INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING
Int. J. Adapt. Control Signal Process. 2010; **24**:467–476
Published online 7 August 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/acs.1137

Semi-blind fast equalization of QAM channels using concurrent gradient-Newton CMA and soft decision-directed scheme

S. Chen*, †

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

SUMMARY

This contribution considers semi-blind adaptive equalization for communication systems that employ high-throughput quadrature amplitude modulation signalling. A minimum number of training symbols, approximately equal to the dimension of the equalizer, are first utilized to provide a rough initial least-squares estimate of the equalizer's weight vector. A novel gradient-Newton concurrent constant modulus algorithm and soft decision-directed scheme are then applied to adapt the equalizer. The proposed semi-blind adaptive algorithm is capable of converging fast and accurately to the optimal minimum mean-square error equalization solution. Simulation results obtained demonstrate that the convergence speed of this semi-blind adaptive algorithm is close to that of the training-based recursive least-square algorithm. Copyright © 2009 John Wiley & Sons, Ltd.

Received 23 September 2008; Revised 30 April 2009; Accepted 16 June 2009

KEY WORDS: channel equalization; quadrature amplitude modulation; semi-blind adaptive algorithm; constant modulus algorithm; soft decision-directed adaptation; stochastic-gradient algorithm; gradient-Newton algorithm; minimum mean-square error

1. INTRODUCTION

In certain communication systems, training is infeasible and blind equalization provides a practical means for combating the detrimental effects of channel dispersion in such systems. Since no training sequence is needed, blind equalization improves system bandwidth efficiency. For systems with constant modulus (CM) signalling, the stochastic-gradient CM algorithm (CMA) [1–3] and block-based CMA [4, 5] offer popular low-complexity equalization schemes. Although the block-based CMA achieves better equalization performance than the stochastic-gradient one, the former introduces excessive processing delay that can cause a serious problem for delay-sensitive applications, such as voice communication. The performance and convergence behaviour of the CMA have been extensively studied [6–8]. Although the CMA is very robust to imperfect carrier recovery, it converges very slowly and introduces undesired phase rotation to the recovered symbol constellation. In addition, the steady-state mean-square error (MSE) achievable by the CMA may not be sufficiently low for the system to achieve an adequate symbol error rate (SER) performance. In particular, in order for the CMA to avoid undesired local minima, it is critical to have a proper initialization of the equaliser's weight vector.

*Correspondence to: S. Chen, School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K.
†E-mail: sqc@ecs.soton.ac.uk

Owing to the ever-increasing demand for wireless communication capacity, high-throughput quadrature amplitude modulation (QAM) schemes [9] have become popular in numerous wireless network standards. For example, the WiMax standard [10] adopts both the 16-QAM and 64-QAM schemes. Although QAM schemes are not CM signals, the stochastic-gradient CMA-based equalizer is shown to work for high-order QAM systems [11]. However, it is even more difficult to obtain a sufficiently low steady-state MSE so that the system achieves an adequate SER performance. A possible solution is to switch to a decision-directed (DD) least mean-square (LMS) adaptation after the convergence of the CMA [12], which should be able to achieve a performance close to the minimum MSE (MMSE) solution. As pointed out in [13], a successful switch to the DD adaptation requires that the CMA's steady-state MSE must be sufficiently small. In practice, such a low level of MSE may not be achievable by the CMA scheme. Alternatively, a block-based composite cost function, which consists of the CMA cost function and the decision-based MSE [14], can be adopted. But the weighting of the two component cost functions must be carefully chosen, which can be difficult to do.

An interesting solution to overcome the above-mentioned problem of the CMA-based blind equalization was suggested in [13]. Instead of switching to a DD adaptation after the CMA has converged, a DD equalizer operates in parallel with a CMA equalizer. The stochastic-gradient-based weight adaptation of the DD equalizer follows that of the CMA equalizer and, to avoid error propagation due to incorrect decisions, the DD adjustment only takes place if the CMA adaptation is deemed to have achieved a successful adjustment of the equalizer weight vector with a high probability. At a cost of slightly more than doubling the complexity of the simple CMA, this combined CMA and DD equalizer is capable of achieving a dramatic improvement in equalization performance over the CMA [13]. More recently, a novel combined CMA and soft DD (SDD) blind equalizer has been proposed [15–17], which achieves a faster convergence and has simpler implementation than the combined CMA and DD scheme of [13]. This stochastic-gradient combined CMA and SDD scheme operates a CMA-based equalizer and the last-stage SDD equalizer of [18] in a truly parallel manner. It is capable of achieving an equalization performance that is close to the MMSE equalization solution based on the perfect channel information and, thus offers a low-complexity high-performance technique for blind equalization of high-order QAM channels.

Blind equalization schemes typically converge much slower than training-based algorithms. For the high-order QAM channels, even the best stochastic-gradient CMA+SDD-based equalizer of [15–17] requires adaptation in tens of thousands of samples to converge. Additionally, blind equalization has an inherent decision ambiguity, which can only be resolved by other means. For example, differential encoding can be employed to resolve this decision ambiguity at the cost of worst-case 3 dB penalty. Alternatively, a few training symbols can be used to remove this decision ambiguity, which leads to semi-blind equalization. The novelty of this contribution is that we propose a gradient-Newton concurrent CMA and SDD scheme for fast semi-blind equalization. Training-based gradient-Newton-type algorithms [19] employ second-order statistics of the input signal to modify the stochastic gradient, which results in much faster convergence and smaller steady-state misadjustment than training-based stochastic-gradient-type algorithms at the cost of increased complexity. A training-based gradient-Newton minimum bit error rate algorithm, for example, is proposed in [20]. In our proposed semi-blind approach, a minimum number of training symbols, approximately equal to the dimension of the equalizer, are first utilized to provide a rough initial least-squares (LS) estimate of the equalizer's weight vector. A gradient-Newton concurrent CMA+SDD algorithm is then applied to adapt the equalizer. The proposed semi-blind adaptive algorithm is shown to be capable of converging fast and accurately to the optimal MMSE equalization solution.

