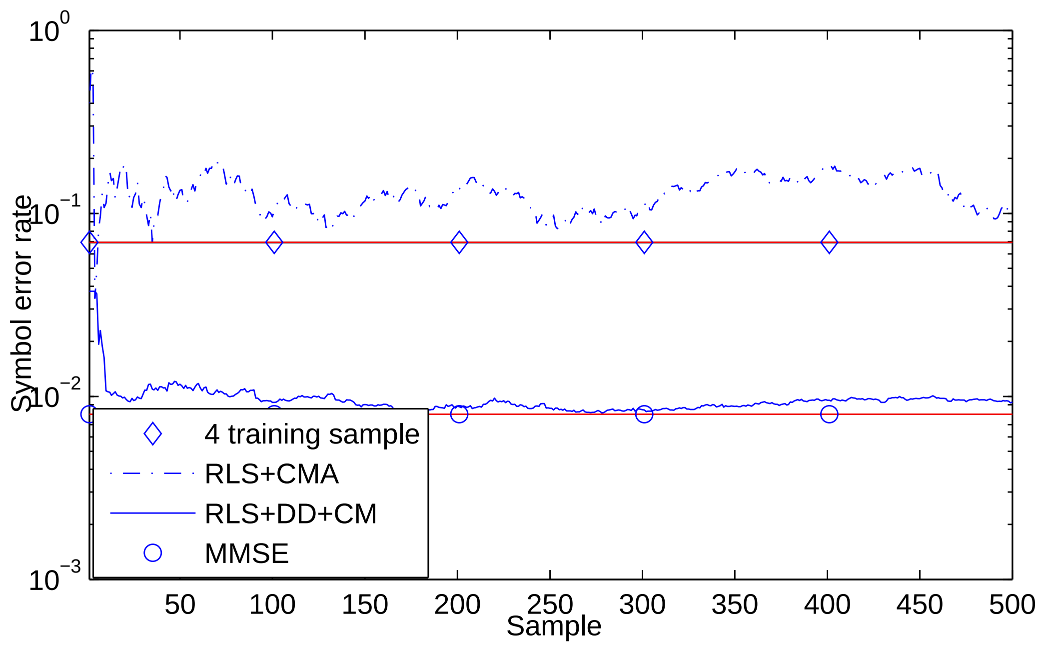

Figure 4: Learning curves, in terms of desired user-one SER, averaged over 200 runs for the stationary system of the four-element ULA supporting four users under the senario of SNR = 20 dB and 16PSK modulation scheme.

