MIMO Capacity under Power Amplifiers Consumed Power and Per-Antenna Radiated Power Constraints

Hei Victor Cheng, Daniel Persson, and Erik G. Larsson
Department of Electrical Engineering (ISY), Linköping University, Sweden
Email: {hei.cheng, daniel.persson, erik.g.larsson}@liu.se

Abstract—We investigate the capacity of the multiple-input-multiple-output channel taking into account the consumed power in the power amplifiers. The mutual information is optimized with a limitation of total consumed power and per-antenna radiated power for a fixed channel with full channel state information at both the transmitter and receiver. The capacity is thus obtained by optimizing the input distribution to maximize the mutual information.

Since the optimization problem is non-convex, direct computation of the capacity suffers from high computational complexity. Hence upper and lower bounds on the capacity are given as benchmarks for different ad-hoc schemes. An efficient suboptimal algorithm is also presented. Numerical results show that the suboptimal algorithm performs close to the capacity.

Keywords—MIMO capacity, power amplifier, consumed power constraint, per-antenna power constraint

I. INTRODUCTION

Recently the use of smartphones and tablets has led to a tremendous increase in the demand for high data rates over the wireless networks. Meanwhile, wireless transmissions are required to be robust to fading and shadowing effects. The idea of using multiple antennas at both the transmitters and the receivers, i.e. multiple-input-multiple-output (MIMO) technology, can be a solution. Comparing to single antenna systems, MIMO offers higher data rates, i.e. the multiplexing gain, meanwhile contributes robustness against fading, i.e. the diversity gain, with the same amount of time and frequency resources. When full knowledge of the channel is available at the transmitter, the transmission can be optimized according to certain different criteria. Telatar shows in [1], with circularly symmetric complex Gaussian noise and full channel state information (CSI) at the transmitter, that the capacity of MIMO channels under a sum radiated power constraint over all the antennas is achieved by eigen-beamforming together with water-filling power allocation over different eigen-modes.

However, the sum power constraint considered in [1] is not realistic as it does not take into account the maximum output constraint of an individual power amplifier (PA). A more realistic setting is given by the per-antenna power constraints, which were first considered in [2]. An algorithm is derived in [3] to obtain the optimal input distribution for a given channel under per-antenna power constraints. Nevertheless in many

applications the power consumed by the power amplifiers consists of both the output power and the power losses in the hardware. Capacity optimization under per-antenna power constraints does not tell us how the power should be optimally allocated to the antennas given a specific channel, total power consumption in each PA, and a total power budget for all antennas.

According to a recent report, power amplifiers are estimated to consume 57% of the energy consumption in macro base stations [4, Figure 12]. Cooling contributes another 10%, which can also be reduced if power dissipation from the PA is reduced. Hence, it is essential to take the PA into account when designing the transmitters at base stations. The importance of carefully taking the hardware into consideration when designing communication systems has recently been put forward by others, e.g. in [5] and [6]. Despite the importance, few efforts are reported on MIMO capacity optimization with a total consumed power budget. Sub-optimal algorithms with a diagonal input covariance matrix are proposed in [7]. In [8] we derived the multiple-input single-input (MISO) capacity with full CSI and provide results on ergodic MISO capacity with per-antenna and total consumed power constraints. In [9] we solved a special case of the MIMO problem with per-antenna and total consumed power constraints, and parallel channels. Until now the capacity of the MIMO channel with per-antenna and total consumed power constraints is unknown and this work will investigate this problem.

The paper is organized as follows: In Section II we describe the system model and establish capacity results of the MIMO system under various power constraints, and we formulate the optimization problem we are going to solve to obtain the capacity under consumed power constraints. In Section III we derive the upper and lower bounds on the capacity. In Section IV we provide an efficient algorithm to find a local optimum point of the optimization problem. In Section V we present the numerical results.

