

High-Order-QAM Blind Equalization for Cognitive Communication Systems

Jin Li, Wei Xing Zheng, *Fellow, IEEE*, Mingqian Liu, *Member, IEEE*, Yunfei Chen, *Senior Member, IEEE*, Nan Zhao, *Senior Member, IEEE*,

Abstract—Blind equalization can not only counteract channel multi-path fading but also guarantee the transmission efficiency of a communication system. However, determining an optimal blind equalizer is typically NP hard. The cost function of such blind equalizer involves an exponentially increasing number of local minima. To reduce the complexity, this study develops an efficient two-stage blind equalization algorithm for quadrature-amplitude-modulation (QAM) systems. To reduce the number of local minima, the multimodulus algorithm (MMA) is implemented in the first stage, and to reduce the steady state error, an improved soft decision-directed algorithm (ISDDA) is implemented in the second stage. A novel modified least squares method (MLSM) is proposed to quickly search for the desired equalizer. The MLSM can converge to the invariance set of the MMA and ISDDA cost functions and has a quadratic termination property. In particular, theoretical analysis shows that the MLSM has a considerably lower computational load than the Newton methods by computing the constant Hessian matrix and its inverse only once. Furthermore, to ensure that the proposed algorithm switches to the second stage as early and suitably as possible, an attainable switching threshold between the two stages is provided on the basis of a hypothesis that the equalizer output error obeys a normal distribution, and the rationale of this hypothesis is also provided. Simulation results illustrate that the proposed algorithm has better convergence stability, superior equalization quality, and considerably faster convergence speed than the traditional stochastic-gradient-type dual-mode blind equalization algorithms.

Index Terms—Blind equalization; modified least squares method; quadrature amplitude modulation; improved soft decision-directed algorithm; switching threshold.

I. INTRODUCTION

Cognitive communication is the trend of future development [1]–[3]. In this system, if the duration of the transmitted symbol becomes smaller than the time dispersion of multipath

propagation, the received symbol will experience the inter-symbol interference (ISI) [4]–[6]. Channel equalizer is an effective approach to compensate such interference [7], [8]. As conventional channel equalizer requires the use of a pre-determined training sequence, it consumes precious spectrum resources. Therefore, it is imperative to reduce the need for training sequence. Specifically, in uncooperative communications without any suitable training sequence, training-based methods are not applicable for a channel equalizer. Owing to these drawbacks of training-based methods [9], [10], blind equalization (BE) techniques, without resorting to training sequences and channel state information, greatly reduces the system overhead [11]. Theoretically, BE can fully utilize the channel bandwidth. Therefore, similar to edge caching [12], BE increases the transmission efficiency in cooperative communications. In particular, because BE does not require any training sequence, it may be the only applicable method to overcome ISI in uncooperative communications. However, BE has drawbacks, such as slow convergence speed and high computational complexity, and it easily falls into a local minimum, resulting in poor equalization quality [13].

On the other hand, high-order QAM signals are frequently adopted in communication systems owing to its high spectral efficiency. However, applying classical BE to high-order QAM systems poses difficulties because of slow convergence rate, high computational complexity, and low equalization accuracy. Therefore, efficient BE for higher-order QAM systems need to be urgently designed.

In classical BE, the exhaustive search method (ESM) is typically used to yield a satisfactory or even a global optimal solution. With ESM, BE can theoretically achieve the same equalization accuracy as the training-based methods, when the same received samples are used by both methods. Unlike the training-based methods that use only the received samples from the training sequence, BE can exploit all the received samples, thereby yielding the best equalization performance. However, because of its high computational load, the ESM is not suitable for real-world communication applications. Thus, the core function of BE is to incorporate an efficient method to search for a satisfactory solution.

The most widely used BE techniques are the property restoral approaches, which can overcome classical BE combinatorial explosion problem; for example, the constant modulus algorithm (CMA) [14], [15], the multimodulus algorithm (MMA) [16], [17], and the soft decision-directed algorithm (SDDA) [18], [19]. The CMA and MMA are perhaps the most popular schemes for BE because of their simplicity and good

J. Li and M. Liu are with the State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China (e-mail: lijn342@163.com; mqliu@mail.xidian.edu.cn).

Wei Xing Zheng is with the Computer, Data and Mathematical Sciences, Western Sydney University, Sydney, NSW 2751, Australia (e-mail: w.zheng@westernsydney.edu.au).

Y. Chen is with Department of Engineering, University of Durham, Durham, UK DH1 3LE (e-mail: Yunfei.Chen@durham.ac.uk).

N. Zhao is with the School of Information and Communication Engineering, Dalian University of Technology, Dalian 116024, China (e-mail: zhaonan@dlut.edu.cn).

This work was supported by the National Natural Science Foundation of China under Grant U2441250, 62231027 and 62071364, Natural Science Basic Research Program of Shaanxi under Grant 2024JC-JCQN-63, the Guangxi Key Research and Development Program under Grant 2022AB46002 and Innovation Capability Support Program of Shaanxi under Grant 2024RS-CXTD-01.

convergence property [20], [21]. These methods reduce the possible number of combinations by merging discrete states of each transmitted code into less states. In the 4-QAM signal, irrespective of which state the signal is, its modulus is always fixed. If there are 1000 samples, the modulus of the transmitted signals has only one possible combination sequence. However, when the CMA and MMA are applied to higher-order constellation systems, its algorithmic performance can be degraded severely because they only use the partial information of constellation points to cause the well-known maladjustment problem [22], [23]. Moreover, the CMA and MMA are adaptive and use a step size significantly smaller than its permissible value, which results in a slower convergence speed. In contrast, decision-type algorithms, such as the SDDA, can noticeably decrease maladjustment in the steady state and improve the convergence speed to a certain extent. However, when the SDDA is applied to higher-order QAM systems, its cost function becomes highly nonconvex and it includes too many local minima, leading to poor convergence performance. In addition, its computational complexity increases as the order increases [24]. To overcome these drawbacks, some researches adopt two-stage algorithms (dual-mode-type algorithms); for example, the dual-mode generalized Sato algorithm (DMGSA) [22], the sinusoidal constellation-matched error minimization algorithm (SCMEMA) [25], the hybrid CMA and radius-directed algorithm [26] and two-stage method [28]. In most of these algorithms, the conventional CMA/MMA is used in the first stage to reduce the local minimum points and to ensure stable convergence. Once the error level in the first stage is reasonably low, a decision-directed algorithm is further utilized to improve the equalization performance. It is worth mentioning that the switching threshold is a crucial parameter for two-stage algorithms. If the switching to the decision-directed algorithm is performed too early, the algorithm may not converge, whereas if it is performed too late, the algorithm may converge slowly and incur a high computational cost. However, to the best of the authors' knowledge, these existing algorithms cannot provide a reasonable switching threshold. Hence, it is essential to give an attainable switching threshold for two-stage BE algorithms. Moreover, the gradient-descent method is used by these methods, which results in a slow convergence speed. Therefore, some Newton-like methods are proposed to celebrate the convergence speed, such as the Newton method-based CMA (CMA-NM) [27].

In this study, we design an efficient two-stage BE algorithm based on the MMA and the proposed improved SDDA (ISDDA), which separately implements soft decision to the real and imaginary parts of the equalizer output. The MMA based on a novel modified least squares method (MLSM) is implemented in the first stage to avoid phase rotation and improve convergence stability. In the second stage, an adaptively selected decision region [28] is adopted for soft decision when the error level is smaller than the given threshold; then the ISDDA based on the proposed MLSM is applied to enhance equalization performance. The primary contributions of this paper are as follows:

- A novel MLSM is proposed to efficiently optimize the

MMA and ISDDA cost functions (CFs). We prove that the MMA CF $J_{\text{MMA}}(\mathbf{w}_k)$ and ISDDA CF $J_{\text{ISDDA}}(\mathbf{w}_k)$ (where \mathbf{w}_k is the iteration sequence of the blind equalizer based on the proposed MLSM) are Lyapunov functions [29], [30] and that \mathbf{w}_k can converge to the invariance set of the MMA CF and ISDDA CF according to the LaSalle invariance principle [31]. Moreover, we demonstrate that the proposed MLSM has a quadratic termination property.

- A theoretical analysis shows that the proposed MLSM reduces the computational load considerably compared with the Newton methods because the MLSM's Hessian matrix and the corresponding inverse matrix remain constant in the iteration.
- We theoretically provide an attainable switching threshold with good error tolerance on the basis of a hypothesis that the error of the BE output is a normal distribution random variable. Moreover, we prove the rationality of this hypothesis and analyze the error tolerance of the switching threshold.
- The ISDDA is proposed for fine equalization. In the adaptively selected decision region, the SDDA is applied to the real and imaginary parts separately and the results are combined to yield the BE output. The equalization accuracy is improved by adaptively selecting the decision region and the computational load is reduced by separately applying the SDDA to the real and imaginary parts.

Finally, we perform simulation to illustrate that the proposed algorithm has favorable equalization accuracy, fast convergence speed, and robust convergence stability.

The rest of the paper is organized as follows. In Section II, the baseband channel equalization model is introduced. The proposed equalization algorithm comprising of the MLSM-based MMA and ISDDA is presented in Section III. In Section IV, simulations results are presented. Finally, conclusions are presented in Section V. Proofs of several key results are collected in appendixes.

Throughout the paper, scalar, vector, and matrix are denoted by plain lowercase, boldface lowercase, and boldface uppercase letters, respectively. $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^{-1}$ denote the complex conjugation, Hermitian matrix, and inverse matrix operators, respectively. Moreover, symbol $\|\cdot\|$ indicates the conventional 2-norm, $|\cdot|$ the magnitude, and $\mathcal{N}(u, \sigma^2)$ the normal distribution function with mean value u and variance σ^2 . Furthermore, $E[\cdot]$ represents the expectation operator, and \otimes is the discrete convolution. In addition, $\text{Re}[\cdot]$ and $\text{Im}[\cdot]$ denote the real and imaginary parts, respectively. Finally, $j = \sqrt{-1}$ is the imaginary unit.

In addition, to make the paper easier to follow, the important variables and notation are given in the table I. (Please the top of the next page.)

