A regional multimodulus algorithm for blind equalization of QAM signals: introduction and steady-state analysis

João Mendes Filho, Maria D. Miranda, and Magno T. M. Silva
Escola Politécnica, University of São Paulo, Brazil

Abstract

It is well-known that constant-modulus-based algorithms present a large mean-square error for high-order quadrature amplitude modulation (QAM) signals, which may damage the switching to decision-directed-based algorithms. In this paper, we introduce a regional multimodulus algorithm for blind equalization of QAM signals that performs similarly to the supervised normalized least-mean-squares (NLMS) algorithm, independently of the QAM order. We find a theoretical relation between the coefficient vector of the proposed algorithm and the Wiener solution and also provide theoretical models for the steady-state excess mean-square error in a nonstationary environment. The proposed algorithm in conjunction with strategies to speed up its convergence and to avoid divergence can bypass the switching mechanism between the blind mode and the decision-directed mode.

Keywords: Adaptive signal processing, blind equalizers, gradient methods, numerical stability, quadrature amplitude modulation, tracking analysis.

1. Introduction

The main challenge of nowadays communications systems is to deliver high amounts of data within short time intervals, pursuing low symbol error rates (SER), and also considering different environment conditions. To accomplish such objective, many modulation schemes were proposed in the literature, as is the case of high-order quadrature amplitude modulation (QAM), which uses the available bandwidth in an efficient manner [3, 4, 5]. This efficient use of bandwidth can be improved with blind adaptive equalizers, which play an important role to remove the intersymbol interference introduced by dispersive channels.

In this context, the constant modulus algorithm (CMA) [6, 7] is the most popular for the blind adaptation of finite impulse response (FIR) equalizers due to its low computational cost. However, it has some drawbacks like the impossibility of solving phase ambiguities introduced by the channel and the possible convergence to undesirable local minima [8, 9]. Additionally, CMA can only achieve a zero steady-state mean-square error for constant modulus signals in a stationary and noiseless environment, and assuming a fractionally-spaced equalizer (FSE) [10]. Therefore, it presents a relatively large misadjustment when used in practical situations to recover nonconstant modulus signals, as is the case of high-order QAM signals (see, e.g., [11] and the references therein).

The multimodulus algorithm (MMA) [12, 13, 14] was proposed to solve the phase ambiguity issue through the minimization of the dispersion of the real and imaginary parts of the equalizer output separately. Although MMA provides better convergence for high-order QAM signals than that of CMA, it still exhibits a relatively large misadjustment in the steady-state since its updating error is zero only when the equalizer output is zero or when its magnitude is equal to the square root of the dispersion constant.

In order to reduce the misadjustment exhibited by CMA and MMA in the equalization of QAM signals, many algorithms have been proposed in the literature (see, e.g., [15, 16, 17, 18, 19, 20, 21, 22] and their references). In particular, some hybrid algorithms have attracted attention due to their relatively simplicity and good performance when compared to CMA or MMA. For instance, the hybrid algorithm of [23] combines CMA concurrently with the soft decision-directed algorithm and therefore, can be interpreted as...
a soft switching blind equalization scheme. The hybrid algorithms of [16, 17, 21] minimize the cost function of CMA (or MMA) added to a penalty term, often referred to as constellation matching error, that takes into account the coordinates of the constellation symbols. These algorithms may achieve a steady-state misadjustment lower than that of CMA and MMA for QAM signals, but which is still relatively large when compared to the misadjustment obtained in the equalization of constant modulus signals with MMA.

In general, to minimize the mean-square error (MSE) achieved by a blind algorithm, a switching mechanism to the decision-directed mode is considered in steady-state. So, the blind algorithm is used during the initial convergence and switched to a decision directed algorithm if an acceptable level of MSE is achieved. This “acceptable level” depends on the QAM order and also on the environment conditions, and may not always be achievable by the existing blind algorithms [24, 23]. Thus, the switching moment has a significant impact on the overall equalization performance, since a high MSE usually leads to an ill-convergence for the decision-directed algorithm [25, 26]. Therefore, many soft switching schemes were proposed in the literature (see, e.g., [27, 24, 23, 28, 29, 30] and their references). To avoid the switching mechanism, the blind algorithm must present a good transient and steady-state performance, independently of the QAM order.

In this paper, we propose a regional multimodulus algorithm (RMA) for blind equalization of low and high-order QAM signals that:

1. can recover simultaneously the modulus and phase of the transmitted signal;
2. has error equal to zero when the equalizer output coincides with the transmitted signal. Due to this property, it treats nonconstant modulus constellations as constant modulus ones, which allows the convergence in the mean to the Wiener solution and avoids the switching to the decision directed algorithm;
3. presents faster convergence than existing blind multimodulus-type algorithms for equalization of QAM signals; and
4. does not diverge if the non-consistent estimates of the transmitted signal are rejected.

With these properties, the proposed blind algorithm tends to perform similarly to a supervised one, independently of the QAM order and therefore, the switching mechanism between the blind mode and the decision-directed mode can be bypassed.

The paper is organized as follows. In the next section, we describe the problem and the general class of algorithms considered. We revisit the multimodulus algorithm and introduce the regional multimodulus algorithm in Sections 3 and 4, respectively. In Section 5, we find a theoretical relation between the coefficient vector of RMA and the Wiener solution. We also provide theoretical models for the steady-state mean-square error in a nonstationary environment. In passing, we should add that the model used in the tracking analysis extends the previous models in order to include the case of baud-rate blind equalization. The strategies to speed up the convergence and avoid divergence of RMA are shown in sections 6 and 7, respectively. In Section 8, we present simulation results in order to compare RMA with some existing algorithms in different situations. Additionally, we also present simulations to compare analytical and experimental results for the steady-state excess mean-square error. Finally, Section 9 provides the main conclusions of the paper.

2. Problem formulation

A simplified baseband communications system is depicted in Fig. 1 [3]. The signal $a(n)$, assumed i.i.d. (independent and identically distributed) and non Gaussian, is transmitted through an unknown channel, whose model is constituted by an FIR filter with impulse response vector $h = [h_0 \ h_1 \ \cdots \ h_{K-1}]^T$ and additive white Gaussian (AWGN) noise, where $(\cdot)^T$ indicates transposition. We assume an $M$-tap FIR equalizer, with input regressor vector $u(n)$ and output $y(n) = u^T(n)w(n-1) = y_a(n) + jy_b(n)$, where $w(n-1)$ is the equalizer weight vector and $y_a(n)$ and $y_b(n)$ are the real and imaginary parts of $y(n)$, respectively. The estimation error $e(n) = e_a(n) + j e_b(n)$ is used to update the equalizer coefficients. From the received signal $u(n)$ and the known statistical properties of the transmitted signal, the equalizer must mitigate the intersymbol interference introduced by the channel and recover a delayed version of the signal $a(n) = a_a(n) + j a_b(n)$, obtaining the estimate $\hat{a}(n - \Delta)$ at the output of the decision device, being $\Delta$ a positive integer.

We focus on a class of normalized blind equalization algorithms based on the stochastic gradient, whose update equation is given by

$$w(n) = w(n-1) + \frac{\mu}{\delta + \|u(n)\|^2} e(n)u^*(n),$$

(1)
where $\mu$ is the step-size, $\delta$ is the regularization factor (small positive constant), $\| \cdot \|$ represents the Euclidean norm, $(\cdot)^*$ stands for the complex-conjugate, and $e(n)$ is the estimation error, whose real and imaginary parts are computed separately as in MMA [12, 13, 14]. This class of algorithms presents three main advantages: i) the normalization improves the convergence rate and facilitates the choice of a proper step-size [31], ii) the definition of the estimation error in terms of its real and imaginary parts may avoid phase ambiguities, as in MMA [14, 32, 33], and iii) it is suitable for practical implementation due to its feasible computational cost.

We assume that the equalization algorithms are implemented in the baud rate or in the $T/2$-fractionally spaced form. The fractionally-spaced implementation is widely considered in the literature since it ensures perfect equalization in a noise-free environment, under certain well-known conditions (see, e.g., [34, 10] and the references therein). We also consider square QAM constellations throughout the paper. However, our approach can be extended straightforwardly to non-square QAM constellations.

