



**QUEEN'S
UNIVERSITY
BELFAST**

Performance Limits of MIMO Systems with Nonlinear Power Amplifiers

Fozooni, M., Matthaiou, M., Björnson, E., & Duong, T. Q. (2015). *Performance Limits of MIMO Systems with Nonlinear Power Amplifiers*. Paper presented at IEEE Global Communications Conference 2015, San Diego, California, United States. <http://globecom2015.ieee-globecom.org/>

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access

This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: <http://go.qub.ac.uk/oa-feedback>

Performance Limits of MIMO Systems with Nonlinear Power Amplifiers

Milad Fozooni*, Michail Matthaiou*[†], Emil Björnson[‡], and Trung Q. Duong*

*School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K.

[†]Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden

[‡]Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden

Email: mfozooni01@qub.ac.uk, m.matthaiou@qub.ac.uk, emil.bjornson@liu.se, trung.q.duong@qub.ac.uk

Abstract—The development of 5G enabling technologies brings new challenges to the design of power amplifiers (PAs). In particular, there is a strong demand for low-cost, nonlinear PAs which, however, introduce nonlinear distortions. On the other hand, contemporary expensive PAs show great power efficiency in their nonlinear region. Inspired by this trade-off between nonlinearity distortions and efficiency, finding an optimal operating point is highly desirable. Hence, it is first necessary to fully understand how and how much the performance of multiple-input multiple-output (MIMO) systems deteriorates with PA nonlinearities. In this paper, we first reduce the ergodic achievable rate (EAR) optimization from a power allocation to a power control problem with only one optimization variable, i.e. total input power. Then, we develop a closed-form expression for the EAR, where this variable is fixed. Since this expression is intractable for further analysis, two simple lower bounds and one upper bound are proposed. These bounds enable us to find the best input power and approach the channel capacity. Finally, our simulation results evaluate the EAR of MIMO channels in the presence of nonlinearities. An important observation is that the MIMO performance can be significantly degraded if we utilize the whole power budget.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) wireless communication systems have been well investigated over the last two decades thanks to their ability to enhance the spectral efficiency and reliability [1], [2]. It is also well known that in a MIMO system, power is mainly consumed by the last parts of the transmitter chain, and in particular by the power amplifiers (PAs). Most contemporary MIMO systems deploy expensive, linear PAs although these components are intimately limited in terms of power efficiency. Yet, with the current strive towards network densification (i.e. the massive MIMO paradigm [3]), future systems are anticipated to deploy inexpensive, nonlinear PAs of high power efficiency.

In this context, only a few publications have studied the impact of nonlinear PAs on the MIMO capacity. For instance, [4] and [5] considered a precise power consumption model of PAs, and then introduced a low complexity algorithm to maximize the sum rate of multiple-input single-output systems. They also extended their beamforming method to the case of parallel MIMO by utilizing a dynamic programming language algorithm. Although the presented model is relatively precise, it does not go beyond the linear region where we can enjoy higher power efficiency. In [6], the authors considered the

impact of PA nonlinearities on channel estimation and then proposed a quantized method to optimize the bit-error-rate and mutual information. Moreover, they proposed a constellation-based compensation method for high-power amplifier nonlinearities in [7]. There are also some other research efforts like [8]–[13], which have dealt with PA nonlinearities in wireless communications, but they either do not focus on MIMO systems or do not evaluate EAR.

Motivated by the above discussion, this paper focuses on studying the performance of a MIMO system deploying nonlinear PAs. In particular, we start our analysis by optimizing the ergodic achievable rate (EAR) over the input power matrix. Then, we simplify this power allocation problem to a power control problem where the only variable is the total consumed power. Finally, we propose lower and upper bounds on the EAR, which can be easily optimized by conventional optimization methods. These analytical results followed by simulations highlight the fact that using the full power budget will reduce the EAR to zero, since PAs nonlinearities become dominant.

Notation: Upper and lower case bold-face letters denote matrices and vectors, respectively. Also, the symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^\dagger$, $\text{Tr}(\cdot)$, and $\text{rank}(\cdot)$ indicate transpose, conjugate, conjugate transpose, trace operator, and matrix rank, respectively. Furthermore, $|\cdot|$ and $\det(\cdot)$ both denote the determinant operator. Finally, $\mathbb{E}[\cdot]$ is the expectation operation, and \mathbf{I}_N refers to an $N \times N$ identity matrix.

II. SIGNAL AND SYSTEM MODELS

The traditional model of flat-fading point-to-point MIMO channels with N_t transmit antennas and N_r receive antennas is

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{w} \quad (1)$$

where $\mathbf{s} = [s_1, s_2, \dots, s_{N_t}]^T \in \mathbb{C}^{N_t \times 1}$ presents the complex Gaussian distributed transmitted signal with zero mean and covariance matrix $\mathbf{K}_s = \mathbb{E}[\mathbf{s}\mathbf{s}^\dagger]$. The received signal is denoted by $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$, while the N_r -dimensional vector \mathbf{w} models the additive circularly symmetric complex Gaussian noise $\mathbf{w} \sim \mathcal{CN}(0, N_0 \mathbf{I}_{N_r})$. Throughout this paper, we assume that the propagation channel coefficients are independently circular symmetric complex Gaussian variable with unit variance. Additionally, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is assumed to be known to the

receiver, but not at the transmitter; however, its statistical characteristics are available at the transmitter. Note that the channel matrix \mathbf{H} is normalized so that N_0 contains both noise variance and pathloss.