II. SYSTEM MODEL

As shown in Fig. 1, the cognitive communication system that experiences multipath fading is considered. The receiving terminal must compensate for signal distortion by channel equalization. An equivalent baseband block diagram is given at the bottom of Fig. 1. The sequence $a(n)$, which is assumed to

TABLE I
COMPUTATIONAL COMPLEXITY OF THE RELATED METHODS

Symbol	\mathbf{w}	$\mathbf{x}(n)$	J	∇	\mathbf{R}	σ_{th}
Meaning	equalizer	sample	cost function	gradient	correlation matrix	threshold

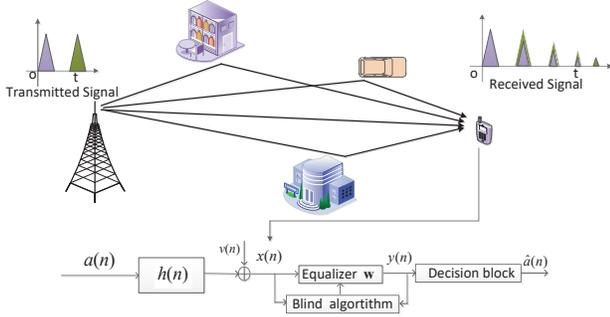


Fig. 1. Baseband model of the blind equalization system.

be QAM modulated and independently identically distributed (i.i.d), is sent through the linear system with impulse response $h(n)$. The QAM symbol takes the value from the set defined by

$$\Phi = \{a_1 + ja_2 \mid a_1, a_2 = \pm 1, \pm 3, \dots\}. \quad (1)$$

Assuming that the channel order is L , the input-output relation can be expressed as

$$x(n) = h(n) \otimes a(n) + v(n) = \sum_{l=0}^{L-1} h(l)a(n-l) + v(n), \quad (2)$$

where the sequence $v(n)$ is the complex additive white Gaussian white noise with mean 0 and variance $E[|v(n)|^2] = \sigma_n^2$.

As shown in Fig. 1, to recover the system input, BE is applied to the system output $x(n)$ to remove channel distortion without using any predetermined training data. If the impulse response of a linear blind equalizer with order \bar{L} is set as $\mathbf{w} = [w(0), w(1), \dots, w(\bar{L}-1)]^T$, then the BE output $y(n)$ satisfies

$$y(n) = \sum_{l=0}^{\bar{L}-1} w^*(l)x(n-l) = \mathbf{w}^H \mathbf{x}(n), \quad (3)$$

where $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-L+1)]^T$ is the sliding-window vector of the received signal. It is worth noting that \bar{L} is usually set to be larger than L to ensure the accurate recovery of the signal. Because the original sequence is unknown (except for its statistical property) to the receiver in BE, the signal recovered from BE will be subject to inherent delay and phase ambiguity. Therefore, the desirable recovered signal of BE is $y(n) \approx Ca(n-\tau)$, where τ is the time delay.

III. TWO-STAGE BLIND EQUALIZATION ALGORITHM BASED ON MODIFIED LEAST SQUARES METHOD

To overcome the problem of poor equalization accuracy as well as unstable and slow convergence of BE, we propose a two-stage BE algorithm based on the MLSM for high-order QAM systems. This can greatly improve the equalization

quality by avoiding maladjustment in the second stage. The MLSM is designed to accelerate convergence owing to its quadratic termination property.

A. MMA Based on MLSM (MMA-MLSM)

The CF of the MMA [16] is described as

$$J(\mathbf{w}) = E[(|\text{Re}(y(n))|^p - R)^2] + E[(|\text{Im}(y(n))|^p - R)^2], \quad (4)$$

where R is the dispersion constant defined by

$$R = \frac{E[|\text{Re}(a(n))|^{2p}]}{E[|\text{Re}(a(n))|^p]} = \frac{E[|\text{Im}(a(n))|^{2p}]}{E[|\text{Im}(a(n))|^p]}$$

and p is generally set to 2. Gradient-type algorithms can be used to optimize a blind equalizer. However, although such algorithms are extremely simple, they converge slowly. Here, we propose a novel MLSM to optimize a blind equalizer. To facilitate the design of the MLSM in the subsequent subsection, the constant p is set as 1 and the MMA CF is formulated as

$$J_{MMA}(\mathbf{w}) = E[(|\text{Re}(y(n))| - R)^2] + E[(|\text{Im}(y(n))| - R)^2]. \quad (5)$$

Remark 1: The actual modulus of the real or imaginary parts of the transmitted signals does not equal to R . Thus, according to (4), we know that, irrespective of the value of the parameter p , the CF given by (4) cannot be zero and there exists maladjustment even when the equalizer converges completely for high-order QAM signals. Moreover, comparing $p = 1$ and $p = 2$, the computational complexities of (5) and its corresponding gradient are slightly lower than those of (4). Furthermore, the size of $|\text{Re}(y(n))|$ or $|\text{Im}(y(n))|$ is similar to $\text{Re}(y(n))$ or $\text{Im}(y(n))$, respectively. Thus, (4) is similar to the mean squared error criterion. This indicates that (5) may lead to a better equalization performance than (4). Hence, when $p = 1$, a slightly better equalization accuracy is expected than when $p = 2$. Therefore, we will use (5) instead of (4).

We only use (5) for coarse equalization for a stable convergence property [20]. To improve the equalization accuracy, a second stage is needed. Now, using the time average to replace the ensemble average, the CF in (5) becomes

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= E[(|\text{Re}(y(n))| - R)^2] + E[(|\text{Im}(y(n))| - R)^2] \\ &= \frac{1}{N} \sum_{n=1}^N (|\text{Re}(y(n))| - R)^2 + (|\text{Im}(y(n))| - R)^2, \end{aligned} \quad (6)$$

where N is the number of available samples. Substituting (3) into (6) and differentiating $J_{MMA}(\mathbf{w})$ with respect to \mathbf{w} yields

the following gradient:

$$\begin{aligned}
& \nabla J_{MMA}(\mathbf{w}) \\
&= \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w} \\
&\quad - \frac{1}{N} \sum_{n=1}^N R \mathbf{x}(n) (\text{sign}(\text{Re}(y(n))) - j \text{sign}(\text{Im}(y(n)))) \\
&= \frac{1}{N} \mathbf{X} \mathbf{X}^H \mathbf{w} - \frac{1}{N} \mathbf{X} \mathbf{z},
\end{aligned} \tag{7}$$

where $\text{sign}(x)$ denotes the signum function, i.e., $\text{sign}(x) = 1$ for $x \geq 0$ and $\text{sign}(x) = -1$ for $x < 0$. Moreover, $\mathbf{X} = [\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(N)]$ is the sample matrix. The vector \mathbf{z} is defined as

$$\begin{aligned}
\mathbf{z} &= \mathbf{z}(\mathbf{y}(\mathbf{w})) \\
&= R[\text{sign}(\text{Re}(y(1))) - j \text{sign}(\text{Im}(y(1))), \\
&\quad \dots, \text{sign}(\text{Im}(y(N))) - j \text{sign}(\text{Im}(y(N)))]^T.
\end{aligned}$$

It can be seen that \mathbf{z} is the function of $\mathbf{y} = [y(1), y(2), \dots, y(N)]^T$ that changes with \mathbf{w} .

The Newton's methods for optimization let the linear approximation to the gradient of an objective function be equal to zero. Basically, Newton's methods belong to stationary-point algorithms. Similarly, if \mathbf{z} is replaced by $\mathbf{z}(\mathbf{y}(\mathbf{w}_k))$, then we get an alternative gradient approximation to (7). By letting such gradient approximation of (7) equal to $\mathbf{0}$, we have

$$\frac{1}{N} \mathbf{X} \mathbf{X}^H \mathbf{w} - \frac{1}{N} \mathbf{X} \mathbf{z}_k = \mathbf{0}, \tag{8}$$

where $\mathbf{z}_k = \mathbf{z}(\mathbf{y}(\mathbf{w}_k))$ and k denotes the iteration number. Let (8) and (7) to be $\mathbf{0}$ respectively, it is very difficult to solve for the optimal \mathbf{w} from (7) since \mathbf{z} changes with \mathbf{w} . By contrast, the optimal \mathbf{w} can be solved though (8) by iteration. Now, the approximate least square solutions of BE is

$$\mathbf{w} = (\mathbf{X} \mathbf{X}^H)^{-1} \mathbf{X} \mathbf{z}_k. \tag{9}$$

We refer to the above scheme as the MLSM. The formula is updated as follows:

$$\mathbf{w}_{k+1} = (\mathbf{X} \mathbf{X}^H)^{-1} \mathbf{X} \mathbf{z}_k = \mathbf{R}^{-1} \mathbf{X} \mathbf{z}_k, \tag{10}$$

where $\mathbf{R} = \mathbf{X} \mathbf{X}^H$.

Remark 2: The MMA-MLSM is a stationary-point algorithm; thus, it has a fast convergent speed like Newton's methods. However, the Hessian matrix and its inverse in Newton's methods vary with iterations and must be computed in each iteration, which results in high computational load. Moreover, the Hessian matrix is sometimes a singular or negative definite one, so that the Newton's methods sometimes are divergent. In contrast, the Hessian matrix $\frac{1}{N} \mathbf{X} \mathbf{X}^H$ in (8) is constant and positive definite, and such a Hessian matrix and its inverse only need to be computed once. Hence, compared with the conventional Newton method, the MMA-MLSM has enhanced stability, fast convergence speed, and reduced computational load.

In the following, we analyze the convergence of the proposed MMA-MLSM.

We use the Lyapunov function and LaSalle invariance principle [29], [31] to analyze the convergence property of

the MMA-MLSM. For the convenience of presentation, we first provide the following definitions.

Definition 1 (Lyapunov function): A function $f(\mathbf{x})$ ($\mathbf{x} \in \mathbb{C}^L$) is called a Lyapunov function with respect to a discrete sequence \mathbf{x}_k when it satisfies the following conditions. 1) $f(\mathbf{x})$ is continuous, 2) $f(\mathbf{x}_{k+1}) \leq f(\mathbf{x}_k)$ for all index k , and 3) the set $\Omega_\varepsilon = \{\mathbf{x} \mid f(\mathbf{x}) < \varepsilon\}$ is bounded for each finite ε .