3. Revisiting the multimodulus algorithm

The multimodulus algorithm minimizes the instantaneous cost function [12, 13, 14]

$$J_{\text{MMA}}(n) = [r_2 - y_2^*(n)]^2 + [r_2 - y_2^2(n)]^2,$$

where $r_2 = E\{a_2^*(n)/E\{a_2^2(n)\} = E\{a_2^*(n)/E\{a_2^2(n)\}$ for square QAM constellations and $E\{\cdot\}$ represents the expectation operator. Its estimation error is given by

$$e_{\text{MMA}}(n) = [r_2 - y_2^*(n)]y_2(n) + j[r_2 - y_2^2(n)]y_1(n).$$

The real part of this error, denoted by $e_{\text{MMA}, R}(n)$, is shown in Fig. 2 as a function of $y_2(n)$, assuming a constant (4-QAM) and a nonconstant (16-QAM) modulus signal (the figure for the imaginary counterpart is identical). For both signals, the real part of the MMA error is equal to zero when $y_2^*(n)$ is null or when $y_2^2(n)$ is equal to the dispersion constant $r_2$. For 4-QAM, $r_2 = 1$ and $e_{\text{MMA}, R}(n) = 0$ when the equalizer output is equal to the coordinates of the constellation symbols, i.e., $y_2(n) = \pm 1$. This behavior does not occur for nonconstant modulus signals. For 16-QAM, $r_2 = 8.2$ with $|e_{\text{MMA}, R}(n)| = 7.2$ and $e_{\text{MMA}, R}(n) = 2.4$ for $y_2(n) = \pm 1$ and $y_2(n) = \pm 3$, respectively. For 64-QAM, $|e_{\text{MMA}, R}(n)|$ assumes a value from the set $\{36, 60, 84\}$, when $y_2(n)$ is equal to one of the symbols coordinates $\{\pm 1, \pm 3, \pm 5, \pm 7\}$. Therefore, similarly to CMA, MMA exhibits a large steady-state mean-square error for nonconstant modulus signals.

Furthermore, the analysis of [33] shows that the MMA cost function presents additional stationary points related to phase rotations multiple of $\pi/2$ and around them it exhibits slow convergence, before converging to the desired minimum. However, this ill convergence occurs very rarely in practical implementations. Occasionally, MMA can also converge to some “offset solutions”, due to phase rotations multiple of $\pi/4$ [14]. These wrong solutions can be avoided through different techniques, as mentioned in [14, Sec. VIII].

4. The regional multimodulus algorithm

In order to treat an $S$-QAM signal ($S \geq 16$) as having constant modulus, inspired in the concurrent algorithm of [23], we divide the real line into $\sqrt{S}/2$ line segments $A_k$ with centers $c_k$, being $k = -\sqrt{S}/4, \ldots, -1, 1, \ldots, \sqrt{S}/4$, as shown in Fig. 3 for the real part of 64-QAM. Each line segment $A_k$, called hereafter as region $A_k$, contains two symbols coordinates denoted by $a_{k,1}$ and $a_{k,2}$, and indicated by asterisks in the figure. Given the equalizer output $y(n)$, with $(\log_2(S) - 2)$ comparisons, it is possible to identify the region $A_k$ to which $y_2(n)$ belongs and the region $A_m$ to which $y_1(n)$ belongs. The index $\ell$
Thus, replacing \( \sqrt{\alpha} \) identified regions. \( \alpha \) and \( \kappa \) transmitted. In other words, the real and imaginary parts of the equalizer output, i.e., we could use the MMA error with a unitary dispersion constant (\( r_a \)) by

\[
\alpha = \begin{cases} 
1 & \text{for } 4\text{-QAM}, \\
2 & \text{for } 16\text{-QAM}. 
\end{cases}
\]

Therefore, we can have the same value in different regions.

Figure 2: Real part of the MMA error as a function of \( y(n) \) for 4-QAM and 16-QAM. The errors at the constellation symbol coordinates are indicated by * (4-QAM) and o (16-QAM).

Table 1: Regions of the real part of 64-QAM; the center of the region \( A_k \) is represented by \( c_k, k = \pm 2, \pm 1 \).

Figure 3: Regions of the real part of 64-QAM; the center of the region \( A_k \) is represented by \( c_k, k = \pm 2, \pm 1 \).

\( (\text{resp., } m) \) is an integer used to specify the region, where the real (resp., imaginary) part of the equalizer output falls in.

In order to treat the symbol coordinates of the identified regions as pertaining to a constant modulus constellation, a translation operation is performed as follows. The centers of the identified regions are translated to the origin of the real line and the rest of the constellation is ignored. Thus, translated versions of the real and imaginary parts of the equalizer output are computed as

\[
\tilde{y}_r(n) = y_r(n) - c_r
\]

and

\[
\tilde{y}_m(n) = y_m(n) - c_m,
\]

where \( c_r \) and \( c_m \) are the centers of the identified regions. Due to the shift of the estimate \( y(n) \) to the origin of the complex plane, everything happens as if the symbols \( \{\pm 1 \pm j\} \) of a 4-QAM constellation had been transmitted. In other words, the real and imaginary parts of the equalizer output, i.e., \( y_r(n) \) and \( y_m(n) \), are identified to belong to regions \( A_r \) and \( A_m \) in a decision-directed manner, respectively. Once inside the identified regions, they are translated to the origin and are treated as being binary signals \( \pm 1 \). Therefore, we could use the MMA error with a unitary dispersion constant (\( r = 1 \)) and with the estimate of the “constant modulus symbol” at iteration \( n \) given by \( \tilde{y}_r(n) + j\tilde{y}_m(n) \). The resulting estimation error is a piecewise function with zeros at the coordinates of the S-QAM constellation symbols. However, as the decision directed algorithm, an algorithm updated with this error function is not able to “open the eye”, since the information of the positions of \( A_r \) and \( A_m \) is lost with the translation operation and the error can have the same value in different regions.

To recover the transmitted sequence, the cost function needs some information about the positions of \( A_r \) and \( A_m \) in the real line. For this purpose, we introduce the following regional cost function

\[
J_{\text{RMA}} = \mathbb{E} \left\{ \alpha_r \left[ 1 - \tilde{y}_r^2(n) \right]^2 \right\} + \mathbb{E} \left\{ \alpha_m \left[ 1 - \tilde{y}_m^2(n) \right]^2 \right\},
\]

where \( \alpha_r \) and \( \alpha_m \) are scale factors, which must carry statistical information about the positions of the identified regions.

Using a procedure similar to that of [20, 22], we now derive the optimum values for the scale factors \( \alpha_r \) and \( \alpha_m \) such that the perfect equalizer is the minimum of the cost function \( J_{\text{RMA}} \) in the absence of noise. Thus, replacing \( \sqrt{\alpha_r} \tilde{y}_r^2(n) \) by \( \kappa_r \sigma^2 r \) in (6) with \( p = 1, 2 \) and \( \kappa_r \) being a constant and calculating

\[
\frac{\partial J_{\text{RMA}}}{\partial \kappa_r} \bigg|_{\kappa_r = 1} = 0,
\]

...
we obtain the optimum value for $\alpha$, i.e.,

$$\alpha_{\ell, o} = \left[ \frac{\mathbb{E}\{a_{\ell,p}^4\}}{\mathbb{E}\{a_{\ell,p}^2\}} \right]^2 = \left[ \frac{a_{\ell,1}^4 + a_{\ell,2}^4}{a_{\ell,1}^2 + a_{\ell,2}^2} \right]^2. \quad (7)$$

We can observe from (7) that it is as if the real part of the QAM constellation symbols had been reduced to $a_{\ell,1}$ and $a_{\ell,2}$, which are the symbol coordinates of the region $A_\ell$ in which $y_\ell(n)$ falls in. To ensure perfect equalization, $\alpha_{\ell, o}$ must be the squared value of the dispersion constant of MMA, assuming the transmission of only two symbols, i.e., $a_{\ell,1}$ and $a_{\ell,2}$. Analogously for the imaginary part, $\alpha_{m, o}$ can be obtained by replacing $\ell$ by $m$ in (7).

Considering 64-QAM, for example, we obtain $\alpha_{-1, o} = \alpha_{1, o} = 2.86^2$ and $\alpha_{-2, o} = \alpha_{2, o} = 6.39^2$. Thus, the farther the region $A_k$ is from the origin, the closer the value of $\alpha_{k, o}$ is to the square of the center of the region, i.e., $c_k^2$.

