On the Achievable Performance of Nonlinear MIMO Systems

João Guerreiro[(1,2)], Rui Dinis[(1,3)], Paulo Montezuma[(1,3)], Mário Marques da Silva[(1,2)]
[(1)] IT, Instituto de Telecomunicações, Lisboa, Portugal.
[(2)] UAL, Universidade Autónoma de Lisboa, Lisboa, Portugal.
[(3)] FCT, Universidade Nova de Lisboa, Monte de Caparica, Portugal.

***Abstract*—Pre-processing techniques employed at the transmitter side of multiple-input, multiple-output (MIMO) systems, such as the ones used by MIMO singular value decomposition (SVD) (MIMO-SVD) techniques, can lead to signals with large envelope fluctuations and high peak-to-average power ratio (PAPR). For this reason, we can have amplification difficulties and increased sensitivity to nonlinear distortion effects. In this work, we consider MIMO-SVD schemes with strong nonlinear distortion effects. It is shown that contrarily to what one could expect, the asymptotic optimum performance of the MIMO system with strong nonlinear distortion effects can be much better than the optimum performance with an ideal, linear transmitter.** [1]

***Index Terms*—MIMO, nonlinear distortion effects, optimum detection, performance evaluation, SVD.**

I. Introduction

In the last decade, multiple-input, multiple-output (MIMO) techniques have contributed to improved broadband wireless communications, since they allow for substantial gains in terms of realiability and/or spectral efficiency [1].

Several MIMO techniques involve some pre-processing. Among them, we have the zero-forcing beamforming [2] and the singular value decomposition (SVD) techniques [3]. With most pre-processing techniques, and with the SVD technique in particular, the precoded signals have high envelope fluctuations, even with single-carrier modulations, which leads to a large peak-to-average (PAPR). Under these conditions, one might expect amplification difficulties and increased sensitivity to nonlinear distortion effects [4], [5]. To avoid this, PAPR-reducing techniques can be employed [6]. However, these techniques can increase the complexity and/or reduce the spectral efficiency. Moreover, the most efficient PAPR-reducing techniques usually involve a clipping operation, which is itself a nonlinear operation. As a consequence, high-PAPR signals are frequently nonlinearly distorted, which might lead to significant performance degradation. To mitigate such degradation, the MIMO-SVD systems can be designed to have more transmit antennas than parallel data streams [7]. However, this cannot be achieved in many practical systems. Moreover, it has been shown that nonlinear distortion effects do not necessarily lead to performance degradation, and they can even lead to capacity gains. In [8] it was shown that there is an optimum back-off level for nonlinearly amplified multicarrier signals that maximizes the mutual information. In addition, for the specific case of single-input, single-output (SISO) orthogonal frequency division multiplexing (OFDM) schemes, it has been shown that with nonlinear transmissions the optimum performance (i.e., the performance of the maximum likelihood receiver) can even be better than the optimum performance obtained with linear transmitters [9]-[11]. Therefore, one might ask if such phenomena also occurs with MIMO techniques in general and with MIMO-SVD techniques in particular.

In this paper, we study the optimum performance of a point-to-point MIMO-SVD system with strong nonlinear effects, which can be associated to, e.g., a clipping operation or a nonlinear power amplifier. Our goal is to investigate the optimum performance of nonlinear MIMO-SVD systems and how it compares with the optimum performance of linear MIMO-SVD systems. However, it is not easy to obtain the optimum performance, since it involves a maximum-likelihood detection. Therefore, we evaluate the potential of the optimum performance through an asymptotic analysis and obtain the corresponding asymptotic optimum performance. More concretely, we start by presenting the asymptotic optimum performance in an idealized MIMO-SVD channel where the singular values have unitary power, showing that there are large potential gains relatively to the optimum performance in a linear MIMO-SVD scenario. After that, we analyze a more realistic case where the power of the singular values is not constant over the different data streams. The analysis of the asymptotic optimum performance allows us to assess the potential of the optimum detection and to conclude that the optimum performance of nonlinear MIMO-SVD systems can be better than the optimum performance with linear transmitters. This is justified by the fact that the nonlinear distortion has useful information for detection purposes.