Definition 2 (Invariance set [31]): If $\mathbf{x}_k \in \mathbb{C}^L$, then the set

$$\Omega = \{\mathbf{x}_k \mid f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) = 0\}$$

is called the invariance set associated with the sequence \mathbf{x}_k .

Based on the definitions of the Lyapunov function and invariance set, the LaSalle invariance principle is given as follows.

Lemma 1 (LaSalle invariance principle [31]): If $f(\mathbf{x}_k)$ is a Lyapunov function w.r.t. the discrete sequence $\mathbf{x}_k \in \mathbb{C}^L$, then \mathbf{x}_k will converge to the invariance set Ω given in Definition 2.

Based on the above theory, we derive the following proposition associated with the MMA-MLSM.

Proposition 1: The CF $J_{MMA}(\mathbf{w}_k)$ is a Lyapunov function w.r.t. the sequence \mathbf{w}_k and the sequence \mathbf{w}_k will converge to the invariance set

$$\tilde{\Omega} = \{\mathbf{w}_k \mid J_{MMA}(\mathbf{w}_{k+1}) - J_{MMA}(\mathbf{w}_k) = 0\}.$$

Proof: See Appendix A. ■

Proposition 1 indicates that the MMA-MLSM can converge to the global minima or local minima. Fortunately, a large number of experiments have shown that the MMA converges to the global minima with a high probability. Ultimately, the proposed MMA-MLSM possesses a stable convergence property.

The convergence stability of the MMA-MLSM is analyzed as above. In the following, the convergence speed of the MLSM is analyzed in terms of the quadratic termination property, which is defined as follows.

Definition 3 (Quadratic termination property [32]): Assume that $g(\mathbf{x})$ is a convex quadratic function, i.e.,

$$g(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} - \mathbf{b}^T \mathbf{x} + c \quad (\mathbf{x} \in \mathbb{R}^n) \tag{11}$$

and matrix \mathbf{A} is positive definite. A method is said to have a quadratic termination property if for every initial point \mathbf{x}_0 , it minimizes $g(\mathbf{x})$ in at most n iterations.

According to this definition, we have the following proposition.

Proposition 2: The proposed MLSM has a quadratic termination property similar to a Newton-type algorithm.

Proof: See Appendix B. ■

Because the single constant modulus used in the MMA does not equal to the multiple-modulus values of the output signals, the output error level in the steady state will be large even though the equalizer converges completely [22], which will affect the performance and convergence speed of the equalizer. The SDDA can improve the convergence speed and the equalization performance. However, SDDA can not converge stably and has a high computational load under the case of

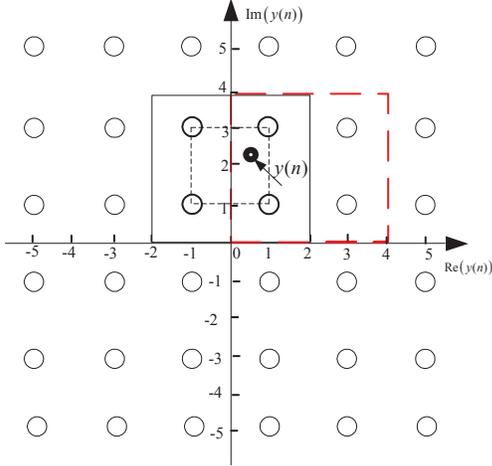


Fig. 2. Adaptive 4-nearest neighbor selection.

high-order signal. Hence, in the second stage, an adaptive 4-nearest-neighbor selection scheme (A4NS) is used for hard decision when the error level reaches the switching threshold. Then, the MLSM-based ISDDA is applied in the 4-nearest neighbors to reduce the computational complexity, improve the equalization performance, and increase the convergence speed.

B. A4NS and Switching Threshold Determination Method

1) A4NS: To illustrate the A4NS, 36-QAM signals are taken as an example. The A4NS of the SDDA depends on all constellations. As shown in Fig. 2, the circles with coordinates $a+bj$ ($a = \pm 1, \pm 3, \pm 5; |b = \pm 1, \pm 3, \pm 5$) are 36-QAM signals and the black annulus with coordinate $y(n)$ is the output signal of the blind equalizer. The fixed-decision region method [33] regards $y(n)$ as one of the four constellations located in the red dotted-line rectangle. A fixed-decision region is not suitable for the equalizer output $y(n)$. To obtain the optimal decision region, $y(n)$ is located in the black solid-line rectangle based on the adaptively selected decision region scheme [28], which is the nearest decision region according to the location of $y(n)$. We can conclude that the mean distance from $y(n)$ to the four constellations in the adaptively selected decision region is smaller than that in the fixed decision region. More importantly, the essential functions of decision are to give an accurate judgment of the equalizer output and then accelerate the convergence and improve the equalization accuracy. According to Fig. 2, if the distance between the real (imaginary) part of the equalizer output $y(n)$ and the real (imaginary) part of the transmitted signals $a(n - \tau)$ is less than 1, the fixed-decision region can give a right decision. In contrast, the A4NS relaxes the distance for giving a right decision to 2. This indicates that the A4NS has the following advantages: (a) If the A4NS is adopted, then BE can switch over to decision early, which accelerates the convergence of BE; (b) The A4NS improves the decision robustness by having a larger error tolerance area.

2) Switching Threshold Determination Method: To improve the equalization quality and accelerate the convergence, BE is switched to the decision-directed mode when the equalization error is small. The switching threshold is very important for such two-stage schemes. However, an attainable switching threshold has not been given in the existing literatures. In this paper, the switching threshold is given theoretically on the basis of a hypothesis that the equalizer output error obeys a normal distribution with a probability density function $\mathcal{N}(0, \sigma^2)$ rather than the residual ISI [26], which is unattainable since the channel state information is unknown in blind scenarios.

When the equalizer \mathbf{w} ideally converges to its optimal value $\tilde{\mathbf{w}}$, the corresponding equalizer output $\tilde{y}(n)$ approximates $a(n - \tau)$ and the CF attains the minimum value. To accelerate the convergence and avoid the steady state maladjustment, the proposed algorithm is switched to the decision-directed algorithm before the MMA CF achieves its minimum. The equalizer output $y(n)$ distributes in a region with $a(n - \tau)$ at the center. The larger the distance between $y(n)$ and $a(n - \tau)$, the smaller is the probability density function. To model this phenomenon, assume that the real (imaginary) part of the equalizer output error $C = \text{Re}(y(n)) - \text{Re}(a(n - \tau))$, with C and C' having the same distribution, is a normal random variable with the probability density function $\mathcal{N}(0, \sigma^2)$ and independent of the transmitted signals. Then, the MMA CF is given as

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= E[(|\text{Re}(y(n))| - R)^2] + E[(|\text{Im}(y(n))| - R)^2] \\ &= E[(|\text{Re}(a(n - \tau)) + C| - R)^2] \\ &\quad + E[(|\text{Im}(a(n - \tau)) + C'| - R)^2]. \end{aligned} \quad (12)$$

Because C and C' have the same distribution, (12) can be simplified as

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= E \left[(|\text{Re}(a(n - \tau)) + C| - R)^2 \right] \\ &\quad + E \left[(|\text{Im}(a(n - \tau)) + C'| - R)^2 \right] \\ &= 2E \left[(|\text{Re}(a(n - \tau)) + C| - R)^2 \right] \\ &= 2E \left[|\text{Re}(a(n))|^2 + 2C\text{Re}(a(n - \tau)) + C^2 \right] \\ &\quad - 4RE [|\text{Re}(a(n)) + C|] + 2R^2 \\ &= 2E \left[|\text{Re}(a(n))|^2 \right] + 4E[C] E[\text{Re}(a(n - \tau))] \\ &\quad + 2E[C^2] - 4RE [|\text{Re}(a(n)) + C|] + 2R^2 \end{aligned} \quad (13)$$

It is well known that $E[\text{Re}(a(n - \tau))] = 0$. Hence, (13) can be reduced to

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2E[C^2] \\ &\quad - 4RE [|\text{Re}(a(n)) + C|] + 2R^2 \\ &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2\sigma^2 \\ &\quad - 4RE [|\text{Re}(a(n)) + C|] + 2R^2 \end{aligned} \quad (14)$$

Moreover, we know that $\text{Re}(a(n))$ is a discrete random variable with probability $P\{\text{Re}(a(n)) = a\} = \frac{1}{\sqrt{Q}} (a = \pm 1, \pm 3, \dots)$ and C is a normal random variable with the probability density function

$\mathcal{N}(0, \sigma^2)$. $\text{Re}(a(n))$ and C are statistically independent. Then, the joint probability density function of $\text{Re}(a(n))$ and C is given as

$$f(\text{Re}(a(n)) = a, C = \bar{C}) = \frac{1}{\sqrt{Q}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{C}^2}{2\sigma^2}}. \quad (15)$$

Substituting (15) into (14) yields the following formula

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2\sigma^2 + 2R^2 \\ &\quad - 4R \frac{1}{\sqrt{Q}} \sum_{i=-\sqrt{Q}/2+1}^{\sqrt{Q}/2} \int_{-\infty}^{+\infty} |(2i-1) + \bar{C}| \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{C}^2}{2\sigma^2}} d\bar{C} \end{aligned} \quad (16)$$

Using the symmetry of the normal probability density function, (16) can be simplified to

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2\sigma^2 + 2R^2 \\ &\quad - 4R \times \frac{2}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \int_{-\infty}^{+\infty} |(2i-1) + \bar{C}| \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{C}^2}{2\sigma^2}} d\bar{C} \\ &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2\sigma^2 - \frac{16R}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(2i-1)^2}{2\sigma^2}} \\ &\quad - \frac{16R}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \int_{-(2i-1)}^0 (2i-1) \times \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\bar{C}^2}{2\sigma^2}} d\bar{C} + 2R^2 \end{aligned} \quad (17)$$

Let $\bar{C}' = \frac{\bar{C}}{\sigma}$. Then (17) is rewritten as

$$\begin{aligned} J_{MMA}(\mathbf{w}) &= 2E \left[|\text{Re}(a(n))|^2 \right] + 2\sigma^2 \\ &\quad - 4R \times \frac{4}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(2i-1)^2}{2\sigma^2}} + 2R^2 \\ &\quad - \frac{16R}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \int_0^{(2i-1)/\sigma} (2i-1) \times \frac{1}{\sqrt{2\pi}} e^{-\frac{\bar{C}'^2}{2}} d\bar{C}' \\ &= \tilde{J}_1(\sigma) \end{aligned} \quad (18)$$

Note that the value of $J_{MMA}(\mathbf{w})$ is translated into a function of variance σ under the assumption that the equalizer output error obeys $\mathcal{N}(0, \sigma^2)$. By common sense, the larger the error between \mathbf{w} and its optimal value $\tilde{\mathbf{w}}$, the larger is the σ value. Furthermore, the equalizer error corresponds to the MMA CF. Hence, if the variance σ increases, the value of $\tilde{J}_1(\sigma)$ increases. So we can conclude that the assumption that C is a normal random variable with the probability density function $\mathcal{N}(0, \sigma^2)$ is reasonable. It is well known that the first-order derivative $\tilde{J}'_1(\sigma) > 0$ implies that $\tilde{J}_1(\sigma)$ increases with increasing σ . In fact, in Appendix C, we will prove that $\tilde{J}'_1(\sigma) > 0$. Finally, we can conclude that the assumption that C is a normal random variable with the probability density function $\mathcal{N}(0, \sigma^2)$ is reasonable.