In the following, scalars are denoted by lower-case letters, vectors are denoted by bold-face lower-case letters, and matrices are denoted by bold-face upper-case letters. A vector $\boldsymbol{x}$ of length N has components x_i, where $i = 1, ..., N$, and a matrix $\boldsymbol{X}$ of size $M \times N$ has elements $X_{m,n}$, with $m = 1, ..., M$ and $n = 1, ..., N$. The determinant of a matrix is denoted as $|\boldsymbol{X}|$. The symbol $\text{diag}(\boldsymbol{X})$ represents a vector formed by the elements on the main diagonal of $\boldsymbol{X}$. Further, $(\cdot)^{\mathrm{T}}$ is the transpose, $(\cdot)^*$ is the conjugate, and $(\cdot)^H$ is the conjugate transpose. The relation $\boldsymbol{X} \succeq \boldsymbol{0}$ tells that matrix $\boldsymbol{X}$ is positive semi-definite.

II. CAPACITY OF MIMO CHANNELS

For the MIMO transmission, we have the baseband channel model given by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}, \tag{1}$$

where $\boldsymbol{y} \in \mathbb{C}^{N_R \times 1}$ is the received signal, $\boldsymbol{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel, and the transmitted symbol vector $\boldsymbol{x} \in \mathbb{C}^{N_T \times 1}$, where N_T is the number of transmit antennas. The channel noise variable $\boldsymbol{n} \in \mathbb{C}^{N_R \times 1}$ is independent identically distributed (i.i.d.) zero-mean circularly symmetric complex Gaussian, i.e. $\boldsymbol{n} \sim \mathcal{CN}(0, \sigma^2\boldsymbol{I})$. In this paper we consider the case with full CSI at the transmitter and at the receiver for a fixed channel $\boldsymbol{H}$. In the case of additive Gaussian noise, it holds that for any power constraints, the signal input giving the largest mutual information (MI) is zero-mean circularly symmetric complex Gaussian with a covariance matrix $\boldsymbol{Q} = E[\boldsymbol{x}\boldsymbol{x}^H]$, which is positive semi-definite by definition [1]. The MI is given as

$$R(\boldsymbol{Q}) = \log_2 \left| \boldsymbol{I} + \frac{1}{\sigma^2}\boldsymbol{H}\boldsymbol{Q}\boldsymbol{H}^H \right| \tag{2}$$

bit/s/Hz, cf. [1]. To find the capacity, the input distribution needs to be optimized subject to all power constraints. In other words, we want to find

$$\sup_{\boldsymbol{Q} \in \mathcal{Q}} R(\boldsymbol{Q}), \tag{3}$$

where $\mathcal{Q}$ is the set of power constraints that are of interest. The radiated power constraint on antenna i is

$$E[|x_i|^2] = Q_{i,i} \le P_{\max}, \tag{4}$$

where $P_{\max}$ is the maximum possible power on antenna i. For the consumed power, following the discussions about amplifier modeling in [9], we set

$$\frac{Q_{i,i}}{P_{\mathrm{cons},i}} = \eta_{\max} \left(\frac{Q_{i,i}}{P_{\max}} \right)^{1-\epsilon}, \tag{5}$$

where ϵ is a parameter with $\epsilon \in [0, 0.5]$, $P_{\mathrm{cons},i}$ is the consumed power on antenna i, and $\eta_{\max} \in [0, 1]$ is the maximum power efficiency obtained only when $Q_{i,i} = P_{\max}$. If $\eta_{\max} = 0$, all power is dissipated. The maximum power efficiency is a fixed parameter, common for all the employed PAs at the transmitter in this paper. From the above, we can write

$$P_{\mathrm{cons},i} = \frac{1}{\eta_{\max}} Q_{i,i}^\epsilon P_{\max}^{1-\epsilon}. \tag{6}$$

The consumed power $P_{\mathrm{cons},i}$ is thus proportional to the ϵ-th power of the radiated power.