Throughout this paper we employ the following notation: bold letters denote matrices or vectors. Italic letters denote scalars. $(\cdot)^T$ and $\mathrm{tr}(\cdot)$ denote the transpose and the trace operator, respectively. $\mathbf{I}_p$ denotes a $p \times p$ identity matrix. $\mathrm{diag}(\mathbf{x})$ is a diagonal matrix whose the diagonal is the vector

[1]This work was supported by Fundação para a Ciência e Tecnologia (FCT) and Instituto de Telecomunicações under projects UID/EEA/50008/2019, PES3N (POCI-01-0145-FEDER-030629) and MASSIVE5G (SAICT-45-2017-02).

$\mathbf{x}$. $\mathbb{E}[\cdot]$ denotes the expected value. $\mathbf{r}_x = \mathbb{E}[\mathbf{x}\mathbf{x}^H]$ and $E_{\mathbf{x}}$ denote the covariance matrix and the energy associated to vector $\mathbf{x}$, respectively.

II. System Characterization

In this paper we consider a point-to-point MIMO-SVD system with G transmit antennas and R receive antennas. We employ a single-carrier transmission over a flat-fading MIMO channel[1] represented by the $R \times G$ channel matrix $\mathbf{h}$. The channel coefficient between the rth receive antenna and the gth transmit antenna is modeled by a complex Gaussian distribution with unitary variance, i.e., $h^{(r,g)} \sim \mathcal{CN}(0,1)$. It should be noted that the assumption of the i.i.d Gaussian entries in the channel matrix does not compromise our analysis, provided that the correlation between them is not too high. With the SVD technique, we can multiplex up to $C = \min(G,R)$ independent spatial data streams over the MIMO channel

$$\mathbf{h} = \mathbf{u}\mathbf{\Lambda}\mathbf{v}^H, \tag{1}$$

where $\mathbf{\Lambda} = \text{diag}\,(\boldsymbol{\lambda}) = \text{diag}\left(\lambda^{(1)}, \lambda^{(2)}, ..., \lambda^{(C)}\right)$ is a $C \times C$ diagonal matrix and $\lambda^{(c)}$ is the singular value associated to the cth stream. Moreover, $\mathbf{u}^H$ and $\mathbf{v}$ denote the decoding and precoding matrices, respectively. The data symbols s_n are selected from a quadrature phase shift keying (QPSK) constellation (the extension to other cases is straightforward). The constellation's alphabet is denoted as $\mathbb{A}$ and its size is $M = 4$. The set of transmitted symbols at the nth time instant is represented by $\mathbf{s}_n = [s_n^{(1)}\ s_n^{(2)}\ ...\ s_n^{(C)}]^T$. The autocorrelation matrix of the data symbols $\mathbf{s}_n$ is $\mathbf{r}_s = \mathbb{E}\left[\mathbf{s}_n\mathbf{s}_n^H\right] = 2\sigma_s^2\mathbf{I}_C$, where σ_s is the amplitude of the QPSK symbols. These symbols are precoded and transmitted by the G antennas. The precoded symbols $\mathbf{x}_n = [x_n^{(1)}\ x_n^{(2)}\ ...\ x_n^{(G)}]^T$ are obtained as

$$\mathbf{x}_n = \mathbf{v}\mathbf{s}_n. \tag{2}$$

After the precoding operation, the precoded symbols can have some correlation. This is particularly true when the number of transmit antennas is much larger than the number of parallel data streams. However, in our scenario, where the number of transmit antennas is equal to the number of receive antennas and the number of data streams, this correlation is negligible. Therefore, we will assume that the precoded symbols are uncorrelated, which means that the average transmit energy is given by[2] $E_{\mathbf{x}} = \text{tr}(\mathbb{E}\left[\mathbf{x}\mathbf{x}^H\right]) = \text{tr}\left(\mathbf{v}\mathbf{r}_s\mathbf{v}^H\right) = \text{tr}\left(\mathbf{r}_s\right)$. In the considered nonlinear transmission scenario, the nonlinearity is represented by $f(\cdot)$ and modeled as a bandpass nonlinearity with amplitude modulation/amplitude modulation (AM/AM) and amplitude modulation/phase modulation (AM/PM) conversion functions $A(\cdot)$ and $\Theta(\cdot)$, respectively [12]. Under these conditions, the nonlinearly distorted signal at the output of the gth antenna is represented by

$$y^{(g)} = f\left(x^{(g)}\right) = A\left(\left|x^{(g)}\right|\right) \exp^{j\left(\Theta\left(\left|x^{(g)}\right|\right)+\arg\left(x^{(g)}\right)\right)}. \tag{3}$$

[1]We assume perfect channel state information (CSI) at the transmitter and receiver.