Now, the BE process is switched to the decision-directed mode when its weight vector becomes \mathbf{w}_{th} , which corresponds to the variance σ_{th} , i.e., $J_{MMA}(\mathbf{w}_{th}) = \tilde{J}_1(\sigma_{th})$. Further, the probability of the corresponding decision-making error should be less than α_{th} to ensure the following ISDDA-MLSM convergence. In other words, the probability $p\{|C| > 2\} = \alpha_{th}$, because a decision error is made when $|C| > 2$ is considered according to the A4NS. Assume that \bar{c}' is a standard normal

random variable with the probability distribution function $F(\bar{C}')$, defined as $F(\bar{C}') = p\{\bar{c}' \leq \bar{C}'\}$. If $\alpha_{th}/2 = F(\alpha)$, then $p\{|C| > 2\} = \alpha_{th}$ when $\sigma_{th} = \frac{-2}{\alpha}$. Finally, when the cost function value is

$$J_{MMA}(\mathbf{w}) \leq J_{MMA}(\mathbf{w}_{th}) = \tilde{J}_1(\sigma_{th}) = \tilde{J}_1\left(\frac{-2}{\alpha}\right),$$

the proposed method switches to the ISDDA-MLSM in order to improve the equalization accuracy and increase the convergence speed.

C. ISDDA Based on MLSM (ISDDA-MLSM)

For a 4-QAM system, the SDDA obtains BE by adjusting the weight vector \mathbf{w} to minimize the CF as follows [19]:

$$J_{SDDA}(\mathbf{w}) = E \left[-\ln \sum_{q=1}^4 \frac{\rho_q}{\sqrt{2\pi}\sigma_q} \exp\left[-\frac{1}{2\sigma_q^2} |y(n) - a_q|^2\right] \right], \quad (19)$$

where ρ_q is the priori probability with respect to $a_q \in \{1 + j, 1 - j, -1 + j, -1 - j\}$ and σ_q^2 is the variance associated with a_q . Generally, \mathbf{w} is updated according to the gradient descent method as follows:

$$\begin{cases} \mathbf{w}_{k+1} = \mathbf{w}_k + \mu e_k^* \mathbf{x}(k) \\ e_k = \frac{1}{Z_k} \sum_{q=1}^4 \exp\left(-\frac{1}{2\sigma_q^2} |y(k) - a_q|^2\right) \times (a_q - y(k)) \\ Z_k = \sum_{q=1}^4 \exp\left(-\frac{1}{2\sigma_q^2} |y(k) - a_q|^2\right) \end{cases} \quad (20)$$

, where μ is the step size and k the iteration index.

For high-order QAM signals, direct application of the SDDA may lead to poor convergence. Hence, we apply the SDDA to the output of BE based on the four constellations contained in the adaptively selected decision region when the MMA CF $J_{MMA}(\mathbf{w}) \leq \tilde{J}_1\left(\frac{-2}{\alpha}\right)$.

To reduce the computational load, we implement the soft decision to the real and imaginary parts of the equalizer output separately. Then, the improved CF can be formulated as

$$\begin{aligned} J_{ISDDA}(\mathbf{w}) &= E \left[-\ln \sum_{p=1}^2 \frac{\rho_p}{\sqrt{2\pi}\sigma_p} \exp\left(-\frac{1}{2\sigma_p^2} (\text{Re}(y(n)) - R_p)^2\right) \right. \\ &\quad \left. \times \sum_{q=1}^2 \frac{\rho_q}{\sqrt{2\pi}\sigma_q} \exp\left(-\frac{1}{2\sigma_q^2} (\text{Im}(y(n)) - I_q)^2\right) \right] \end{aligned} \quad (21)$$

, where R_p (for $p = 1, 2$) and I_q (for $q = 1, 2$) denote the real and imaginary parts of the constellation points in the decision region, respectively. Similarly, with the SDDA, the parameters ρ_p and ρ_q refer to the priori probabilities of $\text{Re}(a(n)) = R_p$ and $\text{Im}(a(n)) = I_q$, respectively. Since $a(n)$ is always i.i.d, the mathematical relation $\rho_p = \rho_q$ holds for all p and q . Moreover, the variance σ_p^2 (σ_q^2) decides the width of Gaussian function corresponding to R_p (I_q) and may considerably affect the equalization quality when it is extremely large or extremely small [34]. In other cases, the variance nearly has no effect to the equalization accuracy. Detailed discussion on the effect of the variance on the performance of the soft decision-directed

algorithm can be found in [34]. Similarly, it is reasonable to set $\sigma_p^2 = \sigma_q^2 = \sigma^2$ because $a(n)$ is i.i.d. Hence, the constants $\frac{\rho_p}{\sqrt{2\pi}\sigma_p}$ and $\frac{\rho_q}{\sqrt{2\pi}\sigma_q}$ are equal and can be neglected. Then, the ISDDA CF can be described as follows:

$$J_{ISDDA}(\mathbf{w}) = E \left[-\ln \sum_{p=1}^2 \exp \left[-\frac{1}{2\sigma^2} [\text{Re}(y(n)) - R_p]^2 \right] \right. \\ \left. \times \sum_{q=1}^2 \exp \left[-\frac{1}{2\sigma^2} [\text{Im}(y(n)) - I_q]^2 \right] \right], \quad (22)$$

which is equivalent to (19) in the physical sense. This implies that the good equalization performance of the SDDA is preserved. However, they differ in mathematical calculation. The CF given by (19) involves four complex multiplications, whereas that by (22) requires four real multiplications, with the computation of other aspects basically being the same. This indicates that the ISDDA has slightly lower computational complexity. Moreover, comparing (22) with the CF given by (3) and (4) in [35], the ISDDA has the following advantages: the proposed algorithm implements the soft decision to the real and imaginary parts separately and then combines the results, whereas the algorithm in [35] implements soft decision to the real or imaginary part individually. The proposed algorithm fully utilizes the comprehensive information of QAM constellation, whereas the algorithm given in [35] uses the information of the real and imaginary parts of QAM constellation separately, possibly causing additional estimation error.

Finally, by replacing statistical average with time average and neglecting the inconsequential constant $\frac{1}{N}$, the CF in (22) can be simplified as follows:

$$J_{ISDDA}(\mathbf{w}) = - \sum_{n=1}^N \ln \sum_{p=1}^2 \exp \left[-\frac{1}{2\sigma^2} [\text{Re}(y(n)) - R_p]^2 \right] \\ \times \sum_{q=1}^2 \exp \left[-\frac{1}{2\sigma^2} [\text{Im}(y(n)) - I_q]^2 \right] \quad (23)$$

. Now, differentiating (23) with respect to \mathbf{w} , we obtain the following gradient expression:

$$\nabla J_{ISDDA}(\mathbf{w}) = \sum_{n=1}^N \sum_{p=1}^2 \frac{1}{Z_r(n)} \frac{1}{2\sigma^2} f_{r,p}(n) [\text{Re}(y(n)) - R_p] \mathbf{x}(n) \\ - j \sum_{n=1}^N \sum_{q=1}^2 \frac{1}{Z_i(n)} \frac{1}{2\sigma^2} f_{i,q}(n) [\text{Im}(y(n)) - I_q] \mathbf{x}(n) \\ = \frac{1}{2\sigma^2} \sum_{n=1}^N \mathbf{x}(n) \mathbf{x}^H(n) \mathbf{w} \\ - \frac{1}{2\sigma^2} \sum_{n=1}^N \left(\sum_{p=1}^2 \frac{1}{Z_r(n)} f_{r,p}(n) R_p - j \sum_{q=1}^2 \frac{1}{Z_i(n)} f_{i,q}(n) I_q \right) \mathbf{x}(n) \quad (24)$$

where $f_{r,p}(n) = \exp \left[-\frac{1}{2\sigma^2} [\text{Re}(y(n)) - R_p]^2 \right]$, $f_{i,q}(n) = \exp \left[-\frac{1}{2\sigma^2} [\text{Im}(y(n)) - I_q]^2 \right]$, $Z_r(n) = \sum_{p=1}^2 f_{r,p}(n)$ and $Z_i(n) = \sum_{q=1}^2 f_{i,q}(n)$. Let

$f(n) = \sum_{p=0}^1 \frac{1}{Z_r(n)} f_{r,p}(n) r_p - j \sum_{q=0}^1 \frac{1}{Z_i(n)} f_{i,q}(n) r_q$ and $\bar{\mathbf{f}} = [f(1), f(2), \dots, f(N)]^T$. It follows from (3) that $\bar{\mathbf{f}}$ is the function of $\mathbf{y} = [y(1), y(2), \dots, y(N)]^T$, i.e., $\bar{\mathbf{f}} = \bar{\mathbf{f}}(\mathbf{y})$. If the inessential constant coefficient $\frac{1}{2\sigma^2}$ is neglected, then (24) can be simplified as

$$\nabla J_{ISDDA}(\mathbf{w}) = \mathbf{X}\mathbf{X}^H \mathbf{w} - \mathbf{X}\bar{\mathbf{f}}(\mathbf{y}). \quad (25)$$

Similarly, we use $\mathbf{y}_k = \mathbf{y}|_{\mathbf{w}=\mathbf{w}_k}$ to replace \mathbf{y} and consider it to be fixed. The updated formula of the ISDDA according to the MLSM is correspondingly designed as

$$\mathbf{w}_{k+1} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\bar{\mathbf{f}}(\mathbf{y}_k) = \mathbf{R}^{-1} \mathbf{X}\bar{\mathbf{f}}_k, \quad (26)$$

where $\bar{\mathbf{f}}_k = \bar{\mathbf{f}}|_{\mathbf{y}=\mathbf{y}_k}$.