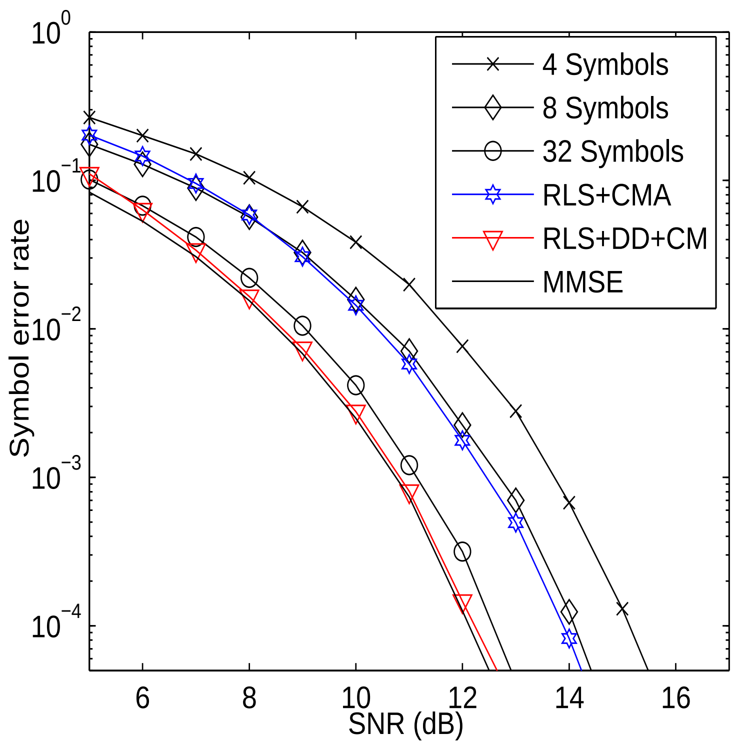

Figure 5: Desired user-one SER performance comparison of the proposed semi-blind RLS+DD+CM algorithm given $K = 4$ training symbols, the semi-blind RLS+CMA [41] given the same $K = 4$ training symbols, the training based beamforming given different numbers of training symbols, and the optimal MMSE beamforming given perfect CSI, for the stationary system of the four-element ULA supporting four users employing signaling of QPSK.

hard DD adaptation is adopted in the proposed approach.

4 Simulation Study

4.1 Stationary System

A ULA with $n_R = 4$ elements and a half-wavelength element spacing was employed to support $n_T = 4$ PSK users. The quadrature PSK (QPSK), 8PSK and 16PSK signalings were considered. The angles of arrival for the four users were $10°$, $40°$, $-15°$ and $-45°$, respectively. The simulated stationary channels were $A_m = 1$, $1 \leq m \leq n_T$. The number of training symbols for a semi-blind scheme was $K = n_R = 4$.

The convergence performance of the proposed semi-blind RLS+DD+CM algorithm was first investigated. Figures 2 to 4 plot the learning curves in terms of desired user-one SER, averaged over 200 runs, in comparison to those of the semi-blind RLS+CMA [41], for the three scenarios of SNR = 12 dB and QPSK, SNR = 15 dB and 8PSK, as well as SNR = 20 dB and 16PSK, respectively, where the SERs of the beamforming receiver based on only 4 training symbols and the MMSE beamforming receiver based on the perfect CSI are also depicted as the benchmarks. We set $\lambda = 0.99$ for both the semi-blind RLS+CMA and proposed RLS+DD+CM. Clearly, with only 4 training symbols, the existing semi-blind RLS+CMA [41] only works reasonably well in the case of QPSK, as can be seen from Figure 2, but it fails for the systems employing 8PSK and 16PSK, as can be observed from Figures 3 and 4. By contrast, our proposed semi-blind RLS+DD+CM algorithm is able to converge rapidly towards the optimal MMSE solutions for all the three systems, adopting QPSK, 8PSK and 16PSK, respectively.

Figures 5 to 7 depict the SER performance of the proposed semi-blind RLS+DD+CM algorithm and the existing semi-blind RLS+CMA [41], both given $K = 4$ training symbols, for the three systems employing 8PSK and 16PSK, respectively, where the beamformers' weight vectors used for SER calculation are those obtained at time index $k = 500$. For comparison purpose, the SERs of the training based beamforming receiver with various training overheads as well as the optimal MMSE solution are also shown in Figures 5 to 7. It can be seen from Figures 5 to 7 that with only four training symbols, our proposed semi-blind RLS+DD+CM algorithm approaches the optimal MMSE solution associated with the perfect MIMO CSI for all the three systems adopting QPSK, 8PSK and 16PSK, respectively. By contrast, with four training symbols, the existing semi-blind RLS+CMA [41] performs poorly. Specifically, for the QPSK system, after the 500 samples of blind RLS+CMA adaptation, its SER is only close to the SER based on 8 training symbols, while for the 8PSK and 16QPSK systems, after the 500 samples of blind RLS+CMA adaptation, its SER is actually worse than the initial SER based on 4 training symbols. We point out that since the basic difference between RLS+DD+CM and RLS+CMA algorithm is to introduce the combined cost function us-