Unfortunately, the above canonical model falls short of describing the nonlinear behavior of PAs and its impact on the end-to-end performance. To this end, in the following we extend the model of (1) to account for these nonlinearities. Note that our analysis remains agnostic to any type of nonlinearity that may be induced by mixers, filters and D/A converters.

In order to incorporate the impact of transceiver impairments, we first need to explain the input/output relation of the power amplifier. In general, for a complex baseband input signal represented as $u_{in}(t) = A(t)e^{j\phi(t)}$, the signal at the output of a PA with amplitude gain $g_A(A(t))$ and phase gain $g_\phi(A(t))$ is

$$u_{out}(t) = g(u_{in}(t)) = g_A(A(t))e^{j(\phi(t)+jg_\phi(A(t)))}. \quad (2)$$

There are various types of PA models for the amplitude and phase gains, such as ideal clipping, traveling wave tube, and solid-state amplifier model [14]. Fortunately, all the models can be encompassed under the umbrella of a polynomial PA model. For the sake of simplicity, we hereafter assume that all PAs have the same nonlinear conversion functions which are known at the transceivers [6], [7]. In general, a polynomial PA model can be easily determined by the following curve fitting of degree N :

$$g_A(A_i) = \sum_{n=0}^{N-1} \beta_{n+1} A_i^{n+1} \quad (3)$$

where A_i is the voltage of the input signal in the i -th PA, $i = 1, 2, \dots, N_t$. Hereafter, we also consider memoryless PAs and ignore the phase distortion [12].¹ Thus, the coefficients, β_{n+1} for $n = 0, 1, \dots, N-1$, are real constant numbers. By this preamble and regarding the Bussgang's theorem [15], the i -th PA output can be expressed in the form of

$$s_i = \alpha_i x_i + d_i \quad i = 1, 2, \dots, N_t \quad (4)$$

where d_i represents the distortion noise which is uncorrelated of the input signal, x_i . Note that, d_i is a zero-mean (not necessarily Gaussian) distribution with power density $\sigma_{d_i}^2$. Since the input signal has a complex Gaussian distribution, its magnitude (A_i) follows a Rayleigh probability density function, such that

$$P(A_i) = \frac{A_i}{\sigma_i^2} \exp\left(-\frac{A_i^2}{2\sigma_i^2}\right) \quad (5)$$

where $\mathbb{E}[x_i^* x_i] = 2\sigma_i^2$. Furthermore, in (4), s_i stands for the amplifier output which will be emitted from the transmit

antennas, and α_i is a constant affected by the PA gain function and its input power. In general, it can be shown that [13]

$$\alpha_i = \frac{\mathbb{E}[x_i^* s_i]}{\mathbb{E}[x_i^* x_i]} \quad (6)$$

$$\sigma_{d_i}^2 = \mathbb{E}[s_i^* s_i] - \alpha_i^2 \mathbb{E}[x_i^* x_i] \quad (7)$$

where we use the fact that $\mathbb{E}[x_i^* d_i] = 0$. Subsequently, these parameters can be easily expressed as

$$\alpha_i = \frac{1}{2\sigma_i^2} \int_0^\infty A_i g_A(A_i) P(A_i) dA_i \quad (8)$$

$$\sigma_{d_i}^2 = \int_0^\infty g_A^2(A_i) P(A_i) dA_i - \frac{1}{2\sigma_i^2} \left(\int_0^\infty A_i g_A(A_i) P(A_i) dA_i \right)^2. \quad (9)$$

Accordingly, it can be shown that the Bussgang's parameters for the polynomial model, can be obtained as [13]

$$\alpha_i = \sum_{n=0}^{N-1} \beta_{n+1} 2^{n/2} \sigma_i^n \Gamma\left(2 + \frac{n}{2}\right) \quad (10)$$

$$\sigma_{d_i}^2 = \sum_{n=2}^{2N} \left(\gamma_n 2^{n/2} \sigma_i^n \Gamma\left(1 + \frac{n}{2}\right) \right) - 2\sigma_i^2 \left(\sum_{n=0}^{N-1} \beta_{n+1} 2^{n/2} \sigma_i^n \Gamma\left(2 + \frac{n}{2}\right) \right)^2 \quad (11)$$

where $\Gamma(\cdot)$ is the Gamma function [16, Eq. (8.310.1)], and γ_n can be defined as follows

$$\gamma_n = \sum_{k=1}^{n-1} \check{\beta}_k \check{\beta}_{n-k}^* \quad (12)$$

and

$$\check{\beta}_k \triangleq \begin{cases} \beta_k, & 1 \leq k \leq N \\ 0, & \text{otherwise.} \end{cases}$$

Now, the impact of PA nonlinearities can be incorporated into (1) based on the Bussgang's model (4):

$$\mathbf{y} = \mathbf{H}(\mathbf{\Lambda}\mathbf{x} + \mathbf{d}) + \mathbf{w} \quad (13)$$

where $\mathbf{\Lambda} = \text{diag}\{\alpha_1, \alpha_2, \dots, \alpha_{N_t}\}$ and $\mathbf{d} = [d_1, d_2, \dots, d_{N_t}]^T$. This equation can be reorganized as

$$\mathbf{y} = \mathbf{H}\mathbf{\Lambda}\mathbf{x} + \underbrace{(\mathbf{H}\mathbf{d} + \mathbf{w})}_{\mathbf{n}} \quad (14)$$

in which the vector \mathbf{n} denotes the aggregated noise at the receiver with the covariance matrix

$$\mathbf{R}_n = \mathbb{E}[\mathbf{nn}^\dagger] = \mathbf{H}\mathbf{D}\mathbf{H}^\dagger + N_0\mathbf{I}_{N_r} \quad (15)$$

where we define $\mathbf{D} \triangleq \mathbb{E}[\mathbf{dd}^\dagger]$.