It is worth emphasizing that a blind gradient-Newton CMA[‡] has been proposed for equalization of CM signals [21]. However, we have found that

[‡] We avoid using the RLS CMA in order not to cause any confusing with the training-based RLS algorithm.

this blind gradient-Newton CMA equalizer cannot achieve the equalization objective for high-order QAM systems. This is one of the motivations for considering semi-blind equalization of high-order QAM systems. With the aid of the information provided by a very short training symbol sequence, in our case in the form of LS initialization of the equalizer's weight vector, it becomes feasible to apply a blind gradient-Newton algorithm. Our simulation results demonstrate that the proposed semi-blind gradient-Newton CMA+SDD adaptive algorithm achieves a convergence speed close to that of the training-based recursive least-square (RLS) algorithm [22]. In terms of computational requirements, the stochastic-gradient CMA+SDD algorithm has a computational complexity of $\mathcal{O}(N_e)$ operations, where N_e is the order of equalizer, similar to that of the training-based LMS algorithm [22], while the gradient-Newton CMA+SDD algorithm has a computational complexity of $\mathcal{O}(N_e^2)$ operations, similar to that of the RLS algorithm [22]. It is also worth pointing out that an alternative semi-blind equalizer has been developed for CM signals [4, 23], which adopts a combined criterion of the block training-based MSE cost function and the block CMA cost function. Our approach is preferred as it leads to a recursive adaptive algorithm. Furthermore, the combined CMA+SDD adaptation is superior over the CMA adaptation for QAM systems.

The remainder of the paper is organized as follows. Section 2 briefly presents the system signal model and equalizer structure. The proposed semi-blind gradient-Newton CMA+SDD adaptive algorithm is detailed in Section 3. Simulation results are presented in Section 4, and the paper concludes in Section 5.

2. SIGNAL MODEL AND EQUALIZER STRUCTURE

Denote the symbol-rate channel impulse response (CIR) as

$$\mathbf{c}^{\mathrm{T}}=[c_0\ c_1\ \ldots\ c_{N_c-1}] \tag{1}$$

where N_c is the length of the CIR. The symbol-rate received signal sample $x(k)$ at receiver can be expressed as [24]

$$x(k)=\sum_{i=0}^{N_c-1} c_i s(k-i)+n(k) \tag{2}$$

where $n(k)$ is a complex-valued Gaussian white noise process with $E[|n(k)|^2]=2\sigma_n^2$, and $s(k)$ is the kth transmitted symbol with the symbol energy $E[|s(k)|^2]=\sigma_s^2$. The modulation scheme is assumed to be the M-QAM and therefore $s(k)$ takes the values from the symbol set

$$\mathcal{S}\triangleq\{s_{i,l}=u_i+ju_l,\ 1\leqslant i,l\leqslant\sqrt{M}\} \tag{3}$$

with the real-part symbol $\Re[s_{i,l}]=u_i=2i-\sqrt{M}-1$ and the imaginary-part symbol $\Im[s_{i,l}]=u_l=2l-\sqrt{M}-1$. We define the receive signal-to-noise ratio (SNR) as

$$\mathrm{SNR}\triangleq\frac{\mathbf{c}^H\mathbf{c}\,\sigma_s^2}{2\sigma_n^2} \tag{4}$$

The equalizer output, given by

$$y(k)=\sum_{i=0}^{N_e-1} w_i^* x(k-i) \tag{5}$$

is passed to the decision device to produce an estimate $\hat{s}(k-\tau)$ of the transmitted symbol $s(k-\tau)$, where N_e is the equalizer order, w_i are the complex-valued equalizer weights, and $0\leqslant\tau\leqslant\tau_{\max}$ is the equalizer's decision delay with $\tau_{\max}\triangleq N_e+N_c-2$.

Define the received signal vector

$$\mathbf{x}(k)=[x(k)x(k-1)\ldots x(k-N_e+1)]^{\mathrm{T}} \tag{6}$$

Then $\mathbf{x}(k)$ can be expressed by the following well-known signal model:

$$\mathbf{x}(k)=\mathbf{C}\ \mathbf{s}(k)+\mathbf{n}(k) \tag{7}$$

where the noise vector is

$$\mathbf{n}(k)=[n(k)n(k-1)\ldots n(k-N_e+1)]^{\mathrm{T}} \tag{8}$$

the transmitted symbol vector is

$$\mathbf{s}(k)=[s(k)s(k-1)\ldots s(k-N_e-N_c+2)]^{\mathrm{T}} \tag{9}$$

DOI: 10.1002/acs

and the $N_{\mathrm{e}}\times(\tau_{\max}+1)$ CIR convolution matrix has the Toeplitz form

$$\mathbf{C}=\begin{bmatrix}\mathbf{c}^{\mathrm{T}} & 0 & \dots & 0\\ 0 & \mathbf{c}^{\mathrm{T}} & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \dots & 0 & \mathbf{c}^{\mathrm{T}}\end{bmatrix}=[\mathbf{c}_0\mathbf{c}_1\cdots\mathbf{c}_{\tau_{\max}}] \tag{10}$$