The minimization of the instantaneous version of $J_{\text{RMA}}$ leads to RMA, whose estimation error is given by

$$e_{\text{RMA}}(n) = \alpha [1 - \hat{y}_\ell^2(n)] \hat{y}_\ell(n) + j \alpha_m [1 - \hat{y}_m^2(n)] \hat{y}_m(n). \quad (8)$$

Fig. 4 shows the real part of this error as a function of $y_\ell(n)$ for 64-QAM, using the optimum scale factors. We can observe that the MMA estimation error is a piecewise function obtained through the repetition of the MMA error shape with $r_2 = 1$ for each region $A_k$, weighted by a scale factor. This error function presents an important characteristic: it is equal to zero at the coordinates of the constellation symbols, which allows RMA to achieve the Wiener solution under some favorable conditions as shown in the next section. To facilitate the visualization, it was normalized by a factor $K$ so that its maximum absolute value is set to unity. This normalization factor can be incorporated to the step-size $\mu$ of the algorithm.

![Figure 4: Real part of the RMA error as a function of $y_\ell(n)$ for 64-QAM. The errors at the constellation symbol coordinates are indicated by $o$, $K = 245$.](image)

To close this section, we should remark that, since the real and imaginary parts of the RMA error are computed separately as in MMA, $J_{\text{RMA}}$ may also present additional stationary points related to phase rotations multiple of $\pi/2$. Furthermore, RMA may occasionally converge to offset solutions, related to phase rotations multiple of $\pi/4$. However, we did not observe any of these behaviors in our simulations.

5. Steady-state analyses

In this section, we obtain a theoretical relation between the coefficient vector of RMA and the Wiener solution, employing the same approach of [35, 36]. Additionally, using the energy conservation arguments of [37] and [31, Ch. 21] and the results of the steady-state analysis of CMA [38, 10, 39, 40, 31, 9], we provide theoretical models for the steady-state mean-square error of MMA and RMA in a nonstationary environment. The analyses are based on the assumptions A1-A8 shown in Appendix A. Some of these assumptions are usual in statistical analysis of blind equalization algorithms and some of them were adjusted for the RMA analysis.

5.1. Relation to the mean-squared error

It is well-known that supervised equalization algorithms (e.g., NLMS filter) seek to minimize the mean-squared error (MSE), defined as

$$J_{\text{MSE}}(w, \Delta) = \frac{1}{5} \mathbb{E}\{|\xi(n)|^2\} = \mathbb{E}\{\xi_n^2(n) + \xi_i^2(n)\}, \quad (9)$$
where
\[ \xi(n) = \xi_\alpha(n) + j\xi_\beta(n) = a(n - \Delta) - y(n). \] (10)

For the sake of comparison, we rewrite \( J_{\text{MSE}} \) by replacing \( y(n) = u^*(n)w \) in (10) and the resulting expression in (9), which leads to
\[ J_{\text{MSE}}(w, \Delta) = \sigma^2 + w^* R w - w^* \hat{p}_a - w^* p_a, \] (11)

where \((\cdot)^*\) stands for the complex conjugate transpose of a vector or a matrix, \( \sigma^2 = E\{a(n)^2\} \), \( R = E\{u^*(n)u^*(n)\} \) is the autocorrelation matrix of the input signal, and \( \hat{p}_a = E\{a(n - \Delta)u^*(n)\} \) is the cross-correlation between the input regressor vector and the transmitted signal \([31]\). Additionally, for any delay \( \Delta \), the minimum of \( J_{\text{MSE}}(w, \Delta) \) occurs for the Wiener-Hopf solution \( w_{\text{wih}} = R^{-1} \hat{p}_a \) and is given by
\[ J_{\text{min}}(\Delta) = \min\{J_{\text{MSE}}(w, \Delta)\} = \sigma^2 - w_{\text{wih}}^T \hat{p}_a. \] (12)

Now, using Assumption A1 of Appendix A and remarking that \( E\{\alpha\} = E\{\alpha_m\} \triangleq \alpha \) for square QAM constellations, (6) can be rewritten at the steady-state as
\[ J_{\text{RMA}} \approx \alpha E\left\{ \left[1 - \tilde{y}(n)\right]^2 \left[1 + \tilde{y}(n)\right]^2 \right\} + E\left\{ \left[1 + y_m(n)\right]^2 [\tilde{a}(n) + \tilde{y}(n)]^2 \right\}. \] (13)

To proceed, we define
\[ \tilde{a}_\ell(n) \triangleq a(n - \Delta) - c_\ell \] (14)
and
\[ \tilde{a}_m(n) \triangleq a(n - \Delta) - c_m. \] (15)
Under Assumption A2 and due to the translation to the origin of the real line, \( \tilde{a}_\ell(n) \) and \( \tilde{a}_m(n) \) always assume a value from the set \{-1, +1\}. Therefore, we obtain the following relations
\[ [1 - \tilde{y}(n)]^2 \left[1 + \tilde{y}(n)\right]^2 = [\tilde{a}(n) - \tilde{y}(n)]^2 [\tilde{a}(n) + \tilde{y}(n)]^2 \]
and
\[ [1 + y_m(n)]^2 [\tilde{a}(n) + y_m(n)]^2 = [\tilde{a}_m(n) - y_m(n)]^2 [\tilde{a}_m(n) + y_m(n)]^2. \]
Replacing these relations in (13), we obtain
\[ J_{\text{RMA}} \approx \alpha E\left\{ \left[\tilde{a}_\ell - \tilde{y}(n)\right]^2 [\tilde{a}_\ell + \tilde{y}(n)]^2 \right\} + E\left\{ [\tilde{a}_m - y_m(n)]^2 [\tilde{a}_m + y_m(n)]^2 \right\}. \] (16)

Remarking that \( \xi_\alpha(n) = a(n - \Delta) - y_m(n) \), using (14), (4), and Assumption A3, we obtain
\[ \tilde{a}_\ell(n) - \tilde{y}(n) \approx \xi_\alpha(n) \] (17)
and
\[ \tilde{a}_\ell(n) + \tilde{y}(n) \approx a(n - \Delta) + y_m(n) - 2c_\ell. \] (18)
Replacing (17) and (18) in (16), using Assumption A4 and the same approach for the imaginary part, we arrive at
\[ J_{\text{RMA}} \approx \alpha E\left\{ \xi_\alpha^2(n)\right\} E\left\{ [a(n - \Delta) + y_m(n) - 2c_\ell]^2 \right\} \]
\[ + E\left\{ \xi_\alpha^2(n)\right\} E\left\{ [a(n - \Delta) + y_m(n) - 2c_\ell]^2 \right\}. \] (19)

For square QAM constellations, \( E\{\xi_\alpha^2\} = E\{\xi_\alpha^2\} \triangleq \sigma_\alpha^2 \) and \( E\{\xi_\alpha^2(n)\} = E\{\xi_\alpha^2\} = 0.5 J_{\text{MSE}}(w, \Delta) \), and therefore the term \( 0.5 J_{\text{MSE}}(w, \Delta) \) can be factorized in (19). Thus, using Assumption A5 the following approximation holds
\[ E\left\{ [a(n - \Delta) + y_m(n) - 2c_\ell]^2 + [a(n - \Delta) + y_m(n) - 2c_\ell]^2 \right\} \approx J_{\text{MSE}}(-w, \Delta) + 8 \sigma_\alpha^2, \] (20)
which leads to
\[ J_{\text{RMA}} \approx 0.5 \hat{\alpha} J_{\text{MSE}}(w, \Delta) \left[ J_{\text{MSE}}(-w, \Delta) + 8 \sigma_\alpha^2 \right]. \] (21)
Replacing (11) into (21), derivating the resulting expression in relation to \( w^* \), and equaling the derivative to zero, we obtain
\[
\mathbf{w}_{\text{RMA}} \approx \mathbf{w}_{\text{RMA}}^0 \mathbf{P}_a^d + \mathbf{w}_{\text{RMA}}^0 R \mathbf{R}_{\text{RMA}} + 4 \sigma_c^2 \mathbf{w}_{\text{WIE}}.
\] (22)

We observe from (22) that \( \mathbf{w}_{\text{RMA}} \) and \( \mathbf{w}_{\text{WIE}} \) are collinear as occurs in [35, 36], where the coefficient vector of CMA was related to the Wiener solution. Taking into account the relation \( \mathbf{w}_{\text{RMA}} \approx \psi \mathbf{w}_{\text{WIE}} \), the scalar \( \psi \) can be expressed in terms of \( J_{\text{min}}(\Delta) \) by identifying (12) in (22), i.e.,
\[
\psi = \frac{2 \psi [\sigma_a^2 - J_{\text{min}}(\Delta)] + 4 \sigma_c^2}{\sigma_a^2 + \psi^2 [\sigma_a^2 - J_{\text{min}}(\Delta)] + 4 \sigma_c^2}.
\] (23)

This relation leads to the following cubic equation in \( \psi \)
\[
\psi^3 + \psi \left( \frac{\sigma_a^2 + 4 \sigma_c^2}{\sigma_a^2 - J_{\text{min}}(\Delta)} - 2 \right) - \frac{4 \sigma_c^2}{\sigma_a^2 - J_{\text{min}}(\Delta)} = 0,
\] (24)
whose roots are complicated expressions. However, if the following approximation holds
\[
\frac{\sigma_a^2}{\sigma_a^2 - J_{\text{min}}(\Delta)} \approx 1,
\] (25)
the real root of (24) is close to one, and in this case, RMA can converge in the mean to the Wiener solution. This enables the steady-state performance of RMA to be close to that of a supervised algorithm, independently of the order of the QAM constellation.