[2]In the following, and for the sake of notation simplicity, we will omit the time dependence in our analysis (i.e., the subscript n).

For the sake of simplicity, we consider an ideal envelope clipping (the generalization to other nonlinear characteristics is straightforward). Therefore, we do not have phase distortion (i.e., $\Theta(\cdot) = 0$) and the AM-AM function is defined as

$$A(|x^{(g)}|) = \begin{cases} |x^{(g)}|, & |x^{(g)}| \leq s_M \\ s_M, & |x^{(g)}| > s_M, \end{cases} \tag{4}$$

where s_M denotes the clipping level. The normalized clipping level is denoted as s_M/σ_x (σ_x^2 denotes the average power of the real and imaginary parts of the precoded symbols). The quantity s_M^2/σ_x^2 can be regarded as the input back-off (IBO), i.e., the ratio of the maximum output power to the input power (although the use of the IBO is more common with power amplifiers, we can use it with any nonlinear characteristic). It should be noted that due to the precoding, a given data symbol affects the signals for all antennas. Therefore, the nonlinear distortion term of each antenna will affect symbols in all data streams. It should also be noted that the power amplifiers are assumed to be equal in all transmit branches, which means that the average output power is the same in all transmit branches. In matrix notation, the nonlinearly distorted signal is denoted as $\mathbf{y} = f(\mathbf{x})$. Although the data symbols that compose $\mathbf{s}_n$ do not have a Gaussian nature, for large values of C, the precoded symbols are well approximated by a zero-mean, complex Gaussian distribution, i.e., $x^{(g)} \sim \mathcal{CN}(0, 2\sigma_x^2)$. This means that the real and imaginary parts of $x^{(g)}$ will be $\sim \mathcal{N}(0, \sigma_x^2)$, and the absolute value of $x^{(g)}$ by r (which will have a Rayleigh distribution, naturally). The Gaussian approximation for the precoded signals is justified by the central limit theorem (CLT) [13] since, in the considered MIMO systems, the precoding matrix $\mathbf{v}$ has large dimensions and the transmitted data symbols $s^{(c)}$ are uncorrelated. By taking advantage of this Gaussian approximation, we can employ the Bussgang's theorem [13], which allows us to decompose a nonlinearly distorted signal as the sum of two uncorrelated components, i.e.,

$$\mathbf{y} = f(\mathbf{x}) = \alpha\mathbf{x} + \mathbf{d}. \tag{5}$$

For the nth time instant, at a particular transmit antenna, we have $y^{(g)} = f\left(x^{(g)}\right) = \alpha x^{(g)} + d^{(g)}$. Since the precoded data symbols $x^{(g)}$ are uncorrelated between the different transmit branches, the nonlinear distortion terms $d^{(g)}$ are also uncorrelated between the different transmit branches. This means that the covariance matrix of the nonlinear distortion terms is $\mathbf{r}_d = \mathbb{E}\left[\mathbf{d}\mathbf{d}^H\right] = 2\sigma_d^2\mathbf{I}_G$, where $\sigma_d^2 = \mathbb{E}[|d^{(g)}|^2]/2$ is the average power of the distortion term. Thus, the average energy of $\mathbf{d}$ is $E_{\mathbf{d}} = \text{tr}\left(\mathbf{r}_d\right) = 2G\sigma_d^2$. Therefore, the average energy associated to the block $\mathbf{y}$ can be computed as $E_{\mathbf{y}} = |\alpha|^2 E_{\mathbf{x}} + E_{\mathbf{d}} = 2G|\alpha|^2\sigma_x^2 + 2G\sigma_d^2$. Under these conditions, the average transmitted bit energy is given by $E_b = \frac{E_{\mathbf{y}}}{2C} = |\alpha|^2 \frac{G}{C}\sigma_x^2 + \frac{G}{C}\sigma_d^2$. It should be noted that the power of the nonlinear distortion term can actually decrease when we decrease the clipping level. However, the total power of the nonlinearly distorted signal and the power of the useful term also decreases with the clipping level. The severity of a given nonlinear characteristic is defined by the ratio between useful and nonlinear distortion powers, which decreases with the clipping level for the ideal envelope clipping characteristic (naturally, this is not necessarily true for all nonlinear characteristics). Therefore, we can say that the lower the clipping level, the more severe the nonlinear distortion effects.