III. ERGODIC ACHIEVABLE RATE ANALYSIS

In this section, we analyze the EAR in the presence of PA nonlinearities by simplifying the power allocation problem to a power control problem.

¹This assumption is widespread in the literature especially for solid state PAs. In other words, memory leads to delay and consequently phase distortion. Therefore, we can ignore the phase distortion in these PAs as they are memoryless.

A. Ergodic Achievable Rate

Based on the channel impairment model in Section II, we are now ready to determine the MIMO achievable rate under Gaussian signaling for an arbitrary number of antennas. When Gaussian symbols are transmitted over the MIMO channel, the capacity (in bits/s) is given by

$$R = \sup_{\text{Tr}(\mathbf{Q}) \leq P_t, \mathbf{Q} \succeq \mathbf{0}} \mathbb{E} \left[B \log_2 \left(\det \left(\mathbf{I}_{N_r} + \mathbf{R}_n^{-1} \mathbf{H} \mathbf{A} \mathbf{Q} \mathbf{A}^\dagger \mathbf{H}^\dagger \right) \right) \right] \quad (16)$$

where B denotes the bandwidth. For the sake of clarity, we will drop B from our subsequent analytical results, but in the numerical results section, we do include the impact of bandwidth. Moreover, $\mathbf{Q} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger]$ and $\text{Tr}(\mathbf{Q}) \leq P_t$ indicates that the transmitter is constrained to its total power. For the purpose of simplification we define $\text{Tr}(\mathbf{Q}) = \sum_{i=1}^{N_t} 2\sigma_i^2 \triangleq P$, and also define the instantaneous MIMO channel correlation matrix as

$$\mathbf{W} \triangleq \begin{cases} \mathbf{H}\mathbf{H}^\dagger, & N_r \leq N_t \\ \mathbf{H}^\dagger\mathbf{H}, & N_r > N_t \end{cases} \quad (17)$$

since its eigenvalues, λ_i , will be often used in our calculations.

Remark 1. In order to reach the ergodic capacity, we need to assume a propagation channel with the maximum uncertainty [17]. For channel with known finite energy, the independent and identically distributed (i.i.d) Gaussian channel provides the maximum entropy. On the other hand, capacity can be achieved by jointly Gaussian input signals [1]. Although we consider the inputs to the PAs to be jointly Gaussian, the outputs of the PAs are not strictly jointly Gaussian. In light of this fact, in the remainder of this paper, we will be referring to (16) as the maximum EAR.

Proposition 1. The EAR is a concave function in its domain with respect to the covariance matrix \mathbf{Q} .

Proof. Note that $\log(\det(\cdot))$ is a concave function in the cones of positive semi-definite matrices [18]. Also \mathbf{R}_n^{-1} , \mathbf{A} , and \mathbf{Q} are all positive semi-definite matrices. Thus, the EAR is a concave function in its domain. \square

Proposition 2. The maximum EAR is achieved when \mathbf{Q} is a scaled identity matrix, i.e. $\mathbf{Q} = \frac{P}{N_t} \mathbf{I}_{N_t}$.

Proof. See Appendix I. \square

It is noteworthy that the PAs nonlinearities are affected by two factors: (i) PA transition function $g_A(\cdot)$, which is assumed to be the same for all the PAs; and (ii) the input power of each PA, i.e. $2\sigma_i^2$. As a consequence, Proposition 2 simplifies extensively our analysis as it allocates equal powers to each PA. In other words, we can conclude that $\sigma_1 = \sigma_2 = \dots = \sigma_{N_t} \triangleq \sigma$, then $\alpha_1 = \alpha_2 = \dots = \alpha_{N_t} \triangleq \alpha$, and $\sigma_{d_1} = \sigma_{d_2} = \dots = \sigma_{d_{N_t}} \triangleq \sigma_d$. Therefore, the power allocation optimization upon the covariance matrix will be reduced to a power control over the scalar, i.e. P . Correspondingly, the noise covariance and EAR are respectively simplified as follows

$$\mathbf{R}_n = \sigma_d^2 \mathbf{H}\mathbf{H}^\dagger + N_0 \mathbf{I}_{N_r} \quad (18)$$

$$R = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \frac{P\alpha^2}{N_t} (N_0 \mathbf{I}_{N_r} + \sigma_d^2 \mathbf{H}\mathbf{H}^\dagger)^{-1} \mathbf{H}\mathbf{H}^\dagger \right| \right]. \quad (19)$$

Note that, in contrast to Telatar's methodology, our objective function in (19) is not necessarily a strictly ascending or descending function of the power P . This is due to the presence of σ_d^2 in (19) which also increases with P .

Corollary 1. The maximum EAR can be rewritten as

$$R = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \det \left(\mathbf{I}_{N_t} + \mathbf{Z} \right) \right] \quad (20)$$

where \mathbf{Z} is a diagonal $N_t \times N_t$ square matrix whose entries are $\zeta_{i,i} = \left(\frac{P\alpha^2 \lambda_i}{N_t \sigma_d^2 \lambda_i + N_t N_0} \right)$.