It is important to emphasize that (22) do not ensure the convergence in the mean of RMA to the Wiener solution. This relation only shows that, under certain conditions, the Wiener solution is a possible solution. This enables the steady-state performance of RMA to be close to that of a supervised algorithm, independently of the order of the QAM constellation. In some situations, RMA can converge to undesired local minima.

### 5.2. Tracking analysis

We assume that in a nonstationary environment, the variation in the optimum solution \( \mathbf{w}_o \) follows a random-walk model [31, Sec. 20.2], i.e.,
\[
\mathbf{w}_o(n) = \mathbf{w}_o(n-1) + \mathbf{q}(n),
\] (26)
where \( \mathbf{q}(n) \) denotes some random perturbation that is independent of the initial conditions \( \{\mathbf{w}_o(-1), \mathbf{w}(-1)\} \) and of \( \{\mathbf{u}(l)\} \) for all \( l < n \). In equalization, \( \mathbf{q}(n) \) models the channel variation and is assumed to be i.i.d., zero-mean, and with nonnegative-definite covariance matrix \( \mathbf{Q} = \mathbb{E} \{ \mathbf{q}^*(n) \mathbf{q}(n) \} \). The optimum solution \( \mathbf{w}_o \) represents one of the global minima of the cost function. When the algorithms are implemented in a baud rate, \( \mathbf{w}_o \) can be close to the Wiener solution (see, e.g., [35] and their references) or can coincide with it as occurs for RMA. In this case, \( \mathbf{w}_o \) does not provide perfect equalization, but can achieve a relatively low MSE. On the other hand, when the algorithms are implemented in a T/2-fractionally spaced form in the absence of noise and under certain well-known conditions, \( \mathbf{w}_o \) represents the zero-forcing solution [34].

One measure of the filter performance at the steady-state is given by the excess mean-square error (EMSE), defined as
\[
\zeta \triangleq \lim_{n \to \infty} \mathbb{E} \{ |e_a(n)|^2 \} = \lim_{n \to \infty} \mathbb{E} \{ e_{a,n}(n)^2 + e_{a,l}(n)^2 \},
\] (27)

where
\[
e_a(n) = e_{a,n}(n) + j e_{a,l}(n) = \mathbf{u}^*(n) \tilde{\mathbf{w}}(n-1)
\] (28)
is the a priori error and
\[
\tilde{\mathbf{w}}(n-1) = \mathbf{w}_o(n-1) - \mathbf{w}(n-1)
\] (29)
is the weight-error vector.

For the class of normalized algorithms considered here, the recurrent equation for the weight-error vector can be obtained by subtracting both sides of (1) from \( \mathbf{w}_o(n) \) and using (26), i.e.,
\[
\tilde{\mathbf{w}}(n) - \mathbf{q}(n) = \tilde{\mathbf{w}}(n-1) - \frac{\mu}{\delta + \| \mathbf{u}(n) \|^2} \mathbf{e}(n) \mathbf{u}^*(n).
\] (30)
Using the energy conservation arguments of [37] and [31, Ch. 21], an analytical expression for the steady-state EMSE of the algorithms of the form (1) can be obtained by equating the squared norms on both sides of (30) and taking the expectation as $n \to \infty$. Since (30) has the same structure as the NLMS algorithm, we obtain the same variance relation of [31, Eq.(21.17)]. Thus, for $\delta \approx 0$, we have

$$
\mu \mathbb{E} \left\{ \frac{|e(n)|^2}{\|u(n)\|^2} \right\} + \frac{\text{Tr}(Q)}{\mu} = 2 \text{Re} \left( \mathbb{E} \left\{ \frac{e^*_n e(n)}{\|u(n)\|^2} \right\} \right),
$$

(31)

where $\text{Tr}(Q)$ stands for the trace of the covariance matrix $Q$. 

In order to simplify (31), we use the separation principle of [31, Ch. 17], in which

$$
\mu \eta_u \mathbb{E} \{ |e(n)|^2 \} + \mu^{-1} \text{Tr}(Q) \approx 2 \eta_u \mathbb{E} \{ e^*_n e(n) \},
$$

(32)

where $\eta_u \triangleq \mathbb{E} \{1/\|u(n)\|^2\}$. For Gaussian inputs and large number of coefficients, $\eta_u$ can be approximated by $1/\sigma_n^2(M-2)$ with $\sigma_n^2 = \mathbb{E} \{ |u(n)|^2 \}$ [41]. The approximation (32) still needs theoretical estimates for $\mathbb{E} \{ |e(n)|^2 \}$ and $\mathbb{E} \{ e^*_n e(n) \}$ at the steady-state. These estimates depend on each algorithm and are obtained in the next sections for MMA and RMA. For this purpose, we present in the sequel a general model used in the analysis.

Using (29), (28), and Assumption A6, we obtain

$$
y(n) = u^T(n)w(n-1) = u^T(n)[w_u(n-1) - \bar{w}(n-1)]
= a(n-\Delta) - e_u(n) - v(n),
$$

(33)

where $v(n)$ plays the role of a disturbance that is assumed to be i.i.d., zero-mean, and independent of $u(n)$, $a(n-\Delta)$, and $e_u(n)$ at the steady-state [42]. The real and imaginary parts of $y(n)$ are given respectively by

$$
y_n(n) = a_n(n-\Delta) - e_{n,u}(n) - v_u(n) \tag{34}
$$

and

$$
y_i(n) = a_i(n-\Delta) - e_{n,i}(n) - v_i(n) \tag{35}$$

This model is used in the following tracking analyses.

5.2.1. Analysis of the multimodulus algorithm

In order to obtain an analytical expression for the EMSE of MMA when it is used to adapt a baud rate equalizer, we first extend the model for the steady-state estimation error of constant-modulus-based algorithms proposed in [10, 40]. For this purpose, we consider the assumptions A6-A8 of Appendix A.

Thus, replacing (34) in the real part of (3) and using Assumption A7, we obtain

$$
e_u(n) \approx \gamma_u(n) e_{n,u}(n) + \beta_u(n),
$$

(36)

where

$$
\gamma_u(n) \triangleq 3a_u^2(n-\Delta) - r_2 - 6v_u(n)a_u(n-\Delta),
$$

(37)

and

$$
\beta_u(n) \triangleq r_2 a_u(n-\Delta) - a_u^3(n-\Delta) + v_u(n) \left[ 3a_u^2(n-\Delta) - r_2 \right].
$$

(38)

The model for $e_i(n)$ can be obtained by replacing the subscript $R$ by the subscript $I$ from (36) to (38).

To calculate the first and second moments of the random i.i.d. variables $\gamma_u(n)$, $\gamma_i(n)$, $\beta_u(n)$, and $\beta_i(n)$, we remark that square QAM constellations have circular symmetry, i.e., $\mathbb{E} \{ a_u^k(n) \} = \mathbb{E} \{ a_i^k(n) \} = 0$ for all odd integers $k > 0$. Therefore, we find that $\mathbb{E} \{ \beta_u(n) \} = \mathbb{E} \{ \beta_i(n) \} = 0$.

$$
\sigma_u^2 \triangleq \mathbb{E} \{ \beta_u^2(n) \} = \mathbb{E} \{ \beta_i^2(n) \} = \mathbb{E} \{ a_u^6(n) - r_2 a_u^4(n) \} + 0.5 \sigma_u^2 \mathbb{E} \{ 3a_i^4(n) + r_2^2 \},
$$

(39)

$$
\tilde{\gamma} \triangleq \mathbb{E} \{ \gamma_u(n) \} = \mathbb{E} \{ \gamma_i(n) \} = 1.5 \sigma_u^2 - r_2,
$$

(40)

and

$$
\tilde{\gamma} \triangleq \mathbb{E} \{ \gamma_i^2(n) \} = \mathbb{E} \{ \gamma_u^2(n) \} = 1.5 \left( r_2 + 6 \sigma_u^2 \right) \sigma_u^2 + r_2^2,
$$

(41)
where \( \sigma_v^2/2 = E\{v_v^2(n)\} = E\{v_v^2(n)\} \) and \( \sigma_d^2/2 = E\{a_d^2(n)\} \).