Similarly, the equalizer can be expressed in the vector form

$$y(k)=\mathbf{w}^H\mathbf{x}(k) \tag{11}$$

where the equalizer's weight vector is defined by

$$\mathbf{w}=[w_0w_1\dots w_{N_{\mathrm{e}}-1}]^{\mathrm{T}} \tag{12}$$

With the perfect channel knowledge, the optimal MMSE solution that minimizes the MSE $J_{\mathrm{MSE}}(\mathbf{w})\triangleq E[|s(k-\tau)-y(k)|^2]$ is given by [22]

$$\mathbf{w}_{\mathrm{MMSE}}=\left(\mathbf{C}\mathbf{C}^H+\frac{2\sigma_n^2}{\sigma_s^2}\mathbf{I}_{N_{\mathrm{e}}}\right)^{-1}\mathbf{c}_\tau \tag{13}$$

where $\mathbf{I}_{N_{\mathrm{e}}}$ denotes the $N_{\mathrm{e}}\times N_{\mathrm{e}}$ dimensional identity matrix and $\mathbf{c}_\tau$ the $(\tau+1)$th column of $\mathbf{C}$.

For the equalizer with the weight vector $\mathbf{w}$ and an unknown decision delay τ, define the combined response of the equalizer and CIR as

$$\mathbf{f}^{\mathrm{T}}=[f_0f_1\dots f_{\tau_{\max}}]\triangleq\mathbf{w}^H\mathbf{C} \tag{14}$$

and let

$$i_{\max}=\arg\max_{0\leqslant i\leqslant\tau_{\max}}|f_i| \tag{15}$$

The equalizer's decision delay is in fact $\tau=i_{\max}$. In simulation, the quality of equalization can be judged using the maximum distortion (MD) measure defined by

$$\mathrm{MD}(\mathbf{w})\triangleq\left(\sum_{i=0}^{\tau_{\max}}|f_i|-|f_{i_{\max}}|\right)\Big/|f_{i_{\max}}| \tag{16}$$

Alternatively, the equalization performance can be assessed using the MSE criterion given by

$$J_{\mathrm{MSE}}(\mathbf{w})\triangleq\sigma_s^2\left((1-\mathbf{w}^H\mathbf{c}_\tau-\mathbf{w}^{\mathrm{T}}\mathbf{c}_\tau^*)+\mathbf{w}^H\left(\mathbf{C}\mathbf{C}^H+\frac{2\sigma_n^2}{\sigma_s^2}\mathbf{I}_{N_{\mathrm{e}}}\right)\mathbf{w}\right) \tag{17}$$

Ultimately, the SER can be simulated to assess the equalization performance.

3. THE PROPOSED SEMI-BLIND ALGORITHM

Assume that the number of available training symbols is K. We denote the available training data as

$$\mathbf{X}_K=[\mathbf{x}(1)\mathbf{x}(2)\dots\mathbf{x}(K)] \tag{18}$$

and

$$\bar{\mathbf{s}}_K=[s(1-\tau)s(2-\tau)\dots s(K-\tau)]^{\mathrm{T}} \tag{19}$$

The LS estimate of the equalizer's weight vector based on the training data $\{\mathbf{X}_K,\bar{\mathbf{s}}_K\}$ is readily given as

$$\mathbf{w}(0)=(\mathbf{X}_K\mathbf{X}_K^H)^{-1}\mathbf{X}_K\bar{\mathbf{s}}_K^* \tag{20}$$

In order to maintain throughput, the number of training pilots should be as small as possible. To ensure that $\mathbf{X}_K\mathbf{X}_K^H$ has a full rank, we will choose K slightly larger than N_{e}, the dimension of $\mathbf{x}(k)$. Because the training data with $K\approx N_{\mathrm{e}}$ are generally insufficient, the initial LS weight vector (20) may not be sufficiently accurate to open the eye. Therefore, decision direct adaptation is generally unsafe. We propose a gradient-Newton CMA+SDD blind scheme to adapt the equalizer (11) with $\mathbf{w}(0)$ of (20) as the initial weight vector, which is capable of converging fast and accurately to the MMSE equalization solution.

A gradient-Newton algorithm [19] uses the inverse of the autocorrelation matrix of $\mathbf{x}(k)$ to modify the stochastic gradient. In reality, the autocorrelation matrix of $\mathbf{x}(k)$ is unknown, and an exponentially weighted time averaging is used to approximate it. As in the training-based RLS algorithm, this inverse matrix can be updated recursively according to [22]

$$\mathbf{P}(k)=\lambda^{-1}\mathbf{P}(k-1)-\lambda^{-1}\mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k-1) \tag{21}$$

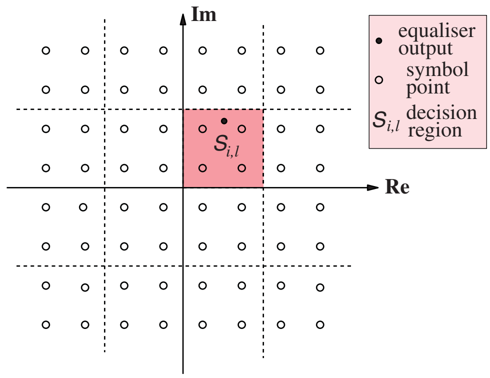

Figure 1. Local decision region partition for soft decision-directed adaptation with 64-QAM constellation.