We are now able to state the rate maximization problem corresponding to (3)

$$\begin{aligned} \underset{\boldsymbol{Q}}{\text{maximize}} \quad & R(\boldsymbol{Q}) \\ \text{subject to} \quad & \frac{P_{\max}^{1-\epsilon}}{\eta_{\max}} \sum_{i=1}^{N_T} Q_{i,i}^\epsilon \le \tilde{P}_{\mathrm{tot}} \\ & 0 \le Q_{i,i} \le P_{\max}, \;\; i = 1, ..., N_T \\ & \boldsymbol{Q} \succeq \boldsymbol{0}, \end{aligned} \tag{7}$$

where $\tilde{P}_{\mathrm{tot}}$ is the total consumed power limitation for all transmit antennas. This means that we aim to find the capacity, subject to both the consumed power constraint and the PA power constraints. Writing

$$P_{\mathrm{tot}} = \frac{\tilde{P}_{\mathrm{tot}} \eta_{\max}}{P_{\max}^{1-\epsilon}}, \tag{8}$$

we can express the optimization problem we aim to solve as follows

$$\begin{aligned} \underset{\boldsymbol{Q}}{\text{maximize}} \quad & R(\boldsymbol{Q}) \\ \text{subject to} \quad & \sum_{i=1}^{N_T} Q_{i,i}^\epsilon \le P_{\mathrm{tot}} \\ & 0 \le Q_{i,i} \le P_{\max}, \;\; i = 1, ..., N_T \\ & \boldsymbol{Q} \succeq \boldsymbol{0}. \end{aligned} \tag{9}$$

III. UPPER BOUND AND LOWER BOUND ON CAPACITY

Problem (9) is non-convex due to the non-convex constraint on consumed power. Therefore finding the globally optimal solution is hard. However we can obtain bounds on the capacity as follows. Consider the constraint

$$f(\boldsymbol{Q}) \triangleq \sum_{i=1}^{N_T} Q_{i,i}^\epsilon \le P_{\mathrm{tot}}. \tag{10}$$

By upper bounding each $Q_{i,i}^\epsilon$ with a linear function we obtain a lower bound on the capacity as we are restricting to a smaller feasible set. As $Q_{i,i}^\epsilon$ is a concave function the first order Taylor expansion will give an upper bound on $Q_{i,i}^\epsilon$. By dividing the region $[0, P_{\max}]$ into K different parts $\{[0, P_{\max}/K], \ldots, [P_{\max}(K-1)/K, P_{\max}]\}$, we can get an upper bound on $Q_{i,i}^\epsilon$ in each interval by picking any point in the sub-region and apply a first order Taylor approximation around that point. Herein we chose the point as the mid-point of each sub-region, i.e. $P_{\max}/2K, 3P_{\max}/2K, \ldots, (2K-1)P_{\max}/2K$. Then in each sub-region, the optimal $\boldsymbol{Q}$ can be obtained by solving the following set of optimization problems

$$\begin{aligned} \underset{\boldsymbol{Q}}{\text{maximize}} \quad & R(\boldsymbol{Q}) \\ \text{subject to} \quad & \sum_{i=1}^{N_T} a_i^j Q_{i,i} + b_i^j \le P_{\mathrm{tot}} \\ & \frac{c_i^j - 1}{K} P_{\max} \le Q_{i,i} \le \frac{c_i^j}{K} P_{\max}, \;\; i = 1, ..., N_T \\ & \boldsymbol{Q} \succeq \boldsymbol{0}, \end{aligned} \tag{11}$$

where $c_i^j = 1, \ldots, K$ indicate the sub-region each antenna i is in, and

$$a_i^j \triangleq \epsilon \left(\frac{2c_i^j - 1}{2K} P_{\max} \right)^{\epsilon-1}, b_i^j \triangleq (1-\epsilon) \left(\frac{2c_i^j - 1}{2K} P_{\max} \right)^\epsilon \tag{12}$$

specify the first order Taylor approximation. To get the lower bound, we choose the $\boldsymbol{Q}$ corresponding to the maximum rate obtained by solving the K^{N_T} problems in different regions.