Using (36), the equivalent model for the imaginary part of the error, (39)-(41), and Assumption A8, we obtain

\[
E\{|e(n)|^2\} \approx \tilde{\gamma} E\{|e_a(n)|^2\} + 2\sigma_d^2
\]

and

\[
E\{e_a^*(n)e(n)\} \approx \tilde{\gamma} E\{|e_a(n)|^2\}.
\]

Recalling that \( \tilde{\gamma} \triangleq E\{|e_a(n)|^2\} \) and replacing (42) and (43) in (32), we arrive at

\[
\zeta_{\text{MMA}} \approx \frac{1}{2\tilde{\gamma} - \mu^2} \left[ 2\mu\sigma_d^2 + \frac{\text{Tr}(Q)}{\mu \eta_v} \right].
\]

From (44), we can observe that \( \zeta_{\text{MMA}} \) is zero, only in a stationary environment \( (Q = 0) \), assuming fractionally-spaced equalization in the absence of noise \( (\nu(n) = 0) \), and for constant modulus signals \( (\sigma_d^2 = 0) \). For nonconstant modulus signals, even when the zero-forcing solution is achieved, the algorithm exhibits a non-zero misadjustment inherent in the constant-modulus-based algorithms [10, 39, 40, 31].

We should notice that the variance of \( \nu(n) \) that appears in (39) can be theoretically estimated by using the following expression

\[
\sigma_v^2 \approx \sigma_d^2 - \mathbf{w}_a^\dagger \mathbf{R} \mathbf{w}_a,
\]

where \( \mathbf{w}_a \triangleq \mathbb{E}\{\mathbf{w}_a(n)\} = \mathbf{w}_a(-1) \). When the optimum filter achieves perfect equalization \( \mathbf{w}_a^\dagger \mathbf{R} \mathbf{w}_a = \sigma_v^2 \) and \( \sigma_v^2 = 0 \). Furthermore, noting that \( \mathbf{w}_a^\dagger \mathbf{P}_a = \mathbf{w}_a^\dagger \mathbf{P}_a \), when \( \mathbf{w}_a = \mathbf{w}_a \), \( \sigma_v^2 = J_{\text{min}}(\Delta) \), defined as in (12).

5.2.2. Analysis of the regional multimodulus algorithm

Subtracting \( c_t \) from both sides of (34), using (4) and (14) and doing the same for the imaginary part, we obtain

\[
\hat{y}_t(n) = \bar{a}_t(n) - e_{a,n}(n) - \nu_t(n)
\]

and

\[
\hat{y}_m(n) = \bar{a}_m(n) - e_{a,l}(n) - \nu_l(n).
\]

To proceed, we should remark that, under Assumption A2, \( \bar{a}_t(n) = \pm 1 \) and \( \bar{a}_m(n) = \pm 1 \) and therefore, the following equalities hold

\[
\bar{a}_t^k(n) = \bar{a}_m^k(n) = 1, \text{ } k \text{ even } (\geq 0)
\]

and

\[
E\{\bar{a}_t^k(n)\} = E\{\bar{a}_m^k(n)\} = \begin{cases} 1, & \text{if } k \text{ even } (\geq 0) \\ 0, & \text{if } k \text{ odd } (>0) \end{cases}.
\]

Thus, replacing (46) and (47) in (8) and using (48) and Assumption A7, we get

\[
e(n) \approx \alpha_t [2\nu_t - 6\bar{a}_t(n)e_{a,n}(n)\nu_t(n) + 2e_{a,n}(n)]
+ j\alpha_m [2\nu_l - 6\bar{a}_m(n)e_{a,l}(n)\nu_l(n) + 2e_{a,l}(n)].
\]

Defining \( \bar{\alpha} \triangleq E\{\alpha_t^2\} = E\{\alpha_m^2\} \), using (50), (48), and (49), we obtain

\[
E\{|e(n)|^2\} \approx \bar{\alpha} [2\theta E\{|e_a(n)|^2\} + 4r_\alpha \sigma_v^2]
\]

and

\[
E\{e_a^*(n)e(n)\} \approx 2\bar{\alpha} E\{|e_a(n)|^2\},
\]

where \( r_\alpha \triangleq E\{\alpha_t^2\}/E\{\alpha_t\} = \bar{\alpha} / \alpha \) and \( \bar{\alpha} \triangleq r_\alpha (9\sigma_v^2 + 2) \).

Replacing (51) and (52) in (32), we arrive at

\[
\zeta_{\text{MMA}} \approx \frac{1}{2 - \mu \bar{\alpha}} \left[ 2\mu r_\alpha \sigma_v^2 + \frac{\text{Tr}(Q)}{2 \mu \bar{\alpha} \eta_v} \right].
\]

Different from MMA, in a stationary environment \( (Q = 0) \) and assuming fractionally-spaced equalization in the absence of noise \( (\nu(n) = 0) \), RMA achieves a zero EMSE, independently of the order of the QAM constellation.
6. Improving the convergence

RMA can take a long time to converge since at the beginning of the convergence the equalizer coefficient vector $w$ can be very distant from the optimum solution. When this condition occurs, the equalizer output can fall in a wrong region, which in turn feeds back wrong updates of $w$, specially when the channel presents difficult equalization. Thus, to improve the convergence of these algorithms, we propose a technique inspired in statistical classification algorithms [43]. It consists basically in taking into account the neighborhood of the main regions $A_\ell$ and $A_m$, where the real and imaginary parts of the equalizer output fall in, respectively. With the additional information from the neighborhood, we conjecture that RMA can learn more quickly the geometry of the constellation, which facilitates its convergence.

Assuming that the real part of the equalizer output falls in the region $A_\ell$ as shown in Fig. 3 for $\ell = -1$, the error is computed using not only the main region $A_\ell$, but also the regions $A_{\ell-1}$ and $A_{\ell+1}$ in its neighborhood (similarly for the imaginary counterpart). Note that if $A_\ell$ is a region of the constellation edges, there will be only inner neighbors. The real part of the RMA error is calculated as

$$e_{RMA,R}(n) = \sum_{k=\ell-1}^{\ell+1} \chi_k \alpha_k \left[ 1 - \bar{y}_k^2(n) \right] \bar{y}_k(n), \quad (54)$$

where $\chi_k = b^{-2}$ with $b = 1$ for $k = \ell$ (main region) and $b = 4$ for $k = \ell \pm 1$ (regions in the neighborhood). The weights $\chi_k$ are important to impose a distinction among the errors calculated in the neighborhood and that of the main region, i.e., the farther the neighbor, the smaller $\chi_k$. Their values were experimentally chosen and lead RMA to a good performance independently of the QAM order.

The neighborhood technique causes a distortion on the RMA error function. To give further insight about this distortion, Fig. 5 shows the real part of the RMA error given by (54) as a function of $y(n)$ for 256-QAM. The real part of the MMA error is also shown in this figure for comparison. These curves were normalized by factors $K$ (indicated in the caption of the figure) so that their maximum absolute values are set to unity. We can observe that with the neighbors, the RMA error no longer contains zeros at the coordinates of the constellation symbols. On the other hand, the error function incorporates an envelope that resembles the MMA error. This suggests that the neighbors provide an estimate of the dispersion factor of $r_2$ of MMA, which in turn, contains statistical information of the constellation. This additional statistical information plays an important role in speeding up the convergence of RMA.

Figure 5: Real part of the RMA error with the neighborhood technique as a function of $y(n)$ for 256-QAM ($K=239$). The real part of the MMA error for 256-QAM ($K=1661$) is shown as reference. The errors at the constellation symbol coordinates are indicated by $\circ$.

Although the convergence rate of RMA can be improved with the aid of the neighbors, its EMSE increases since its error function becomes closer than that of MMA. Thus, the aid of the neighbors should be disregarded when RMA achieves the steady-state. For this purpose, instead of weighting the neighbor errors by $b^{-2}$, we consider a time function $\chi_k(n) = b^{p(n)}$, where $k = \ell \pm 1$ for the real part and $k = m \pm 1$ for the imaginary part. The exponent $p(n)$ should assume the form of the function shown in Fig. 6, where $\xi(n) = \lambda \xi(n-1) + (1-\lambda) |\hat{u}(n-D) - y(n)|^2$ is an estimate of the mean-squared decision error and $0 < \lambda < 1$ is a forgetting factor. It should be notice that $-10 \leq p(n) \leq 2$ and that the smaller the EMSE, the smaller is the value of $p(n)$, and consequently the smaller the weight $\chi_k(n)$. One possible expression for
$p(n)$ that matches with Fig. 6 is given by

$$p(n) = 7.15 \exp \left[ \frac{8 \left( \xi(n) - 0.03 \right)}{\exp \left[ 8 \left( \xi(n) - 0.03 \right) \right] + 1} - 9.15, \right] \quad (55)$$

Through simulations, we observe that $p(n)$ does not depend on the QAM order and is important to make the EMSE of the algorithms smaller at the steady-state. Instead of (55), other functions could be used provided they have approximately the form of the function shown in Fig. 6.

