HAL
open science

Predistortion Techniques for Vector Perturbation Precoding of One-Bit Massive-MIMO

Inbar Fijalkow, A Lee Swindlehurst

To cite this version:

Inbar Fijalkow, A Lee Swindlehurst. Predistortion Techniques for Vector Perturbation Precoding of One-Bit Massive-MIMO. ASILOMAR, 2017, Pacic Grove, CA, United States. hal-01704849

HAL Id: hal-01704849

https://hal.science/hal-01704849

Submitted on 8 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Predistortion Techniques for Vector Perturbation Precoding of One-Bit Massive-MIMO

Inbar Fijalkow[1], A. Lee Swindlehurst[2]
[1] ETIS, UMR 8051, Université Paris Seine, Université Cergy-Pontoise, ENSEA, CNRS, F-95000 Cergy, France, inbar.fijalkow@ensea.fr
[2] Center for Pervasive Communications and Computing, University of California Irvine, Irvine, CA 92697, USA, swindle@uci.edu

Abstract - Massive MIMO aims to build wireless base stations with hundreds of coherently operating antennas serving tens of single antenna users in order to improve the capacity by a factor of 10-50 and the energy-efficiency trade-off by up to a thousand times. One-bit quantized precoding has been proposed to efficiently handle the RF front-end when the array is implemented with so many antennas. In this paper, we modify signal dynamic range reduction techniques such as active constellation extension (ACE) to perform symbol vector perturbation prior to the one-bit quantization. The resulting one-bit quantization of the perturbed vector provides a non-linear precoding with low complexity. Our simulations show that the proposed vector precoding outperforms both one-bit quantized Zero-Forcing and the much more complex vector perturbation precoding.

I. Introduction

Massive MIMO aims to build wireless base stations with hundreds of coherently operating antennas serving tens of single antenna users in order to improve the capacity by a factor of 10-50 and the energy-efficiency trade-off by up to a thousand times [1], [2]. However, the resulting power consumption and hardware costs increase with the number of antennas [3]. A fundamental question is how to handle the RF front-end with so many antennas. The answer will determine whether it is feasible to build high-end multi-user massive MIMO base stations using low-cost RF devices. We believe that the dirty-RF paradigm (see for instance [4]) will help the design of waveforms and signal processing algorithms to simplify the RF front-end implementation. We focus on an approach that has gained attention recently, namely constant envelope signals using one-bit quantization on the in-phase (I) and quadrature (Q) dimensions that maximize the power amplifier efficiency, eliminate the adaptive gain controller and minimize the power consumed by analog-to-digital conversion (ADC) and digital-to-analog conversion (DAC) [5].

Direct nonlinear designs that take the one-bit constraint into account when designing the transmit signal vector have been shown to perform very well [6]–[8], but their complexity is high and may not be suitable for real implementations. On the other hand, a very simple approach is to quantize the output of a standard linear precoder [9], but it exhibits an error floor at moderate to high SNRs. An improved version of the technique employing vector perturbation has been proposed in [10], but it still requires a greedy search of dimension equal to the number of antennas, which is large in massive MIMO applications. In this paper, we propose to revisit the active constellation extension approach of [11] to reduce the signal dynamic range before quantization. This method has the advantage of operating on the symbols prior to the linear precoding, which is only of dimension equal to the number of users, typically an order of magnitude smaller than the number of antennas.

The paper is organized as follows. Section II introduces the one-bit downlink model, and Section III presents existing one-bit quantized precoding techniques. In Section IV, we propose the generalization of ACE for vector perturbation and show the results of simulations that indicate its excllent performance. In what follows, uppercase boldface letters, $\mathbf{A}$, indicate a matrix, with $[\mathbf{A}]_{kl}$ and a_{kl} interchangeably denoting the element at the kth row and l^{th} column. Lower boldface letters, $\mathbf{a}$, indicate a column vector, with a_k denoting the kth element of the column vector. The symbols $(.)^*$, $(.)^T$ and $(.)^H$ denote the complex conjugate, matrix transpose and the transpose-conjugate of the argument respectively.

II. One-Bit Downlink System Model

A. Mathematical Notation and Assumptions

We assume a flat-fading downlink scenario in which an M-antenna BS is attempting to send QPSK symbols s_k to $k = 1, \cdots, K$ single-antenna users. The BS transmits an $M \times 1$ vector $\sqrt{\rho}\mathbf{x}$, where $\sqrt{\rho}$ is a fixed gain and the elements of $\mathbf{x}$ are constrained to be equal to $\pm 1 \pm j$, i.e., one-bit in the I-Q components of the signal at the BS.