Similarly, if we lower bound each $Q_{i,i}^\epsilon$ with a linear function, we obtain an upper bound on the capacity as we are relaxing the constraints. Again from the fact that $Q_{i,i}^\epsilon$ is concave, the lower bounds can be constructed using the quantized

regions above. Specifically we join all the neighbouring vertices of each boundary $0, P_{\max}/K, \ldots, P_{\max}(K-1)/K, P_{\max}$ by straight lines. Hence the upper bound on capacity can be obtained by solving (11) with the variables in the first constraint defined as

$$
\begin{aligned}
a_i^j &\triangleq \left[\left(\frac{c_i^j}{K} P_{\max} \right)^{\epsilon} - \left(\frac{c_i^j - 1}{K} P_{\max} \right)^{\epsilon} \right] / (P_{\max}/K), \\
b_i^j &\triangleq c_i^j \left(\frac{c_i^j - 1}{K} P_{\max} \right)^{\epsilon} - (c_i^j - 1) \left(\frac{c_i^j}{K} P_{\max} \right)^{\epsilon}.
\end{aligned}
\tag{13}
$$

To get the upper bound, we choose the $\boldsymbol{Q}$ corresponding to the maximum rate obtained by solving the K^{N_T} problems in different regions. The upper and lower bounds will converge if the number K of quantization regions of $Q_{i,i}$ increases and thus give us the desired capacity.

IV. EFFICIENT ALGORITHM FOR SUBOPTIMAL SOLUTION

The computational complexity of obtaining the upper and lower bounds given in Section III is polynomial in K and exponential in N_T. The computational complexity increases polynomially with K, however with large N_T the exponential complexity is not feasible. Therefore a practical algorithm which can perform well with acceptable complexity is preferred when N_T is large. We now propose a successive convex approximation algorithm to serve the purpose.

The successive convex approximation is an algorithm framework [10] used in power control applications to handle non-convex constraints [11, Sec IV. A]. The idea is to solve a series of approximated problems where the non-convex constraint is approximated with a convex function in each problem. For our problem (9), we approximate the non-convex function $f(\boldsymbol{Q})$ with a convex function $f_k(\boldsymbol{Q})$ in the k-th iteration. The convex optimization problem to be solved in the k-th iteration is

$$
\begin{aligned}
\underset{\boldsymbol{Q}}{\text{maximize}} \quad & R(\boldsymbol{Q}) \\
\text{subject to} \quad & f_k(\boldsymbol{Q}) \leq P_{\text{tot}} \\
& 0 \leq Q_{i,i} \leq P_{\max}, \quad i = 1, \ldots, N_T \\
& \boldsymbol{Q} \succeq \mathbf{0}.
\end{aligned}
\tag{14}
$$

If we construct a family of functions $f_k(\mathbf{Q})$ in each iteration k satisfying the conditions

1) $f(\mathbf{Q}) \leq f_k(\mathbf{Q})$, $\forall \mathbf{Q}$ in the feasible set,
2) $f(\mathbf{Q}^{(k-1)}) = f_k(\mathbf{Q}^{(k-1)})$, where $\mathbf{Q}^{(k-1)}$ is the solution from the previous iteration,
3) $\nabla f(\mathbf{Q}^{(k-1)}) = \nabla f_k(\mathbf{Q}^{(k-1)})$,

the algorithm will give a solution satisfying the Karush-Kuhn-Tucker (KKT) conditions for the original problem [10].

The first condition is to ensure that the solution we get is feasible for the original problem. The second condition ensures that the solution from the previous iteration is feasible for the current iteration. As a result the objective value of the original problem increases after every iteration since the solution from the previous iteration is a feasible solution to the problem in the current iteration. The second and third conditions together guarantee that the KKT conditions for the original problem

will be satisfied at convergence. Since each $\mathbf{Q}_{i,i}^{\epsilon}$ is concave in $\mathbf{Q}$, we can easily verify that approximating them with the first order Taylor expansion will give a function satisfying all the above conditions. Therefore we choose the $f_k(\mathbf{Q}^{(k-1)})$ as follows

$$
f_k(\mathbf{Q}^{(k-1)}) = \sum_{i=1}^{N_T} \epsilon (Q_{i,i}^{(k-1)})^{\epsilon-1} (Q_{i,i} - Q_{i,i}^{(k-1)}) + (Q_{i,i}^{(k-1)})^{\epsilon}
\tag{15}
$$

To conclude, we can obtain a suboptimal solution to (9) with the procedure described in Algorithm 1.