![Figure 6: The exponent $p(n)$ as function of $\xi(n)$](image)

Depending on the order of the QAM constellation, more than two regions can be considered in the neighborhood of the main region of each real and imaginary part of the equalizer output. However, we observed through simulations from 64 to 1024-QAM that two neighbors for the real part and two for the imaginary part are sufficient to improve significantly the convergence of RMA. Since this technique involves only computations of scalars, RMA still maintains a computational cost that increases linearly with $M$. In a practical implementation, the value of $\theta^{p(n)}$ could be read from a look-up table.

### 7. Avoiding divergence

When $e(n)$ is based on the constant-modulus cost function, the conditions to ensure the stability of stochastic gradient-type algorithms, as is the case of (1), are not evident. Therefore, the convergence and stability of constant-modulus-based algorithms have been the subject of research for many years (see, e.g., [8, 10, 44, 45] and their references). For a range of step-sizes, [45] observed that CMA may diverge or not in a given run, with a probability of divergence that depends on how close the initial condition is to a local minimum, on the step-size, and on the noise level. Although these results are important to understand the CMA behavior, they do not solve the divergence problem in practical situations, since the local minima of the constant-modulus cost-function are unknown.

In order to avoid divergence, [46] proposed a modified version of CMA with two distinct operation modes. In the first mode, the algorithm works as a normalized CMA and in the second mode, it rejects inconsistent estimates of the transmitted signal. In spite of the lack of a proof for the numerical stability of this algorithm, [46] observed through several simulations that it never diverges. The same philosophy was used in [47] to avoid divergence in the Shalvi-Weinstein algorithm, but with a theoretical proof of its stability.

Using the results of [46], to avoid divergence in MMA and RMA, we rewrite the step-size and the error of these algorithms as $\mu = \bar{\mu} / \bar{\gamma}$ and $e(n) = \bar{\gamma} \tilde{e}(n)$, where

$$\tilde{e}(n) = e(n) / \bar{\gamma} = d(n) - y(n), \quad (56)$$

and $\bar{\gamma}$ is defined in (40).

Writing the error as in $d(n) - y(n)$, all the nonlinearity is included only in $d(n)$, which, as well as $y(n)$, is an estimate of the transmitted signal. Thus, [46] conjectured that the consistency between the estimates $d(n) = \tilde{e}(n) + y(n)$ and $y(n)$ will be ensured if their real parts have the same sign (similarly for the imaginary parts). In this case, the algorithm works in what is called region of interest (ROI). If the real parts or the imaginary parts of these estimates do not have the same sign, the estimate $d(n)$ is simply rejected, which leads to $\tilde{e}(n) = -y(n)$ and $e(n) = -\bar{\gamma} y(n)$. In this case, the algorithm leaves the ROI and enters the second operation mode. In [46], it was shown for scalar CMA filters that $e(n) = -\bar{\gamma} y(n)$ makes the algorithm return to the ROI. In the vector case, the good performance of the algorithm was confirmed through numerical simulations. It is important to notice that this mechanism can also be used in conjunction with the neighborhood technique of Section 6, since the only difference is the computation of the error $e(n)$ considered in (56).

The constant $\bar{\gamma}$ plays an important role here since it is related to the definition of the ROI. For instance, using (56) in MMA, $d_\alpha(n)$ and $y_\alpha(n)$ will have the same sign for $0 \leq |y_\alpha(n)| \leq \sqrt{1.5} \sigma_a$. In this case, we
always have \( \max(|a_n(n)|) \leq \sqrt{1.5} \sigma_a = \sqrt{3E\{a_n^2(n)\}} \) (similarly for the imaginary part). Thus, the bound \( \sqrt{1.5} \sigma_a \) allows to include all the symbol coordinates in the ranges of \( y_u(n) \) and \( y_l(n) \) when the algorithm operates inside the ROI, independently of the QAM order.

Outside the ROI, (1) reduces to

\[
\mathbf{w}(n) = \left[ \mathbf{I} - \frac{\mu^2}{\delta + \|\mathbf{u}(n)\|^2} \mathbf{u}^*(n)\mathbf{u}^T(n) \right] \mathbf{w}(n-1).
\]

(57)

Thus, we can guarantee that the dual-mode algorithms are stable by noting that the matrix between brackets has all eigenvalues with absolute values less than or equal to one if \( 0 < \mu < 2/\bar{\gamma} \). This result does not depend on the use of the neighbors, since this affect only the definition of \( d(n) \), which is disregarded outside the ROI. Inside the ROI, the stability analysis of MMA and RMA is not simple and we intend to pursue this matter elsewhere.

To close this section, we now obtain an expression for the steady-state EMSE of dual-mode MMA and RMA outside the ROI. Since (57) makes the algorithms return to the ROI, MMA and RMA operate in this mode only for a finite-time interval. Therefore, the analytical expression obtained here is the result of a worst-case analysis. Using energy conservation arguments, the steady-state variance relation outside the ROI can be obtained replacing \( e(n) \) and \( \mu \) in (32) by \( -y(n) \) and \( \mu \bar{\gamma} \), respectively, which leads to

\[
\mu \bar{\gamma} \eta_n \mathbb{E}\{|y(n)|^2\} + \frac{\text{Tr}(\mathbf{Q})}{\mu \bar{\gamma}} \approx -2\eta_n \Re\{\mathbb{E}\{e_n^*(n)y(n)\}\}.
\]

(58)

Using (33), the steady-state EMSE outside the ROI for MMA and RMA can be approximated by

\[
\zeta_{\text{outROI}} \approx \frac{1}{2 - \mu \bar{\gamma}} \left[ \mu \bar{\gamma}(\sigma_a^2 + \sigma_l^2) + \frac{\text{Tr}(\mathbf{Q})}{\mu \bar{\gamma} \eta_n} \right].
\]

(59)

A similar expression was obtained in [48] for the dual-mode CMA, assuming fractionally-spaced equalizer in the absence of noise.

8. Simulation results

The simulations are divided into four parts. First, we verify that RMA provides perfect equalization, considering the transmission of a 1024-QAM signal and assuming a FSE in the absence of noise. In the second part, we show some results concerning the ability of RMA to achieve the Wiener solution. We also show that the technique of Section 6 greatly improves its convergence rate. In the third part, we evaluate the influence of noise in terms of the symbol error rate (SER) as a function of the signal-to-noise ratio (SNR). In the last part, we verify the accuracy of the steady-state analysis, considering a stationary and a nonstationary environment.

For comparison, we consider a concurrent algorithm similar to that of [23] and the hybrid algorithm of [21] that uses a Gaussian function as penalty. Both algorithms were updated in a normalized manner as (1) to facilitate the comparison with MMA and RMA. The normalized hybrid algorithm is referred here to as g-ECMA (extended CMA with Gaussian penalty). In the case of the concurrent algorithm, instead of CMA considered in [23], we combined MMA concurrently with the soft decision-directed algorithm, whose real and imaginary parts of its estimation error are computed separately. The resulting algorithm is referred here to as CSD (concurrent soft-decision) algorithm. We should notice that the good behavior of both algorithms depends on the proper choice of the step-size of MMA and of an additional adaptation parameter, which in turn is related to the soft decision-directed algorithm in the case of CSD and to the penalty term in the case of g-ECMA. The tuning of these parameters depends of the QAM order and can be quite complicated. Furthermore, the error functions of these algorithms do not present zeros at the symbol coordinates, and therefore, they may not achieve the Wiener solution.

