Research Article

Edgeworth Expansion Based Model for the Convolutional Noise pdf

Yonatan Rivlin and Monika Pinchas

Department of Electrical and Electronic Engineering, Ariel University, 40700 Ariel, Israel

Correspondence should be addressed to Monika Pinchas; monika.pinchas@gmail.com

Recently, the Edgeworth expansion up to order 4 was used to represent the convolutional noise probability density function (pdf) in the conditional expectation calculations where the source pdf was modeled with the maximum entropy density approximation technique. However, the applied Lagrange multipliers were not the appropriate ones for the chosen model for the convolutional noise pdf. In this paper we use the Edgeworth expansion up to order 4 and up to order 6 to model the convolutional noise pdf. We derive the appropriate Lagrange multipliers, thus obtaining new closed-form approximated expressions for the conditional expectation and mean square error (MSE) as a byproduct. Simulation results indicate hardly any equalization improvement with Edgeworth expansion up to order 4 when using optimal Lagrange multipliers over a nonoptimal set. In addition, there is no justification for using the Edgeworth expansion up to order 6 over the Edgeworth expansion up to order 4 for the 16QAM and easy channel case. However, Edgeworth expansion up to order 6 leads to improved equalization performance compared to the Edgeworth expansion up to order 4 for the 16QAM and hard channel case as well as for the case where the 64QAM is sent via an easy channel.

1. Introduction

In this work, we deal with the convolutional noise arising at the output from a blind deconvolutional process. A blind deconvolution process arises in many applications such as seismology, underwater acoustic, image restoration, and digital communication [1]. Consider the digital communication case. During transmission, a source signal undergoes a convolutive distortion between its symbols and the channel impulse response. This distortion is referred to as intersymbol interference (ISI) [2]. Thus, a blind adaptive filter is used to remove the convolutive effect of the system to produce the source signal [2]. This process is called blind deconvolution. Since the updated coefficients used in the blind adaptive filter are not the ideal values, a noise named convolutional noise occurs at the output of the deconvolution process in addition to the source signal. Blind deconvolution algorithms based on adaptive filtering techniques generate an estimate of the desired response by applying a nonlinear transformation to sequences involved in the adaptation process [1, 3]. The

Bussgang algorithm is one of the three important families of blind equalization algorithms, where the nonlinearity is in the output of the adaptive equalization filter [1, 3]. According to [4, 5], we may find among the traditional Bussgang-type methods the Sato's [6], Godard's [7], Benveniste et al. [8], Besneveniste-Goursat's [9], and the Stop-and-Go [10] algorithm. For Bussgang-type methods, the nonlinearity is designed to minimize a cost function based on high-order statistics (HOS) according to one approach [6, 7], or calculated directly according to the Bayes rules [11–15]. According to [1, 3], the main difference between the Bussgang type algorithms lies in the choice of the memoryless nonlinearity. Obviously, the performance of this kind of blind equalizer depends substantially on the memoryless nonlinearity [1]. Chapter eight in [5] deals with the question whether the chosen equalizer leads to perfect equalization performance from the MSE point of view. From the work in [1] we know that if the MSE tends to zero, then also the residual ISI tends to zero (perfect equalization). Furthermore, chapter eight in [5] distinguishes between the MSE obtained from the cost

function approach and the MSE obtained from the Bayesian approach. According to chapter eight in [5], the MSE related to the Bayesian approach valid in the convergence state, where the convolutional noise is very small, is approximately given by the convolutional noise power (variance of the convolutional noise) multiplied by a constant. To show this outcome the approximated MSE obtained in [1] was recalled in this chapter [5]. Also the approximated MSE obtained in [16] could have been used for showing that the MSE related to the Bayesian approach valid in the convergence state, where the convolutional noise is very small, is approximately given by the convolutional noise power multiplied by a constant. The only difference between the approximated expression for the MSE obtained in [16] and that in [1] is the value of the constant multiplying the convolutional noise power. In [1] this constant is approximately equal to one regardless of the constellation input, while in [16] this constant is constellation input dependent and is approximately equal to two for the 16QAM input. This paper also adheres to the Bayesian approach where the approximated MSE is developed as a byproduct. We will show in this paper that the approximated MSE valid in the convergence state, where the convolutional noise is very small, is approximately given by the convolutional noise power multiplied by a constant. This constant is constellation input independent, and is approximately equal to one in addition to higher order statistics of the convolutional noise such as the kurtosis. For the case where the convolutional noise pdf is assumed to be Gaussian, our new approximated MSE tends to the expression obtained in [1].

Next we turn to the approximated MSE related to the cost function approach where the derivation of the cost function with respect to the equalized output signal is a polynomial function of the equalized output signal as is in Godard's case [7], for example. For that case, the approximated MSE is approximately equal to a constant in addition to another constant multiplied with the convolutional noise power [5]. Since those constants are constellation input dependent, the approximated MSE may not tend to zero when the convolutional noise power tends to zero unlike in the Bayesian approach. For example, let us consider the 16QAM input and Godard's [7] algorithm (cost function approach). The residual MSE for that case is not zero even when the convolutional noise power tends to zero while this is not the case for the Bayesian approach [1].

In the literature, we find many blind adaptive algorithms based on the cost function approach (e.g [7, 17–20]), while only a few are based on Bayes rules. The reason may lie in the fact that for the latter case (algorithm based on Bayes rules), the conditional expectation (the expectation of the source signal given the equalized output signal) has to be calculated. But, this task might be difficult because: (i) the source pdf might be non-Gaussian, and might not be known. In addition, the expression for the conditional expectation should hold for a wide range of source pdfs and not only for a specific source. (ii) No model is available for the convolutional noise pdf that is valid for the whole deconvolution process. Thus, without having the source and convolutional noise pdfs the derivation of the conditional expectation seems to be impossible. According to [1] the conditional expectation was derived for non-Gaussian sources by Bellini [11, 12], Fiori [13, 14] and Haykin [15]. However, all the mentioned expressions for the conditional expectation [11–15] are suitable only for uniformly distributed source signals. Thus they cannot cope with a source having a general pdf shape [1]. In addition, the works in [11–15] modeled the convolutional noise pdf as Gaussian. Recently ([1, 16]), two closed-form approximated expressions were obtained for the conditional expectation where the source pdf was approximated with the maximum entropy density estimation technique and Edgeworth Expansion series, respectively. Those expressions for the conditional expectation ([1, 16]) do not impose any restrictions (except of even symmetric) on the pdf of the unobserved input sequence. Hence they are suitable for a wider range of source pdf compared with Bellini's [11, 12], Fiori's [13, 14], or Haykin's [15] expression. In [1, 16], the Laplace integral method was needed for approximating the integrals involving in the conditional expectation calculations. In addition, in both cases ([1, 16]), the convolutional noise pdf was assumed to be Gaussian for the whole deconvolution process as was assumed in [11–15]. But, according to Haykin [15], in the early stages of the iterative deconvolution process, the ISI is typically large with the result that the data sequence and the convolutional noise are strongly correlated. Furthermore, the convolutional noise sequence is more uniform than Gaussian [21]. The maximum entropy algorithm [1], where the conditional expectation was calculated with the maximum entropy density approximation technique and Gaussian model for the source signal and convolutional noise pdf, respectively, has shown to have improved equalization performance compared with the classical methods ([22] (RCA algorithm), [7, 9, 13, 17, 18, 23, 24]). However, probably a more appropriate model for the convolutional noise pdf than the choice of the Gaussian one might lead to improved equalization performance compared to the results presented in [1]. The first attempt to characterize the convolutional noise pdf different than the Gaussian model was introduced in [5]. Here the conditional expectation was approximately obtained by approximating the source and convolutional noise pdf with the maximum entropy density approximation technique and Edgeworth expansion series up to order four, respectively. An unknown pdf may be approximated with three types of orthogonal expansions using the Hermite polynomials, namely, the Gram-Charlier of type A, Gauss-Hermite, and Edgeworth expansions [4, 25]. According to [4, 25–27], the Edgeworth expansion is much more useful in many applications, since it is directly connected to the moments and cumulants of a pdf (the property which is lost in the Gauss-Hermite series). It is also a true asymptotic expansion, so that the error of the approximation is controlled (a property which is not found in the Gram-Charlier of type A) [4, 25]. Thus, Edgeworth expansion up to order 6 will describe the unknown pdf better than Edgeworth Expansion up to order four. The idea of approximating a non-Gaussian signal is not new in the literature. As a matter of fact, we may find several works [28–32] dealing with pdfs applicable for the non-Gaussian case but encompass also the Gaussian model. However, the idea of modeling the convolutional noise pdf,

which is changing in time, different than the Gaussian pdf is quite new and was introduced at the first time in [5]. Here the convolutional noise pdf was modeled with the Edgeworth expansion series up to order four and the source signal was approximated with the maximum entropy density approximation technique.

The maximum entropy density approximation technique involves Lagrange multipliers which have to be defined; otherwise, this approximation is not applicable. However, according to [1], finding the Lagrange multipliers is not an easy task. In some cases, an analytical solution for the Lagrange multipliers does not exist (please refer to [1] for more details). In order to overcome the problem, the approximated MSE was derived in [1] and the required Lagrange multipliers were those Lagrange multipliers that lead the approximated MSE to minimum. Note that those Lagrange multipliers [1] were obtained assuming a Gaussian model for the convolutional noise pdf, but were also applied in [5] where the conditional expectation was calculated using the Maximum entropy density approximation technique and Edgeworth expansion series up to order four for the source signal and convolutional noise pdf, respectively even though those Lagrange multipliers are not the optimal set for that case. Although those Lagrange multipliers were not the appropriate set for [5], improved equalization performance was obtained with the new approximated expression for the conditional expectation introduced in [5] compared to the conditional expectation obtained in [1]. Note that the only difference between those expressions ([1, 5]) was the chosen model for the convolutional noise pdf that was taken in the conditional expectation calculations. Thus, it is probable that even more improved equalization performance may be obtained, if the appropriate Lagrange multipliers are applied in [5] and the Edgeworth expansion series up to order six are used instead of order four for the convolutional noise pdf. But to apply the appropriate Lagrange multiplier in [5], we must first derive the MSE related to the conditional expectation that uses the maximum entropy density approximation technique for the source signal pdf and the Edgeworth expansion series up to order four and six for the convolutional noise pdf. Hitherto, this derivation has not been done. Furthermore, the approximated conditional expectation has not been derived where the source signal and the convolutional noise pdf are approximated with the maximum entropy density approximation technique and Edgeworth expansion series up to order six, respectively.

In this paper we derive the appropriate set of Lagrange multipliers where the source signal and convolutional noise pdf are approximated with the maximum entropy density approximation technique and Edgeworth expansion series up to order four and six, respectively. We derive new approximated expressions for the conditional expectation and MSE where the source signal and convolutional noise pdf are approximated with the maximum entropy density approximation technique and Edgeworth expansion series up to order six, respectively. In this paper we test the effect of the use of the optimal set of Lagrange multipliers to the equalization performance. Furthermore, we examine whether increasing the series order in the Edgeworth Expansion can supply
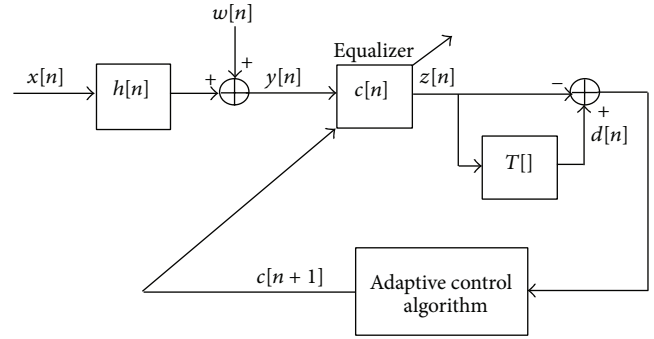

FIGURE 1: System description.

even more improved equalization performance compared to those already presented in [5] where we saw significant equalization performance improvement compared to the maximum entropy algorithm [1]. We also assess whether the increase in the series order is justified given the increase in the complexity of the algorithm.