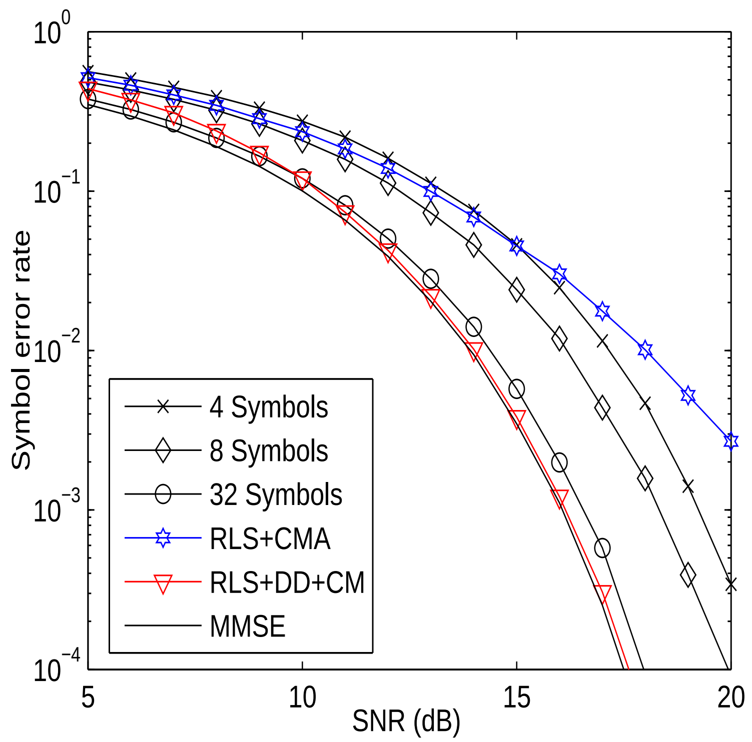

Figure 6: Desired user-one SER performance comparison of the proposed semi-blind RLS+DD+CM algorithm given $K = 4$ training symbols, the semi-blind RLS+CMA [41] given the same $K = 4$ training symbols, the training based beamforming given different numbers of training symbols, and the optimal MMSE beamforming given perfect CSI, for the stationary system of the four-element ULA supporting four users employing signaling of 8PSK.

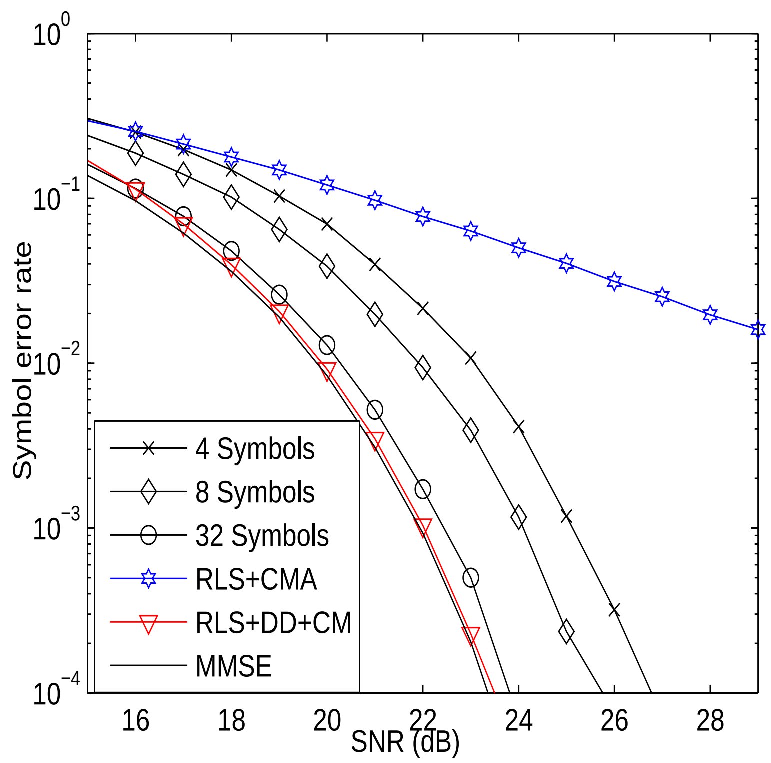

Figure 7: Desired user-one SER performance comparison of the proposed semi-blind RLS+DD+CM algorithm given $K = 4$ training symbols, the semi-blind RLS+CMA [41] given the same $K = 4$ training symbols, the training based beamforming given different numbers of training symbols, and the optimal MMSE beamforming given perfect CSI, for the stationary system of the four-element ULA supporting four users employing signaling of 16PSK.