Proof. See Appendix II. \square

Proposition 3. Assuming i.i.d Rayleigh fading channels, the maximum EAR under the proposed PAs nonlinearities model is

$$R = \sup_{0 \leq P \leq P_t} \frac{r}{\ln 2} K \sum_{m=1}^r \sum_{n=1}^r (-1)^{n+m} \det(\mathbf{\Omega}) \Gamma(t+1) \sum_{k=1}^{t+1} \left(e^{1/f} E_{t+2-k} \left(\frac{1}{f} \right) - e^{1/g} E_{t+2-k} \left(\frac{1}{g} \right) \right) \quad (21)$$

where we define $q \triangleq \max\{N_t, N_r\}$, $r \triangleq \min\{N_t, N_r\}$, $t \triangleq n + m + q - r - 2$, and $K \triangleq \left[\prod_{i=1}^r (q-i)! \prod_{j=1}^r (r-j)! \right]^{-1}$ is a constant. Moreover, $f \triangleq \frac{P\alpha^2 + N_t \sigma_d^2}{N_0 N_t}$, $g \triangleq \frac{\sigma_d^2}{N_0}$, and also $E_n(x) = x^{n-1} \Gamma(1-n, x)$. Here, $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$ is incomplete Gamma function [16, Eq. (8.350.2)]. Finally, $\mathbf{\Omega}$ is an $(r-1) \times (r-1)$ matrix whose (i, j) -th entry is given by

$$\mathbf{\Omega} = (\phi_{ij}^{(n)(m)} + q - r)! r^{-\frac{1}{r-1}} \quad (22)$$

for which,

$$\phi_{ij}^{(n)(m)} \triangleq \begin{cases} i + j - 2, & \text{if } i \leq n, \text{ and } j \leq m \\ i + j, & \text{if } i \geq n, \text{ and } j \geq m \\ i + j - 1, & \text{otherwise.} \end{cases} \quad (23)$$

Proof. Assuming i.i.d. Rayleigh fading channels, \mathbf{W} is full-rank with probability one [19]. Recalling Corollary 1, the maximum EAR can be expressed as

$$R = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \prod_{i=1}^{\text{rank}(\mathbf{Z})} (1 + \zeta_{i,i}) \right] = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[r \log_2 \left(1 + \frac{P\alpha^2 \lambda}{N_t \sigma_d^2 \lambda + N_t N_0} \right) \right]. \quad (24)$$

Note that \mathbf{W} is an $r \times r$ random, positive semi-definite matrix following the complex Wishart distribution. Therefore, it has

real non-negative eigenvalues, and the probability density function of its unordered eigenvalue, λ , can be found in [20]

$$P_\lambda(\lambda) = K \sum_{m=1}^r \sum_{n=1}^r (-1)^{n+m} \lambda^{n+m+q-r-2} e^{-\lambda} \det(\mathbf{\Omega}). \quad (25)$$

We can now define $a \triangleq P\alpha^2$, $b \triangleq N_t \sigma_d^2$, $c \triangleq N_0 N_t$, and then proceed along with some integral techniques

$$R = \sup_{0 \leq P \leq P_t} \int_0^\infty r \log_2 \left(1 + \frac{a+b}{c} \lambda \right) P_\lambda(\lambda) d\lambda - \int_0^\infty r \log_2 \left(1 + \frac{b}{c} \lambda \right) P_\lambda(\lambda) d\lambda. \quad (26)$$

Then, the final result can be easily obtained following the methodology of [21]. \square

B. Asymptotic Analysis

The maximum EAR may behave as a non-increasing (non-monotonic) function in the problem on hand. This is due to the presence of the nonlinearity distortion power σ_d^2 in (19), which also scales with the input power (see (11)). Thus, we seek to work out the EAR in the asymptotic regime. First, we obtain the EAR when we use the whole power budget, i.e. $P = P_t \rightarrow \infty$. After some manipulations, it can be shown that the EAR approaches to a saturation point, according to

$$\lim_{\sigma \rightarrow \infty} \alpha = \beta_N 2^{N-1/2} \sigma^{N-1} \Gamma \left(\frac{3+N}{2} \right) \quad (27)$$

$$\lim_{\sigma \rightarrow \infty} \sigma_d^2 = \gamma_2 N 2^N \sigma^{2N} \Gamma(1+N) - N_t 2^N \beta_N^2 \sigma^{2N} \Gamma^2 \left(\frac{3+N}{2} \right) \quad (28)$$

where $\sigma^2 = \frac{P}{2N_t}$ is the power of input signal (real/imaginary part). Thus,

$$R_{\text{high}} = \lim_{P \rightarrow \infty} R = r \log_2 \left(1 + \frac{\beta_N^2 \Gamma^2 \left(\frac{3+N}{2} \right)}{\gamma_2 N \Gamma(1+N) - N_t \beta_N^2 \Gamma^2 \left(3 + \frac{n}{2} \right)} \right). \quad (29)$$

The high-power asymptote in (29) reveals that increasing the input power with no bound, leads to a saturated value for the EAR. This observation is in sharp contrast with the classical MIMO results [1], [2], which however, consider perfect linear PAs.² At the other extreme, we find a closed-form expression for the EAR in the low SNR regime where $P \rightarrow 0$. In this case, we use the approximation $\ln(1+u) \sim u$ for small u to get

$$R_{\text{low}} = \lim_{P \rightarrow 0} R = \sup_{0 \leq P \leq P_t} \beta_1^2 \frac{(\log_2 e) P}{N_0 N_t} \mathbb{E}[\text{Tr}(\mathbf{W})] = \beta_1^2 \frac{(\log_2 e) P N_r}{N_0}. \quad (30)$$

²Interestingly, our result is in line with [22] which used a Gaussian model for the hardware residual distortions.