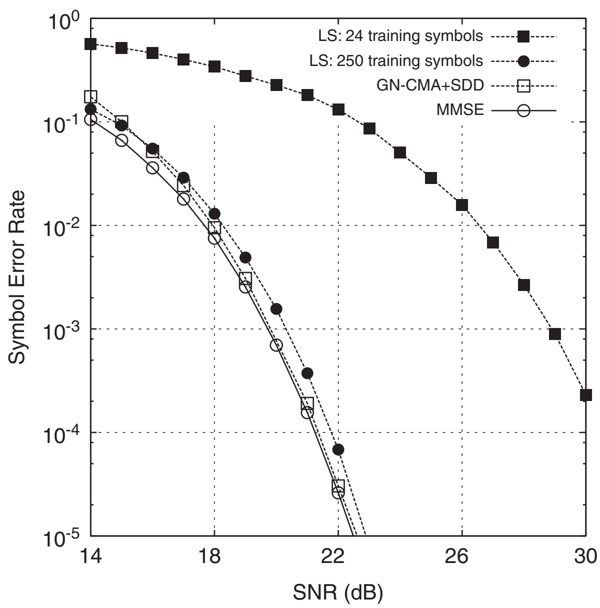

Figure 2. SER comparison of the three equalizers (the block LS equalizer with training symbols $K=24$ and 250, the semi-blind gradient-Newton CMA+SDD equalizer, and the MMSE equalizer) for the 5-tap 16-QAM example.

with

$$\mathbf{g}(k)=\frac{\lambda^{-1}\mathbf{P}(k-1)\mathbf{x}(k)}{1+\lambda^{-1}\mathbf{x}^H(k)\mathbf{P}(k-1)\mathbf{x}(k)} \tag{22}$$

where $\lambda\leqslant 1$ is the forgetting factor [22]. For stationary channels, $\lambda=1$ is appropriate. The initial $\mathbf{P}(0)$ can be set to $\mathbf{P}(0)=(\mathbf{X}_K\mathbf{X}_K^H)^{-1}$. Let the equalizer's weight vector be split into two parts, yielding $\mathbf{w}=\mathbf{w}_c+\mathbf{w}_d$. The initial $\mathbf{w}_c$ and $\mathbf{w}_d$ are simply set to $\mathbf{w}_c(0)=\mathbf{w}_d(0)=0.5\mathbf{w}(0)$. Denote the equalizer's output at sample k as $y(k)=\mathbf{w}^H(k)\mathbf{x}(k)$.

Table I. Algorithmic parameters of the gradient-Newton and stochastic-gradient CMA+SDD algorithms for the 5-tap 16-QAM example given SNR=21 dB.

Algorithm	μ_{CMA}	μ_{SDD}	ρ
Gradient-Newton	0.006	0.99	0.2
Stochastic-gradient	6×10^{-6}	10^{-3}	0.2

The weight vector $\mathbf{w}_c$ is updated using the gradient-Newton CMA according to

$$\mathbf{w}_c(k+1)=\mathbf{w}_c(k)+\mu_{\text{CMA}}\mathbf{P}(k)\varepsilon^*(k)\mathbf{x}(k) \tag{23}$$

with

$$\varepsilon(k)=y(k)(\Delta-|y(k)|^2) \tag{24}$$

where $\Delta=E[|s(k)|^4]/E[|s(k)|^2]$ and μ_{CMA} is the step size of the CMA. It is obvious that this gradient-Newton CMA algorithm reduces to the conventional stochastic-gradient CMA [1, 2] if $\mathbf{P}(k)$ is replaced with an identity matrix. Note that for semi-blind block-based CMA equalization of CM signals, an optimal step size can be found at the expense of considerably increased complexity [23] but this optimal step size is inapplicable here for the recursive gradient-Newton CMA. In practice, the step size μ_{CMA} in (23) can be chosen empirically to achieve a best performance in terms of convergence speed and steady-state misadjustment. Note that the step size of the gradient-Newton CMA algorithm can be set to a value much larger than the step size of the stochastic-gradient CMA counterpart.

The weight vector $\mathbf{w}_d$ is updated using the gradient-Newton SDD scheme, which is now described. The complex phasor plane is divided into the $M/4$ rectangular regions, and each region $\mathscr{S}_{i,l}$ contains four symbol points as defined in the following:

$$\mathscr{S}_{i,l}=\{s_{r,m},\ r=2i-1,2i, m=2l-1,2l\} \tag{25}$$

where $1\leqslant i,l\leqslant\sqrt{M}/2$. If the equalizer output $y(k)\in\mathscr{S}_{i,l}$, a local approximation of the marginal probability

DOI: 10.1002/acs

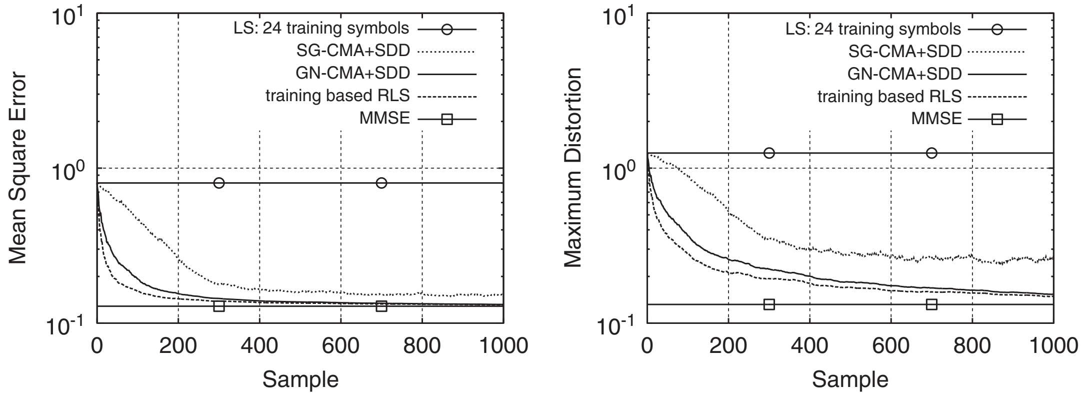

Figure 3. Convergence performance of the stochastic-gradient and gradient-Newton CMA+SDD algorithms as well as the training-based RLS algorithm, in terms of the MSE and MD measures averaged over 10 runs, for the 5-tap 16-QAM example given SNR = 21 dB.