ing DD, it can be concluded from the results that this novel cost function combining the constant modulus criterion with decision-directed adaptation has helped.

4.2 Flat Fading System

A beamforming receiver assisted communication system with $n_T = 4$, $n_R = 5$ and 8PSK signaling was simulated. The system's channel impulse response taps $h_{l,m}$ for $1 \leq l \leq 5$ and $1 \leq m \leq 4$ were the uncorrelated complex-valued Gaussian processes with zero mean and $\mathrm{E}\left[\left|h_{l,m}\right|^2\right] = 1$. 100 random MIMO channel matrices were generated and for each random MIMO channel matrix, the performance was averaged over 100 system realisations.

The average SER performance for the purely training based scheme with 6, 16 and 32 training symbols, respectively, the existing semi-blind RLS+CMA algorithm [41] initialized by training using $K = n_R + 1 = 6$ symbols, and our proposed semi-blind RLS+DD+CM algorithm with the aid of the same 6 training symbols are shown in Figure 8, in comparison with the achievable performance of the MMSE beamforming receiver solution given the perfect MIMO CSI. Similar to the previous stationary example, $\lambda = 0.99$ was set for both the semi-blind beamforming receiver schemes, and the beamformers' weight vectors used for SER calculation are those obtained at time index $k = 500$. It can be observed from Figure 8 that in order to achieve a similar performance as the proposed semi-blind beamforming scheme, the purely training based scheme requires 32 training symbols, and our proposed semi-blind scheme converges to the optimal MMSE solution associated with the perfect MIMO CSI when the system's SNR is sufficiently high. By contrast, the existing semi-blind RLS+CMA algorithm [41] fails to work and exhibits a high SER floor, as can be seen clearly from Figure 8.

5 Conclusions

In this paper we have introduced a novel semi-blind adaptive beamforming receiver scheme for SDMA induced MIMO communication systems that employ high-order PSK signaling. To resolve the unavoidable estimation and decision ambiguities associated with pure blind adaptive schemes, we have only used a minimum number of training symbols to provide a rough initial least squares estimate of the beamformer's weight vector, which is equal to the number of receiver antennas. The beamforming receiver weight vector is obtained by minimizing a composite cost function based on the constant modulus criterion and decision-directed adaptation. The proposed scheme is a type of RLS algorithm since this cost function can be approximated as a quadratic form with a closed-form solution. This semi-blind adaptive beamforming receiver scheme is capable of converging fast to the optimal MMSE beamforming solution associated with the perfect MIMO CSI. The simulation results have also demonstrated that our proposed semi-blind RLS+DD+CM algorithm significantly outperforms the existing semi-blind RLS+CMA algorithm, in terms of both convergence rate and achievable SER performance.

References

[1] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2016 White Paper. Available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ visual-networking-index-vni/mobile-white-paper-c11-520862.html

[2] 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, nos. 2/3/4, pp. 1740–1751, 1994.

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

[4] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Communications*, vol. 5, no. 1, pp. 28–35, Feb. 1998.

[5] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communication*, vol. 5, no. 1, pp. 23–27, Feb. 1998.

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

[7] P. Vandenameele, L. van Der Perre, and M. Engels, *Space Division Multiple Access for Wireless Local Area Networks*. Boston: Kluwer Academic Publishers, 2001.

[8] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge: Cambridge University Press, 2003.