The paper is organized as follows. In Section 2 we describe the system under consideration, and in Section 3 we present the new model for the convolutional noise pdf, closed-form approximated expressions for the conditional expectation, MSE, and Lagrange multipliers. Simulation results are given in Section 4, and Section 5 is our conclusion.

2. System Description

The system is illustrated in Figure 1, where we make the same assumptions as in [1]:

(1) the input sequence $x[n]$ has an even symmetric probability distribution function with zero mean. $x[n]$ may be a real or a complex variables (where real and imaginary parts of $x[n]$ are independent);

(2) the unknown channel $h[n]$ can be a nonminimum phase linear time-invariant filter with no deep zeros (the zeros lie sufficiently far from the unit circle);

(3) $c[n]$ is a tap-delay equalizer;

(4) $w[n]$ is an additive Gaussian white noise;

(5) $T[]$ is a memoryless nonlinear function.

The input sequence $x[n]$ is transmitted through the channel $h[n]$ and is corrupted with noise $w[n]$. Therefore, the equalizer's input sequence $y[n]$ may be written as

$$y[n] = x[n] * h[n] + w[n], \tag{1}$$

where "$*$" denotes the convolution operation.

From [1], the equalized output signal may be defined as

$$z[n] = x[n] + p[n] + \widetilde{w}[n], \tag{2}$$

where $p[n]$ is the convolution noise arising from the difference between the ideal value of $c[n]$ and the initial guess of $c[n]$, $c_g[n]$ and $\widetilde{w}[n]$ represents the convolution between $w[n]$ and $c_g[n]$.

The ISI is often used as a measure of performance in equalizer applications, as defined by Pinchas and Bobrovsky [1]:

$$\text{ISI} = \frac{\sum_n |\tilde{s}_n|^2 - |\tilde{s}|_{\max}^2}{|\tilde{s}|_{\max}^2}, \tag{3}$$

where $|\tilde{s}|_{\max}$ is the maximum absolute value of the convolution between $h[n]$ and $c_g[n]$ ($\tilde{s}_n = c_g[n] * h[n]$).

Next we consider the adaptation mechanism of the equalizer. According to Figure 1, we define $d[n]$ as an estimator of $x[n]$, which is produced by the function $T[z[n]]$.

Thus the error signal is

$$\tilde{e}[n] = T[z[n]] - z[n]. \tag{4}$$

The adaptive mechanism uses this error to update the equalizers taps [33]:

$$\underline{c}_{\text{eq}}[n+1] = \underline{c}_{\text{eq}}[n] + \mu \cdot \tilde{e}[n] \, \underline{y}^*[n], \tag{5}$$

where $(\cdot)^*$ is the conjugate operation on $(\cdot)$, μ is the step-size parameter, and $\underline{c}_{\text{eq}}[n]$ is the equalizer vector where the input vector is $\underline{y}[n] = [y[n] \cdots y[n-N+1]]^T$ and N is the equalizer's tap length. The operator $(\cdot)^T$ denotes the transpose of the function $[\cdot]$. The conditional expectation $[E[x[n] \mid z[n]]]$, where $E[\cdot]$ stands for the expectation operation, is considered as a good estimate of $T[z[n]]$ [3]. In the following, $T[z[n]]$ is represented by the conditional expectation as was done in [1].

3. New Model for the Convolutional Noise pdf

In this section we describe the proposed model for the convolutional noise pdf which is based on the Edgeworth expansion. It differs from the Gaussian model but can turn under some conditions back to the Gaussian model as will be shown in this section. In Section 3.1 we derive the conditional expectation based on the new proposed model for the convolutional noise pdf. In Section 3.2 we present the new Lagrange multipliers via the MSE expression. It should be pointed out that we are not trying to find or propose a new expression for the conditional expectation but rather trying to show via the conditional expectation whether the new proposed model for the convolutional noise pdf (Edgeworth expansion of orders four and six) can lead to improved equalization performance with optimal Lagrange multipliers. The convolutional noise pdf for the real valued case may be defined as

$$f_p(p) \cong \frac{\exp\left(-(p)^2/2\sigma_p^2\right)}{\sqrt{2\pi}\sigma_p} \left[1 + \sum_{r=3}^{r} \frac{K_r\left(p/\sigma_p\right) H_r\left(p/\sigma_p\right)}{r!}\right], \tag{6}$$

where ! represents factorial, the convolution noise $p[n]$ is defined as p, and σ_p^2 is the variance of the convolutional noise. $H_r(p/\sigma_p)$ are normalized Hermite Polynomials of order r described in [34] and presented in Appendix A as $H_r(\zeta)$. $K_r(p/\sigma_p)$ are the one-dimensional central cumulants

in terms of central moments described in [34, Table 3], and are also presented in Appendix A.

Thus, (6) up to order 6 becomes

$$f_p(p) \cong \frac{\exp\left(-p^2/2\sigma_p^2\right)}{\sqrt{2\pi}\sigma_p} \times \left[1 + \frac{K_4\left(p/\sigma_p\right) H_4\left(p/\sigma_p\right)}{4!} + \frac{K_6\left(p/\sigma_p\right) H_6\left(p/\sigma_p\right)}{6!}\right], \tag{7}$$

where according to Appendices A and B we have

$$\frac{K_4\left(p/\sigma_p\right) H_4\left(p/\sigma_p\right)}{4!}$$
$$= \left(\frac{E\left[p^4\right] - 3(\sigma_p^2)^2}{4!(\sigma_p^2)^2}\right) \left(\frac{p^4}{(\sigma_p^2)^2} - \frac{6p^2}{\sigma_p^2} + 3\right),$$

$$\frac{K_6\left(p/\sigma_p\right) H_6\left(p/\sigma_p\right)}{6!} \tag{8}$$
$$= \left(\frac{E\left[p^6\right] - 15\left(\sigma_p^2\right) E\left[p^4\right] + 30(\sigma_p^2)^3}{6!(\sigma_p^2)^3}\right)$$
$$\times \left(\frac{p^6}{(\sigma_p^2)^3} - \frac{15p^4}{(\sigma_p^2)^2} + \frac{45p^2}{(\sigma_p^2)} - 15\right).$$

So, with the help of (8), (7) becomes

$$f_p(p) \cong \frac{\exp\left(-p^2/2\sigma_p^2\right)}{\sqrt{2\pi}\sigma_p}$$
$$\times \left[1 + \left(\frac{E\left[p^4\right] - 3(\sigma_p^2)^2}{4!(\sigma_p^2)^2}\right) \left(\frac{p^4}{(\sigma_p^2)^2} - \frac{6p^2}{\sigma_p^2} + 3\right)\right.$$
$$+ \left(\frac{E\left[p^6\right] - 15\left(\sigma_p^2\right) E\left[p^4\right] + 30(\sigma_p^2)^3}{6!(\sigma_p^2)^3}\right)$$
$$\left. \times \left(\frac{p^6}{(\sigma_p^2)^3} - \frac{15p^4}{(\sigma_p^2)^2} + \frac{45p^2}{(\sigma_p^2)} - 15\right)\right]. \tag{9}$$

According to (7) and (8) the convolutional noise pdf based on the Edgeworth expansion of order 4 can be obtained (by resetting the last product in (7)) as

$$
\begin{aligned}
f_p(p) \\
\cong \frac{\exp\left(-p^2/2\sigma_p^2\right)}{\sqrt{2\pi}\sigma_p} \\
\times \left[1 + \left(\frac{E[p^4] - 3\left(\sigma_p^2\right)^2}{4!\left(\sigma_p^2\right)^2}\right)\left(\frac{p^4}{\left(\sigma_p^2\right)^2} - \frac{6p^2}{\sigma_p^2} + 3\right)\right].
\end{aligned}
$$

(10)

According to Haykin [15] a Gaussian convolutional noise pdf model is only applicable in the latter stages of the deconvolution process when the process is close to optimal. This is reflected in our model (6) when the expressions in the parentheses are close to one thus getting back to the expression that describes a Gaussian distribution:

$$
f_p(p) \cong \frac{\exp\left(-(p)^2/2\sigma_p^2\right)}{\sqrt{2\pi}\sigma_p}.
$$

(11)

3.1. The Approximated Expression for the Conditional Expectation. In this subsection we use our new model for the convolutional noise pdf to calculate the conditional expectation.

Theorem 1. *For the following assumptions:*

(1) *the source signal $x[n]$ is an independent signal with known variance and higher moments. In the following we denote $x[n]$ as x,*

(2) *no noise is added,*

(3) *the convolutional noise $p[n]$ and the source signal $x[n]$ are independent; thus, $\sigma_z^2 = \sigma_x^2 + \sigma_p^2$ where σ_z^2 represents the variance of the equalizer's output,*

(4) *the convolutional noise $p[n]$ has zero mean and variance $\sigma_p^2 = E[p[n]p^*[n]]$,*

the conditional expectation is given by

$$
\begin{aligned}
E[x \mid z] \cong &\left(z_1 + \frac{g_1''(z_1)}{2g(z_1)}\left(\sigma_{z_1}^2 - \sigma_{x_1}^2\right)\right. \\
&+ \frac{g_1''''(z_1)}{8g(z_1)}\left(\sigma_{z_1}^2 - \sigma_{x_1}^2\right)^2\right) \\
&\times \left(1 + \frac{g''(z_1)}{2g(z_1)}\left(\sigma_{z_1}^2 - \sigma_{x_1}^2\right)\right. \\
&\left.+ \frac{g''''(z_1)}{8g(z_1)}\left(\sigma_{z_1}^2 - \sigma_{x_1}^2\right)^2\right)^{-1}
\end{aligned}
$$

$$
\begin{aligned}
&+ j\left(z_2 + \frac{g_2''(z_2)}{2g(z_2)}\left(\sigma_{z_2}^2 - \sigma_{x_2}^2\right)\right. \\
&+ \frac{g_2''''(z_2)}{8g(z_2)}\left(\sigma_{z_2}^2 - \sigma_{x_2}^2\right)^2\right) \\
&\times \left(1 + \frac{g''(z_2)}{2g(z_2)}\left(\sigma_{z_2}^2 - \sigma_{x_2}^2\right)\right. \\
&\left.+ \frac{g''''(z_2)}{8g(z_2)}\left(\sigma_{z_2}^2 - \sigma_{x_2}^2\right)^2\right)^{-1},
\end{aligned}
$$