Algorithm 1 Successive convex optimization for problem (9)

1: choose $\mathbf{Q}^{(0)}$ satisfying the constraints and initialize $k = 1$
2: **repeat**
3:     form the k-th approximated problem (14) of (9) by approximating $f(\mathbf{Q})$ with $f_k(\mathbf{Q})$ around $\mathbf{Q}^{(k-1)}$ as in (15),
4:     solve the k-th approximated problem to obtain $\mathbf{Q}^{(k)}$,
5:     $k \leftarrow k + 1$
6: **until** convergence
7: **return** $\mathbf{Q}^{(k)}$

V. NUMERICAL EXPERIMENTS

We choose the experimental parameters as follows. The parameter $\eta_{\max} = 0.55$ which is a realistic value, cf. [12, eq. (6)] and $\epsilon = 0.5$,. We want to investigate the system operating with a total consumed power less than the maximum possible total consumed power (when all antennas are operating at $P_{\max}$). With maximum possible total consumed power, our solution and the solution in [3] are the same. Hence we choose $P_{\max} = 1$, $\tilde{P}_{\text{tot}} = 1.818$ such that $P_{\text{tot}} = 1$ for simplicity, i.e. the ratio between $P_{\max}$ and P_{tot} is constant, while the ratio between P_{tot} and σ^2 is varied. We define the signal-to-noise ratio (SNR) to be P_{tot}/σ^2, and rate is measured in bits per channel use (bpcu).

A. Schemes for Comparison

For comparison with the proposed methods in Section IV, we employ the following methods.

1) Uniform Power Allocation: The power is uniformly allocated among all the transmit antennas, i.e. $Q_{i,i} = (P_{\text{tot}}/N_T)^{1/\epsilon}$, for $i = 1, \ldots, N_T$.
2) Random Antenna Selection: We randomly choose a subset of the antennas and allocate all power among them.
3) Antenna Selection: We use antenna selection with our algorithm in Section IV. We choose the antennas corresponding to the columns of $\boldsymbol{H}$ with largest Euclidean norm. Thereafter, we run the algorithm in Section IV only among these antennas. This can greatly reduce the complexity of the optimization procedure since we reduce the number of dimensions.