[9] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications – A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[10] S. Sugiura, S. Chen, and L. Hanzo, "A unified MIMO architecture subsming space shift keying, OSTBC, BLAST and LDC," in *Proc. VTC 2010-Fall* (Ottawa, Canada), Sep. 6-9, 2010, pp. 1–5.

[11] S. Sugiura, S. Chen, and L. Hanzo, "MIMO-aided near-capacity turbo transceivers: taxonomy and performance versus complexity," *IEEE Communications Surveys and Tutorials*, vol. 14, no. 2, pp. 421–442, 2012.

[12] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[13] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

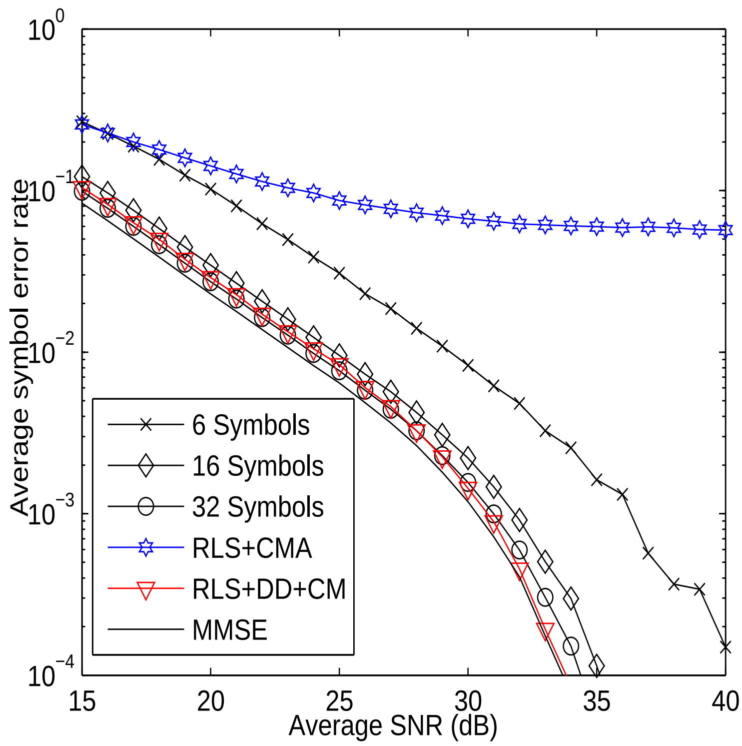

Figure 8: Desired user-one average SER performance of the proposed semi-blind RLS+DD+CM algorithm given $K = 6$ training symbols, the RLS+CMA given the same $K = 6$ training symbols, the training based beamforming given different number of training symbols, and the MMSE beamforming given perfect CSI, for the flat fading 5×4 8PSK beamforming system.

[14] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Technical J.*, vol. 20, pp. 11–22, 2015.

[15] S. Chen, N. N. Ahmad, and L. Hanzo, "Adaptive minimum bit error rate beamforming," *IEEE Trans. Wireless Communications*, vol. 4, no. 2, pp. 341–348, Mar. 2005.

[16] S. Chen, L. Hanzo, N. N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted receiver for QPSK wireless communication," *Digital Signal Processing*, vol. 15, no. 6, pp. 545–567, 2005.

[17] S. Chen, L. Hanzo, and A. Livingstone, "MBER Space-time decision feedback equalization assisted multiuser detection for multiple antenna aided SDMA systems," *IEEE Trans. Signal Processing*, vol. 54, no. 8, pp. 3090–3098, Aug. 2006

[18] S. Chen, "Adaptive beamforming assisted receiver," Chapter 3 in: C. Sun, J. Cheng and T. Ohira, Eds. *Handbook on Advancements in Smart Antenna Technologies for Wireless Networks*, Information Science Reference, 2008, pp. 68–92.

[19] X. Hong and S. Chen, "A minimum approximate-BER beamforming approach for PSK modulated wireless systems," *Int. J. Automation and Computing*, vol. 5, no. 3, pp. 284–289, 2008.