(12)

where

$$
z = z_1 + jz_2; \qquad x = x_1 + jx_2
$$

$$
s = 1, 2; \quad g(z_s) = \exp\left(\sum_{k=0}^{K} \lambda_k z_s^k\right)
$$

$$
\times \left[1 + \frac{K_4\left(p/\sigma_p\right)H_4\left(p/\sigma_p\right)}{4!}\right.
$$

$$
\left.+ \frac{K_6\left(p/\sigma_p\right)H_6\left(p/\sigma_p\right)}{6!}\right],
$$

$$
g''(z_s) = \left\{\frac{d^2}{dx_s^2}\left[\exp\left(\sum_{k=0}^{K}\lambda_k x_s^k\right)\right.\right.
$$

$$
\times \left[1 + \frac{K_4\left(p/\sigma_p\right)H_4\left(p/\sigma_p\right)}{4!}\right.
$$

$$
\left.\left.\left.+ \frac{K_6\left(p/\sigma_p\right)H_6\left(p/\sigma_p\right)}{6!}\right]\right]\right\}_{x_s=z_s},
$$

$$
g''''(z_s) = \left\{\frac{d^4}{dx_s^4}\left[\exp\left(\sum_{k=0}^{K}\lambda_k x_s^k\right)\right.\right.
$$

$$
\times \left[1 + \frac{K_4\left(p/\sigma_p\right)H_4\left(p/\sigma_p\right)}{4!}\right.
$$

$$
\left.\left.\left.+ \frac{K_6\left(p/\sigma_p\right)H_6\left(p/\sigma_p\right)}{6!}\right]\right]\right\}_{x_s=z_s},
$$

$$
g_1''(z_s) = \left\{\frac{d^2}{dx_s^2}\left[x_s\exp\left(\sum_{k=0}^{K}\lambda_k x_s^k\right)\right.\right.
$$

$$
\times \left[1 + \frac{K_4\left(p/\sigma_p\right)H_4\left(p/\sigma_p\right)}{4!}\right.
$$

$$
\left.\left.\left.+ \frac{K_6\left(p/\sigma_p\right)H_6\left(p/\sigma_p\right)}{6!}\right]\right]\right\}_{x_s=z_s},
$$

$$g_1''''(z_s) = \left\{ \frac{d^4}{dx_s^4} \left[x_s \exp\left(\sum_{k=0}^{K} \lambda_k x_s^k\right) \right.\right.$$

$$\times \left[1 + \frac{K_4\left(p/\sigma_p\right) H_4\left(p/\sigma_p\right)}{4!} \right.$$

$$\left.\left.\left.\left. + \frac{K_6\left(p/\sigma_p\right) H_6\left(p/\sigma_p\right)}{6!} \right]\right]\right\}\right|_{x_s = z_s}, \tag{13}$$

where $()''$ and $()''''$ denote the second and fourth derivative of $()$, respectively, $z = z[n]$, z_1, and x_1 are the real parts of z and x, respectively. z_2 and x_2 are the imaginary parts of z and x, respectively. $\sigma_{x_1}^2$, $\sigma_{x_2}^2$ are the variances of the real and imaginary parts of the source signal, respectively, and $\sigma_{z_1}^2$, $\sigma_{z_2}^2$ are the variances of the real and imaginary parts of the equalized output signal, respectively. The Lagrange multipliers λ_k are given in Section 3.2. For more details concerning $g''(z)$, $g''''(z)$, $g_1''(z)$, and $g_1''''(z)$, please refer to Appendix B.

Comments. Please note that (12) is similar to the expression for the conditional expectation derived in [1]. However, $g(z)$, $g''(z)$, $g''''(z)$, $g_1(z)$, $g_1''(z)$, and $g_1''''(z)$ are very different from those obtained in [1].

Proof. We start our derivation for the real valued case and then we extend them to the two independent quadrature carrier one. By using Bayes' rule we obtain

$$E[x \mid z] = \frac{\int_{-\infty}^{\infty} x f_{z/x}(z/x) f_x(x) dx}{\int_{-\infty}^{\infty} f_{z/x}(z/x) f_x(x) dx}, \tag{14}$$

where in our case $f_{z/x}(z/x)$ is the convolutional noise pdf and the unknown source pdf, $f_x(x)$, will be estimated with the maximum entropy density approximation technique as was done in [1] and defined as $\widehat{f}_x(x)$

$$\widehat{f}_x(x) \cong \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right), \tag{15}$$

where λ_k $(k = 0, 1, 2, \ldots, K)$ are the Lagrange multipliers given in Section 3.2.

Substituting (6) and (15) into (14) and by using (2) we obtain

$$E[x \mid z] = \left(\int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left(\frac{-(z-x)^2}{2\sigma_p^2} \right) \right.$$

$$\times \left[1 + \frac{K_4\left((z-x)/\sigma_p\right) H_4\left((z-x)/\sigma_p\right)}{4!} \right.$$

$$\left. + \frac{K_6\left((z-x)/\sigma_p\right) H_6\left((z-x)/\sigma_p\right)}{6!} \right]$$

$$\left. \cdot \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right) dx \right)$$

$$\times \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left(\frac{-(z-x)^2}{2\sigma_p^2} \right) \right.$$

$$\times \left[1 + \frac{K_4\left((z-x)/\sigma_p\right) H_4\left((z-x)/\sigma_p\right)}{4!} \right.$$

$$\left. + \frac{K_6\left((z-x)/\sigma_p\right) H_6\left((z-x)/\sigma_p\right)}{6!} \right]$$

$$\left. \cdot \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right) dx \right)^{-1}. \tag{16}$$

Next we define

$$g(x) = \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right)$$

$$\times \left[1 + \frac{K_4\left((z-x)/\sigma_p\right) H_4\left((z-x)/\sigma_p\right)}{4!} \right.$$

$$\left. + \frac{K_6\left((z-x)/\sigma_p\right) H_6\left((z-x)/\sigma_p\right)}{6!} \right], \tag{17}$$

$$g_1(x) = x g(x), \tag{18}$$

$$\rho = 2\sigma_p^2; \qquad \psi(x) = (z-x)^2; \qquad \sigma_p^2 = \sigma_z^2 - \sigma_x^2, \tag{19}$$

so (16) becomes

$$E[x \mid z] = \frac{\int_{-\infty}^{\infty} g_1(x) \cdot \exp\left(-\psi(x)/\rho\right) dx}{\int_{-\infty}^{\infty} g(x) \cdot \exp\left(-\psi(x)/\rho\right) dx}. \tag{20}$$

This integral (20) may be approximated with the Laplace integral method according to [1, 16, 35]. Thus, we may write

$$\int_{-\infty}^{\infty} g_1(x) \cdot \exp\left(\frac{-\psi(x)}{\rho} \right) dx$$

$$\cong \exp\left(\frac{-\psi(x)}{\rho} \right) \sqrt{ \frac{2\pi\rho}{\psi''(x_0)} } \left(g_1(x_0) + \frac{g_1''(x_0)}{2} \frac{\rho}{\psi''(x_0)} \right.$$

$$+ \frac{g_1''''(x_0)}{8} \left(\frac{\rho}{\psi''(x_0)} \right)^2$$

$$\left. + O\left(\frac{\rho}{\psi''(x_0)} \right)^3 \right)$$

$$\int_{-\infty}^{\infty} g(x) \cdot \exp\left(\frac{-\psi(x)}{\rho}\right) dx$$

$$\cong \exp\left(\frac{-\psi(x)}{\rho}\right) \sqrt{\frac{2\pi\rho}{\psi''(x_0)}} \left(g(x_0) + \frac{g''(x_0)}{2} \frac{\rho}{\psi''(x_0)}\right.$$

$$+ \frac{g''''(x_0)}{8}\left(\frac{\rho}{\psi''(x_0)}\right)^2$$

$$\left. + O\left(\frac{\rho}{\psi''(x_0)}\right)^3\right),$$

(21)

where $O(x)$ is defined as $\lim_{x \to 0}(O(x)/x) = c$, and c is a constant and

$$\psi'(x) = -2(z - x); \qquad \psi''(x) = 2;$$
$$\psi'(x) = 0 \implies x_0 = z.$$

(22)

By substituting (21) into (20) and dividing the numerator and denominator by $g(x_0)$, using (22), we obtain

$$E[x \mid z] \cong \left(z + \frac{g_1''(z)}{2g(z)}\left(\sigma_z^2 - \sigma_x^2\right)\right.$$

$$+ \frac{g_1''''(z)}{8g(z)}\left(\sigma_z^2 - \sigma_x^2\right)^2\right)$$

$$\times \left(1 + \frac{g''(z)}{2g(z)}\left(\sigma_z^2 - \sigma_x^2\right)\right.$$

$$\left. + \frac{g''''(z)}{8g(z)}\left(\sigma_z^2 - \sigma_x^2\right)^2\right)^{-1}.$$

(23)

Note that (23) is very similar to the approximated expression for the conditional expectation derived in [1]; however, $g(z)$ and its derivatives are very different. According to [11], the conditional mean estimate of the complex datum x ($x = x_1 + jx_2$) given the complex observation z ($z = z_1 + jz_2$) can be written as

$$E[x \mid z] = E[x_1 \mid z_1] + jE[x_2 \mid z_2].$$

(24)

This completes our Proof. □

3.2. MSE and Lagrange Multipliers. Since we deal with the real valued or two independent quadrature carrier case, it is enough to derive the MSE as well as the expression for the Lagrange multipliers for the real valued case only. Thus, we consider in the following only the real valued case.

In this section we search for those Lagrange multipliers that bring the MSE to a minimum as was done in [1]:

$$\frac{d}{d\lambda_k}\left(E[\hat{x} - x]^2\right) = 0,$$

(25)

where $\hat{x}$ is the conditional expectation given in (23).

The approximated MSE according to [1] is given by

$$E[\hat{x} - x]^2 \cong \frac{\sigma_p^2}{\left(1 + \sigma_p^2\left(E[g''(x)/2g(x)]\right)\right)^2},$$

(26)

where $\hat{x}$ is the conditional expectation given in [1].

Although $g(z)$, $g''(z)$, $g''''(z)$, $g_1(z)$, $g_1''(z)$, and $g_1''''(z)$ are very different from those obtained in [1], (23) looks quite similar to the approximated expression for the conditional expectation derived in [1]. Thus, it make sense to use in this paper the same technique for deriving the MSE as was carried out in [1]. The convolutional noise pdf (9) is for $p = 0$ the pdf used in [1] (for $p = 0$) multiplied by a constant parameter $(1+3A-15B)$, where A and B are given in Appendix B in (B.6) and (B.7), respectively. Therefore, based on (94), (95), and (96) in [1], the MSE for our case is the same MSE given in (26) multiplied by $(1 + 3A - 15B)$ with our expression for $E[(g''(x)/2g(x))_{p=0}]$. Thus, we obtain

$$E[\hat{x} - x]^2 \cong \frac{\sigma_p^2(1 + 3A - 15B)}{\left(1 + \sigma_p^2\left(E[g''(x)/2g(x)]_{p=0}\right)\right)^2}.$$

(27)

According to [36], for small values of σ_p^2 such that $|\sigma_p^2(E[g''(x)/2g(x)]_{p=0})| < 1$ we obtain

$$E[\hat{x} - x]^2 \cong (1 + 3A - 15B)$$

$$\times \frac{\sigma_p^2}{\left(1 + \sigma_p^2\left(E[g''(x)/2g(x)]_{p=0}\right)\right)^2}$$

$$\cong (1 + 3A - 15B)$$

$$\times \left(\sigma_p^2\left(1 - \sigma_p^2\left(2E\left[\frac{g''(x)}{2g(x)}\right]_{p=0}\right)\right)\right).$$

(28)

Next we minimize (28) with respect to the Lagrange multipliers, λ_k,

$$\frac{d}{d\lambda_k}(1 + 3A - 15B)\sigma_p^2\left(1 - \sigma_p^2\left(2E\left[\frac{g''(x)}{2g(x)}\right]_{p=0}\right)\right) = 0$$

$$\implies \frac{d}{d\lambda_k}\left(E\left[\frac{g''(x)}{2g(x)}\right]_{p=0}\right) = 0.$$

(29)

In Appendix E we derive (29) and obtain

$$m_{k-2}k(k-1) + \frac{2}{(1 + 3A - 15B)}$$

$$\times \sum_{l=2, l \neq k}^{K}\left(\lambda_l klm_{k+l-2}\right) + \frac{2\lambda_k k^2 m_{2k-2}}{1 + 3A - 15B} = 0,$$

(30)

where $m_k = E[x^k]$.