Remark 3: We can see that the iteration formula of the proposed ISDDA-MLSM is similar to that of the MMA-MLSM. Hence, the ISDDA-MLSM also has a lower computational complexity because of the constant correlation matrix \mathbf{R} . We only need to compute $\bar{\mathbf{f}}_k$ and implement two multiplications of matrix and vector, i.e.,

$$\underline{\mathbf{f}} = \mathbf{X}\bar{\mathbf{f}}_k, \quad \mathbf{w}_{k+1} = \mathbf{R}^{-1} \underline{\mathbf{f}}.$$

Thereby, the computational load is significantly reduced.

In addition to the advantages mentioned above, the ISDDA-MLSM has a stable convergence property, based on which the following conclusion can be derived.

Proposition 3: The function $J_{ISDDA}(\mathbf{w}_k)$ is also a Lyapunov function and the equalizer weight vector \mathbf{w}_k converges to the invariance set

$$\widehat{\Omega} = \{\mathbf{w}_k | J_{ISDDA}(\mathbf{w}_{k+1}) - J_{ISDDA}(\mathbf{w}_k) = 0\}.$$

Proof: See Appendix D. ■

To sum up, the detailed rationale of the proposed TSBEA-MLSM is summarized in Algorithm 1, where the accuracy parameter ε is a sufficiently small positive constant and K is the maximum iteration times. (Please see the top of the next top.)

IV. NUMERICAL RESULTS AND DISCUSSION

Simulation is conducted to investigate the performance of the proposed TSBEA-MLSM, which is composed of the MMA-MLSM and ISDDA-MLSM as described in Section III, and compare it with the performance of the CMA, DMGSA, SCMEMA and CMA-NM in terms of equalization quality and convergence speed. The former is accessed by the symbol error rate (SER) and the mean squared error (MSE) between $y(n)$ and $a(n - \tau)$. The latter is measured by the residual ISI [26]. The MSE and ISI are defined, respectively, as

$$MSE = E[|Cy(n) - a(n - \tau)|^2] \quad (27)$$

$$ISI = \frac{\sum_{n=0, n \neq n_{\max}}^{L+\bar{L}-2} |\tilde{h}(n)|^2}{|\tilde{h}(n_{\max})|^2}, \quad (28)$$

where $\tau = \arg \max_{\tau} E[|y(n)a^H(n - \tau)|]$, $C = \frac{\sum_{n=1}^N a(n - \tau)y^*(n)}{\sum_{n=1}^N y(n)y^*(n)}$, $\tilde{h}(n)$ is the combined impulse response of

Algorithm 1: The Proposed TSBEA-MLSM

1: **Initialization of MMA:** We find the optimal solution by using the proposed TSBEA-MLSM. First, the MMA IS initialized. The \mathbf{w}_0 is initialized by unit center criteria, i.e. $\mathbf{w}_0 = [0, \dots, 1, \dots, 0]^H$, where the element 1 is in the center of the vector. The iteration index k is set to be 0, initial error e is set larger than then accuracy parameter $\tilde{J}_1(\sigma_{th})$.

2: **While** $e > \tilde{J}_1(\sigma_{th})$ and $k < K$ **do**

3: Calculate $y(n) n = 1, \dots, N$ by using $\mathbf{w} = \mathbf{w}_k$. Then, we update \mathbf{w}_{k+1} by using iteration formula (10).

4: Update iteration error $e = J(\mathbf{w}_k)$.

5: Update iteration index $k = k + 1$ and let $\mathbf{w}_k = \mathbf{w}_{k+1}$.

6: **End While**

7: **Return** \mathbf{w}_k

8: **Initialization of ISDDA:** The blind equalizer is initialized by $\mathbf{w}_0 = \mathbf{w}_k$. The iteration index k is set to be 0, initial error e is set larger than then accuracy parameter ε .

9: **While** $e > \eta$ and $k < K$ **do**

10: Calculate $y(n) n = 1, \dots, N$ by using $\mathbf{w} = \mathbf{w}_k$. Finally, we update \mathbf{w}_{k+1} by using iteration formula (26).

11: Update iteration error $e = \|\mathbf{w}_{k+1} - \mathbf{w}_k\|_2$.

12: Update iteration index $k = k + 1$ and let $\mathbf{w}_k = \mathbf{w}_{k+1}$.

13: **End While**

14: **Return** \mathbf{w}_k

$h(n)$ and the equalizer $w(n)$, $\tilde{h}(n) = h(n) \otimes w(n) = [\tilde{h}(0), \tilde{h}(1), \dots, \tilde{h}(\bar{L} + L - 2)]$, and $n_{\max} = \arg \max_n |\tilde{h}(n)|^2$.

Dispersive 16-QAM system is considered. The order of BE is chosen as 16, initialized with a unitary value in the center position and zeros elsewhere. The TSBEA-MLSM is switched to the ISDDA when

$$J_{MMA}(\mathbf{w}) \leq J_{MMA}(\mathbf{w}_{th}) = \tilde{J}_1(\sigma_{th} = \frac{-2}{\alpha})$$

and the switching time of the SCMEMA is chosen to be 30 iterations. The sample number N is set to be 1000 except for the third experiment. The variance σ^2 is suitably taken as 0.8. The signal-to-noise ratio (SNR) is defined as

$$J_{MMA}(\mathbf{w}) \leq J_{MMA}(\mathbf{w}_{th}) = \tilde{J}_1(\sigma_{th} = \frac{-2}{\alpha}).$$

Finally, the channel impulse response (CIR) $\mathbf{h}(n)$ is set to be $\mathbf{h}_1(n) = [1.0, 0.5 \exp(-j3/4\pi)]$ (which is the same as the simulation environments considered in [25]) in the first few experiments and then set to be $\mathbf{h}_2(n) = [0.005 - 0.004j, 0.1 + 0.003j, -0.24 - 0.104j, 0.854 + 0.520j, -0.218 + 0.273j, 0.19 - 0.14j, -0.06 + 0.020j]$ to measure the performances of the methods.

When the CIR is given by $\mathbf{h}_1(n)$, the switching threshold is $\sigma_{th} = 0.7765, 0.8597, 0.9216, 0.9739, 1.0204$, and the corresponding decision-making error is $\alpha_{th} = 1\%, 2\%, 3\%, 4\%, 5\%$. The MSE of the TSBEA-MLSM in the steady state is shown in Fig. 3 and Fig. 4 for $SNR = 12dB, 20dB, 28dB, 36dB, 44dB$. We can observe that: If the threshold is too large, i.e. the TSBEA-MLSM switched to the second stage (ISDDA) too early, the MSE and SER are too large, which indicates the BE fails. Moreover, We expect the method to turn to the second stage as early as possible to accelerate the convergence. According to Fig. 3 and Fig. 4, it is fairly reasonable to consider $\sigma_{th} = 0.7765$ as

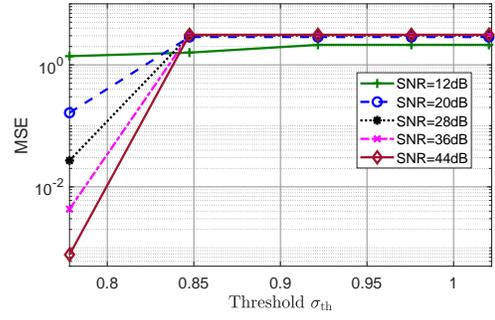


Fig. 3. MSE of the TSBEA-MLSM for different threshold σ_{th} values when $SNR = 12dB, 20dB, 28dB, 36dB, 44dB$ and CIR is $\mathbf{h}_1(n)$.

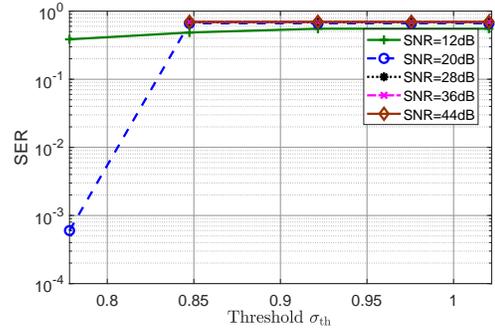


Fig. 4. SER of the TSBEA-MLSM for different threshold σ_{th} values when $SNR = 12dB, 20dB, 28dB, 36dB, 44dB$ and CIR is $\mathbf{h}_1(n)$.

the switching threshold and the corresponding function value $\tilde{J}_1(0.7765) = 3.2319$.

In the following, we first compare the performance of all related algorithms under the case of channel $\mathbf{h}_1(n)$.

A. Simulation results under the case of channel $\mathbf{h}_1(n)$.

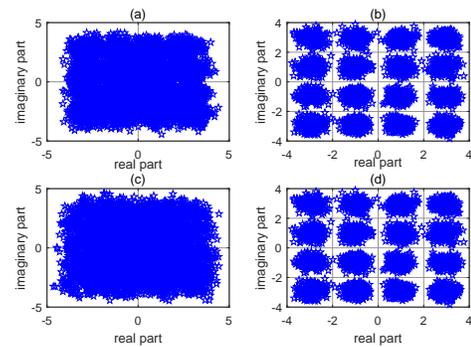


Fig. 5. Constellation diagrams: the equalizer output of (a) the first stage of the SCMEMA, (b) the second stage of the SCMEMA, (c) the first stage of the TSBEA-MLSM with $\sigma_{th} = 0.7765$, and (d) the second stage of the TSBEA-MLSM with $\sigma_{th} = 0.7765$, when $SNR = 20dB$ and CIR is $\mathbf{h}_1(n)$.