[20] B. R. Vojčić and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Communications*, vol. 46, no. 10, pp. 1346–1355, Oct. 1998.

[21] W. Yao, S. Chen, S. Tan, and L. Hanzo, "Minimum bit error rate multiuser transmission designs using particle swarm optimisation," *IEEE Trans. Wireless Communications*, vol. 8, no. 10, pp. 5012–5017, Oct. 2009.

[22] W. Yao, S. Chen, and L. Hanzo, "Generalised MBER-based vector precoding design for multiuser transmission," *IEEE Trans. Vehicular Technology*, vol. 60, no. 2, pp. 739–745, Feb. 2011.

[23] W. Yao, S. Chen, and L. Hanzo, "Particle swarm optimisation aided MIMO multiuser transmission designs," *J. Computational and Theoretical Nanoscience*, Special Issue on A New Frontier of Cognitive Informatics and Cognitive Computing, vol. 9, no. 2, pp. 266–275, 2012.

[24] X. Zhu, Z. Wang, C. Qian, L. Dai, J. Chen. S. Chen, and L. Hanzo, "Soft pilot reuse and multicell block diagonalization precoding for massive MIMO systems," *IEEE Trans. Vehicular Technology*, vol. 65, no. 5, pp. 3285–3298, May 2016.

[25] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: a study of estimator trade-

offs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, no. 3, pp. 884–893, Mar. 2006.

[26] B. Widrow, P. E. Mantey, L. J. Griffiths, and B. B. Goode, "Adaptive antenna systems," *Proc. IEEE*, vol. 55, no. 12, pp. 2143–2159, Dec. 1967.

[27] L. J. Griffiths, "A simple adaptive algorithm for real-time processing in antenna arrays," *Proc. IEEE*, vol. 57, no. 10, pp. 1696–1704, Oct. 1969.

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

[29] J. J. Shynk and R. P. Gooch, "The constant modulus array for cochannel signal copy and direction finding," *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 652–660, Mar. 1996.

[30] J. Sheinvald, "On blind beamforming for multiple non-Gaussian signals and the constant-modulus algorithm,"*IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1878-1885, Jul. 1998.

[31] K. Yang, T. Ohira, Y. Zhang, and C. Y. Chi, "Super-exponential blind adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1549–1563, Jun. 2004.

[32] E. de Carcalho and D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (global) identifiability conditions," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1053–1064, Apr. 2004.

[33] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Vehicular Technology*, vol. 56, no. 2, pp. 670–685, Mar. 2007.

[34] L. Tong, R. Liu, V. C. Soon, and Y.-F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits and Systems*, vol. 38, no. 5, pp. 499-509, May 1991.

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

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

[37] S. Chen, S. Sugiura, and L. Hanzo, "Semi-blind joint channel estimation and data detection for space-time shift keying systems," *IEEE Signal Processing Letters*, vol. 17, no. 12, pp. 993–996, Dec. 2010.

[38] P. Zhang, S. Chen, and L. Hanzo, "Reduced-complexity near-capacity joint channel estimation and three-stage turbo detection for coherent space-time shift keying," *IEEE Trans. Communications*, vol 61, no. 5, pp. 1902–1913, May 2013.

[39] P. Zhang, S. Chen, and L. Hanzo, "Embedded iterative semi-blind channel estimation for three-stage-concatenated MIMO-aided QAM turbo-transceivers," *IEEE Trans. Vehicular Technology*, vol. 63, no. 1, pp. 439–446, Jan. 2014.

[40] S. Chen, Y. Wang, and L. Hanzo, "Semi-blind adaptive beamforming for high-throughput quadrature amplitude modulation systems," *Int. J. Automation and Computing*, vol. 7, no. 4, pp. 565–570, 2010.

[41] Y. Chen, L. Tho, B. Champagne, and C. Xu, "Recursive least squares constant modulus algorithm for blind adaptive array," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1452–1456, May 2004.