Please refer to Appendices C and D for explicit expressions for the Lagrange multipliers valid for the 16QAM and 64QAM input constellation case, respectively. It should be pointed out that (30) is the same expression obtained in [1] for $A = 0$ and $B = 0$.

4. Simulation

In this section we test and compare the equalization performance obtained for the 16QAM and 64QAM input case with the convolutional noise pdf (6) to the case where the convolutional noise pdf is modeled as Gaussian. In addition, we examine whether the optimal Lagrange multipliers have a significant effect on the quality of the equalization performance and determine whether the transition from order four to order six in the Edgeworth Expansion series is worthwhile despite the increase in complexity in the algorithm.

According to [1], where the convolutional noise pdf was modeled as a Gaussian pdf, the equalizer's taps were updated as

$$c_l[n+1] = c_l[n] + \mu W y^*[n-l] \quad \text{with}$$

$$W = \left[\left(E[x_1 \mid z_1] \left[\frac{z_1(n) E[x_1 \mid z_1]}{\langle (z_1)^2 \rangle_n} \right] + j E[x_2 \mid z_2] \right) \right.$$

$$\left. \times \left[\frac{z_2(n) E[x_2 \mid z_2]}{\langle (z_2)^2 \rangle_n} \right] \right) - z(n) \right] \tag{31}$$

with (12) and (30) for $A = 0$, $B = 0$, and

$$\langle z_s^2 \rangle_n = (1 - \beta) \langle z_s^2 \rangle_{n-1} + \beta (z_s)_n^2, \tag{32}$$

where $\langle \rangle$ stands for the estimated expectation operator, β and μ are positive step-size parameters, $\langle z_s^2 \rangle_0 > 0$, and l stands for lth tap of the equalizer and $s = 1, 2$. In the following, this equalizer [1] will be denoted as "MaxEnt." To illustrate the equalization performance with the convolutional noise pdf (6) we used the equalizer's update mechanism given in (31) and (32) with the conditional expectation presented in (12) for the following settings.

For $A \neq 0$, $B = 0$ the equalizer will be called "Edgeworth order 4."

For $A \neq 0$, $B \neq 0$ the equalizer will be called "Edgeworth order 6."

For "Edgeworth order 4" we need (B.6) and for "Edgeworth order 6" we also need (B.7). In this work we used for $E[p^4]$ and $E[p^6]$ the following settings:

$$E[p^4] = a(\sigma_p^2)^2, \qquad E[p^6] = b(\sigma_p^2)^3, \tag{33}$$

where a and b were chosen for fast convergence speed and low steady state ISI.

The parameters a and b are denoted in the following as a_{EE4} for "Edgeworth order 4" ($b = 0$ for "Edgeworth order 4") and a_{EE6}, b_{EE6} for "Edgeworth order 6."

Table 1

Method	"Godard"	"MaxEnt"	"Edgeworth order 4"	"Edgeworth order 6"
Convolutional noise pdf		Gaussian	According to (10)	According to (9)

The step-size parameter μ is denoted in the following as μ_{MaxEnt}, μ_{EE4}, and μ_{EE6} for "MaxEnt," "Edgeworth order 4," and "Edgeworth order 6," respectively. In the following, the β parameter is denoted as the β_{MaxEnt}, β_{EE4}, and β_{EE6} for "MaxEnt," "Edgeworth order 4," and "Edgeworth order 6," respectively. According to [1] the denominator of (23) cannot be zero. Therefore, during our simulations, the equalizer's taps were updated only if the denominator was greater than ε ($\varepsilon > 0$). For the equalization performance comparison we also used the algorithm defined by Godard [7]. The equalizer's taps for Godard's algorithm [7] were updated according to

$$c_l[n+1] = c_l[n] - \mu_G \left(|z[n]|^2 - \frac{E\left[|x[n]|^4\right]}{E\left[|x[n]|^2\right]} \right)$$

$$\times z[n] y^*[n-l], \tag{34}$$

where μ_G is the step-size and $| \, |$ is the absolute operator.

Two different sources were considered:

16QAM and 64QAM sources (modulations using $\pm\{1, 3\}$, $\pm\{1, 3, 5, 7\}$ levels respectively, for in-phase and quadrature components).

Two different channels were considered:

channel case 1 (initial ISI = 0.44), where the channel parameters were taken according to Shalvi and Weinstein [17]: $h_n = \{0 \text{ for } n < 0; -0.4 \text{ for } n = 0; 0.84 \cdot 0.4^{n-1} \text{ for } n > 0\}$,

channel case 2 (initial ISI = 1.402), where the channel parameters were taken according to [18]: $h_n = \{0.2258, 0.5161, 0.6452, 0.5161\}$.

For channel case 1 and channel case 2 we used an equalizer with 13 and 21 taps, respectively. The equalizers where initialized by setting the center tap equal to one and all others to zero.

The step-size parameters, μ_G, μ_{MaxEnt}, β_{MaxEnt}, μ_{EE4}, β_{EE4}, μ_{EE6}, and β_{EE6} were chosen for fast convergence speed and low steady state ISI.

Figures 2–4 show the simulated equalization performance (ISI as a function of iteration number) of three equalization methods: "Godard," "MaxEnt," "Edgeworth order 4," and "Edgeworth order 6" for channel case 1, 16QAM, and SNR = 30 dB.

Table 1 lists the various equalization methods and the corresponding model for the convolutional noise pdf used in the algorithm.

Two cases were considered for "Edgeworth order 4":

Case a. Nonoptimal Lagrange multipliers (Lagrange multipliers similar to those used in "MaxEnt" [1]).

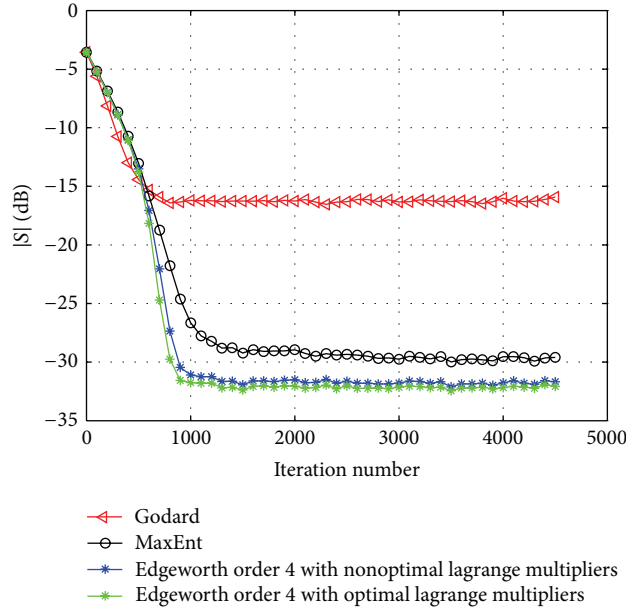*Case b.* Optimal Lagrange multipliers (according to Appendix C).

FIGURE 2: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to $\mu_G = 1 \times 10^{-4}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-4}$. For Edgeworth order 4 with nonoptimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 4.3 \times 10^{-4}$, $\beta_{\text{EE4}} = 4.5 \times 10^{-4}$, and for Edgeworth order 4 with optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 4.3 \times 10^{-4}$, and $\beta_{\text{EE4}} = 4.5 \times 10^{-4}$. In addition we set for MaxEnt and Edgeworth order 4 with nonoptimal Lagrange multipliers $\varepsilon = 0$. For Edgeworth order 4 with optimal Lagrange multipliers we set $\varepsilon = 0.5$.

Based on the simulation results (Figures 2–4), "Edgeworth order 4" and "Edgeworth order 6" achieve better equalization performance in terms of convergence speed and lower residual ISI compared to the other methods.

According to Figure 2, "Edgeworth order 4" with optimal Lagrange multipliers has a slightly faster convergence speed and a slightly lower residual ISI compared to the non-optimal Lagrange multipliers case. According to Figure 3, "Edgeworth order 6" with optimal Lagrange multipliers leads to a faster convergence speed compared with "Edgeworth order 4" with nonoptimal Lagrange multipliers but leads the system to a slightly higher residual ISI. Figure 4 shows that "Edgeworth order 4" and "Edgeworth order 6" with optimal Lagrange multipliers have both approximately the same convergence speed while "Edgeworth order 4" leads the system to a lower residual ISI compared to "Edgeworth order 6". Figures 2–4 indicate that there is no justification using "Edgeworth order 6" over "Edgeworth order 4" and "Edgeworth order 4" with optimal Lagrange multipliers over "Edgeworth order 4" with nonoptimal Lagrange multipliers for channel case 1.

Figures 5, 6, and 7 indicate that our previously made conclusion based on Figures 2–4 also holds when the SNR is decreased.

Figures 8–10 show the simulated equalization performance (ISI as a function of iteration number) of four equalization methods: "Godard," "MaxEnt," "Edgeworth order 4," and "Edgeworth order 6" for channel case 2, 16QAM, and SNR = 30 dB. Based on Figures 8–10, "Edgeworth order 4" and "Edgeworth order 6" have better equalization performance in terms of convergence speed compared to "MaxEnt" but lead the system to a slightly higher residual ISI compared to the "MaxEnt" method.

Figure 8 indicates that "Edgeworth order 4" with optimal Lagrange multipliers leads to a slightly faster convergence speed compared to the nonoptimal Lagrange multipliers case while leaving the system with approximately the same residual ISI.

Figures 9 and 10 show that "Edgeworth order 6" with optimal Lagrange multipliers leads to a much faster convergence speed compared to the optimal and nonoptimal Lagrange multiplier case with "Edgeworth order 4," while having approximately the same residual ISI for both cases.

Figure 11 shows the simulated equalization performance (ISI as a function of iteration number) of three equalization methods: "MaxEnt," "Edgeworth order 4," and "Edgeworth order 6" for channel case 1, 64QAM constellation input and for the noiseless case.

Three optimal Lagrange multipliers were used for "Edgeworth order 4" and for "Edgeworth order 6" (please refer to Appendix D).

Based on the simulation results (Figure 11), a much faster convergence speed is obtained by using "Edgeworth order 6" with optimal Lagrange multipliers compared to the case where "Edgeworth order 4" with optimal Lagrange multipliers or the "MaxEnt" method is used.