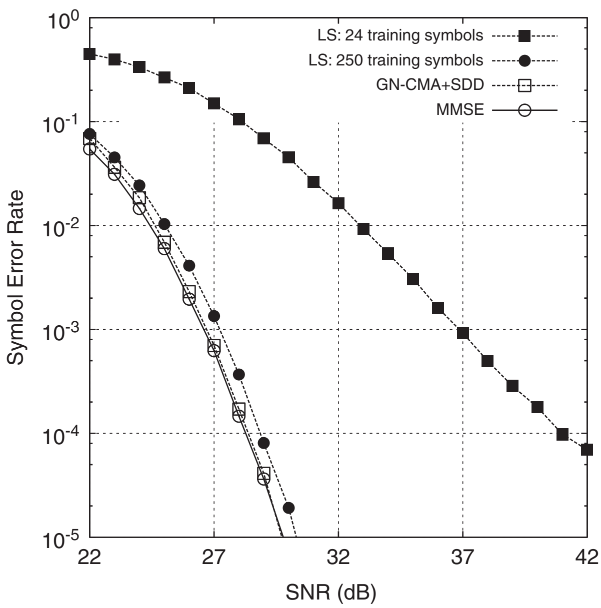

Figure 4. SER comparison of the three equalizers (the block LS equalizer with training symbols $K=24$ and 250, the semi-blind gradient-Newton CMA+SDD equalizer, and the MMSE equalizer) for the 5-tap 64-QAM example.

density function (PDF) of $y(k)$ is given by [15, 16]

$$\hat{p}(\mathbf{w}, y(k)) \approx \sum_{r=2i-1}^{2i} \sum_{m=2l-1}^{2l} \frac{1}{8\pi\rho} \mathrm{e}^{-(|y(k)-s_{r,m}|^2)/2\rho} \quad (26)$$

where ρ is the cluster width associated with the four clusters of each $\mathscr{S}_{i,l}$. An illustration of decision region partition is given in Figure 1. The stochastic-gradient SDD algorithm [15, 16] is designed to maximize the log of the local marginal PDF criterion $E[J_{\mathrm{LMAP}}(\mathbf{w},k)]$, where $J_{\mathrm{LMAP}}(\mathbf{w},k) = \rho \log(\hat{p}(\mathbf{w}, y(k)))$, via a stochastic gradient optimization. By contrast, the proposed gradient-Newton SDD algorithm uses $\mathbf{P}(k)$ to modify the stochastic gradient and updates $\mathbf{w}_d$ according to

$$\mathbf{w}_d(k+1) = \mathbf{w}_d(k) + \mu_{\mathrm{SDD}} \mathbf{P}(k) \frac{\partial J_{\mathrm{LMAP}}(\mathbf{w}(k), k)}{\partial \mathbf{w}_d} \quad (27)$$

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

$$\frac{\partial J_{\mathrm{LMAP}}(\mathbf{w}, k)}{\partial \mathbf{w}_d} = \frac{1}{Z_N} \sum_{r=2i-1}^{2i} \sum_{m=2l-1}^{2l} \mathrm{e}^{-(|y(k)-s_{r,m}|^2)/2\rho} \times (s_{r,m} - y(k))^* \mathbf{x}(k) \quad (28)$$

with the normalization factor

$$Z_N = \sum_{r=2i-1}^{2i} \sum_{m=2l-1}^{2l} \mathrm{e}^{-(|y(k)-s_{r,m}|^2)/2\rho} \quad (29)$$

This gradient-Newton SDD algorithm reduces to the stochastic-gradient SDD algorithm of [15, 16] by replacing $\mathbf{P}(k)$ with an identity matrix.

Soft decision nature can be seen explicitly in (28). Rather than committing to a single hard decision

DOI: 10.1002/acs

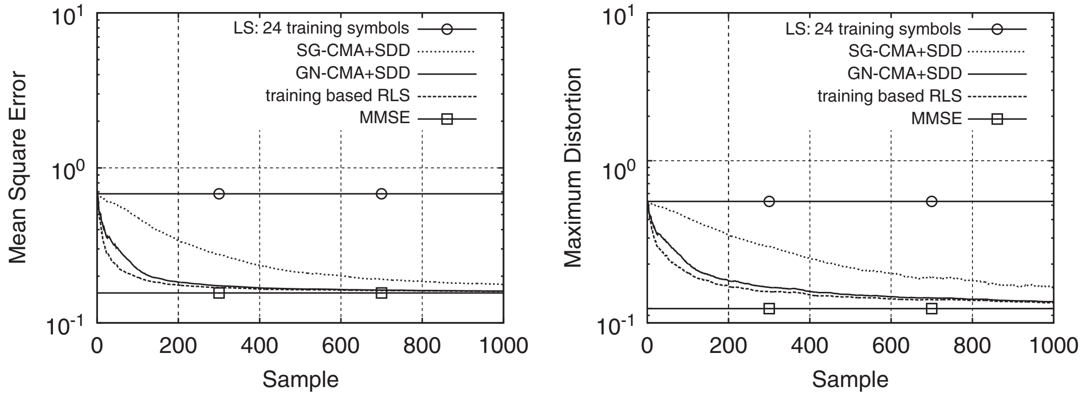

Figure 5. Convergence performance of the stochastic-gradient and gradient-Newton CMA+SDD algorithms as well as the training-based RLS algorithm, in terms of the MSE and MD measures averaged over 10 runs, for the 5-tap 64-QAM example given SNR = 27 dB.