Let $r_k = \mathbf{h}_k\mathbf{x} + n_k$ be the signal received by user k where $\mathbf{h}_k$ represents the $1 \times M$ channel to user k. We write the received signals from all users in a single vector $\mathbf{r} = [r_1 \ \cdots \ r_K]^T$ so that the overall system model becomes

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbf{x} + \mathbf{n} \ , \tag{1}$$

where the $K \times 1$ vector $\mathbf{n}$ represents a vector of independent Gaussian noise terms of variance σ_n^2 at each user, and $\mathbf{H}$ is the $K \times M$ matrix with k^{th} row $\mathbf{h}_k$. We will assume that the channel $\mathbf{H}$ is known at the BS.

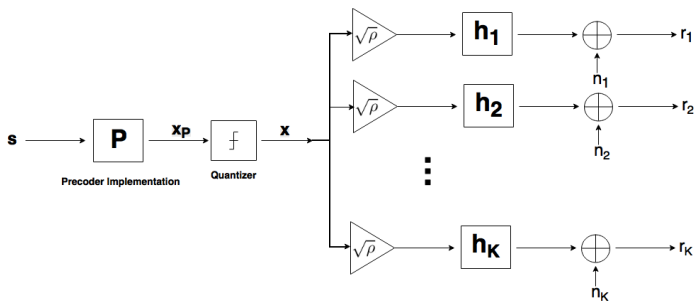

Fig. 1: System Model

For the downlink, the BS designs the precoded vector $\mathbf{x}$ such that the single antenna user k correctly detects the QPSK symbol s_k from r_k without supplementary processing. We denote $\mathbf{s} = [s_1 \ \cdots \ s_K]^T$ and $\mathcal{S} = \{1+j, 1-j, -1+j, -1-j\}$ as the set of QPSK constellation points. We assume that $\rho = \frac{\rho_0}{M}$, where ρ_0 is defined to be the transmit SNR.

III. One-Bit Downlink Precoding

A. One-bit Quantized Linear Precoding

The simplest one-bit vector precoder consists of the output of a standard linear precoder quantized by one-bit DACs prior to transmission. In particular, assuming a linear precoding matrix $\mathbf{P}$, the transmitted signal is $\mathbf{x} = \mathbb{Q}(\mathbf{P}\mathbf{s})$, where the one-bit quantization operation is defined as

$$\mathbb{Q}(\mathbf{x}) = \text{sign}(\Re(\mathbf{x})) + j.\text{sign}(\Im(\mathbf{x})), \tag{2}$$

with $\Re(\cdot)$ representing the real part, $\Im(\cdot)$ the imaginary part, and sign$(\cdot)$ the sign of their arguments.

The most commonly used precoder is the Zero-Forcing (ZF) precoder where $\mathbf{P} = \mathbf{H}^\dagger(\mathbf{H}\mathbf{H}^\dagger)^{-1}$ and $\mathbf{x}_{ZF}$ denotes the one-bit quantized ZF precoded vector. The work of [9] provides an asymptotic analysis of its performance. It provided a good approximation of the BER under the hypotheses that $M \gg K \gg 1$ and $\mathbf{H} = \Sigma\tilde{\mathbf{H}}$ is a $K \times M$, full row-rank matrix where $\tilde{\mathbf{H}}$ is composed of zero-mean, unit-variance i.i.d. Gaussian elements, and Σ is a diagonal matrix of path gains for the different users. For high SNR and equal path gains, [9] derived the high SNR error floor to be

$$P_e \approx 2Q\left(\sqrt{\frac{\frac{2}{\pi}}{1-\frac{2}{\pi}}(\frac{M}{K}-1)}\right) \tag{3}$$

illustrating how the error rate decreases with the ratio $\frac{M}{K}$. The noise-free received vector, $\mathbf{H}\mathbf{x}$, is shown to become propotional to $\mathbf{s}$ with a scaling factor increasing with $\frac{M}{K}$, and hence the multi-user interference vanishes in the asymptotic regime. These theoretical values match very well with the simulation results even for non-asymptotic values of K and M. However, the performance of the one-bit quantized Zero-Forcing precoder degrades, showing the need of improved precoders for moderate to large values of K and M.

B. MSE Precoding

Assuming Gaussian noise, one might choose to implement the maximum likelihood encoder, which attempts to solve the minimum squared error (MSE) [12]

$$\mathbf{x} = \arg \min_{\mathbf{v}\in\mathcal{S}^M} ||\alpha\mathbf{s} - \mathbf{H}\mathbf{v}||^2 , \tag{4}$$

for some value α. In general, (4) requires on exhaustive search of order $O(4^M)$, which is prohibitively expensive even for relatively small values of M, let alone in the massive MIMO case. In [6], a method based on semi-definite relaxation is proposed as an approximate solution.