5. Conclusion

In this paper we used the Edgeworth expansion up to order 4 and up to order 6 to model the convolutional noise pdf
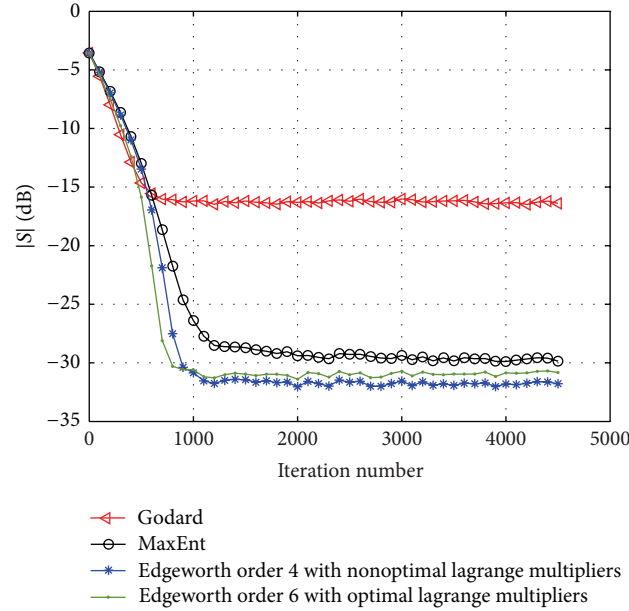
FIGURE 3: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to: $\mu_G = 1 \times 10^{-4}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-4}$. For Edgeworth order 4 with non-optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 4.3 \times 10^{-4}$, $\beta_{\text{EE4}} = 4.5 \times 10^{-4}$, and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 4.5$, $b_{\text{EE6}} = 38$, $\mu_{\text{EE6}} = 5.2 \times 10^{-4}$, and $\beta_{\text{EE6}} = 4.5 \times 10^{-4}$. In addition we set for MaxEnt and Edgeworth order 4 with non-optimal Lagrange multipliers $\varepsilon = 0$. For Edgeworth order 6 with optimal Lagrange multipliers we set $\varepsilon = 0.5$.
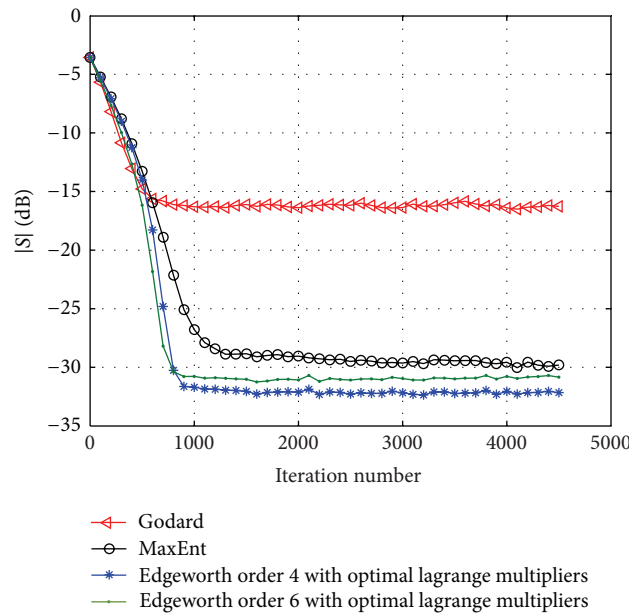

FIGURE 4: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to: $\mu_G = 1 \times 10^{-4}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-4}$. For Edgeworth order 4 with optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 4.3 \times 10^{-4}$, and $\beta_{\text{EE4}} = 4.5 \times 10^{-4}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 4.5$, $b_{\text{EE6}} = 38$, $\mu_{\text{EE6}} = 5.2 \times 10^{-4}$, and $\beta_{\text{EE6}} = 4.5 \times 10^{-4}$. In addition we set for MaxEnt $\varepsilon = 0$, and for Edgeworth order 4 and order 6 $\varepsilon = 0.5$.
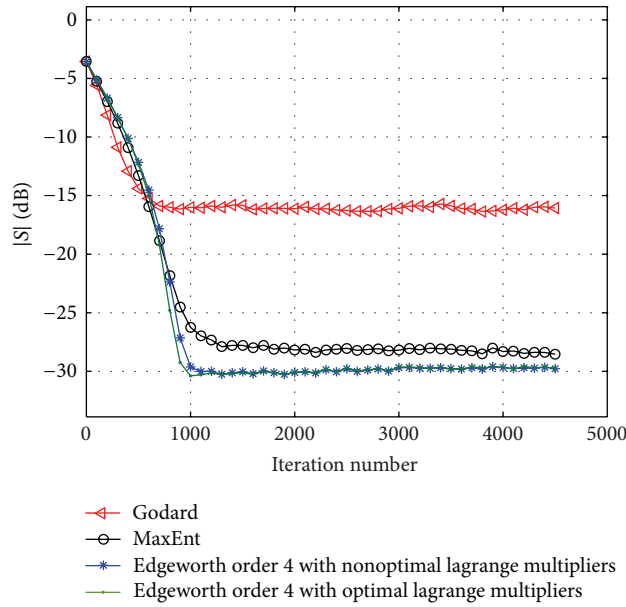
FIGURE 5: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 26 dB. The step-size parameters were set to: $\mu_G = 1 \times 10^{-4}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-4}$. For Edgeworth order 4 with non-optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE4}} = 4 \times 10^{-4}$ and for Edgeworth order 4 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 4.5$, $b_{\text{EE6}} = 32$, $\mu_{\text{EE6}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE6}} = 4 \times 10^{-4}$. In addition we set for MaxEnt and Edgeworth order 4 with optimal and nonoptimal Lagrange multipliers $\varepsilon = 0.5$.
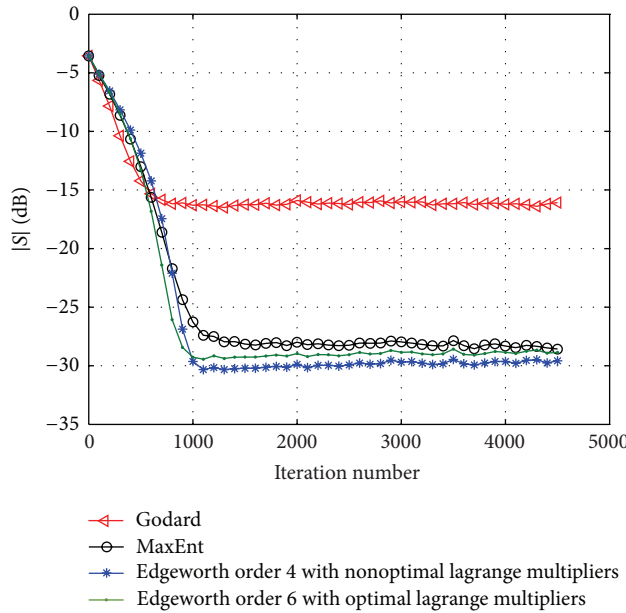

FIGURE 6: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 26 dB. The step-size parameters were set to: $\mu_G = 1 \times 10^{-4}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-4}$. For Edgeworth order 4 with non-optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE4}} = 4 \times 10^{-4}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 4.5$, $b_{\text{EE4}} = 32$, $\mu_{\text{EE6}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE6}} = 4 \times 10^{-4}$. In addition we set for MaxEnt, Edgeworth order 4, and Edgeworth order 6 $\varepsilon = 0.5$.
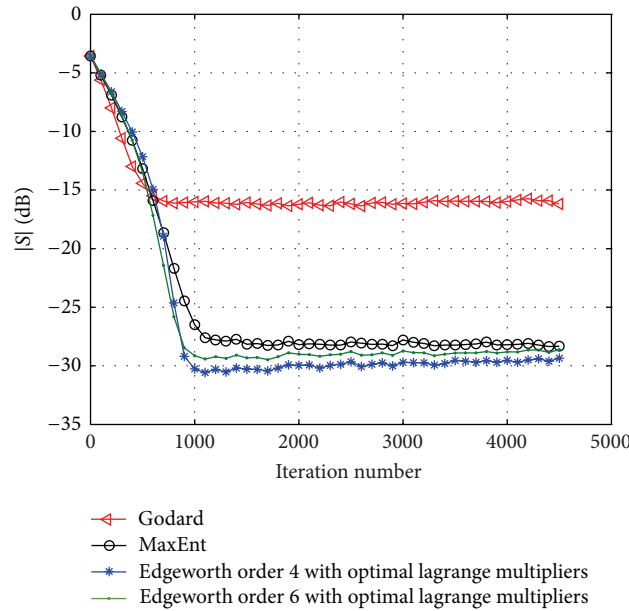
FIGURE 7: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for SNR of 26 dB. The step-size parameters were set to: $\mu_G = 1 \times 10^{-5}$, $\mu_{\text{MaxEnt}} = 4 \times 10^{-4}$, $\beta_{\text{MaxEnt}} = 2 \times 10^{-6}$. For Edgeworth order 4 with non-optimal Lagrange multipliers we used $a_{\text{EE4}} = 4.5$, $\mu_{\text{EE4}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE4}} = 4 \times 10^{-4}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 4.5$, $b_{\text{EE6}} = 32$, $\mu_{\text{EE6}} = 3.7 \times 10^{-4}$, and $\beta_{\text{EE6}} = 4 \times 10^{-4}$. In addition we set for MaxEnt, Edgeworth order 4, and Edgeworth order 6 $\varepsilon = 0.5$.
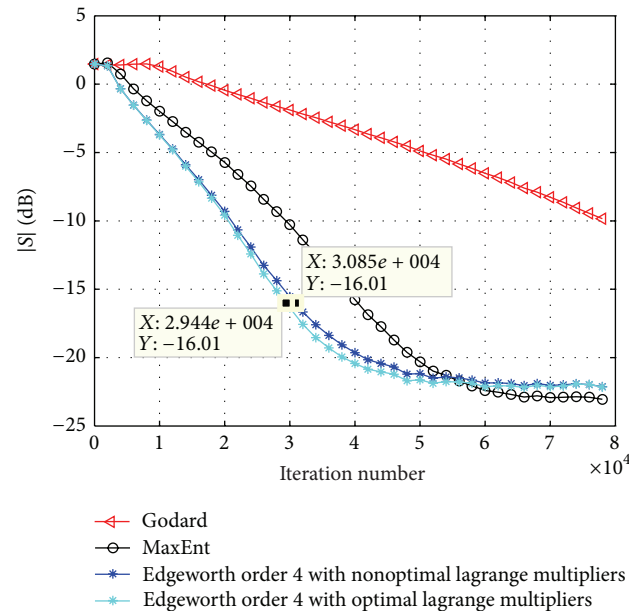