As shown in Fig. 5, both the SCMEMA and the proposed method avoid the local minimum points of the constellation-matched function of the SCMEMA and the ISDDA of the TSBEA-MLSM, respectively, by its equalization of the first stage. Further, blind equalizers converge well when the CIR is simply set as $\mathbf{h}_1(n)$.

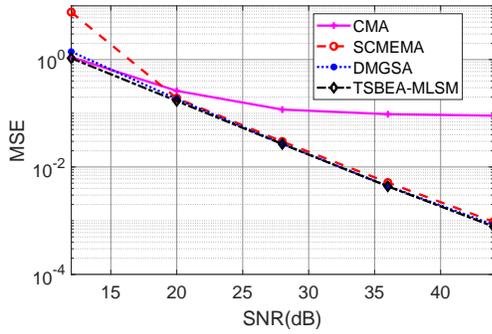


Fig. 6. Performance comparison of the TSBEA-MLSM, SCMEMA, DMGSA, and CMA in terms of MSE when CIR is $\mathbf{h}_1(n)$.

As shown in Fig. 6, when the SNR is large and the CIR is $\mathbf{h}_1(n)$, the MSE of the TSBEA-MLSM is almost equal to those of the SCMEMA and DMGSA, and they are considerably lower than that of the CMA. This superior performance of the two-stage methods is attributed to the constellation-matched error method in the second stage, which will avoid the maladjustment of the CMA (MMA) when the equalizer is nearly convergent.

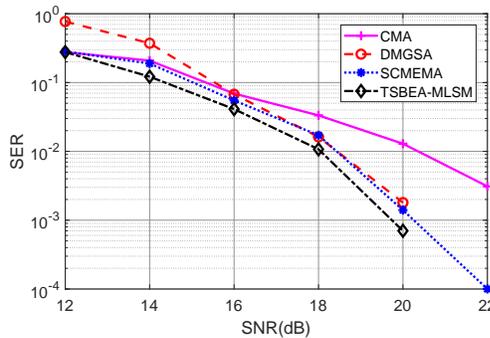


Fig. 7. Performance comparison of the TSBEA-MLSM, SCMEMA, DMGSA, and CMA in terms of SER when CIR is $\mathbf{h}_1(n)$.

Fig. 7 displays the SER curves with respect to the SNR of the four algorithms when the CIR is $\mathbf{h}_1(n)$. The figure shows that the TSBEA-MLSM, SCMEMA, and DMGSA have much better equalization performance than the CMA overall, because the constellation-matched error method eliminates the problem of steady state maladjustment. Moreover, the TSBEA-MLSM has a fairly smaller SER than the SCMEMA and DMGSA, because (i) the TSBEA-MLSM adopts the batch processing method, avoiding extreme errors of adaptive methods; and (ii) the adaptively selected decision region improves the reliability of the TSBEA-MLSM.

Fig. 8 displays the residual ISI of the TSBEA-MLSM, SCMEMA, and DMGSA versus iterations when the CIR is $\mathbf{h}_1(n)$. The number of iterations required to approach convergence is less than 20 for the TSBEA-MLSM, but approximately 4000 for the SCMEMA, approximately 12000 for the DMGSA, and 10000 for the CMA. This phenomenon indicates the proposed algorithm has a considerably faster convergence speed thanks to its quadratic termination property.

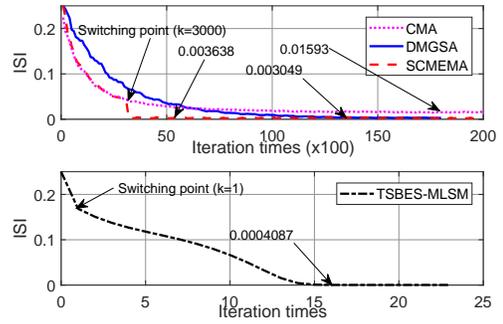


Fig. 8. Residual ISI of the TSBEA-MLSM, SCMEMA, and DMGSA with iterations when $SNR = 20dB$ and CIR is $\mathbf{h}_1(n)$.

In addition, the ISI value of the TSBEA-MLSM in the steady state is 0.0005058, the CMEMA is 0.003638, the DMGSA is 0.003049, and the CMA is 0.01593. This implies that the proposed algorithm has the best equalization accuracy owing to the batch processing technology and the adaptively selected decision region it adopts. Furthermore, the two-stage methods SCMEMA and TSBEA-MLSM have a switching point of $k = 3000$ (Which is same as the the reference [25]) and $k = 1$, respectively, as shown in Fig. 8. It can be seen that the SCMEMA already has a low ISI of approximately 0.04 when $k = 3000$. In this case, implementing a constellation-matched error method, like in the SCMEMA, can greatly accelerate the convergence as shown in Fig. 8. By contrast, the ISI is approximately 0.17 when $k = 1$ for the TSBEA-MLSM, which indicates that the TSBEA-MLSM switches to the second stage early and ISDDA-MLSM has a robust error tolerance. The acceleration of the convergence is not observed probably because the gradient of the MMA is large and the MMA-MLSM also has a fast convergence speed when $k = 1$ and the CIR is set as $\mathbf{h}_1(n)$.

Unless stated, in the following experiments, the CIR is set otherwise as $\mathbf{h}_2(n)$.

B. Simulation results under the case of channel $\mathbf{h}_2(n)$.

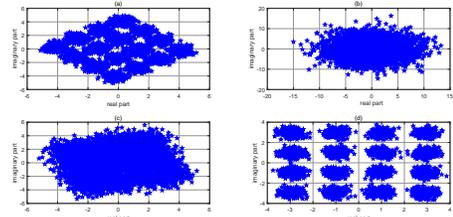


Fig. 9. Constellation diagrams: the equalizer output of (a) the first stage of the SCMEMA, (b) the second stage of the SCMEMA, (c) the first stage of the TSBEA-MLSM with $\sigma_{th} = 0.7765$, and (d) the second stage of TSBEA-MLSM with $\sigma_{th} = 0.7765$, when $SNR = 20dB$ and CIR is $\mathbf{h}_2(n)$.

Fig. 9(a) and (b) suggest the following observations. 1) The transmission signal profile has a phase rotation in subgraph Fig. 9(a), indicating that 3000 iterations of the CMA provide a good initial value to the SCMEMA. 2) The SCMEMA fails when the CIR is $\mathbf{h}_2(n)$. This is mainly due to the following

reasons. When the CIR is simply $\mathbf{h}_1(n)$, the equalizer output of the first stage of the SCMEMA does not revolve as a whole, whilst the CIR $\mathbf{h}_2(n)$ causes the phase rotation of the equalizer output signals; this is a significant feature of the CMA. However, the SCMEMA is constructed on the basis of the constellation coordinate. The real part (imaginary part) of these revolved constellation points are not located at the minimum of the sine function. Hence, the initialization is invalid, resulting in the failure of equalization. By contrast, the proposed algorithm converges well and has a good equalization performance, because BE with the MMA provides the output signals without rotation and the coordinates of these output constellation points are near the minimum of the cost function of the improved soft decision. Moreover, the proposed algorithm does not need an excellent initial value because of the robustness of the decision owing to the large error tolerance area of the A4NS. The constellation diagram shown in subgraph Fig. 9(c) is quite complex, indicating that the TSBEA-MLSM can be transferred to the second stage early and fast convergence can be realized, avoiding the maladjustment.

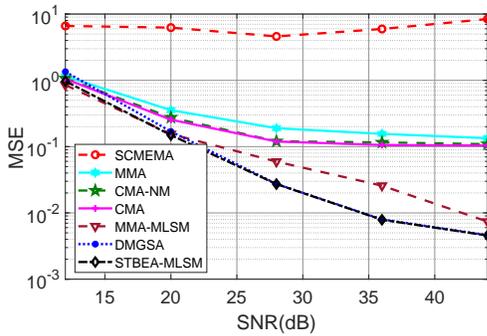


Fig. 10. Performance comparison of the TSBEA-MLSM, SCMEMA, DMGSA, CMA, CMA-NM, MMA and MMA-MLSM in terms of MSE when CIR is $\mathbf{h}_2(n)$.

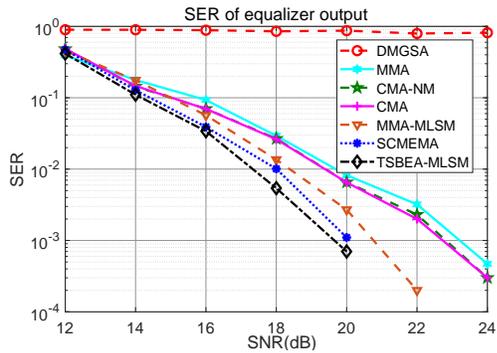


Fig. 11. Performance comparison of the TSBEA-MLSM, SCMEMA, DMGSA, CMA, CMA-NM, MMA and MMA-MLSM in terms of SER when CIR is $\mathbf{h}_2(n)$.

As shown in Fig. 10 and Fig. 11, the TSBEA-MLSM has the best equalization performance in terms of MSE and SER when the CIR is $\mathbf{h}_2(n)$. This is because the two-stage method avoids the maladjustment of the MMA, the batch-processing technology removes the inaccuracy of the adaptive methods,

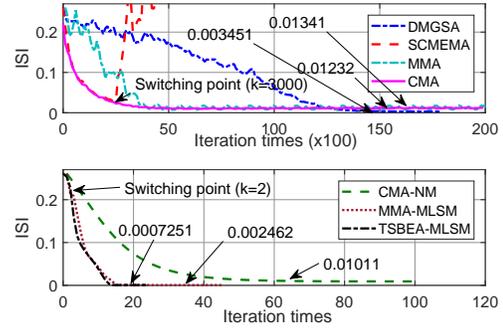


Fig. 12. ISI with iterations for the TSBEA-MLSM, DMGSA, SCMEMA, CMA, CMA-NM, MMA and MMA-MLSM when $SNR = 20dB$ and CIR is $\mathbf{h}_2(n)$.

and the adaptively selected decision region scheme improves the reliability of the TSBEA-MLSM. Note that the SCMEMA fails for the phase rotation of the equalizer output signals caused by the CMA. More spatially, the the MSE and SER OF CMA-NM is similar to that of MMA. The performance of MMA-MLSM is superior to MMA due to adoption of batch processing method. Moreover, the MSE and SER of TSBEA-MLSM is lower than that of MMA-MLSM since the two-stage method (ISDDA) avoids the maladjustment of the MMA.