The received signal $\mathbf{z} = [z_n^{(1)}\ z_n^{(2)}\ ...\ z^{(R)}]^T$ is given by $\mathbf{z} = \mathbf{h}\mathbf{y} + \boldsymbol{\nu}$, where $\boldsymbol{\nu} = [\nu_n^{(1)}\ \nu_n^{(2)}\ ...\ \nu_n^{(R)}]^T$ denotes the set of additive white Gaussian noise (AWGN) samples, with $\nu^{(r)} \sim \mathcal{CN}(0, 2\sigma_\nu^2)$ and $\sigma_\nu^2 = N_0/2$ denoting the two-sided noise power spectral density (PSD). The decoded signal is obtained by completing the SVD decomposition, as follows:

$$\begin{aligned}\mathbf{r} &= \mathbf{u}^H\mathbf{z} = \mathbf{u}^H\mathbf{u}\boldsymbol{\Lambda}\mathbf{v}^H\mathbf{y} + \mathbf{u}^H\boldsymbol{\nu} \\ &= \boldsymbol{\Lambda}\mathbf{v}^H\mathbf{y} + \mathbf{u}^H\boldsymbol{\nu} = \alpha\boldsymbol{\Lambda}\mathbf{s} + \boldsymbol{\Lambda}\underbrace{\mathbf{v}^H\mathbf{d}}_{\mathbf{D}} + \mathbf{u}^H\boldsymbol{\nu}.\end{aligned} \tag{6}$$

III. Optimum Detection

As is widely known, the optimum performance can be obtained with a receiver that implements the maximum a posterior (MAP) criterion, which can be formulated as $\hat{\mathbf{s}}_n = \underset{\mathbf{s}_n \in \mathbb{A}^C}{\text{argmax}}\ p(\mathbf{s}_n|\mathbf{r}_n)$. By considering equal a priori probabilities, the MAP criterion reduces to the maximum likelihood (ML) criterion, which is given by $\hat{\mathbf{s}}_n = \underset{\mathbf{s}_n \in \mathbb{A}^C}{\text{argmax}}\ p(\mathbf{r}_n|\mathbf{s}_n)$. If the transmitted sequences are equiprobable, the optimum receiver that minimizes the error rate is based on the minimum squared Euclidean distance (SED) criterion [11], i.e.,

$$\hat{\mathbf{s}} = \underset{\mathbf{s} \in \mathbb{A}^C}{\text{argmin}}\ \left|\left|\mathbf{r} - \mathbf{u}^H\mathbf{h}f(\mathbf{v}\mathbf{s})\right|\right|^2. \tag{7}$$

Clearly, the optimum detector needs to perform a large set of calculations, since there are M^C possible transmitted sequences $\mathbf{s}$ (4^C for the case of QPSK constellations). In fact, the nonlinear function can be regarded as a shaping effect on the multidimensional equivalent constellation (for the case where we have C QPSK streams, we are working with an equivalent constellation with 4^C points over a size-$2C$ space), and we are studying its optimum performance. Due to the high complexity of this scenario, it is not easy to obtain the optimum performance. Indeed, insights on the optimum performance are usually obtained by considering some approximations such as the so-called union bound, which constitutes an upper (pessimistic) bound on the BER. This upper bound is based on pairwise error probability (PEP) between two sequences, i.e., the probability associated to the error event of detecting $\mathbf{s}'$ given that $\mathbf{s}$ was transmitted. We define the PEP between two sequences as $P\left(\mathbf{s} \rightarrow \mathbf{s}'|\mathbf{s}\right)$. Here, it should be noted that the only noise term in the PEP calculation is the channel noise, since the nonlinear distortion is an intrinsic part of the transmitted signal. Under these conditions, the so-called union bound can be formulated as