B. Simulation Results

Fig. 1 shows rate for varying P_{tot}/σ^2 in a 2×2 MIMO system averaged over 100 channel realizations. The number of sub-regions K for generating the upper and lower bounds is
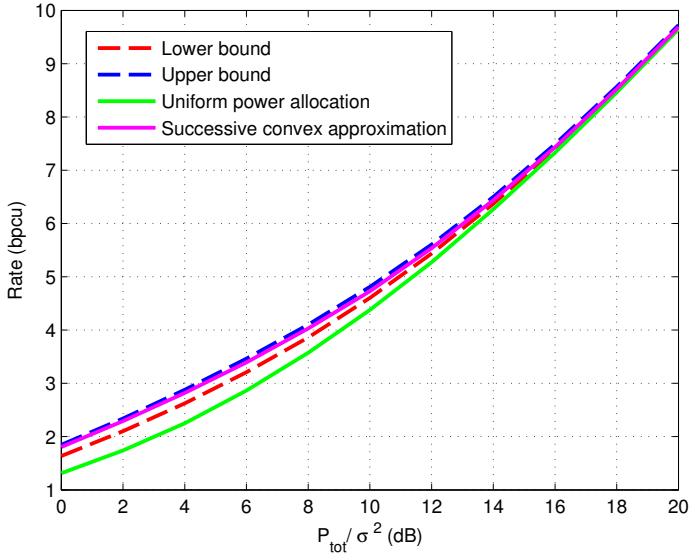
Fig. 1. Rate against SNR for a 2x2 MIMO system averaged over 100 channel realizations, $P_{\max} = 1$, and $P_{\mathrm{tot}} = 1$. The number of sub-regions K for generating the upper and lower bounds is 10.
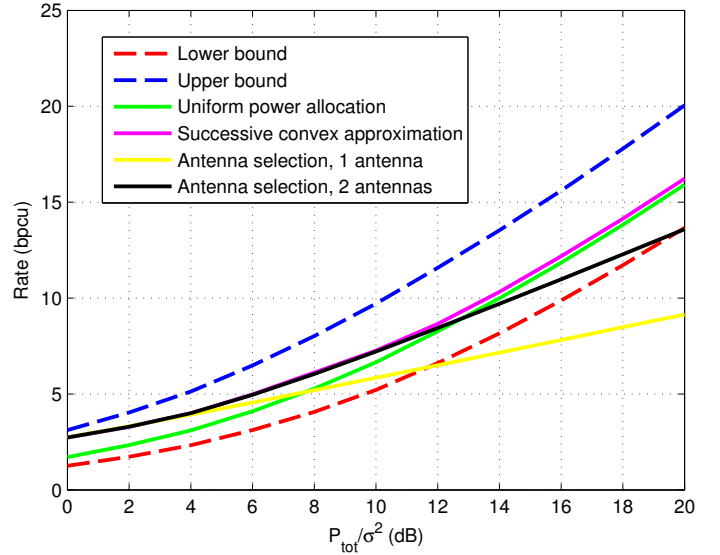

Fig. 2. Rate against SNR for a 4x4 MIMO system averaged over 100 channel realizations, $P_{\max} = 1$, and $P_{\mathrm{tot}} = 1$. The number of sub-regions K for generating the upper and lower bounds is 3.

10. We observe that at low SNR, the same rate is achieved with a SNR gain of 2.5 dB from optimizing the input distribution compared with the uniform power allocation scheme. We note that at low SNR, the optimal strategy is to choose a single antenna to which all power is allocated. At high SNR, all 4 curves converge, which suggests that uniform power allocation is close to optimal in that regime. This is analogous to the water-filling results in traditional MIMO capacity with a sum radiated power constraint. The difference is that in our case the power is allocated among the antennas, not the eigen-modes. This suggest that we can turn off the hardware associated with the antennas that are not in use and save power [13], [14].

Similar results are observed in a 4×4 MIMO system as illustrated in Fig. 2. The number of sub-regions K for generating the upper and lower bounds is 3. We compare to the antenna selection schemes suggested above with 1 and 2 antennas, which are close to optimal at low SNR, but non-optimal at high SNR. On the other hand, uniform power allocation is close to optimal at high SNR, but sub-optimal at low SNR. The proposed algorithm in Section IV performs best over the whole range of investigated SNR.

In a scenario when there are more transmit antennas than receive antennas, which is a practical case in downlink cellular communication, there is always a gain from optimizing the input distribution compared to uniform power allocation, as shown in Fig. 3. Moreover we observe that in this case using antenna selection scheme with 2 antennas is close to optimal in most cases. This suggests that using the same number of transmit and receive antennas does not lose much if we choose appropriate antennas. We also compare to the random antenna selection scheme with 1 antenna which performs poorly in all regimes. Choosing the best antennas for antenna selection has a modest computational complexity of $O(N_T \log(N_T))$.