Table II. Algorithmic parameters of the gradient-Newton and stochastic-gradient CMA+SDD algorithms for the 5-tap 64-QAM example given SNR = 27 dB.

Algorithm	μ_{CMA}	μ_{SDD}	ρ
Gradient-Newton	5×10^{-4}	0.99	0.1
Stochastic-gradient	10^{-7}	10^{-4}	0.1

$\mathcal{Q}[y(k)] = \hat{s}(k-\tau)$, where $\mathcal{Q}[\bullet]$ denotes the quantization operator, as the hard DD scheme would, alternative decisions are also considered in the local region $\mathcal{S}_{i,l}$ that includes $\mathcal{Q}[y(k)]$, and each tentative decision is weighted by an exponential term $e^{\{\bullet\}}$, which is a function of the distance between the equalizer's soft output $y(k)$ and the tentative decision $s_{r,m}$. This soft decision nature substantially reduces the risk of error propagation and achieves faster and more accurate convergence, compared with the hard DD scheme [15, 16]. The choice of the cluster width ρ, defined in the context of the local PDF (26), should ensure a proper separation of the four clusters of $\mathcal{S}_{i,l}$. As 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 too small a ρ value is used, 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 algorithm is insensitive to the value of ρ employed and an appropriate ρ can easily be chosen from a large range of values. More specifically, when the equalization objective is accomplished, $y(k) \approx s(k-\tau) + e(k)$, where $e(k)$ is Gaussian distributed with zero mean. Therefore, the value of ρ is related to the variance of $e(k)$, which is $2\sigma_n^2 \mathbf{w}^H \mathbf{w}$. Thus, for high SNR cases, small ρ may be desired, while for low SNR cases, large ρ may be preferred. The choice of μ_{SDD} for the gradient-Newton SDD algorithm is particularly simple, as it can be set to 1.0 or a value smaller than but close to 1.0.[§]

It is also clear that the proposed gradient-Newton CMA+SDD algorithm has a computational complexity of $\mathcal{O}(N_e^2)$ operations per update, similar to that of the training-based RLS algorithm, while the

[§]Because the SDD adaptation is more like the true training-based adaptation, the step size for the gradient-Newton SDD algorithm can be set to $\mu_{\text{SDD}} = 1.0$ just as in the case of the training-based RLS algorithm. But there is no reason against using a more cautious value, say, $\mu_{\text{SDD}} = 0.99$.

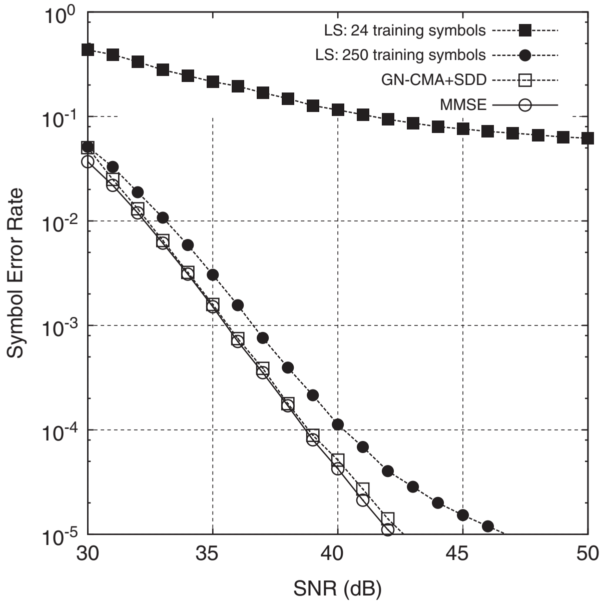

Figure 6. SER comparison of the three equalizers (the block LS equalizer with training symbols $K=24$ and 250, the semi-blind gradient-Newton CMA+SDD equalizer, and the MMSE equalizer) for the 5-tap 256-QAM example.

stochastic-gradient CMA+SDD algorithm has a simpler complexity of $\mathcal{O}(N_e)$ operations per update, similar to that of the training-based LMS algorithm.

4. SIMULATION STUDY

A simulated QAM communication system was considered, whose CIR was defined by

$$\mathbf{c}^{\mathrm{T}} = [(-0.2+j0.3)\ (-0.5+j0.4)\ (0.7-j0.6)\ (0.4+j0.3)\ (0.2+j0.1)] \tag{30}$$

This CIR had a length of $N_c=5$. The equalizer order was set to $N_e=23$, and the optimal decision delay was found empirically to be $\tau=17$. The proposed semi-blind gradient-Newton CMA+SDD algorithm was compared with the semi-blind stochastic-gradient CMA+SDD algorithm, using the block training-based LS estimator (20) and the training-based RLS algorithm as the benchmarks. Since the equalizer length was $N_e=23$, the initial training data length was chosen as $K=24$ for the two semi-blind adaptive equalizers. The step sizes μ_{CMA} and μ_{SDD} as well as the cluster width ρ for the stochastic-gradient CMA+SDD algorithm were chosen empirically to ensure a best performance in terms of convergence speed and steady-state misadjustment. For the gradient-Newton CMA+SDD algorithm, the step size μ_{CMA} and the cluster width ρ were chosen similarly, while the set size μ_{SDD} was simply set to the value of 0.99.