The result in (30) is consistent with some classical MIMO results where PAs are assumed to be perfectly linear [19]. This is due to the fact that in the low SNR regime, PAs are still operating in their linear regime.

C. Bounds

In the previous subsections, we have reduced the problem from a power allocation optimization to a power control based on only one variable, P . However, (21) and (24) are inherently complicated formulas. Motivated by this, we confine the results between an upper and two lower bounds such that $R_{\text{lower}} \leq R \leq R_{\text{upper}}$. We start with the upper bound using Jensen's inequality for the concave function $\log_2(1+x)$ in (a) and for the concave function $\frac{1}{1+\frac{1}{x}}$ in (b):

$$\begin{aligned} R &= \sup_{0 \leq P \leq P_t} \mathbb{E} \left[r \log_2 \left(1 + \frac{P\alpha^2 \lambda}{N_t \sigma_d^2 \lambda + N_t N_0} \right) \right] \\ &\stackrel{(a)}{\leq} \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \mathbb{E} \left[\frac{P\alpha^2 \lambda}{N_t \sigma_d^2 \lambda + N_t N_0} \right] \right) \\ &\stackrel{(b)}{\leq} \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2 \left(1 + \frac{N_0}{\sigma_d^2 \mathbb{E}[\lambda]} \right)} \right) \\ &= \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2 q}{N_t \sigma_d^2 q + N_t N_0} \right) \\ &\triangleq R_{\text{upper}}. \end{aligned} \quad (31)$$

The EAR can also be lower bounded by recalling $\mathbb{E}[\log_2(1+\rho(\lambda))] \geq \log_2(1 + \exp[\mathbb{E}[\ln \rho(\lambda)]])$ in (c), to get

$$\begin{aligned} R &= \sup_{0 \leq P \leq P_t} \mathbb{E} \left[r \log_2 \left(1 + \frac{P\alpha^2 \lambda}{N_t \sigma_d^2 \lambda + N_t N_0} \right) \right] \\ &\stackrel{(c)}{\geq} \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2} \exp \left(\mathbb{E} \left[\ln \left(\frac{\lambda}{\lambda + \frac{N_0}{\sigma_d^2}} \right) \right] \right) \right) \\ &\geq \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2} \exp \left(\mathbb{E}[\ln(\lambda)] - \ln \left(\mathbb{E}[\lambda] + \frac{N_0}{\sigma_d^2} \right) \right) \right) \\ &= \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2 \exp(\mathbb{E}[\ln(\lambda)])}{N_t \sigma_d^2 q + N_t N_0} \right) \\ &\triangleq R_{\text{lower}1}. \end{aligned} \quad (32)$$

It is known that $\mathbb{E}[\ln(\det(\mathbf{W}))] = \sum_{l=0}^{r-1} \psi(q-l)$, where $\psi(\cdot)$ represents the Euler digamma function [23]. Hence, we can represent $\mathbb{E}[\ln(\lambda)]$ in the following way

$$\begin{aligned} \mathbb{E}[\ln(\lambda)] &= \frac{1}{r} \mathbb{E} \left[\sum_{i=1}^r \ln(\lambda_i) \right] = \frac{1}{r} \mathbb{E}[\ln(\det(\mathbf{W}))] \\ &= \frac{1}{r} \sum_{l=0}^{r-1} \psi(q-l). \end{aligned} \quad (33)$$

$$\sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2 \exp\left(\frac{1}{r} \sum_{l=0}^{r-1} \psi(q-l)\right)}{N_t \sigma_d^2 q + N_t N_0} \right) \leq R \leq \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2 q}{N_t \sigma_d^2 q + N_t N_0} \right). \quad (34)$$

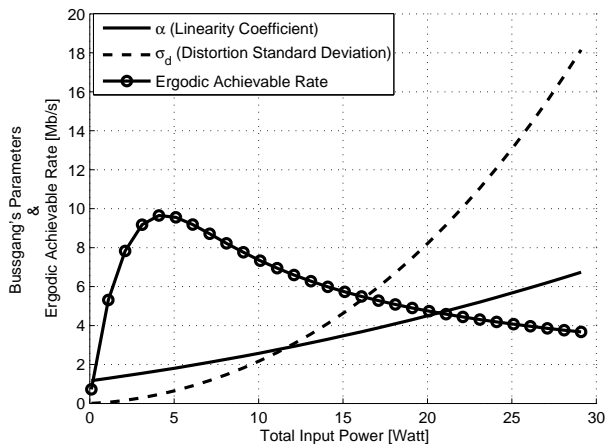


Fig. 1. Role of Bussgang's theory parameters on the EAR ($N_t = 3$, $N_r = 4$, $A_{os} = 1$, $v = 1$, $N_0 = 1$, $B = 1$ MHz)

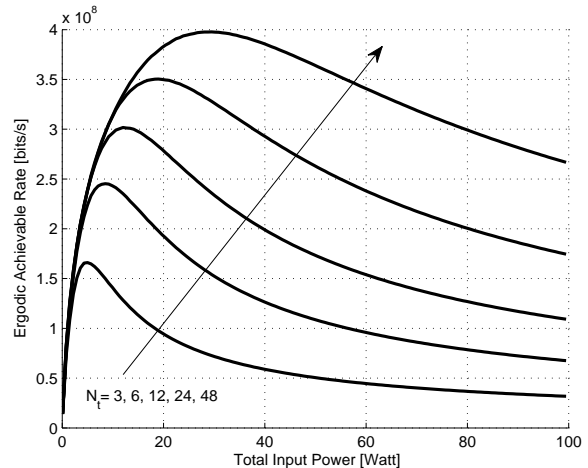


Fig. 2. Role of transmitter antennas on EAR ($N_r = 4$, $A_{os} = 1$, $v = 1$, $N_0 = 1$, $B = 20$ MHz)

To sum up, the EAR can be bounded by (34), shown at the top of this page.