Finally we are looking at a large MIMO system $N_T = N_R = 10$. Fig. 4 shows the rate in such system for varying $P_{\mathrm{tot}}/\sigma^2$ and averaging over 100 channel realizations. We
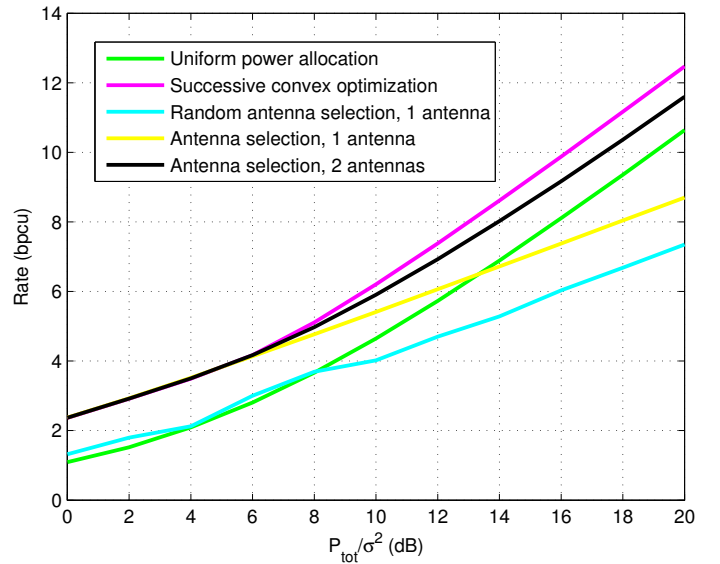

Fig. 3. Rate against SNR for a 10x2 MIMO system averaged over 100 channel realizations, $P_{\max} = 1$, and $P_{\mathrm{tot}} = 1$.

can see that in this case using a small number of antennas pays off more compared to the case with 2 receive antennas. There are still gains of antenna selection over random antenna selection in all SNR regimes but the gap is getting smaller with increasing number of receive antennas. The gains reduce from around 1 bpcu in a 10×2 system to around 0.6 bpcu in a 10×10 system.

C. Discussion

In all of our investigated cases, antenna-selection results are showing up under the sum-consumed power constraint at low SNR, but not under a sum- radiated power constraint. This phenomenon can be intuitively explained by that the consumed
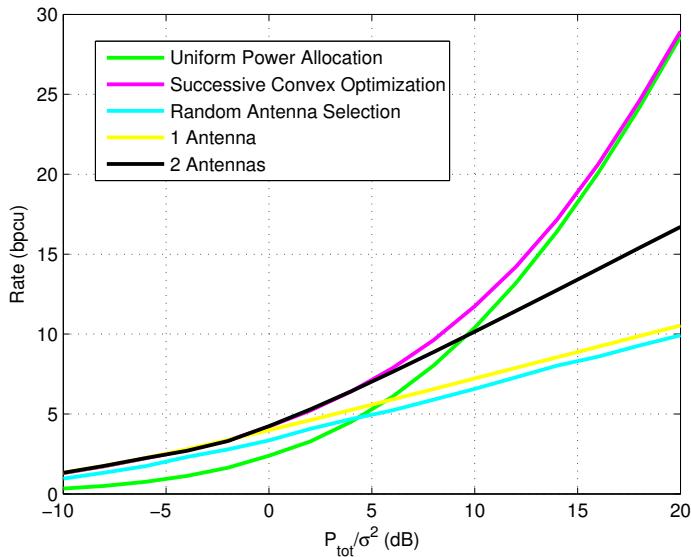
Fig. 4. Rate against SNR for a 10x10 MIMO system averaged over 100 channel realizations, $P_{\max} = 1$, and $P_{\text{tot}} = 1$.

power can be expressed as a constant times the ϵ-th power of the radiated power. Low radiated power will be penalized more compared to high radiated power, i.e. a lot of consumed power gives very little radiated power in this regime. This effect promotes allocating either a lot of power, or no power, to an antenna. Furthermore we observe that the constraint (10) is equivalent to

$$||\text{diag}(\boldsymbol{Q})||_\epsilon = \left(\sum_{i=1}^{N_t} Q_{i,i}^\epsilon \right)^{\frac{1}{\epsilon}} \leq P_{\text{tot}}^{1/\epsilon} \qquad (16)$$

This provides a link to the results in the compressed sensing literature. It is known that the ℓ_p-norm with $0 < p < 1$ enhances sparsity, and therefore it is used to replace the ℓ_0 norm, c.f. [15] and [16].