$$P_b \leq \sum_{\text{all }\mathbf{s}} P\left(\mathbf{s}\right) \sum_{l} P\left(\mathbf{s} \xrightarrow{l} \mathbf{s}'|\mathbf{s}\right), \tag{8}$$

where l is the number of bit errors and $P\left(\mathbf{s} \xrightarrow{l} \mathbf{s}'|\mathbf{s}\right)$ is the probability that the detected sequence differ in l bits from the transmitted sequence. In ideal AWGN channels, the PEP between two sequences is

$$P\left(\mathbf{s} \xrightarrow{l} \mathbf{s}'|\mathbf{s}\right) = Q\left(\sqrt{\frac{\mathcal{D}^2(\mathbf{s},\mathbf{s}')}{2N_0}}\right), \tag{9}$$

where

$$\begin{aligned}\mathcal{D}^2(\mathbf{s},\mathbf{s}') &= \left|\left|\boldsymbol{\Lambda}(\alpha(\mathbf{s}-\mathbf{s}') + (\mathbf{D}-\mathbf{D}'))\right|\right|^2 \\ &= \sum_c \left|\lambda^{(c)}\left(\alpha\left(s^{(c)} - {s'}^{(c)}\right) + \left(D^{(c)} - {D'}^{(c)}\right)\right)\right|^2,\end{aligned} \tag{10}$$

is the SED between $\mathbf{s}$ and $\mathbf{s}'$, which is intimately related with the power of the different singular values, i.e., $\left|\lambda^{(c)}\right|^2$. Another common approximation is to consider that the PEP is dominated by single bit errors, which constitutes an optimistic approximation. This approximation is valid in the asymptotic region, i.e., when the SNR is large. In that case, we have

$$P_b \approx \sum_{\text{all }\mathbf{s}} P\left(\mathbf{s}\right) P_1\left(\mathbf{s} \xrightarrow{1} \mathbf{s}'|\mathbf{s}\right). \tag{11}$$

Since for QPSK constellations there is a total of $2C$ sequences $\mathbf{s}'$ in those conditions, the computation of the approximate BER of (11) might still involve a heavy computational complexity. To overcome this, we can calculate an histogram of the SED and obtain an approximation of the BER by considering only a reduced set of $N_{\text{Seq}} << M^C$ sequences (and its 1 bit variations). With such histogram, we can write the another BER approximation based on a weighted sum of the different SED values, i.e.,

$$P_b \approx \sum_i \frac{f_{\text{abs}}\left(\mathcal{D}_i^2\right)}{2N_{\text{Seq}}C} Q\left(\sqrt{\frac{\mathcal{D}_i^2}{2N_0}}\right), \tag{12}$$

where $f_{\text{abs}}(\mathcal{D}_i^2)$ represents a counter with the number of times that the ith SED value $\mathcal{D}_i^2$ was observed in the SED histogram. As our study regarding the optimum performance is based on the PEP for sequences that differ in a single bit, and this approximation is valid in the asymptotic region (i.e., for large values of SNR), we will refer to this approximation of the optimum performance as the "asymptotic optimum performance" in the remaining of the paper.

IV. Asymptotic Optimum Performance

In the following, we present a set of performance results regarding the asymptotic optimum performance of the MIMO-SVD system introduced in Section II. Unless stated otherwise, we consider $C = 128$ data streams, $G = 128$ transmit antennas and $R = 128$ receive antennas (similar results were observed for other values of R, G and C, provided that they are much larger than 1). The nonlinear characteristic is assumed to be known at the receiver[3].