Hereafter, we also use another lower bound, named second lower bound, that can be especially useful when the number of receive antennas is much higher than the number of transmit antennas.³ We follow the same approach to prove this bound.

$$\begin{aligned} R &\geq \sup_{0 \leq P \leq P_t} \mathbb{E} \left[r \log_2 \left(1 + \exp \left(\ln \left(\frac{P\alpha^2}{N_t \sigma_d^2} \frac{\lambda}{\lambda + \frac{N_0}{\sigma_d^2}} \right) \right) \right) \right] \\ &\geq \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2} \exp \left(\mathbb{E} \left[\ln \left(\frac{\lambda}{\lambda + \frac{N_0}{\sigma_d^2}} \right) \right] \right) \right) \\ &= \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2} \exp \left(\ln \left(\frac{1}{1 + \frac{N_0}{\sigma_d^2} \mathbb{E} \left[\frac{1}{\lambda} \right]} \right) \right) \right) \\ &= \sup_{0 \leq P \leq P_t} r \log_2 \left(1 + \frac{P\alpha^2}{N_t \sigma_d^2 + N_t N_0 \frac{1}{N_r - N_t}} \right) \\ &= R_{\text{lower2}} \end{aligned} \quad (35)$$

where we have used the central Wishart matrix properties

$$\mathbb{E} \left[\text{Tr} \left(\mathbf{W}^{-1} \right) \right] = \frac{N_t}{N_r - N_t} \quad N_r \geq N_t + 1. \quad (36)$$

³According to (36), the second lower bounds becomes loose whenever the numbers of transmit and receive antennas are close to each other. However, we can introduce a new lower bound like $R_{\text{lower}} = \max\{R_{\text{lower1}}, R_{\text{lower2}}\}$ to always guarantee a tight lower bound.

IV. SIMULATION RESULTS

In this section, we present simulation results illustrating the EAR of MIMO systems in the presence of PA nonlinearities by generating 10^4 Monte-Carlo realizations of the flat fading matrix \mathbf{H} . Herein, we choose the solid state PAs, whose AM/AM and AM/PM functions are specified by [24]

$$g_A(A) = \frac{A}{\left[1 + \left(\frac{A}{A_{os}} \right)^{2v} \right]^{\frac{1}{2v}}} \quad (37)$$

$$g_\phi(A) = 0 \quad (38)$$

where A_{os} denotes the output saturation voltage and v sets the smoothness of transition from the linear region to the saturation part. In particular, for large v this model approaches the ideal clipping PA model which is commonly used to represent the hard clipping effect [6]. Furthermore, we approximate the solid state model with a polynomial of degree 9, and ignore the even orders since they contribute with the out-band distortion [14]. Figure 1 depicts the role of the Bussgang's parameters on the EAR. Although, the linearity coefficient (α), is dominant in the low-power regime, distortion (σ_d) dominates in the higher input power. By this observation, a non-monotonic behavior of EAR function is anticipated.

The performance of MIMO system under PA nonlinearities model for different transceiver antennas is shown in Fig. 2 and 3, respectively. It can be easily observed that an increase in the number of transmit antennas pushes the maximum of the EAR into higher input powers. On the other hand, when we increase the number of antennas on the receiver side, the

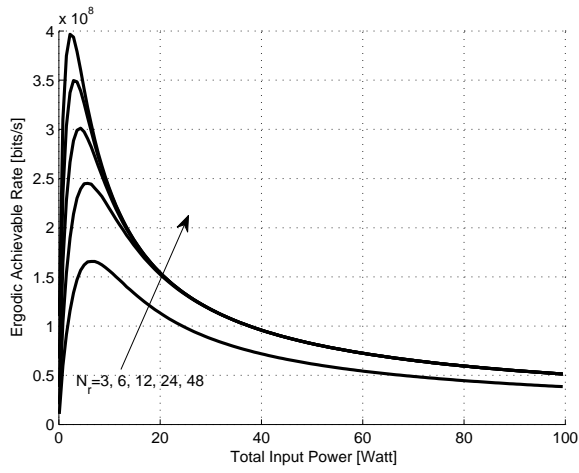


Fig. 3. Role of receiver antennas on ergodic achievable rate ($N_r = 4$, $A_{os} = 1$, $v = 1$, $N_0 = 1$, $B = 20$ MHz)

maximum point is obtained in the lower input power. It is best suitable to justify this behavior by the fact that PAs, as a major source of nonlinearities, only exist in the transmitter side. So, as a practical result, we are interested in utilizing more antennas in the receiver side to capture the maximum EAR by a lower input power. Finally, Fig. 4 illustrates how tight the suggested bounds are. These bounds seem to be a very good approximation of the EAR for any total input power. However, approximating the best total input power that leads to the maximum EAR is more important. Figure 4 confirms that the best power of lower/upper bounds leads to an EAR which is very close to the actual maximum EAR.

V. CONCLUDING REMARKS

Working with inexpensive nonlinear PAs seems to be a viable solution for the next generation of wireless systems. This nonlinear behavior distorts the transmitted signal and can effectively reduce the achievable rate in any communication system. This performance degradation becomes substantial when the power fed into the PAs is high. On the other hand, PAs offer their best efficiency in their nonlinear regime. Motivated by the above fundamental tradeoff, we have analytically quantified the impact of PAs nonlinearities on the achievable rate of MIMO systems. Our analysis derived closed-form exact expressions along with tractable asymptotic approximations. Our future work will include the determination of optimal operating points to achieve a predetermined rate constraint.