For a smaller ϵ, the sparsity enhancement effect is more significant, as we observed in preliminary simulations. Values of ϵ close to 0 are important for some classes of PAs, for example, [7] states that class A PAs have ϵ close to 0. We also note that there are other ways of performing antenna selection than by the maximum column norm described in Section V-A, i.e. antenna selection may be an effective solution more often than what is observed in Fig. 1 to Fig. 4.

VI. CONCLUSION

In this work we considered the capacity of the MIMO channel taking into account both a limitation of total consumed power and per-antenna radiated power constraints. For a fixed channel with full CSI at both the transmitter and the receiver, maximization of the mutual information was formulated as an optimization problem. Lower and upper bounds on the capacity were provided by numerical algorithms based on partitioning of the feasible region. Both bounds were shown to converge and give the exact capacity when number of regions increases. An efficient suboptimal algorithm performing close to the capacity was also presented.

Simulation results showed that in the low SNR regime, antenna selection was the optimal scheme while at high SNR uniform power allocation was close to optimal. The sparsity of antennas being used increased when the parameter ϵ was decreased further. Our results indicated that antenna selection may play an important role in some future wireless systems.

REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[2] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO transmit beamforming under uniform elemental power constraint," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5395–5406, Nov. 2007.

[3] M. Vu, "The capacity of MIMO channels with per-antenna power constraint," *IEEE Transactions on Information Theory, submitted*, Jun. 2011. [Online]. Available: http://arxiv.org/abs/1106.5039

[4] "EARTH - energy aware radio and network technologies deliverable d2.3: Energy efficiency analysis of the reference systems, areas of improvements and target breakdown," 2012. [Online]. Available: https://bscw.ict-earth.eu/pub/bscw.cgi/d71252/EARTH_WP2_D2.3_v2.pdf

[5] M. Dohler, R. Heath, A. Lozano, C. Papadias, and R. Valenzuela, "Is the PHY layer dead?" *IEEE Communications Magazine*, vol. 49, no. 4, pp. 159 –165, Apr. 2011.

[6] G. Fettweis, M. Lohning, D. Petrovic, M. Windisch, P. Zillmann, and W. Rave, "Dirty RF: a new paradigm," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 4, Sep. 2005, pp. 2347–2355.

[7] A. He, S. Srikanteswara, K. K. Bae, T. Newman, J. Reed, W. Tranter, M. Sajadieh, and M. Verhelst, "Power consumption minimization for MIMO systems - a cognitive radio approach," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 2, pp. 469–479, 2011.

[8] D. Persson, T. Eriksson, and E. Larsson, "Amplifier-aware multiple-input single-output capacity," *IEEE Transaction on Communications, accepted for publication*, 2013.

[9] ——, "Amplifier-aware multiple-input multiple-output power allocation," *IEEE Communications Letters*, vol. 17, no. 6, pp. 1112–1115, Jun. 2013.

[10] B. R. Marks and G. P. Wright, "A general inner approximation algorithm for nonconvex mathematical programs," *Operations Research*, vol. 26, no. 4, pp. 681–683, 1978.

[11] M. Chiang, C. wei Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *IEEE Transactions on Wireless Communications*, vol. 6, no. 7, pp. 2640–2651, July 2007.

[12] G. Tsouri and D. Wulich, "Impact of linear power amplifier efficiency on capacity of OFDM systems with clipping," in *IEEE Convention of Electrical and Electronics Engineers in Israel*, Dec. 2008, pp. 134–136.

[13] C. Yu and A. Zhu, "A single envelope modulator-based envelope-tracking structure for multiple-input and multiple-output wireless transmitters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 10, pp. 3317–3327, Oct. 2012.

[14] "Texas Instruments RF front end power solutions for 2G, 3G, and 4G enabled devices," 2013. [Online]. Available: http://www.ti.com/ww/en/analog/power_management/envelope_tracking.htm

[15] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, Oct 2007.

[16] S. Foucart and M.-J. Lai, "Sparsest solutions of underdetermined linear systems via ℓ_q-minimization for $0 < q \leq 1$," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 395 – 407, 2009.