A. Equal Power Singular Values

Let us start by considering an idealized random channel where the singular values of all streams have unitary power,

[3]This can be justified by the fact that the envelope clipping is a nonlinear characteristic that solely depends on one parameter - the normalized clipping level. Therefore, it is not difficult to share this knowledge between the transmitter and the receiver.

i.e., $\left|\lambda^{(c)}\right|^2 = 1$. It should be noted that if we had a linear transmission, $\alpha = 1$ and there would be no distortion. As a result, the minimum squared Euclidean distance would be $\mathcal{D}^2(\mathbf{s},\mathbf{s}') = 4E_b$ and the asymptotic BER would be $P_b = Q\left(\sqrt{\frac{2E_b}{N_0}}\right)$, which is the conventional BER associated to ideal AWGN channels. Let us now consider a nonlinear transmission. Compared to the BER with linear transmissions defined above, there will be a performance gain if the normalized SED $\mathcal{D}^2/4E_b$ increases in the presence of nonlinear distortion effects (i.e., if $\mathcal{D}^2 > 4E_b$). Since we are concerned with sequences that differ in only 1 bit, i.e., with the error events that dominate the asymptotic behavior of the BER curve, we defined an asymptotic gain associated to the optimum detection as

$$\mathcal{G} = \frac{\mathcal{D}^2}{4E_b}. \tag{13}$$

Fig. 1 shows the PDF of the asymptotic gain considering different values of s_M/σ_x and different values of C. From

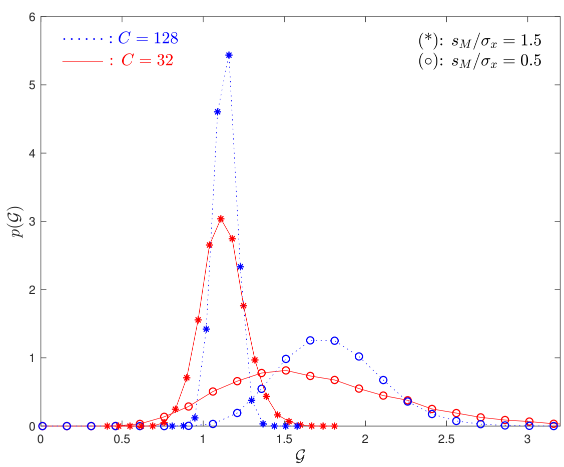

Fig. 1. PDF of the asymptotic gain $\mathcal{G}$ considering different values of C.

the figure, it can be observed that the asymptotic gain $\mathcal{G}$ tends to be higher than one, especially when s_M/σ_x decreases. Note that this does not mean that the SED increases when the normalized clipping level decreases. Actually, for the envelope clipping characteristics considered in this work, the SED (given by $\mathcal{D}^2$) decreases with s_M/σ_x. However, the asymptotic optimum performance is driven by the pairwise error probability between sequences, which is a function of the normalized SED (given by $\mathcal{D}^2/4E_b$). What can be observed is that when we normalize the SED by the average energy of the transmitted signals (which is proportional to the power of the signals at the nonlinearities' output), we have a gain[4] relatively the linear case, i.e. $\mathcal{G} > 1$. Therefore, one could expect that the optimum performance of nonlinear MIMO-SVD systems can be better than the optimum performance of linear MIMO-SVD systems. Moreover, the gains tend to increase when the normalized clipping level s_M/σ_x decreases. In addition, as C increases, the variance of the asymptotic gain tends to reduce, which means that for very large values of C, the asymptotic performance of the optimum detector can be approximated by

$$P_b \approx Q\left(\sqrt{\mathbb{E}[\mathcal{G}]\frac{2E_b}{N_0}}\right), \tag{14}$$

where $\mathbb{E}[\mathcal{G}]$ is the average asymptotic gain. Note that the asymptotic optimum performance represented in (14) is asymptotic in two senses, i.e., it is valid: (i) for large SNR values (see sec. III); and (ii) for large values of C. In [10] it is presented a formula for obtaining the average asymptotic gain for nonlinear OFDM transmissions in ideal AWGN channels. By following an approach similar to the one of [10] for our nonlinear MIMO-SVD scenario, it can be shown that the theoretical average asymptotic gain can be calculated as

$$\mathbb{E}[\mathcal{G}] = \frac{\int_0^{+\infty}\left(A'^2(r) + \frac{A^2(r)}{r^2}\right)p(r)\mathrm{d}r}{\int_0^{+\infty} A^2(r)p(r)\mathrm{d}r}, \tag{15}$$

where $p(r)$ denotes the PDF of the random variable that models the absolute values of the precoded symbols and $A'(r)$ is the derivative of the AM-AM characteristic defined in (4). Fig. 2 presents the theoretical and simulated average asymptotic gain of our idealized MIMO-SVD system considering different IBO values and $C = 128$. From the figure, it can be seen that