APPENDIX I PROOF OF PROPOSITION 2

The optimality of this result is an advanced consequence of [22, Corollary 1] or following Telatar's methodology in [1]. A standard Gaussian random matrix, \mathbf{H} , is a bi-unitarily invariant matrix. It means that the joint distribution of its entries equals that of $\mathbf{U}\mathbf{H}\mathbf{V}^\dagger$ for any unitary matrices \mathbf{U} and \mathbf{V} independent of \mathbf{H} . Now, by following Telatar's approach, we can limit our attention only to a diagonal \mathbf{D} , and diagonal $\mathbf{\Lambda}\mathbf{Q}\mathbf{\Lambda}^\dagger$ in (15),

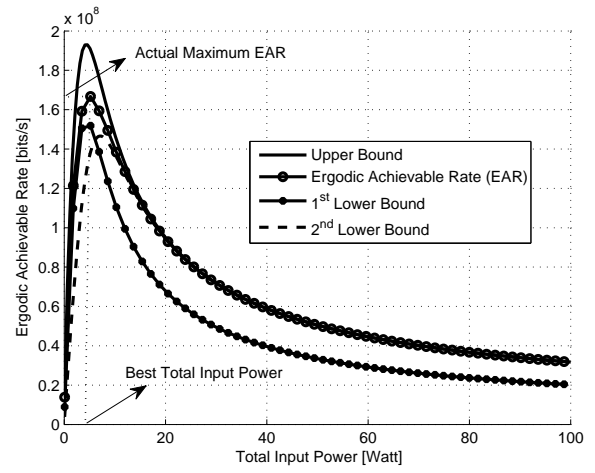


Fig. 4. Lower and upper bounds for the ergodic achievable rate of MIMO systems under PA nonlinearity assumption ($N_t = 3$, $N_r = 4$, $A_{os} = 1$, $v = 1$, $N_0 = 1$, $B = 20$ MHz)

(16). According to the definition $\mathbf{\Lambda}$ is diagonal, hence, we simply conclude that \mathbf{Q} must be diagonal as well.

Assume that $\hat{\mathbf{Q}}$ is the best power allocation, and also $\mathbf{\Pi}_i$, $i = 1, 2, \dots, N_t!$, is a permutation matrix which has exactly one "1" in each row and each column and zeros elsewhere. Since $\hat{\mathbf{Q}}$ is a diagonal matrix which satisfies $\text{Tr}(\hat{\mathbf{Q}}) \leq P_t$ and $\hat{\mathbf{Q}} \succeq \mathbf{0}$, it can be easily derived that $\tilde{\mathbf{Q}} = \frac{1}{N_t!} \sum_{i=1}^{N_t!} \mathbf{\Pi}_i \hat{\mathbf{Q}} \mathbf{\Pi}_i^\dagger$ satisfies both constraints as well. Now given the following function

$$\Psi(\mathbf{Q}) \triangleq \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + (\mathbf{H}\mathbf{D}\mathbf{H}^\dagger + N_0\mathbf{I}_{N_r})^{-1} \mathbf{H}\mathbf{\Lambda}\mathbf{Q}\mathbf{\Lambda}^\dagger \mathbf{H}^\dagger \right| \right] \quad (39)$$

we can demonstrate that $\Psi(\tilde{\mathbf{Q}}) \geq \Psi(\hat{\mathbf{Q}})$. By starting from the left hand side, and by taking into account the concavity of the function Ψ in (d) below, we get

$$\begin{aligned} \Psi(\tilde{\mathbf{Q}}) &= \Psi \left(\frac{1}{N_t!} \sum_{i=1}^{N_t!} \mathbf{\Pi}_i \hat{\mathbf{Q}} \mathbf{\Pi}_i^\dagger \right) \\ &\stackrel{(d)}{\geq} \frac{1}{N_t!} \sum_{i=1}^{N_t!} \Psi \left(\mathbf{\Pi}_i \hat{\mathbf{Q}} \mathbf{\Pi}_i^\dagger \right) \\ &= \frac{1}{N_t!} \sum_{i=1}^{N_t!} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \left(\mathbf{H} \left(\mathbf{\Pi}_i \mathbf{D} \mathbf{\Pi}_i^\dagger \right) \mathbf{H}^\dagger + N_0 \mathbf{I}_{N_r} \right)^{-1} \right. \right. \\ &\quad \left. \left. \left(\mathbf{H} \left(\mathbf{\Pi}_i \mathbf{\Lambda} \mathbf{\Pi}_i^\dagger \right) \left(\mathbf{\Pi}_i \hat{\mathbf{Q}} \mathbf{\Pi}_i^\dagger \right) \left(\mathbf{\Pi}_i \mathbf{\Lambda}^\dagger \mathbf{\Pi}_i^\dagger \right) \mathbf{H}^\dagger \right) \right| \right]. \end{aligned} \quad (40)$$

Note that any permutation on the input covariance matrix, leads to a same permutation on the matrices \mathbf{D} and $\mathbf{\Lambda}$ as we have already done in (40). Considering the fact that $\mathbf{\Pi}\mathbf{\Pi}^\dagger = \mathbf{I}$,

the last equation is simplified to

$$\Psi(\tilde{\mathbf{Q}}) \geq \frac{1}{N_t!} \sum_{i=1}^{N_t!} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \left(\mathbf{H} \mathbf{D} \mathbf{H}^\dagger + N_0 \mathbf{I}_{N_r} \right)^{-1} \right. \right. \\ \left. \left. \mathbf{H} \mathbf{A} \hat{\mathbf{Q}} \mathbf{A}^\dagger \mathbf{H}^\dagger \right| \right] = \Psi(\hat{\mathbf{Q}}). \quad (41)$$