However, minimizing (4) does not minimize the bit error rate (BER), an observation that has been exploited in the development of precoders based on the idea of "constructive interference" [13]–[15]. For the case where the elements of the desired vector $\mathbf{s}$ are themselves drawn from a QPSK alphabet, the ML encoder over-constrains the problem by attempting to force $\mathbf{H}\mathbf{x}$ to be close to $\mathbf{s}$, when in fact all that is necessary is that its elements lie within the correct decision regions so that the users can properly decode them as the desired constellation points s_k. The noise-free components of $\mathbf{H}\mathbf{x}$ can in principle be far away from $\mathbf{s}$ and still be decoded correctly; in fact, often the farther $\mathbf{H}\mathbf{x}$ is away from $\mathbf{s}$ in the correct region, the farther the received signal is from the decision boundaries and hence the more resilient to noise. This idea has been applied to the problem of one-bit precoding for the massive MIMO downlink in [8] and shown to provide excellent results, albeit with a relatively high computational cost.

C. Greedy Vector Pertubation Precoding

Our recent paper [10] proposed to perturb the quantized ZF transmit vector $\mathbf{x}_{ZF}$ in order to reduce the metric used in the above constructive interference work, which corresponds to minimizing the BER for the worst-case user. This vector perturbation method is performed either directly on the size-M vector $\mathbf{x}_{ZF}$ or on the size-K symbol vector $\mathbf{s}$, yielding similar performance. The method that performs the perturbation the length M vector generally outperforms the other approach, but requires greater computation in massive MIMO scenarios since M is much larger than K. The motivation for the method proposed in this paper is to find a low-complexity perturbation method with improved performance that works in the K-dimensional space of the signals.

IV. ACE for Vector Perturbation

Vector symbol perturbation has been proposed for OFDM systems in a systematic way in order to perform signal dynamic range reduction in [11]. This method is referred to as Active Constellation Extension (ACE). In what follows, we show how this method may be modified in order to address the one-bit precoding problem described above. Interestingly, we find that the adaptive update to the signals for the one-bit problem should in fact be the opposite of that used in [11].

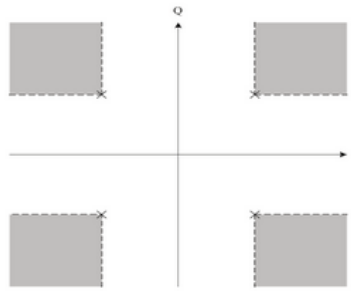

Fig. 2: Region of interest of QPSK, $\mathcal{F}(\mathbf{s})$.

A. Generalized ACE

In OFDM, the transmitter side performs an inverse DFT to process the symbol vector before transmission. It can be viewed as a square precoder with $P_{m,k} = \frac{1}{\sqrt{M}} e^{2\pi jmk/M}$. The DFT performed at the receiver can be viewed as an operation similar multiplication by the channel matrix $\mathbf{H}$ in our application. Generalizing ACE to a non-square linear precoder $\mathbf{P}$ instead of the inverse DFT, we propose to find $\underline{\epsilon}$ so that:

$$\underline{\epsilon} = \arg \min_{\underline{\epsilon} \in \mathcal{F}(\mathbf{s})} \max_{n} |(\mathbf{Ps})_n + \sum_{k=1}^{K} \mathbf{P}_{n,k} \underline{\epsilon}_k|^2 \tag{5}$$

where the region of interest $\mathcal{F}(\mathbf{s})$ is defined in gray in Figure 2. The motivation for this approach is an attempt to eliminate the amplitude variation in the transmit vector $\mathbf{x}$ prior to application of the one-bit quantization.

The proposed method operates as follows. Given an initial estimate of the unquantized transmit vector $\mathbf{x}$ (e.g., from the ZF precoder), we project the elements of $\mathbf{x}$ onto the circle in the complex plane of radius a, and then multiply the result by the channel $\mathbf{H}$. We then project the resulting perturbation $\underline{\epsilon}$ to the received symbols, and project it onto the set $\mathcal{F}(\mathbf{s})$, effectively eliminating any perturbations that decrease the distance of the worst-case user to the decision boundaries of the received symbols. We alternate between these two projections for a fixed number of iterations T_{max}. At that point the one-bit quantization is applied to the result in order to obtain the transmit vector. The details of the algorithm are outlined below.