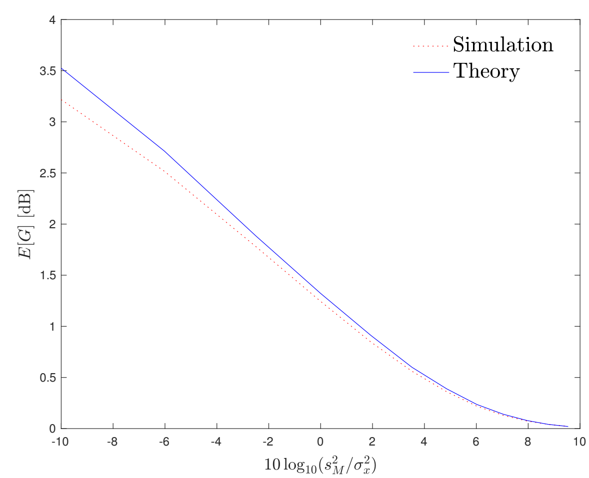

Fig. 2. Evolution of the average asymptotic gain with the IBO.

(15) constitutes a very tight approximation for obtaining the average asymptotic gain for nonlinear MIMO-SVD systems. As expected, the average asymptotic gain tends to 0 dB as s_M/σ_x increases.

B. Conventional MIMO-SVD

Common MIMO channels have singular values with different powers. Therefore, the distribution power of the singular values differ among the streams, which means that each stream will have a different BER, associated to the corresponding

[4]Note that a fair comparison between the linear and nonlinear transmissions should be made in terms of the same output (transmitted) power.

singular value. Since the theoretical analysis of this scenario is not simple, we employed a semi-analytical approach where we start by obtaining the distribution of the normalized SED given by $\gamma = \mathcal{D}^2/4E_b$ over the different channel realizations. This can be regarded as an "equivalent fading factor", since its distribution will give an insight on the asymptotic optimum performance, i.e.,

$$P_b = \int Q\left(\sqrt{2\gamma\frac{E_b}{N_0}}\right) p\left(\gamma\right) d\gamma. \tag{16}$$

Fig. 3 shows the distribution of the equivalent fading factor considering $C = 128$ and different values of s_M/σ_x.

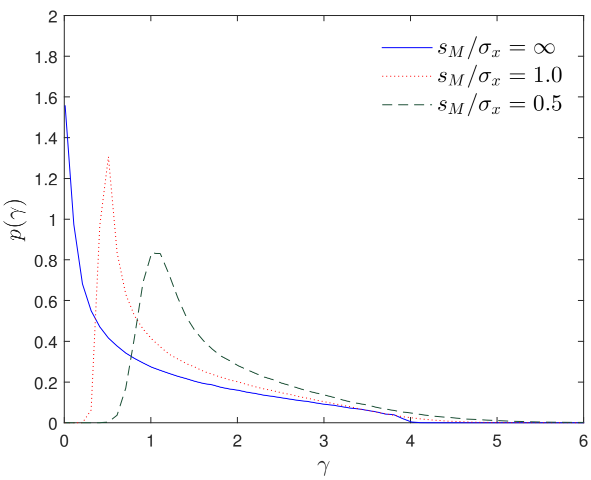

Fig. 3. PDF of the equivalent fading factor considering flat-fading channels.