As seen in Fig. 12, the SCMEMA diverges at the switching point for the phase rotation of the equalizer output signals caused by the CMA. DMGSA, CMA and MMA show a similar convergence performance as the case when the CIR is $\mathbf{h}_1(n)$. The CMA-NM has much faster convergence speed than the DMGSA and CMA due to adoption of Newton method. The proposed TSBEA-MLSM has the best steady-state performance and a considerably fast convergence speed (even faster than the CMA-NM) due to its quadratic termination property. Furthermore, the convergence speed of the TSBEA-MLSM increases after switching to the second stage of the TSBEA-MLSM when the CIR is $\mathbf{h}_2(n)$. Therefore the convergence speed of the TSBEA-MLSM is also faster than that of MMA-MLSM. Comparison of the ISI in the cases when the CIR is $\mathbf{h}_1(n)$ and $\mathbf{h}_2(n)$ indicates that the TSBEA-MLSM has a robust convergence stability owing to the good error tolerance of the ISDDA and the un-revolved output signals of the first stage.

V. CONCLUSION

In this paper, we have proposed a two-stage BE algorithm for cognitive communication systems. A novel MLSM has been developed to quickly search the optimal blind equalizer. The proposed MMA-MLSM and ISDDA-MLSM can converge stably. Furthermore, we have proposed a practical method to provide a switching threshold for the two-stage BE algorithm, which greatly improves the effectiveness and reliability of the algorithm. Compared with the other two-stage algorithms, the TSBEA-MLSM has superior equalization quality, faster convergence speed, and better convergence stability, and it also decreases the total computational load with the use of the constant correlation matrix.

APPENDIX A
PROOF OF PROPOSITION 1

According to the definition of the Lyapunov function, we can prove that the function $J_{MMA}(\mathbf{w}_k)$ is a Lyapunov function through the following three parts.

Part 1: $J_{MMA}(\mathbf{w}_k)$ is composed of the linear function ($y(n) = \mathbf{w}^H \mathbf{x}(n)$), absolute value function ($|\operatorname{Re}(y(n))|$), and quadratic function ($(|\operatorname{Re}(y(n))| - R)^2$), which are all basic functions. Consequently, the function $J_{MMA}(\mathbf{w}_k)$ is continuous.

Part 2: Based on the iteration formula (10), we can obtain the iteration direction of the proposed MMA-MLSM as

$$\mathbf{d}_{k+1} = (\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{z}_k - \mathbf{w}_k. \quad (\text{A.1})$$

The condition $J_{MMA}(\mathbf{w}_{k+1}) \leq J_{MMA}(\mathbf{w}_k)$ can be proved by the fact that \mathbf{d}_{k+1} is a descent direction. It is well known that $\nabla J_{MMA}^H(\mathbf{x}_k)(\mathbf{d}_{k+1}) < 0$ implies that \mathbf{d}_{k+1} is a descent direction when $\nabla J_{MMA}(\mathbf{x}_k) \neq 0$. Because \mathbf{R}^{-1} is a positive definite matrix, the following mathematical relation holds:

$$\begin{aligned} & \nabla J_1^H(\mathbf{w}_k) \mathbf{d}_{k+1} \\ &= \nabla J_1^H(\mathbf{w}_k) \left((\mathbf{X}\mathbf{X}^H)^{-1} \mathbf{X}\mathbf{z}_k - \mathbf{w}_k \right) \\ &= \nabla J_1^H(\mathbf{w}_k) (\mathbf{X}\mathbf{X}^H)^{-1} (\mathbf{X}\mathbf{z}_k - \mathbf{X}\mathbf{X}^H \mathbf{w}_k) \\ &= -N \nabla J_1^H(\mathbf{w}_k) \mathbf{R}^{-1} \nabla J_1^H(\mathbf{w}_k) < 0 \end{aligned} \quad (\text{A.2})$$

Hence, we can conclude that $J_{MMA}(\mathbf{w}_k + \lambda \mathbf{d}_{k+1}) \leq J_{MMA}(\mathbf{w}_k)$.

Part 3: For any finite constant ζ , the set $\{\mathbf{w}_k | J_{MMA}(\mathbf{w}_k) < \zeta\}$ is obviously bounded, because $J_{MMA}(\mathbf{w}_k)$ is unbounded when \mathbf{w}_k is unbounded.

Through the analysis of the above three parts, we prove that $J_{MMA}(\mathbf{w}_k)$ is a Lyapunov function. Therefore, according to the LaSalle invariance principle (Lemma 1), the discrete sequence \mathbf{w}_k will converge to the invariance set

$$\tilde{\Omega} = \{\mathbf{w}_k | J_{MMA}(\mathbf{w}_{k+1}) - J_{MMA}(\mathbf{w}_k) = 0\}.$$

This completes the proof of Proposition 1.

APPENDIX B
PROOF OF PROPOSITION 2

Differentiating $g(\mathbf{x})$ given in (11) with respect to \mathbf{w} yields the gradient

$$\nabla g(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}. \quad (\text{B.1})$$

According to the MLSM, let $\nabla g(\mathbf{x}) = 0$ and we have the MLSM iteration formula

$$\mathbf{x}_{k+1} = \mathbf{A}_k^{-1} \mathbf{b}_k = \mathbf{A}^{-1} \mathbf{b}. \quad (\text{B.2})$$

We notice that the proposed MLSM degenerate into a general least square method and it is obvious that \mathbf{x}_{k+1} is independent of \mathbf{x}_k . In other words, \mathbf{x}_{k+1} can be convergent with only one iteration, irrespective of the value of the initial point \mathbf{x}_0 . This completes the proof of Proposition 2.

APPENDIX C
PROOF OF CONDITION $\tilde{J}'_1(\sigma) > 0$

The derivative $\tilde{J}'_1(\sigma)$ is given by

$$\begin{aligned} & \tilde{J}'_1(\sigma) \\ &= 4\sigma - 4R \times \frac{4}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2i-1)^2}{2\sigma^2}} + \frac{\sigma}{\sqrt{2\pi}} \frac{(2i-1)^2}{\sigma^3} e^{-\frac{(2i-1)^2}{2\sigma^2}} \\ &+ 4R \times \frac{4}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{(2i-1)}{\sqrt{2\pi}} e^{-\frac{(2i-1)^2}{2\sigma^2}} \frac{(2i-1)}{\sigma^2} \\ &= 4\sigma - 4R \times \frac{4}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{1}{\sqrt{2\pi}} e^{-\frac{(2i-1)^2}{2\sigma^2}} \end{aligned} \quad (\text{C.1})$$

To prove the inequality $\tilde{J}'_1(\sigma) > 0$, we prove that $\frac{\tilde{J}'_1(\sigma)}{4\sigma} > 0$, i.e.,

$$1 - 4 \times \frac{R}{\sqrt{Q}} \sum_{i=1}^{\sqrt{Q}/2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2i-1)^2}{2\sigma^2}} > 0. \quad (\text{C.2})$$

Here, $R < \max\{|\operatorname{Re}(a(n))|\} = \sqrt{Q} - 1$, so $\frac{R}{\sqrt{Q}} < 1$. According to the amplification and minification method, if the inequality

$$\sum_{i=1}^{\sqrt{Q}/2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2i-1)^2}{2\sigma^2}} \leq 0.25 \quad (\text{C.3})$$

holds, then inequality (C.2) holds and so $\tilde{J}'_1(\sigma) > 0$ holds. The left-hand side expression of inequality (C.3) is the function of variance σ . Let function $f(\sigma)$ be

$$f(\sigma) = \sum_{i=1}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(2i-1)^2}{2\sigma^2}}. \quad (\text{C.4})$$

Now, the original problem is transformed into proving $f(\sigma) \leq 0.25$. Let $\frac{\sqrt{2\pi}\sigma}{2}$ be s . Then the function $f(\sigma)$ can be re-expressed as

$$f(\sigma) = f(s) = \frac{1}{2s} \sum_{i=1}^{\infty} e^{-\frac{\pi(i-1/2)^2}{s^2}} = \frac{1}{4s} \sum_{i=-\infty}^{\infty} e^{-\frac{\pi(i-1/2)^2}{s^2}}. \quad (\text{C.5})$$

According to Poisson's sum formula [31], we have

$$f(s) = \frac{1}{4s} \sum_{i=-\infty}^{\infty} e^{-\frac{\pi(i-1/2)^2}{s^2}} = \frac{1}{4s} \sum_{i=-\infty}^{\infty} e^{-j(\pi i)} e^{-\pi i^2 s^2} s. \quad (\text{C.6})$$

Simplifying the above formula yields

$$\begin{aligned} f(s) &= \frac{1}{4} \sum_{i=-\infty}^{\infty} e^{-j(\pi i)} e^{-\pi i^2 s^2} \\ &= \frac{1}{4} + \frac{1}{4} \times 2 \sum_{i=1}^{\infty} (-1)^i e^{-\pi i^2 s^2} \\ &= 0.25 - 0.5 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-\pi i^2 s^2} \\ &\leq 0.25 - 0.5(e^{-\pi s^2} - e^{-4\pi s^2}) < 0.25. \end{aligned} \quad (\text{C.7})$$

Finally, we can conclude that the maximum value of function $f(\sigma)$ ($f(s)$) is not larger than 0.25. According to the above analysis, formula (C.3) holds and then $\tilde{J}'_1(\sigma) > 0$. This complete the proof of $\tilde{J}'_1(\sigma) > 0$.

APPENDIX D
PROOF OF PROPOSITION 3

The proof of Proposition 3 is similar to that of Proposition 1. Here, we only give a simple proof of this proposition.

(i) Clearly, $J_{ISDDA}(\mathbf{w})$ is continuous, because function $J_{ISDDA}(\mathbf{w})$ is constructed by several basic functions.