To verify the performance of MMA in a wide range of situations, we assumed different simulation scenarios throughout this section. As benchmarks, we consider the supervised NLMS algorithm and the Wiener solution. In all simulations, the equalizers were initialized with the typical center spike and we used the normalized versions of all algorithms with \( \delta = 10^{-8} \) (see Eq. (1)). Instead of using the optimum scale factors \( \alpha_{r_o} \) and \( \alpha_{m_o} \) in RMA (see Eq. (7)), we used the absolute values of the centers of the regions, i.e., \( \alpha_r = |c_r| \) and \( \alpha_m = |c_m| \). This does not affect the performance and facilitates its implementation since the distance between centers of adjacent regions is always equal to 4. Furthermore, the approach of Section 7 was considered in MMA and RMA in order to avoid divergence.
8.1. Perfect equalization

Fig. 7 shows the mean-squared decision error (MSE) estimated from the ensemble-average of 500 independent runs. To facilitate the visualization, the MSE curves were filtered by a moving-average filter with 128 taps. This procedure is repeated throughout this section. We assume a T/2-FSE with \( M = 20 \) coefficients and the transmission of a 1024-QAM signal through the channel

\[
h^*_7 = [-0.2 + j0.3 \ -0.5 + j0.4 \ 0.7 - j0.6 \ 0.4 + j0.3 \ 0.2 + j0.1 \ -0.1 + j0.2]
\]

in the absence of noise [28, Table 2]. Using (44) with \( \sigma^2 = \text{Tr}(Q) = 0 \), we obtain \( \text{EMSE} \approx -9.27 \text{ dB} \), which agrees with the experimental result, since we can observe in Fig. 7-(a) that MMA achieves a steady-state MSE slightly superior to \(-10 \text{ dB}\). This behavior is expected for nonconstant modulus signals as is the case of 1024-QAM. The CSD algorithm presents a performance superior than that of MMA, but does no achieve perfect equalization, since its steady-state MSE is approximately \(-15 \text{ dB}\). A similar behavior is observed for g-ECMA, which achieves a steady-state MSE of \(-23 \text{ dB}\) at iteration \( n = 7 \times 10^5 \). On the other hand, RMA achieves an MSE slightly inferior to \(-15 \text{ dB}\) at iteration \( n = 10^6 \) (Fig. 7-(a)) and converges to the Wiener solution like the supervised NLMS algorithm, which provides perfect equalization in this case \((-300 \text{ dB} \text{ in the Matlab precision})\), as we can observe in Fig. 7-(b). The experimental results of RMA agree with our analytical results, since the MSE predicted by (53) is zero for a FSE in a stationary environment \( (\sigma^2 = \text{Tr}(Q) = 0) \). The main limitation for using RMA in practical situations is its slow convergence. Note that it takes approximately \( 2.5 \times 10^6 \) iterations to achieve perfect equalization. This drawback can be overcome by using the technique of Section 6 as we show next.

![MSE for (a) MMA (\( \mu = 2 \times 10^{-6} \)), CSD (\( \mu = 10^{-1}, \mu/\mu_{\text{SSE}} = 10^{-4}, \rho = 0.6 \)), NLMS (\( \mu = 2 \times 10^{-7}, \beta_G = 10^3, \sigma = 0.2 \)), g-ECMA (\( \mu = 4 \times 10^{-3} \)), NLMS (\( \mu = 5 \times 10^{-1}, \Delta_1 = 2 \)), WIE (\( \Delta_1 = 2 \)); (b) RMA (\( \mu = 4 \times 10^{-3} \)), NLMS (\( \mu = 5 \times 10^{-1} \)), g-ECMA (\( \mu = 2 \times 10^{-7} \)), WIE.](image)

8.2. Wiener solution and neighborhood

To illustrate the advantage of the technique proposed in Section 6, Fig. 8-(a) shows the MSE for RMA with neighbors (solid curves) and without neighbors (dotted curves), considering the transmission of a 256-QAM signal through two different channels. From \( n = 0 \) to \( n = 10^6 \), we consider the voice-band telephonic channel \( h_2 \) used in [49, Fig. 2] and at \( n = 10^6 \), it is changed abruptly to channel \( h_3 = [0.36 \ 0.86 \ 0.36]^T \). In both cases, we assume absence of noise and a baud-rate equalizer with \( M = 21 \) coefficients. We can observe in Fig. 8-(a) that RMA converges to the Wiener solution like a supervised algorithm, as observed in Section 5.1. Additionally, the neighbors greatly improve its convergence rates, being essential in nonstationary environments. For instance, after the abrupt change in the channel and taking into account the neighbors, RMA takes approximately less \( 1.5 \times 10^6 \) iterations to converge to the Wiener solution. For reference, we also show the curves for NLMS and MMA. Fig. 8-(b) shows the ensemble-average of the exponent \( p(n) \). When \( p(n) \approx -10 \), the aid of the neighbors could be disregarded without performance degradation. Note that only two neighbors for the real part and two for the imaginary part were used. It is important to notice that sometimes RMA does not converge to the best Wiener solution. This occurred in the case of channel \( h_2 \), in which the blind equalizers recovered the transmitted signal with a delay of \( \Delta_2 = 11 \) samples, while the best Wiener solution occurs for a delay of 13 samples. However, this issue can be bypassed by adjusting the initialization. For comparison, we also show the MSE curve of g-ECMA, whose step-size and weight of the Gaussian penalty were adjusted in order to ensure a good tradeoff between convergence rate and steady-state MSE. Again, we can observe that it outperforms MMA, but,

\footnote{Note that, since \( \sigma^2 = 0 \), the values of EMSE and MSE are the same.}
unlike RMA, it does not achieve the Wiener solution and converges more slowly than RMA with neighbors. Since CSD presents an intermediate behavior between MMA and g-ECMA, we disregard its MSE curve hereafter to simplify the figures.

\[
\begin{array}{c|c}
\text{MSE (dB)} & \text{MMA} & \text{RMA} & \text{NLMS} & \text{WIE} \\
\end{array}
\]

\begin{align*}
\text{with neighbors} \\
\text{MMA} & \text{ NLMS} & \text{WIE} \\
\end{align*}

\[
E\{p(n)\} \text{ iterations}
\]

Figure 8: (a) MSE for MMA (\(\mu = 2 \times 10^{-5}\)), g-ECMA (\(\mu = 1.5 \times 10^{-5}\), \(\beta_G = 10^{3}\), \(\sigma = 0.2\)), RMA (\(\mu = 10^{-3}\)), NLMS (\(\mu = 5 \times 10^{-1}\), \(\Delta_2 = \Delta_3 = 11\)), WIE (\(\Delta_2 = \Delta_3 = 11\)); (b) ensemble average of \(p(n)\); 2 neighbors, channels \(h_2\) and \(h_3\), \(M = 21\), 256-QAM, baud-rate equalizer, absence of noise, average of 500 independent runs.

Considering again the voice-band telephonic channel \(h_2\), but assuming the transmission of a 4096-QAM signal, we obtain the MSE curves of Fig. 9-(a). In this case, we used four neighbors for the real part and four for the imaginary part, since less neighbors are not sufficient to speed up the convergence of RMA. We can observe that the neighbors provide a substantial reduction in the convergence time of RMA. Again, it achieves approximately the Wiener solution at the steady-state, which, in this case, leads to an MSE of \(J_{\text{min}}(13)\) is \(\sigma_v^2 = -41.37\) dB. On the other hand, MMA presents an MSE slightly superior to \(-10\) dB. The experimental results can be predicted by our analysis, since the EMSE predicted by (44) and (53) are respectively \(\zeta_{\text{MMA}} \approx -10.20\) dB and \(\zeta_{\text{RMA}} \approx -57.99\) dB. Recalling that MSE = \(\zeta + \sigma_v^2\), we obtain \(\text{MSE}_{\text{MMA}} \approx -10.19\) dB and \(\text{MSE}_{\text{RMA}} \approx -41.28\) dB, which agree with the experimental results. Fig. 9-(b) shows the ensemble-average of the exponent \(p(n)\). The aid of the neighbors could be disregarded for \(p(n) \approx -10\), which slightly decreases the computational cost of the proposed algorithms in the steady-state.

8.3. Performance with noise

Fig. 10 shows curves of symbol error rate as a function of the signal-to-noise ratio, assuming the cable TV channel \(h_4\) with 128 taps, obtained from “data2.mat” of the database available at http://spib.rice.edu/spib/cable.html. We also assume a T/2-FSE with \(M = 52\) coefficients. In practical situations, uncoded 1024-QAM and 4096-QAM signals require high SNR (superior to 40 dB) for attaining low symbol error rates. The curve for the AWGN channel is considered as a benchmark. RMA was implemented with neighbors to accelerate its convergence rate. For 1024-QAM (Fig. 10-a), RMA performs close to NLMS and outperforms MMA. For 4096-QAM (Fig. 10-b), we can observe that MMA presents the worst performance since in the best case, it leads to an SER of approximately \(10^{-3}\) and RMA performs close to NLMS for SNR > 40 dB. Possibly, to take the estimate of the RMA coefficient vector closer to the Wiener solution for SNR < 40 dB, the number of neighbors should be increased. We can observe from these simulations that RMA is able to perform close to a supervised algorithm, even in presence of noise and independently of the QAM order. We should notice that g-ECMA may perform close to RMA if their adaptation parameters are properly adjusted. However, depending on SNR, to obtain a SER similar to that of RMA, it can be necessary to use a very small step-size, which makes its convergence rate not competitive when compared to that of RMA with neighbors.