FIGURE 8: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 2. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to $\mu_G = 1 \times 10^{-5}$, $\mu_{\text{MaxEnt}} = 2 \times 10^{-4}$, and $\beta_{\text{MaxEnt}} = 2 \times 10^{-6}$. For Edgeworth order 4 with nonoptimal Lagrange multipliers we used $a_{\text{EE4}} = 5$, $\mu_{\text{EE4}} = 3 \times 10^{-4}$, and $\beta_{\text{EE4}} = 2 \times 10^{-6}$ and for Edgeworth order 4 with optimal Lagrange multipliers we used $a_{\text{EE4}} = 4$, $\mu_{\text{EE4}} = 3 \times 10^{-4}$, and $\beta_{\text{EE4}} = 2 \times 10^{-6}$. In addition we set for MaxEnt, Edgeworth order 4 with optimal and nonoptimal Lagrange multipliers $\varepsilon = 0.5$.
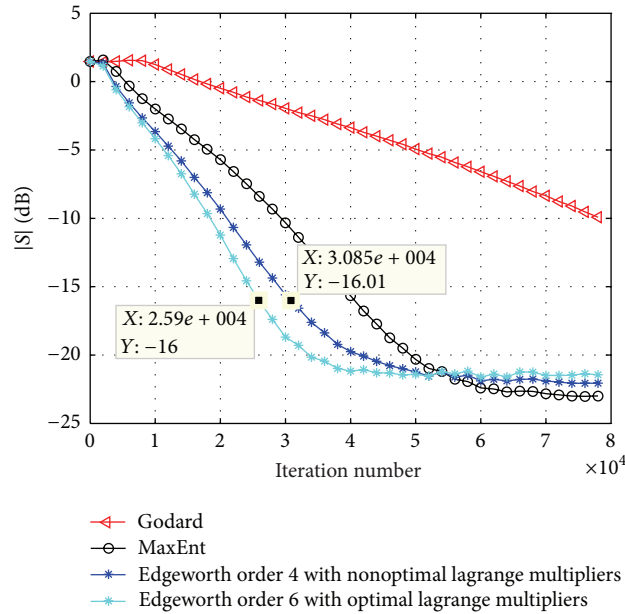
Figure 9: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 2. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to $\mu_G = 1 \times 10^{-5}$, $\mu_{MaxEnt} = 2 \times 10^{-4}$, and $\beta_{MaxEnt} = 2 \times 10^{-6}$. For Edgeworth order 4 with nonoptimal Lagrange multipliers we used $a_{EE4} = 5$, $\mu_{EE4} = 3 \times 10^{-4}$, and $\beta_{EE4} = 2 \times 10^{-6}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{EE6} = 4.1$, $b_{EE6} = 25$, $\mu_{EE6} = 3.3 \times 10^{-4}$, and $\beta_{EE6} = 2.1 \times 10^{-6}$. In addition we set for MaxEnt, Edgeworth order 4, and Edgeworth order 6 $\varepsilon = 0.5$.

Figure 10: Performance comparison between equalization algorithms simulated for a 16QAM input source going through channel case 2. The averaged results were obtained in 100 Monte Carlo trials for SNR of 30 dB. The step-size parameters were set to $\mu_G = 1 \times 10^{-5}$, $\mu_{MaxEnt} = 2 \times 10^{-4}$, and $\beta_{MaxEnt} = 2 \times 10^{-6}$. For Edgeworth order 4 with optimal Lagrange multipliers we used $a_{EE4} = 4$, $\mu_{EE4} = 3 \times 10^{-4}$, and $\beta_{EE4} = 2 \times 10^{-6}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{EE6} = 4.1$, $b_{EE6} = 25$, $\mu_{EE6} = 3.3 \times 10^{-4}$, and $\beta_{EE6} = 2.1 \times 10^{-6}$. In addition we set for Edgeworth orders 4 and 6 $\varepsilon = 0.5$.
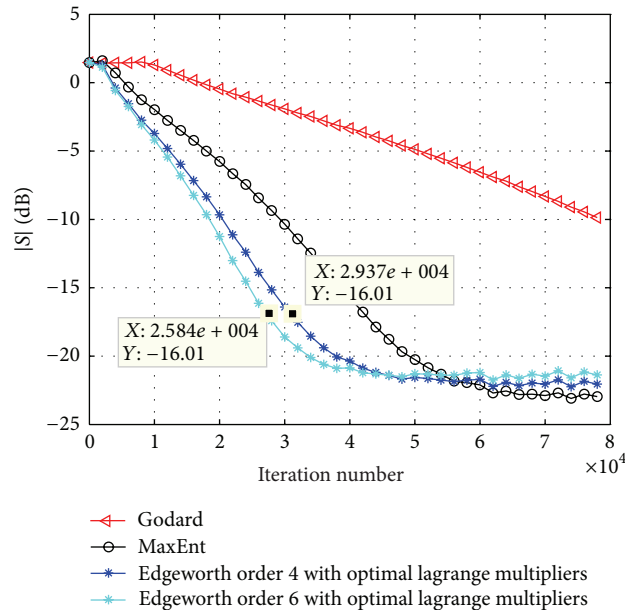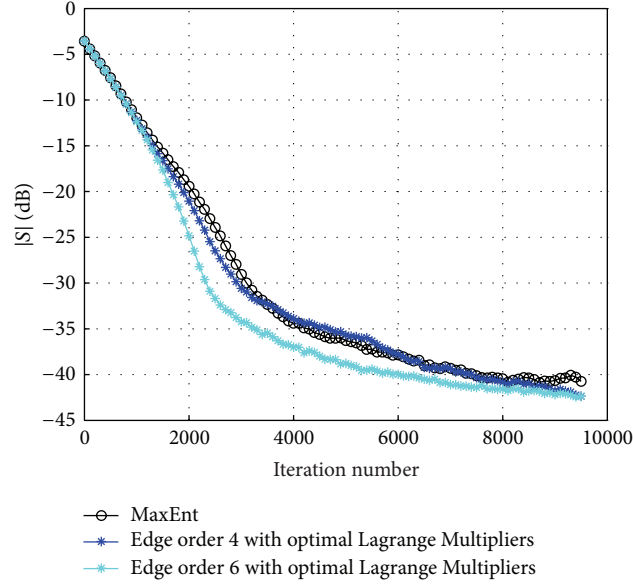
FIGURE 11: Performance comparison between equalization algorithms simulated for a 64QAM input source going through channel case 1. The averaged results were obtained in 100 Monte Carlo trials for the noiseless case. The step-size parameters were set to: $\mu_{\text{MaxEnt}} = 5.5 \times 10^{-5}$, $\beta_{\text{MaxEnt}} = 6 \times 10^{-5}$. For Edgeworth order 4 with optimal Lagrange multipliers we used $a_{\text{EE4}} = 2.4$, $\mu_{\text{EE4}} = 6.2 \times 10^{-5}$, and $\beta_{\text{EE4}} = 6 \times 10^{-5}$ and for Edgeworth order 6 with optimal Lagrange multipliers we used $a_{\text{EE6}} = 2.4$, $b_{\text{EE6}} = 4$, $\mu_{\text{EE6}} = 6.2 \times 10^{-5}$, and $\beta_{\text{EE6}} = 6 \times 10^{-5}$. In addition we set for all methods $\varepsilon = 0.7$.

in the conditional expectation calculations where the source pdf was approximated according to the maximum entropy density approximation technique. We derived new Lagrange multipliers and obtained new closed-form approximated expressions for the conditional expectation and MSE as byproducts. According to simulation results for the 16QAM constellation input, there is no justification for using "Edgeworth order 4" with optimal Lagrange multipliers over "Edgeworth order 4" with nonoptimal Lagrange multipliers when dealing with easy or hard channels. In addition, hardly any equalization performance improvement was observed with "Edgeworth order 6" (with optimal Lagrange multipliers) compared to "Edgeworth order 4" with optimal and nonoptimal Lagrange multipliers for the easy channel case (channel case 1). However, a much faster convergence speed was observed with "Edgeworth order 6" with optimal Lagrange multipliers for the hard channel case (channel case 2) compared to "Edgeworth order 4" with optimal and nonoptimal Lagrange multipliers. Furthermore, a much faster convergence speed was obtained for the 64QAM input and easy channel case with "Edgeworth order 6" compared to "Edgeworth order 4" both with optimal Lagrange multipliers.

Appendices

A. Hermite Polynomials and the One Dimensional Central Cumulants

A closed-form expression for the Hermite Polynomials can be defined as in [37]

$$H_r(\zeta) = \sum_{k=0}^{r/2} \frac{(-1)^k r!}{k! \, (r - 2k)} (2\zeta)^{r-2k}. \tag{A.1}$$

Specifically, the first Hermite Polynomials up to order 6 are

$$
\begin{aligned}
H_0(\zeta) &= 1, \\
H_1(\zeta) &= 2\zeta, \\
H_2(\zeta) &= 4\zeta^2 - 2, \\
H_3(\zeta) &= \zeta^3 - 3\zeta, \\
H_4(\zeta) &= \zeta^4 - 6\zeta^2 + 3, \\
H_5(\zeta) &= \zeta^5 - 10\zeta^3 + 15\zeta, \\
H_6(\zeta) &= \zeta^6 - 15\zeta^4 + 45\zeta^2 - 15.
\end{aligned}
\tag{A.2}
$$

According to [34, Table 3], the one-dimensional central cumulants in terms of central moments are

$$
\begin{aligned}
K_3(\zeta) &= E\left[\zeta^3\right], \\
K_4(\zeta) &= E\left[\zeta^4\right] - 3\left(E\left[\zeta^2\right]\right)^2, \\
K_5(\zeta) &= E\left[\zeta^5\right] - 10 E\left[\zeta^2\right] E\left[\zeta^3\right], \\
K_6(\zeta) &= E\left[\zeta^6\right] - 15 E\left[\zeta^2\right] E\left[\zeta^4\right] \\
&\quad + 30\left(E\left[\zeta^2\right]\right)^3 - 10\left(E\left[\zeta^3\right]\right)^2.
\end{aligned}
\tag{A.3}
$$

B. Expressions for $g''(z)$, $g''''(z)$, $g_1''(z)$, $g_1''''(z)$, A, B, $K_4(p/\sigma_p)$, $K_6(p/\sigma_p)$, $H_4(p/\sigma_p)$, and $H_6(p/\sigma_p)$

In order to facilitate the derivation process of (17) and (18) we divide (17) into two parts:

$$g(x) = h(x) t(x), \tag{B.1}$$

where

$$t(x) = \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right), \tag{B.2}$$

and

$$
\begin{aligned}
h(x) &= \left(1 + \frac{K_4\left(p/\sigma_p\right) H_4\left(p/\sigma_p\right)}{4!} \right. \\
&\quad \left. + \frac{K_6\left(p/\sigma_p\right) H_6\left(p/\sigma_p\right)}{6!}\right) \\
&= \left(1 + \frac{K_4\left((z-x)/\sigma_p\right) H_4\left((z-x)/\sigma_p\right)}{4!} \right. \\
&\quad \left. + \frac{K_6\left((z-x)/\sigma_p\right) H_6\left((z-x)/\sigma_p\right)}{6!}\right).
\end{aligned}
\tag{B.3}
$$