(ii) Let $\widehat{\mathbf{d}}_{k+1} = (\mathbf{X}\mathbf{X}^H)^{-1}\mathbf{X}\widehat{\mathbf{r}}_k - \mathbf{w}_k$. Then we can obtain $J_{ISDDA}(\mathbf{w}_k + \lambda\widehat{\mathbf{d}}_{k+1}) \leq J_{ISDDA}(\mathbf{w}_k)$ ($\nabla J_{ISDDA}(\mathbf{x}_k) \neq 0$)

(where parameter λ is a small positive step size) and $J_{ISDDA}(\mathbf{w}_{k+1}) \leq J_{ISDDA}(\mathbf{w}_k)$ by the same method used in the proof given in Appendix A.

(iii) If \mathbf{w}_k is unbounded, then $J_{MMA}(\mathbf{w}_k)$ is unbounded. Hence, the set

$$\{\mathbf{w}_k \mid J_{ISDDA}(\mathbf{w}_k) < \xi\}$$

is bounded for all finite constant ξ .

Through the above analysis, it is known that $J_{ISDDA}(\mathbf{w}_k)$ is a Lyapunov function. Finally, we can conclude that the discrete sequence \mathbf{w}_k converges to the invariance set

$$\widehat{\Omega} = \{\mathbf{w}_k \mid J_{ISDDA}(\mathbf{w}_{k+1}) - J_{ISDDA}(\mathbf{w}_k) = 0\}$$

according to the LaSalle invariance principle (Lemma 1). This completes the proof of Proposition 3.

REFERENCES

- [1] S. Ghose, D. Mishra, S. P. Maity and G. C. Alexandropoulos, "Jointly Optimal RIS Placement and Power Allocation for Underlay D2D Communications: An Outage Probability Minimization Approach," *IEEE Transactions on Cognitive Communications and Networking*, vol. 10, no. 2, pp. 622-633, April 2024.
- [2] Q. Wu, S. Shi, Z. Wan, Q. Fan, P. Fan, and C. Zhang, "Towards V2I Age-Aware Fairness Access: A DQN Based Intelligent Vehicular Node Training and Test Method," *Chinese Journal of Electronics*, vol. 32, no. 6, pp. 1230-1244, November 2023.
- [3] M. Liu, Z. Zhang, Y. Chen, J. Ge and N. Zhao, "Adversarial Attack and Defense on Deep Learning for Air Transportation Communication Jamming," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 1, pp. 973-986, Jan. 2024.
- [4] A. Caciularu and D. Burshtein, "Unsupervised Linear and Nonlinear Channel Equalization and Decoding Using Variational Autoencoders," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 1003-1018, Sept. 2020.
- [5] K. Moshksar, "On a Class of Time-Varying Gaussian ISI Channels," *IEEE Transactions on Information Theory*, vol. 70, no. 2, pp. 1147-1166, Feb. 2024.
- [6] J. T. Wang, "Oversampling for Colored Noise ISI Channels," *IEEE Wireless Communications Letters*, vol. 13, no. 3, pp. 652-655, March 2024.
- [7] N. Gautam, S. V. Pendem, B. Lall and A. Choudhary, "Transformer-Based Nonlinear Equalization for DP-16QAM Coherent Optical Communication Systems," *IEEE Communications Letters*, vol. 28, no. 3, pp. 577-581, March 2024.
- [8] X. You, J. Chen and C. Yu, "Performance of Location-Based Equalization for OFDM Indoor Visible Light Communications," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 4, pp. 1229-1243, Dec. 2019.
- [9] L. Yang, Q. Zhao and Y. Jing, "Channel Equalization and Detection With ELM-Based Regressors for OFDM Systems," *IEEE Communications Letters*, vol. 24, no. 1, pp. 86-89, Jan. 2020.
- [10] L. M. Marrero, J. T. Gomez, F. Dressler and M. J. F. G. Garcla, "Adaptive Blind MPSK Constellation Recovery and Equalization for Cognitive Radio Applications," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 11, pp. 11988-12000, Nov. 2022.
- [11] J. Mao et al., "A GAN-Based Semantic Communication for Text Without CSI," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 14498-14514, Oct. 2024.
- [12] Q. Wu, W. Wang, P. Fan, Q. Fan, H. Zhu and K. B. Letaief, "Co-operative Edge Caching Based on Elastic Federated and Multi-Agent Deep Reinforcement Learning in Next-Generation Networks," in *IEEE Transactions on Network and Service Management*, vol. 21, no. 4, pp. 4179-4196, Aug. 2024.
- [13] B. Liu, C. Gong, J. Cheng, Z. Xu and J. Liu, "Blind and Semi-Blind Channel Estimation/Equalization for Poisson Channels in Optical Wireless Scattering Communication Systems," *IEEE Transactions on Wireless Communications*, vol. 21, no. 8, pp. 5930-5946, Aug. 2022.
- [14] J. Li, W. X. Zheng, M. Liu, Y. Chen and N. Zhao, "Robust Blind Equalization for NB-IoT Driven by QAM Signals," *IEEE Internet of Things Journal*, vol. 11, no. 12, pp. 21499-21512, June, 2024.
- [15] K. Fana, H. Hou, Y. Tang, P. Liu, "Orthogonality constrained analytic CMA for blind signal extraction improvement," *Signal Processing*, vol. 205, pp.1-8, Feb. 2023.
- [16] Q. Mayyalaa, K. Abed-Meraim, and A. Zerguine, "A class of multi-modulus blind deconvolution algorithms using hyperbolic and gives rotations for MIMO systems," *Signal Processing*, vol. 183, art. 107895, Jun. 2021.
- [17] S. Zhou, R. Chu, G. Liu, J. Zou and L. Sun, "Design and Implementation of High-Speed and Low-Complexity Blind Equalization Algorithm," *2023 8th International Conference on Signal and Image Processing (ICSIP)*, Wuxi, China, 2023, pp. 597-602.
- [18] C. Yin, W. Feng, J. Li, X. Bao and G. Li, "Soft-Input Soft-Output Block Decision Feedback Equalization for ISI Channels," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6213-6224, Sept. 2021.
- [19] S. L. Wang and H. Wen, "Soft Decision Adjusted Modulus Algorithm for Blind Equalization," *2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*, Chongqing, China, 2022, pp. 1881-1884.
- [20] Y. Chen, Y. Wang, K. Wang, L. Zhang and J. Yu, "FPGA-Based Demonstration of WDM-PAM4-PON System With Low-Complexity CMA," *IEEE Photonics Technology Letters*, vol. 35, no. 11, pp. 621-624, June, 2023.
- [21] Y. Chen, B. Dong, P. Gao and W. Xiong, "Fast MIMO Blind Detection via Modified MMA Approach Over the Stiefel Manifold," *IEEE Wireless Communications Letters*, vol. 13, no. 5, pp. 1483-1487, May, 2024.
- [22] C. P. Fan, C. H. Fang, H. J. Hu and W. N. Hsu, "Design and analyses of a fast feed-forward blind equalizer with two-stage generalized multilevel modulus and soft decision-directed scheme for high-order QAM cable downstream receivers," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2132-2140, November, 2010.
- [23] J. Li, D.-Z. Feng and W. X. Zheng, "A Robust Decision Directed Algorithm for Blind Equalization Under -Stable Noise," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4949-4960, 2021.
- [24] Z. Lv, B. Feng and L. Tan, "An Improved Dual-mode Blind Equalization Algorithm for QAM Signals," *2022 6th International Conference on Robotics and Automation Sciences (ICRAS)*, Wuhan, China, 2022, pp. 283-287.
- [25] L. He and M. Amin, "A dual-mode technique for improved blind equalization for QAM signals," *IEEE Signal Process. Lett.*, vol. 10, no. 2, pp. 29-31, Feb. 2003.
- [26] G. Di Rosa and A. Richter, "Likelihood-Based Selection Radius Directed Equalizer With Time-Multiplexed Pilot Symbols for Probabilistically Shaped QAM," *Journal of Lightwave Technology*, vol. 39, no. 19, pp. 6107-6119, Oct, 2021.
- [27] K. Kreutz-Delgado and Y. Isukapalli, "Use of the Newton Method for Blind Adaptive Equalization Based on the Constant Modulus Algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3983-3995, Aug. 2008.
- [28] H. J. Hu and C. P. Fan, "Two-stage generalized multilevel modulus and soft decision-directed based fast blind equalization for high-order QAM cable systems," *Proc. 2009 IEEE Region 10 Conf. (TENCON'2009)*, pp. 1-5, Singapore, Nov. 2009.
- [29] B. Chen, J. Cao, G. Lu and L. Rutkowski, "Lyapunov Functions for the Set Stability and the Synchronization of Boolean Control Networks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 11, pp. 2537-2541, Nov. 2020.
- [30] Q. Wu, W. Wang, P. Fan, Q. Fan, J. Wang and K. B. Letaief, "URLLC-Aware Resource Allocation for Heterogeneous Vehicular Edge Computing," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 8, pp. 11789-11805, Aug. 2024.
- [31] D. Gerbet and K. R?benack, "Application of LaSalles Invariance Principle on Polynomial Differential Equations Using Quantifier Elimination," *IEEE Transactions on Automatic Control*, vol. 67, no. 7, pp. 3590-3597, July, 2022.

- [32] L. Zhao, D. Wang and Y. Yang, "The quadratic property of the L-MBFGS methods for training neural networks," *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, Jilin, China, 2011, pp. 849-852.
- [33] S. Chen, W. Yao and L. Hanzo, "Semi-Blind Adaptive Spatial Equalization for MIMO Systems with High-Order QAM Signalling," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, pp. 4486-4491, November, 2008.
- [34] S. A. Alvarez, "Gaussian RBF Centered Kernel Alignment (CKA) in the Large-Bandwidth Limit," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 5, pp. 6587-6593, May, 2023.
- [35] R. Mitra, S. Singh, and A. Mishra, "Improved multi-stage clustering-based blind equalization," *IET Commun.*, vol. 5, no. 9, pp. 1255-1261, Jun. 2011.



Citation on deposit: Li, J., Zheng, W. X., Liu, M., Chen, Y., & Zhao, N. (in press). High-Order-QAM Blind Equalization for Cognitive Communication Systems. *IEEE Transactions on Cognitive Communications and Networking*

For final citation and metadata, visit Durham Research Online URL:

<https://durham-repository.worktribe.com/output/3314512>

Copyright statement: This accepted manuscript is licensed under the Creative Commons Attribution 4.0 licence.

<https://creativecommons.org/licenses/by/4.0/>