8.4. Accuracy of the steady-state analysis

To verify the accuracy of the steady-state analysis, we assume the transmission of a 256-QAM signal through the channel \(h_5 = [0.1925 \ 0.9623 \ 0.1925]^T\) in the absence of noise. We also assume a baud-rate...
Figure 9: (a) MSE for MMA ($\mu = 10^{-7}$), RMA ($\mu = 5 \times 10^{-4}$), RMA with neighbors ($\mu = 10^{-4}$), NLMS ($\mu = 5 \times 10^{-1}$), WIE ($\Delta = 13$); (b) ensemble average of $\{p(n)\}$ for 4 neighbors, channel $h_2$, $M = 21$, 4096-QAM, baud-rate equalizer, absence of noise, and ensemble average of 500 independent runs.

Figure 10: Log of SER as a function of SNR for (a) 1024-QAM, MMA ($\mu = 2 \times 10^{-6}$), RMA ($\mu = 4 \times 10^{-3}$); (b) 4096-QAM, MMA ($\mu = 10^{-7}$), RMA ($\mu = 3 \times 10^{-3}$); NLMS ($\mu = 5 \times 10^{-1}$, $\Delta_4 = 44$); 4 neighbors; channel $h_4$, $M = 52$, $T/2$-FSE.

equalizer with $M = 21$ coefficients. Fig. 11 shows the theoretical curves of the EMSE predicted by (44), (53), and (59), and also the experimental values estimated through an ensemble-average of 500 independent runs. We should notice that the neighbors are used during the transient for accelerating the convergence and disregarded in steady-state. Therefore, RMA can achieve quickly the steady-state and the neighbors do not bias the solution.

Fig. 11-(a) shows the EMSE curves assuming a stationary environment ($\text{Tr}(Q) = 0$). We can observe that the experimental results agree with our analysis for a wide range of step-size. Note that the vertical lines represent a limit of stability. The dual-mode MMA works inside the ROI for lower step-sizes ($10^{-5} \leq \mu \leq 10^{-3}$) and in this case, its behavior is well predicted by (44). In the interval of $\mu$ where the probability of divergence is higher, MMA works outside the ROI during a larger number of iterations. In this case, the experimental EMSE is bounded by the theoretical curve predicted by (59), assuming the operation outside the ROI, as we can observe for $3 \times 10^{-3} \leq \mu \leq 2 \times 10^{-2}$. The steady-state EMSE for RMA is well predicted by (53) for $10^{-5} \leq \mu \leq 10^{-2}$. For $\mu$ around $10^{-2}$ its behavior is also bounded by (59) due to the mechanism used to avoid divergence. In this case, we can observe that RMA presents a much better performance than that of MMA.

Fig. 11-(b) shows the EMSE curves assuming a nonstationary environment ($Q = 10^{-7}$I). As in the stationary case, the experimental EMSE of MMA agrees with (44) for $10^{-5} \leq \mu \leq 3 \times 10^{-3}$ and for $\mu > 3 \times 10^{-3}$,
it is bounded by (59). On the other hand, we can identify intervals of \( \mu \) where the steady-state EMSE of RMA can be well predicted by the analysis. For larger step-sizes, the behavior is similar to the stationary case, i.e., the EMSE of RMA is bounded by (59). For smaller step-sizes, RMA gets stuck in local minima, thus providing solutions worse than predicted.

\[
\text{EMSE (dB)}
\]

\[
\begin{array}{ccc}
\text{Analysis} & \text{Simulation} & \text{MMA} & \text{MMA} \\
\text{RMA} & \text{RMA} & \text{MMA} & \text{MMA} \\
\end{array}
\]

(a)\[0 \quad -50 \quad -100 \quad -150 \]

(b)\[0 \quad -20 \quad -40 \quad -60 \]

Figure 11: EMSE as a function of step-size \( \mu \): (a) stationary environment (b) nonstationary environment \( Q = 10^{-7}I; \) 256-QAM, channel \( h_n \), absence of noise, baud-rate equalizer, no neighbors, \( M = 21 \), ensemble-average of 500 independent runs.

9. Conclusion

The simulation results suggest that RMA can be a good candidate for blind equalization of QAM signals. It is able to perform close to a supervised algorithm, even in presence of noise and independently of the QAM order. Using the techniques of sections 6 and 7, it presents a relatively fast convergence and does not diverge. Furthermore, the model obtained for the steady-state EMSE of Section 5.2 shows a reasonable agreement with simulations. It is important to notice that the neighborhood technique plays a role of a soft switching and is important to ensure a good transient behavior of RMA. This technique can also be straightforwardly extended to speed up the convergence of algorithms whose estimation error is a piecewise function, as the decision-direct-based algorithms proposed in the literature (see, e.g., [50, 19, 2, 51]).

Appendix A. Assumptions of the steady-state analysis

The assumptions used in the steady-state analysis are shown below. Assumptions A1-A5 are used to obtain the relation between the coefficient vector of RMA and the Wiener solution. Assumptions A6, A7, and A8 are used in the tracking analysis of MMA and RMA.

A1) The scale factors \( \alpha_\ell \) and \( \alpha_m \) are independent of \([1 - \tilde{y}_\ell(n)]^2\) and \([1 - \tilde{y}_m(n)]^2\). This assumption essentially requires the steady-state fluctuations in the quadratic errors \([1 - \tilde{y}_\ell(n)]^2\) and \([1 - \tilde{y}_m(n)]^2\) to be insensitive to \( \alpha_\ell \) and \( \alpha_m \), respectively.

A2) The signal-to-noise ratio at the steady-state is high enough such that \( a_\ell(n - \Delta) \) is equal to one of the symbols \( a_\ell,1 \) or \( a_\ell,2 \) of the region \( A_\ell \) in which \( y_n(n) \) belongs (similarly for the imaginary part). In other words, the regions \( A_\ell \) and \( A_m \) contain \( a_\ell(n - \Delta) \) and \( a_m(n - \Delta) \), respectively.

A3) The output of the equalizer adapted with RMA is close to that obtained with a supervised equalizer (e.g., NLMS filter), whose coefficients are updated using (10). This assumption is reasonable when A1 is satisfied at the steady-state and the delay \( \Delta \) is the same for the blind and supervised estimates [35, 36].

A4) \([a_\ell(n - \Delta) + y_n(n) - 2e]\) and \([a_m(n - \Delta) + y_n(n) - 2e]\) are independent of \( \xi_\ell^2(n) \) and \( \xi_m^2(n) \), respectively [35].

A5) \( c_\ell \) is independent of \( a_\ell(n - \Delta) \) and \( y_n(n) \), which implies \( E\{c_\ell a_\ell(n - \Delta)\} = E\{c_\ell y_n(n)\} = 0 \) (similarly for the imaginary part).
A6) The transmitted signal can be written as
\[ a(n - \Delta) = u^T(n)w_a(n - 1) + v(n), \] (60)
where \( v(n) = v_a(n) + jv_i(n) \) plays the role of a disturbance that is assumed to be i.i.d., zero-mean, and independent of \( u(n) \), \( a(n - \Delta) \), and \( e_a(n) \) at the steady-state. With the inclusion of \( v(n) \) in (60), the case of a baud rate equalization is covered since \( v(n) \neq 0 \) means that the optimum filter does not achieve perfect equalization. This model is most commonly used in the context of system identification being refereed as linear regression model [31], but it can also be used in the analysis of adaptive equalizer algorithms as shown in [42].

A7) In steady-state, terms depending on \( e_a^k(n), e_i^k(n), v_a^k(n), \) and \( v_i^k(n), k \geq 2 \) can be disregarded since they are sufficiently small when compared to terms depending on \( e_a(n), e_i(n), v_a(n), \) and \( v_i(n) \), respectively. In other words, the equalizer can not achieve perfect equalization, but sufficiently mitigate the intersymbol interference introduced by the channel [10, 40, 45].

A8) The real part of the \( a \) \( a \) priori error is independent of its imaginary part at the steady-state.

References