As can be seen in the figure, the distribution of the "equivalent fading factor" assumes higher values in the presence of stronger nonlinear distortion effects. Fig. 4 shows the asymptotic optimum performance (obtained with (16)) considering different values of s_M/σ_x. From the figure, it can be

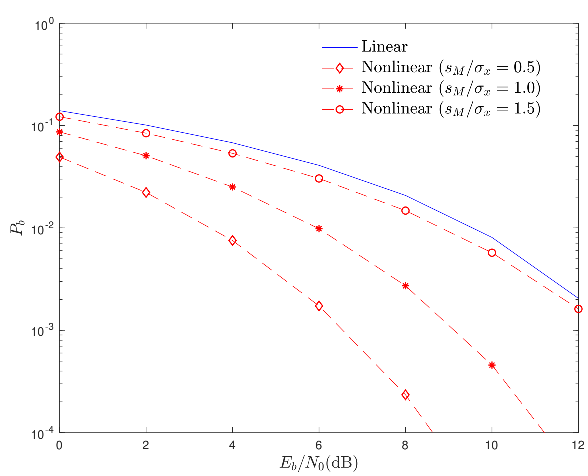

Fig. 4. Asymptotic optimum performance.

observed that as the normalized clipping level increases, the magnitude of the nonlinear distortion effects decreases and the asymptotic optimum performance approaches the average BER of a MIMO-SVD system with linear transmitters. However, in the presence of strong nonlinear distortion effects, one can observe large potential gains, which means that the optimum performance can be much better than the performance with linear transmissions. This can be explained by the fact that the nonlinear effects give rise to a kind of diversity effect, since the precoding combined with the nonlinear operation makes a given symbol being affected by all data streams, which reduces the likelihood that this symbol being affected by a low power singular value.

V. Conclusions

In this paper we studied the potential optimum performance of nonlinear MIMO-SVD systems. Contrarily to what one could expect, it is shown that when the number of data streams is large, the optimum performance of nonlinear MIMO-SVD systems can be better than the optimum performance of linear MIMO-SVD systems. This can be explained by the fact that the nonlinear distortion term has some information on the transmitted signals that can be used for detection purposes, provided that an optimum detection is employed.

References

[1] G. Foschini and M. Gans "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[2] H. Yang and T. L. Marzetta,"Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013

[3] G. Lebrun, J. Gao, and M. Faulkner, "MIMO transmission over a time-varying channel using SVD," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 757–764, Mar. 2005.

[4] I. Iofedov and D. Wulich, "MIMO-OFDM with nonlinear power amplifiers", *IEEE Trans. Commun.*, vol. 63, no. 12, pp. 4894-4904, Dec. 2015.

[5] J. Guerreiro, R. Dinis and P. Montezuma, "Massive MIMO with nonlinear amplification: signal characterization and performance evaluation", in *Proc. IEEE GLOBECOM'16*, Washington DC, USA, Dec. 2016.

[6] Y. Rahmatallah and S. Mohan, "Peak-to-average power ratio reduction in OFDM systems: a survey and taxonomy," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1567–1592, Mar. 2013.

[7] J. Guerreiro, R. Dinis and P. Montezuma, "On the impact of strong nonlinear effects on massive MIMO SVD systems with imperfect channel estimates," *IEEE VTC'2017 (Fall)*, Toronto, Canada, Sept. 2017.

[8] I. Gutman and D. Wulich, "On achievable rate of multicarrier with practical high power amplifier," *European Wireless Conference 2012*, Poland, Poznan, April 17-20, 2012.

[9] F. Peng and W. Ryan, "MLSD bounds and receiver designs for clipped OFDM channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3568–3578, Sept. 2008.

[10] J. Guerreiro, R. Dinis and P. Montezuma, "Optimum and sub-optimum receivers for OFDM signals with strong nonlinear distortion effects," *IEEE Trans. Commun.*, vol. 91, no. 9, pp. 3830–3840, Sept. 2013.

[11] J. Guerreiro, R. Dinis and P. Montezuma, "On the optimum multicarrier performance with memoryless nonlinearities," *IEEE Trans. on Communications*, vol. 63, no. 2, pp. 498– 509, Feb. 2015.

[12] A. Kaye, D. George and M. Eric, "Analysis and compensation of bandpass nonlinearities for communications," *IEEE Trans. Commun.*, vol. 20, no. 5, pp. 965–972, Oct. 1972.

[13] A. Papoulis and S. Pillai, "Probability, random variables and stochastic processes," McGraw-Hill, 2002.