As we assumed $\hat{\mathbf{Q}}$ is the best power allocation, so $\Psi(\tilde{\mathbf{Q}}) \geq \Psi(\hat{\mathbf{Q}})$ would be valid only by equality. In other words, $\tilde{\mathbf{Q}}$ is the best power allocation which can be written in terms of following matrix transformation

$$\tilde{\mathbf{Q}} = \frac{1}{N_t!} \sum_{i=1}^{N_t!} \mathbf{\Pi}_i \hat{\mathbf{Q}} \mathbf{\Pi}_i^\dagger = \frac{1}{N_t} \text{Tr}\{\hat{\mathbf{Q}}\} \mathbf{I}_{N_t} \triangleq \frac{P}{N_t} \mathbf{I}_{N_t} \quad (42)$$

which is indeed a scaled identity matrix. It is obvious that P can be interpreted as the total input power.

APPENDIX II PROOF OF COROLLARY 1

Let $\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\dagger$ be a singular value decomposition of the Rayleigh fading channel. Applying this singular value decomposition, and using the fact that $\det(\mathbf{I} + \mathbf{A} \mathbf{B}) = \det(\mathbf{I} + \mathbf{B} \mathbf{A})$ we derive a simpler closed-form expression for the maximum EAR

$$R = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \frac{P \alpha^2}{N_t} (N_0 \mathbf{I}_{N_r} + \sigma_d^2 \mathbf{H} \mathbf{H}^\dagger)^{-1} \mathbf{H} \mathbf{H}^\dagger \right| \right] \\ = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \frac{P \alpha^2}{N_t} (N_0 \mathbf{I}_{N_r} + \sigma_d^2 \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\dagger)^{-1} \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\dagger \right| \right] \\ = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_r} + \frac{P \alpha^2}{N_t} (\mathbf{U} (N_0 \mathbf{I}_{N_r} + \sigma_d^2 \mathbf{\Sigma}) \mathbf{U}^\dagger)^{-1} \mathbf{U} \mathbf{\Sigma} \mathbf{U}^\dagger \right| \right] \\ = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_t} + \frac{P \alpha^2}{N_t} (N_0 \mathbf{I}_{N_r} + \sigma_d^2 \mathbf{\Sigma})^{-1} \mathbf{\Sigma} \right| \right] \\ = \sup_{0 \leq P \leq P_t} \mathbb{E} \left[\log_2 \left| \mathbf{I}_{N_t} + \mathbf{Z} \right| \right]. \quad (43)$$

REFERENCES

- [1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecom*, vol. 10, no. 6, pp. 585–595, 1999.
- [2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, 1998.
- [3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [4] D. Persson, T. Eriksson, and E. G. Larsson, "Amplifier-aware multiple-input multiple-output power allocation," *IEEE Commun. Lett.*, vol. 17, no. 6, pp. 1112–1115, June 2013.
- [5] —, "Amplifier-aware multiple-input single-output capacity," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 913–919, Mar. 2014.
- [6] J. Qi and S. Aïssa, "On the power amplifier nonlinearity in MIMO transmit beamforming systems," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 876–887, Mar. 2012.
- [7] —, "Analysis and compensation of power amplifier nonlinearity in MIMO transmit diversity systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2921–2931, July 2010.
- [8] D. Dardari, V. Tralli, and A. Vaccari, "A theoretical characterization of nonlinear distortion effects in OFDM systems," *IEEE Trans. Commun.*, vol. 48, no. 10, pp. 1755–1764, Oct. 2000.

- [9] P. Banelli and S. Cacopardi, "Theoretical analysis and performance of OFDM signals in nonlinear AWGN channels," *IEEE Trans. Commun.*, vol. 48, no. 3, pp. 430–441, Mar. 2000.
- [10] P. Banelli, "Theoretical analysis and performance of OFDM signals in nonlinear fading channels," *IEEE Trans. Wireless Commun.*, vol. 2, no. 2, pp. 284–293, Mar. 2003.
- [11] G. K. Psaltopoulos and A. Wittneben, "Nonlinear MIMO: Affordable MIMO technology for wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 824–832, Feb. 2010.
- [12] F. H. Gregorio, *Analysis and compensation of nonlinear power amplifier effects in multi-antenna OFDM systems*. Helsinki University of Technology, Nov. 2007.
- [13] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer, 2008.
- [14] S. C. Cripps, *Advanced Techniques in RF Power Amplifier Design*. Artech House, 2002.
- [15] J. J. Bussgang, "Crosscorrelation functions of amplitude-distorted Gaussian signals," 1952.
- [16] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, 2007.
- [17] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [18] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [19] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [20] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.
- [21] X. Zhang, M. Matthaiou, E. Björnson, M. Coldrey, and M. Debbah, "On the MIMO capacity with residual transceiver hardware impairments," in *Proc. IEEE ICC*, June 2014, pp. 5299–5305.
- [22] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, Jan. 2013.
- [23] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Communications and Information theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [24] C. Rapp, "Effects of HPA-nonlinearity on a 4-DPSK/OFDM-signal for a digital sound broadcasting signal," in *1991 European Conf. Satellite Commun.*, vol. 1, Oct. 1991, pp. 179–184.