According to (A.3), the values for $H_4(p/\sigma_p)$ and $H_6(p/\sigma_p)$ are

$$
H_4\left(\frac{p}{\sigma_p}\right) = \left(\frac{(z-x)^4}{\left(\sigma_p^2\right)^2} - \frac{6(z-x)^2}{\sigma_p^2} + 3\right);
$$

$$
\left[H_4\left(\frac{p}{\sigma_p}\right)\right]_{x=z} = 3,
$$

$$
H_6\left(\frac{p}{\sigma_p}\right) = \left(\frac{(z-x)^6}{\left(\sigma_p^2\right)^3} - \frac{15(z-x)^4}{\left(\sigma_p^2\right)^2} + \frac{45(z-x)^2}{\left(\sigma_p^2\right)} - 15\right);
$$

$$
\left[H_6\left(\frac{p}{\sigma_p}\right)\right]_{x=z} = -15.
\tag{B.4}
$$

Their derivatives may be expressed as

$$
\left[\frac{d^2}{dx^2}\left(H_4\left(\frac{p}{\sigma_p}\right)\right)\right]_{x=z} = -\frac{12}{\sigma_p^2};
$$

$$
\left[\frac{d^4}{dx^4}\left(H_4\left(\frac{p}{\sigma_p}\right)\right)\right]_{x=z} = \frac{24}{\sigma_p^4}
$$

$$
\left[\frac{d^2}{dx^2}\left(H_6\left(\frac{p}{\sigma_p}\right)\right)\right]_{x=z} = \frac{90}{\sigma_p^2};
$$

$$
\left[\frac{d^4}{dx^4}\left(H_6\left(\frac{p}{\sigma_p}\right)\right)\right]_{x=z} = \frac{-360}{\sigma_p^4}.
\tag{B.5}
$$

For $K_4(p/\sigma_p)$ and $K_6(p/\sigma_p)$ as described in (A.3) we define the following constants (results after normalization):

$$
A = \frac{K_4\left(p/\sigma_p\right)}{4!} = \frac{E\left[p^4\right] - 3\left(\sigma_p^2\right)^2}{4!\left(\sigma_p^2\right)^2}, \tag{B.6}
$$

$$
B = \frac{K_6\left(p/\sigma_p\right)}{6!} = \frac{E\left[p^6\right] - 15\left(\sigma_p^2\right) E\left[p^4\right] + 30\left(\sigma_p^2\right)^3}{6!\left(\sigma_p^2\right)^3}. \tag{B.7}
$$

We set $g(z)$ as follows:

$$
g(z) = [g(x)]_{x=z} = \exp\left(\sum_{k=0}^{K} \lambda_k z^k\right)(1 + 3A - 15B). \tag{B.8}
$$

The second and fourth derivatives of $g(z)$ are as follows:

$$
\begin{aligned}
g''(z) &= \left[h''(p) t(x) + 2h'(p) t'(x) + h(p) t''(x)\right]_{x=z}, \\
g''''(z) &= \left[h''''(p) t(x) + 4h'''(p) t'(x) + 6h''(p) t''(x)\right. \\
&\quad \left. + 4h'(p) t'''(x) + h(p) t''''(x)\right]_{x=z} \\
&= \left[h''''(p) t(x) + 4h'''(p) t'(x) + 6h''(p) t''(x)\right. \\
&\quad \left. + 4h'(p) t'''(x) + h(p) t''''(x)\right]_{x=z}.
\end{aligned}
\tag{B.9}
$$

The definitions for $g_1(x)$ and its derivatives are

$$
\begin{aligned}
g_1(x) &= xg(x) \implies g_1(z) = zg(z), \\
g_1''(z) &= 2g'(z) + zg''(z), \\
g_1''''(z) &= 4g'''(z) + zg''''(z).
\end{aligned}
\tag{B.10}
$$

By substituting (B.9), (B.10) in (23) the closed-form approximated expression for the conditional expectation is obtained.

C. Lagrange Multipliers Equations for 16QAM

For the 16QAM case we use two Lagrange multipliers (λ_2, λ_4); thus, we have $K = 4$.

In order to get the required Lagrange multipliers we recall (30)

$$
\begin{aligned}
&m_{k-2}(k-1)k + \frac{2}{(1 + 3A - 15B)} \\
&\quad \times \sum_{l=2,l\neq k}^{K}\left(\lambda_l klm_{k+l-2}\right) + \frac{2\lambda_k k^2 m_{2k-2}}{1 + 3A - 15B} = 0,
\end{aligned}
\tag{C.1}
$$

where we use A and B from (B.6) and (B.7), respectively. For λ_2 we have $k = 2$; thus, when substituting $k = 2$ and $K = 4$ into (30) we get

$$m_0 (2 - 1) 2 + \frac{2\lambda_2 m_{4-2} 2^2}{1 + 3A - 15B} + \frac{2\lambda_4 m_{2+4-2} \cdot 2 \cdot 4}{1 + 3A - 15B} = 0$$

$$\Downarrow \qquad\qquad \text{(C.2)}$$

$$2 + 8\lambda_2 \frac{m_2}{1 + 3A - 15B} + 16\lambda_4 \frac{m_4}{1 + 3A - 15B} = 0.$$

For λ_4 we have $k = 4$; thus, when substituting $k = 4$ and $K = 4$ into (30) we get

$$m_{4-2} (4 - 1) 4 + \frac{2\lambda_4 m_6 4^2}{1 + 3A - 15B} + \frac{2\lambda_2 m_{4+2-2} \cdot 4 \cdot 2}{1 + 3A - 15B} = 0$$

$$\Downarrow$$

$$2m_2 + 32\lambda_4 \frac{m_6}{1 + 3A - 15B} + 16\lambda_2 \frac{m_4}{1 + 3A - 15B} = 0. \qquad \text{(C.3)}$$

$$\left(-\frac{1}{4} \frac{-15m_6 B - 3m_4 m_2 - 9m_4 m_2 A + 45m_4 m_2 B + m_6 + 3m_6 A}{m_2 m_6 - m_4^2} \right.$$
$$\left. \frac{1}{8} \frac{-15m_4 B - 3m_2^2 - 9m_2^2 A + 45m_2^2 B + m_4 + 3m_4 A}{m_2 m_6 - m_4^2} \right). \qquad \text{(C.6)}$$

which can be written as:

$$\lambda_2 = -\frac{625}{96}B + \frac{125}{288} + \frac{125}{96}A,$$

$$\lambda_4 = \frac{85}{192}B - \frac{17}{576} - \frac{17}{192}A. \qquad \text{(C.8)}$$

By substituting $A = 0$ and $B = 0$ into (C.8), we get the Lagrange multipliers for the 16QAM input as appeared in [1].

From (C.2) and (C.3) we get a linear system with two equations:

$$8\lambda_2 \frac{m_2}{1 + 3A - 15B} + 16\lambda_4 \frac{m_4}{1 + 3A - 15B} + 2 = 0,$$

$$16\lambda_2 \frac{m_4}{1 + 3A - 15B} + 32\lambda_4 \frac{m_6}{1 + 3A - 15B} + 12m_2 = 0, \qquad \text{(C.4)}$$

which can be written by

$$\begin{pmatrix} 8\dfrac{m_2}{1 + 3A - 15B} & 16\dfrac{m_4}{1 + 3A - 15B} \\ 16\dfrac{m_4}{1 + 3A - 15B} & 32\dfrac{m_6}{1 + 3A - 15B} \end{pmatrix} \begin{pmatrix} \lambda_2 \\ \lambda_4 \end{pmatrix}$$
$$= \begin{pmatrix} -2 \\ -12m_2 \end{pmatrix}, \qquad \text{(C.5)}$$

where the solution is

Next, we substitute the known moments $m_2 = 5$, $m_4 = 41$, and $m_6 = 365$ into (C.5) and obtain

$$\begin{pmatrix} \lambda_2 \\ \lambda_4 \end{pmatrix} = \left(-\frac{1}{4} \frac{-15m_6 B - 3m_4 m_2 - 9m_4 m_2 A + 45m_4 m_2 B + m_6 + 3m_6 A}{m_2 m_6 - m_4^2} \right.$$
$$\left. \frac{1}{8} \frac{-15m_4 B - 3m_2^2 - 9m_2^2 A + 45m_2^2 B + m_4 + 3m_4 A}{m_2 m_6 - m_4^2} \right) \qquad \text{(C.7)}$$

$$= \begin{pmatrix} -\dfrac{625}{96}B + \dfrac{125}{288} + \dfrac{125}{96}A \\ \dfrac{85}{192}B - \dfrac{17}{576} - \dfrac{17}{192}A \end{pmatrix}$$

D. Lagrange Multipliers Equations for 64QAM

For the 64QAM case we use three Lagrange multipliers $(\lambda_2, \lambda_4, \lambda_6)$; thus, we have $K = 6$.

In order to get the required Lagrange multipliers we recall (30)

$$m_{k-2} (k - 1) k + \frac{2}{(1 + 3A - 15B)}$$
$$\times \sum_{l=2, l \neq k}^{K} (\lambda_l k l m_{k+l-2}) + \frac{2\lambda_k k^2 m_{2k-2}}{1 + 3A - 15B} = 0, \qquad \text{(D.1)}$$

where we use A and B from (B.6) and (B.7), respectively.

For λ_2 we have $k = 2$; thus, when substituting $k = 2$ and $K = 6$ into (30) we get

$$m_0 (2 - 1) 2 + \frac{2\lambda_2 m_{4-2} 2^2}{1 + 3A - 15B} + \frac{2\lambda_4 m_{2+4-2} \cdot 2 \cdot 4}{1 + 3A - 15B}$$

$$+ \frac{2\lambda_6 m_{2+6-2} \cdot 2 \cdot 6}{1 + 3A - 15B} = 0$$

$$\Downarrow$$

$$2m_0 + 8\lambda_2 \frac{m_2}{1 + 3A - 15B}$$

$$+ 16\lambda_4 \frac{m_4}{1 + 3A - 15B} + 24\lambda_6 \frac{m_6}{1 + 3A - 15B} = 0.$$

$$\text{(D.2)}$$

For λ_4 we have $k = 4$; thus, when substituting $k = 4$ and $K = 6$ into (30) we get

$$m_{4-2} (4 - 1) 2 + \frac{2\lambda_4 m_{8-2} 4^2}{1 + 3A - 15B} + \frac{2\lambda_2 m_{4+2-2} \cdot 4 \cdot 2}{1 + 3A - 15B}$$

$$+ \frac{2\lambda_6 m_{4+6-2} \cdot 4 \cdot 6}{1 + 3A - 15B} = 0$$

$$\Downarrow$$

$$6m_2 + 32\lambda_4 \frac{m_6}{1 + 3A - 15B} + 16\lambda_2 \frac{m_4}{1 + 3A - 15B}$$

$$+ 48\lambda_6 \frac{m_8}{1 + 3A - 15B} = 0.$$

$$\text{(D.3)}$$

For λ_6 we have $k = 6$; thus, when substituting $k = 6$ and $K = 6$ into (30) we get

$$m_{6-2} (6 - 1) 6 + \frac{2\lambda_6 m_{2\cdot 6-2} 6^2}{1 + 3A - 15B} + \frac{2\lambda_2 m_{6+2-2} \cdot 6 \cdot 2}{1 + 3A - 15B}$$

$$+ \frac{2\lambda_4 m_{6+4-2} \cdot 6 \cdot 4}{1 + 3A - 15B} = 0$$

$$\Downarrow$$

$$30m_4 + 72\lambda_6 \frac{m_{10}}{1 + 3A - 15B} + 24\lambda_2 \frac{m_6}{1 + 3A - 15B}$$

$$+ 48\lambda_4 \frac{m_8}{1 + 3A - 15B} = 0.$$

$$\text{(D.4)}$$

From (D.2), (D.3), and (D.4) we get a linear system with three equations:

$$8\lambda_2 \frac{m_2}{1 + 3A - 15B} + 16\lambda_4 \frac{m_4}{1 + 3A - 15B}$$

$$+ 24\lambda_6 \frac{m_6}{1 + 3A - 15B} = -2m_0,$$

$$16\lambda_2 \frac{m_4}{1 + 3A - 15B} + 32\lambda_6 \frac{m_6}{1 + 3A - 15B}$$

$$+ 48\lambda_8 \frac{m_8}{1 + 3A - 15B} = 6m_2,$$

$$24\lambda_2 \frac{m_6}{1 + 3A - 15B} + 48\lambda_4 \frac{m_8}{1 + 3A - 15B}$$

$$+ 72\lambda_6 \frac{m_{10}}{1 + 3A - 15B} = 30m_4.$$

$$\text{(D.5)}$$

By substituting the known moments,

$$m_2 = 21, \qquad m_4 = 777, \qquad m_6 = 33501,$$

$$m_8 = 1540497, \qquad m_{10} = 73074981,$$

$$\text{(D.6)}$$

into (D.5) we obtain

$$\lambda_2 = -\frac{1417409}{2745600} A - \frac{1417409}{8236800} + \frac{1417409}{549120} B,$$

$$\lambda_4 = \frac{13573}{1647\,360} + \frac{13573}{549120} A - \frac{13573}{109824} B,$$

$$\lambda_6 = -\frac{211}{2246400} - \frac{211}{748800} A + \frac{211}{149760} B.$$

$$\text{(D.7)}$$

E. Derivation of $(d/d\lambda_k)E[(g''(x)/2g(x))_{p=0}] = 0$

By using (B.8) and (B.9) we can write

$$E\left[\left(\frac{g''(x)}{2g(x)} \right)_{p=0} \right]$$

$$= E\left[\frac{\left(-12\left(A/\sigma_p^2\right) - 90\left(B/\sigma_p^2\right)\right) \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right)}{2\left(1 + 3A - 15B\right) \exp\left(\sum_{k=0}^{K} \lambda_k x^k\right)} \right]$$

$$
\begin{aligned}
&+ E\Bigg[[1 + 3A - 15B] \\
&\quad \times \Bigg(\Bigg(\sum_{k=0}^{K} \left(\lambda_k x^{k-2} k^2 - \lambda_k x^{k-2} k \right) \Bigg) \exp\Bigg(\sum_{k=0}^{K} \lambda_k x^k \Bigg) \Bigg) \\
&\quad \times \Bigg(2(1 + 3A - 15B) \exp\Bigg(\sum_{k=0}^{K} \lambda_k x^k \Bigg) \Bigg)^{-1} \Bigg] \\
&+ E\Bigg[\frac{\left(\sum_{k=0}^{K} \lambda_k x^{k-1} k \right)^2 \exp\left(\sum_{k=0}^{K} \lambda_k x^k \right)}{2\left(1 + 3A - 15B \right) \exp\left(\sum_{k=0}^{K} \lambda_k x^k \right)} \Bigg] \\
&= E\Bigg[\frac{\left(-12\left(A/\sigma_p^2 \right) - 90\left(B/\sigma_p^2 \right) \right)}{2\left(1 + 3A - 15B \right)} \Bigg] \\
&\quad + E\Bigg[\frac{\left(\sum_{k=0}^{K} \left(\lambda_k x^{k-2} k^2 - \lambda_k x^{k-2} k \right) \right)}{2} \Bigg] \\
&\quad + E\Bigg[\frac{\left(\sum_{k=0}^{K} \lambda_k x^{k-1} k \right)^2}{2\left(1 + 3A - 15B \right)} \Bigg].
\end{aligned}
$$

(E.1)

Thus we have

$$
\begin{aligned}
&\frac{d}{d\lambda_k} E\Bigg[\left(\frac{g''(x)}{2g(x)} \right)_{p=0} \Bigg] \\
&= \frac{d}{d\lambda_k} \Bigg[E\Bigg[\frac{\left(\sum_{k=0}^{K} \left(\lambda_k x^{k-2} k^2 - \lambda_k x^{k-2} k \right) \right)}{2} \Bigg] \Bigg] \\
&\quad + \frac{d}{d\lambda_k} \Bigg[E\Bigg[\frac{\left(\sum_{k=0}^{K} \lambda_k x^{k-1} k \right)^2}{2\left(1 + 3A - 15B \right)} \Bigg] \Bigg].
\end{aligned}
$$

(E.2)

Next we set (E.2) to zero and obtain

$$
\begin{aligned}
&\frac{m_{k-2} k(k-1)}{2} + \frac{1}{2\left(1 + 3A - 15B \right)} \sum_{l=2, l \neq k}^{K} \left(\lambda_l k l m_{k+l-2} \right) \\
&\quad + \frac{\lambda_k k^2 m_{2k-2}}{2\left(1 + 3A - 15B \right)} = 0,
\end{aligned}
$$

(E.3)

or

$$
\begin{aligned}
&m_{k-2} k(k-1) + \frac{2}{\left(1 + 3A - 15B \right)} \\
&\quad \times \sum_{l=2, l \neq k}^{K} \left(\lambda_l k l m_{k+l-2} \right) + \frac{2\lambda_k k^2 m_{2k-2}}{1 + 3A - 15B} = 0.
\end{aligned}
$$

(E.4)

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] M. Pinchas and B. Z. Bobrovsky, "A maximum entropy approach for blind deconvolution," *Signal Processing*, vol. 86, no. 10, pp. 2913–2931, 2006.

[2] M. Pinchas, "Two blind adaptive equalizers connected in series for equalization performance improvement," *Journal of Signal and Information Processing*, vol. 4, pp. 64–71, 2013.

[3] C. L. Nikias and A. P. Petropulu, Eds., *Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework*, chapter 9, Prentice-Hall, 1993.

[4] M. Pinchas, *Blind Equalizers by Techniques of Optimal Non-Linear Filtering Theory*, VDM Verlagsservice gesellschaft mbH, 2009.

[5] M. Pinchas, *The Whole Story Behind Blind Adaptive Equalizers/Blind Deconvolution*, Bentham eBooks, 2012.

[6] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation systems," *IEEE Transactions on Communications*, vol. 23, no. 6, pp. 679–682, 1975.

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

[8] A. Benveniste, M. Goursat, and G. Ruget, "Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications," *IEEE Transactions on Automatic Control*, vol. 25, no. 3, pp. 385–399, 1980.

[9] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Transactions on Communications*, vol. 32, no. 8, pp. 871–883, 1984.

[10] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a "stop-and-go" decision-directed algorithm," *IEEE Transactions on Communications*, vol. 35, no. 9, pp. 877–887, 1987.

[11] S. Bellini, "Bussgang techniques for blind equalization," in *IEEE Global Telecommunication Conference Records*, pp. 1634–1640, December 1986.

[12] S. Bellini, "Blind equalization," *Alta Frequenza*, vol. 57, pp. 445–450, 1988.

[13] S. Fiori, "A contribution to (neuromorphic) blind deconvolution by flexible approximated Bayesian estimation," *Signal Processing*, vol. 81, no. 10, pp. 2131–2153, 2001.

[14] S. Fiori, "Fast fixed-point neural blind-deconvolution algorithm," *IEEE Transactions on Neural Networks*, vol. 15, no. 2, pp. 455–459, 2004.

[15] S. Haykin, "Blind deconvolution," in *Adaptive Filter Theory*, S. Haykin, Ed., vol. 20, chapter 20, Prentice-Hall, Englewood Cliffs, NJ, USA, 1991.

[16] M. Pinchas and B. Z. Bobrovsky, "A novel HOS approach for blind channel equalization," *IEEE Transactions on Wireless Communications*, vol. 6, no. 3, pp. 875–886, 2007.

[17] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Transactions on Information Theory*, vol. 36, no. 2, pp. 312–321, 1990.

[18] M. Lazaro, I. Santamaria, D. Erdogmus, K. E. Hild, C. Pantaleon, and J. C. Principe, "Stochastic blind equalization based on PDF

fitting using Parzen estimator," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, part 1, pp. 696–704, 2005.

[19] M. Pinchas, "A MSE optimized polynomial equalizer for 16QAM and 64QAM constellation," *Signal, Image and Video Processing*, vol. 5, no. 1, pp. 29–37, 2011.

[20] M. Pinchas, "What are the analytical conditions for which a blind equalizer will loose the convergence state?" *Signal, Image and Video Processing*, vol. 6, no. 2, pp. 325–340, 2012.

[21] R. Godfrey and F. Rocca, "Zero memory non-linear deconvolution," *Geophysical Prospecting*, vol. 29, no. 2, pp. 189–228, 1981.

[22] R. D. Gitlin, J. F. Hayes, and S. B. Weinstein, *Automatic and Adaptive Equalization, Data Communications Principles*, vol. 8, chapter 8, Plenum, New York, NY, USA, 1992.

[23] G. Im, C. Park, and H. Won, "A blind equalization with the sign algorithm for broadband access," *IEEE Communications Letters*, vol. 5, no. 2, pp. 70–72, 2001.

[24] O. Shalvi and E. Weinstein, "Super-exponential methods for blind deconvolution," *IEEE Transactions on Information Theory*, vol. 39, no. 2, pp. 504–519, 1993.

[25] S. Blinnikov and R. Moessner, "Expansions for nearly Gaussian distributions," *Astronomy and Astrophysics Supplement Series*, vol. 130, no. 1, pp. 193–205, 1998.

[26] H. Cramér, Ed., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, NJ, USA, 1951.

[27] M. G. Kendall and A. Stuart, Eds., *The Advanced Theory of Statistics*, Charles Griffin and Company Limited, 1958.

[28] N. C. Beaulieu and D. J. Young, "Designing time-hopping ultrawide bandwidth receivers for multiuser interference environments," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 255–284, 2009.

[29] M. Novey, T. Adali, and A. Roy, "A complex generalized Gaussian distribution—characterization, generation, and estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, part 1, pp. 1427–1433, 2010.

[30] Q. Z. Ahmed, M. Alouini, and S. Aissa, "Bit error-rate minimizing detector for amplify-and-forward relaying systems using generalized gaussian kernel," *IEEE Signal Processing Letters*, vol. 20, no. 1, pp. 55–58, 2013.

[31] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor, "Complex elliptically symmetric distributions: survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, 2012.

[32] W. Chen, Q. Gao, H. Xiong, L. Fei, and Q. Li, "TH-UWB receiver based on two PDFS approximation in multiuser systems," *Progress in Electromagnetics Research*, vol. 137, pp. 101–106, 2013.

[33] A. K. Nandi, Ed., *Blind Estimation Using Higher-Order Statistics*, chapter 2, Kluwer Academic Publishers, Boston, Mass, USA, 1999.

[34] D. C. C. Bover, "Moment equation methods for nonlinear stochastic systems," *Journal of Mathematical Analysis and Applications*, vol. 65, no. 2, pp. 306–320, 1978.

[35] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*, International Series in Pure and Applied Mathematics, Chapter 6, McGraw-Hill, New York, NY, USA, 1978.

[36] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*, Dover Publications, Dover, NJ, USA, 1965.

[37] M. Abramowitz and I. A. Stegun, "Orthogonal polynomials," in *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, U. W. Hochstrasser, Ed., chapter 22, pp. 773–775, Dover, New York, NY, USA, 1965.