16-QAM Example. The modulation scheme was 16-QAM, and the SER performance of the optimal MMSE equalizer is depicted in Figure 2. The block training-based LS equalizer was next tested. Given the training data $\{\mathbf{X}_K, \bar{\mathbf{s}}_K\}$, the LS estimate of the equalizer weight vector was provided by (20), and the SER performance of the block training-based LS equalizer was also depicted in Figure 2, given $K=24$ and 250, respectively. It can be seen that $K=24$ was insufficient for the training-based equalizer to achieve an adequate SER performance and at least $K=250$ training symbols were required by the equalizer to approximate the optimal MMSE equalization solution.

The convergence performance of the proposed semi-blind gradient-Newton CMA+SDD algorithm was investigated, using the semi-blind stochastic-gradient CMA+SDD algorithm and the training-based RLS algorithm as the two benchmarks. Given SNR = 21 dB, appropriate algorithmic parameters found empirically for the two blind algorithms are given in Table I. The initial weight vector was obtained as the block LS estimate (20) with $K=24$ for all the three adaptive equalizers, and Figure 3 shows the learning curves of the three adaptive algorithms in terms of the MSE and MD measures, respectively, where the performance measures were averaged over 10 different random runs. It can be seen from Figure 3 that the convergence speed of the gradient-Newton CMA+SDD algorithm, which was observed to be close to the training-based RLS algorithm, was much faster than that of the stochastic-gradient CMA+SDD algorithm. The SER of the proposed semi-blind gradient-Newton CMA+SDD equalizer, illustrated in Figure 2, closely matched that of the optimal MMSE equalizer.

64-QAM Example. For the 64-QAM signalling, the SER performance of the optimal MMSE equalizer and

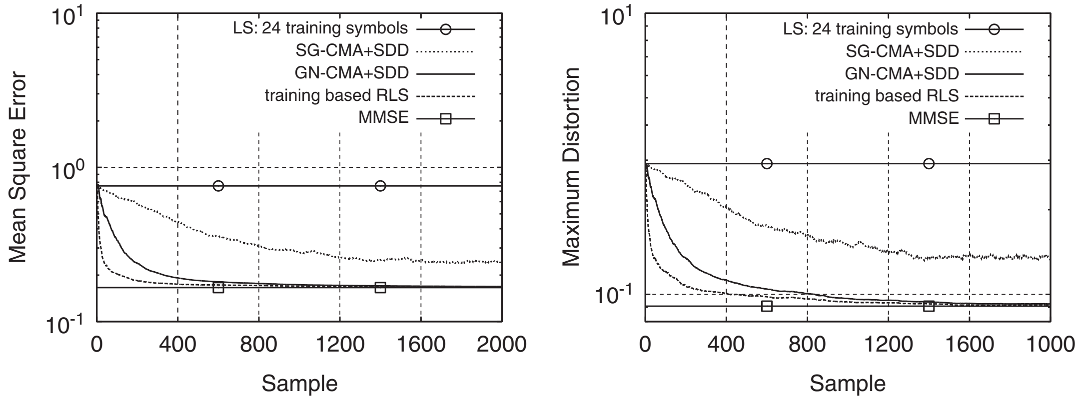

Figure 7. Convergence performance of the stochastic-gradient and gradient-Newton CMA+SDD algorithms as well as the training-based RLS algorithm, in terms of the MSE and MD measures averaged over 10 runs, for the 5-tap 256-QAM example given SNR = 36 dB.

Table III. Algorithmic parameters of the gradient-Newton and stochastic-gradient CMA+SDD algorithms for the 5-tap 256-QAM example given SNR = 36 dB.

Algorithm	μ_{CMA}	μ_{SDD}	ρ
Gradient-Newton	2×10^{-5}	0.99	0.05
Stochastic-gradient	10^{-8}	10^{-5}	0.05

the block training-based LS equalizer given $K=24$ and $K=250$, respectively, are compared in Figure 4. Again it is seen that $K=24$ was insufficient for the training-based equalizer to achieve an adequate SER performance and at least $K=250$ training symbols were required by the training-based equalizer to approximate the optimal MMSE equalization solution. The convergence performance of the semi-blind gradient-Newton and stochastic-gradient CMA+SDD algorithms as well as the training-based RLS algorithm are shown in Figure 5, given SNR = 27 dB and averaged over 10 runs, where the initial weight vector was obtained as the LS estimate (20) with $K=24$. The appropriate algorithmic parameters for the two semi-blind algorithms were found empirically and they are listed in Table II. It can be seen from Figure 5 that the convergence speed of the gradient-Newton CMA+SDD algorithm was very close to that of the training-based RLS algorithm, and it was much faster than that of the stochastic-gradient CMA+SDD algorithm. Accurate convergence of the proposed semi-blind gradient-Newton CMA+SDD equalizer to the optimal MMSE solution was also confirmed by its SER performance depicted Figure 4.

256-QAM Example. For this 256-QAM example, the SER performance of the optimal MMSE equalizer and the block training-based LS equalizer given $K=24$ and $K=250$, respectively, as well as the semi-blind gradient-Newton CMA+SDD equalizer are compared in Figure 6. The initial weight vector of the semi-blind gradient-Newton CMA+SDD algorithm was set to the LS estimate (20) with $K=24$. It can be seen from Figure 6 that the proposed semi-blind gradient-Newton CMA+SDD equalizer converged to the optimal MMSE solution accurately. Figure 7 depicts the learning curves in terms of the MSE and MD measures, respectively, for the training-based RLS algorithm, the semi-blind stochastic-gradient and gradient-Newton CMA+SDD algorithms given SNR = 36 dB, where the performance measures were

DOI: 10.1002/acs

averaged over 10 different random runs. The appropriate algorithmic parameters for the two semi-blind algorithms found empirically are listed in Table III.

5. CONCLUSIONS

Fast semi-blind adaptive equalization has been considered for communication systems that employ high-throughput QAM signalling. The scheme is semi-blind as it employs a minimum number of training symbols, approximately equal to the dimension of the equalizer, to provide a rough initial LS estimate of the equalizer's weight vector. This enables the utilization of a novel gradient-Newton concurrent CMA and SDD scheme to adapt the equalizer. The proposed semi-blind adaptive algorithm is capable of converging fast and accurately to the optimal MMSE equalization solution. Specifically, it has a computational complexity similar to that of the training-based RLS algorithm and the simulation results obtained have demonstrated that the convergence speed of this semi-blind adaptive algorithm is very close to that of the training-based RLS algorithm.

REFERENCES

1. Godard D. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications* 1980; **COM-28**:1867–1875.
2. Treichler JR, Agee BG. A new approach to multipath correction of constant modulus signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 1983; **ASSP-31**(2):459–472.
3. Treichler JR. Application of blind equalization techniques to voiceband and RF modems. *Preprints 4th IFAC International Symposium on Adaptive Systems in Control and Signal Processing*, France, 1992; 705–713.
4. Zarzoso V, Comon P. Blind and semi-blind equalization based on the constant power criterion. *IEEE Transactions on Signal Processing* 2005; **53**(11):4363–4375.
5. Zarzoso V, Comon P. Optimal step-size constant modulus algorithm. *IEEE Transactions on Communications* 2008; **56**(1):10–13.
6. Ding Z, Kennedy RA, Anderson BDO, Johnson Jr CR. Ill-convergence of Godard blind equalizers in data communication systems. *IEEE Transactions on Communications* 1991; **39**(9):1313–1327.
7. Ding Z, Johnson Jr CR, Kennedy RA. On the (non)existence of undesirable equilibria of Godard blind equalisers. *IEEE Transactions on Signal Processing* 1992; **40**(10):2425–2432.
8. Johnson Jr R, Schniter P, Endres TJ, Behm JD, Brown DR, Casas RA. Blind equalization using the constant modulus criterion: a review. *Proceedings of the IEEE* 1998; **86**(10):1927–1950.
9. Hanzo L, Ng SX, Keller T, Webb W. *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-coded, Turbo-equalised and Space–Time Coded OFDM, CDMA and MC-CDMA Systems*. Wiley: Chichester, U.K., 2004.
10. IEEE 802.16. Air Interface For Fixed Broadband Wireless Access System, Section 8: PHY, 2004.
11. Jablon NK. Joint blind equalization, carrier recovery, and timing recovery for high-order QAM signal constellations. *IEEE Transactions on Signal Processing* 1992; **40**(6): 1383–1398.
12. Macchi O, Eweda E. Convergence analysis of self-adaptive equalizers. *IEEE Transactions on Information Theory* 1984; **IT-3**(2):161–176.
13. De Castro FCC, De Castro MCF, Arantes DS. Concurrent blind deconvolution for channel equalization. *Proceedings of the ICC'2001*, Helsinki, Finland, vol. 2, 11–15 June 2001; 366–371.
14. Benveniste A, Goursat M. Blind equalizers. *IEEE Transactions on Communications* 1984; **COM-32**(8):871–883.
15. Chen S. Low complexity concurrent constant modulus algorithm and soft decision directed scheme for blind equalisation. *IEE Proceedings of the Vision, Image and Signal Processing* 2003; **150**(5):312–320.
16. Chen S, Chng ES. Concurrent constant modulus algorithm and soft decision directed scheme for fractionally-spaced blind equalization. *Proceedings of the ICC 2004*, Paris, France, vol. 4, 20–24 June 2004; 2342–2346.
17. Chen S, Wolfgang A, Hanzo L. Constant modulus algorithm aided soft decision directed scheme for blind space–time equalisation of SIMO channels. *Signal Processing* 2007; **87**(11):2587–2599.
18. Chen S, McLaughlin S, Grant PM, Mulgrew B. Multi-stage blind clustering equaliser. *IEEE Transactions on Communications* 1995; **43**(3):701–705.
19. Diniz PSR. *Adaptive Filtering: Algorithms and Practical Implementation* (2nd edn). Kluwer Academic Publishers: Boston, MA, 2003.
20. de Lamare RC, Sampaio-Neto R. Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels. *IEEE Communications Letters* 2003; **7**(2):73–75.
21. Doğançay K, Kennedy RA. Least squares approach to blind channel equalization. *IEEE Transactions on Communications* 1999; **47**(11):1678–1687.
22. Haykin S. *Adaptive Filter Theory* (3rd edn). Prentice-Hall: Upper Saddle River, NJ, 1996.
23. Zarzoso V, Comon P. Semi-blind constant modulus equalization with optimal step size. *Proceedings of the ICASSP 2005*, Philadelphia, PA, vol. 3, 18–23 March 2005; 577–580.
24. Proakis JG. *Digital Communications* (3rd edn). McGraw-Hill: New York, 1995.

DOI: 10.1002/acs