1) initialization: $\mathbf{x}^{(0)} = \mathbf{x}_{ZF} = \mathbf{Ps}$
2) at iteration m:
   (a) Projection on radius a sphere, the clipped-off signal:
   - $e_n^{(m)} = \begin{cases} x_n^{(m)} \left(\frac{a}{|x_n^{(m)}|} - 1 \right) & \text{when } |x_n^{(m)}| > a \\ 0 \text{ otherwise} \end{cases}$

   (b) Projection on the region of interest $\mathcal{F}(\mathbf{s})$:
   - $\tilde{\underline{\epsilon}}^{(m)} = \mathbf{He}^{(m)}$
   - $\mathcal{R}(\epsilon_k^{(m)}) = \mathcal{R}(\tilde{\epsilon}_k^{(m)}) \mathbf{1}_{[0,\infty[} \left(sign(\mathcal{R}(s_k)\mathcal{R}(\tilde{\epsilon}_k^{(m)})) \right)$ and $\mathcal{I}(\epsilon_k^{(m)}) = \mathcal{I}(\tilde{\epsilon}_k^{(m)}) \mathbf{1}_{[0,\infty[} \left(sign(\mathcal{I}(s_k)\mathcal{I}(\tilde{\epsilon}_k^{(m)})) \right)$, for $k = 1, ..., K$

   (c) $\mathbf{x}^{(m+1)} = \mathbf{x}^{(0)} + \mathbf{P}\underline{\epsilon}^{(m)}$
3) repeat **2)** while $m < T_{max}$
4) finally $\mathbf{x} = \mathbb{Q}(\mathbf{x}^{(0)} + \mathbf{P}\underline{\epsilon}^{(m)})$

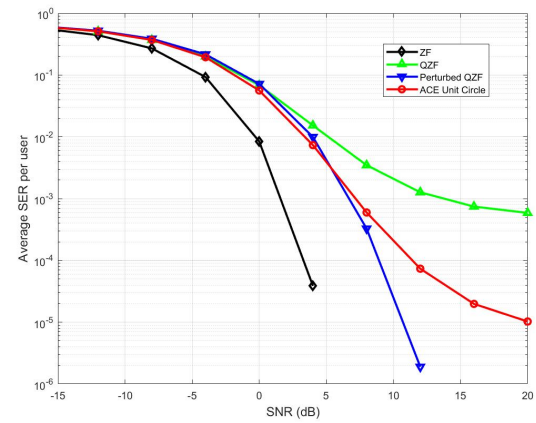

Fig. 3: Average SER vs. SNR, $K = 16$, $M = 128$, average over 50000 realizations of i.i.d channel and noise.

Unexpectedly, simulations based on this approach showed no improvement with respect to one-bit quantized ZF. This led us to investigate the "Anti-ACE" method described next.

B. Anti-ACE

We discovered that when changing the sign of the clipped-off signal $\mathbf{e}$ to its opposite in 2)(a):

$$e_n^{(m)} = -x_n^{(m)} \left(\frac{a}{|x_n^{(m)}|} - 1 \right) \text{ when } |x_n^{(m)}| > a, \tag{6}$$

the "anti-ACE" precoder outperforms the one-bit quantized ZF precoder, as shown in Figure 3. In effect, instead of reducing the amplitude variation in the elements of the vector prior to the one-bit quantization, this amplifies the difference between the large and small elements.

We try to understand this behavior by observing the symbol constellation $s_k + \epsilon_k^{(m)}$ (before precoding and quantization) as the number of iterations increase, as shown in Fig. 4. Firstly, we observe that the amplitude of $\underline{\epsilon}^{(m)}$ increases with the number of iterations, i.e. there is no convergence. Secondly, the phases seem to align in order to create constructive interference. A further refinement of the algorithm is presented next.

C. Anti-ACE Box Constraint

In Fig. 3, we note that in spite of an improved performance for low to intermediate SNR, the "anti-ACE" precoder still shows an error floor for high SNR. This error floor may be due to the fact that in our precoding scheme, we do not want to influence the dynamic range of the modulus of $\mathbf{x}$ as in [11], but rather its real and imaginary parts. Therefore, we propose

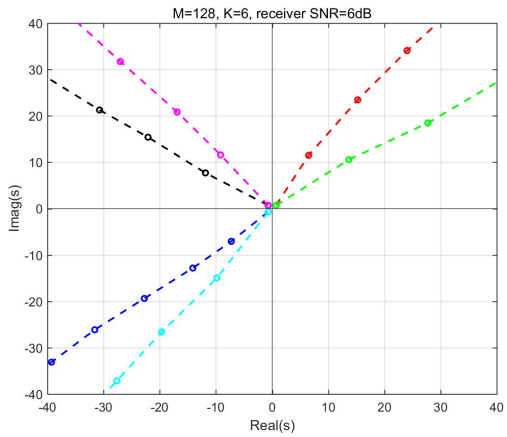

Fig. 4: Constellation $s_k + \epsilon_k^{(m)}$ as iterations increase.

to modify (5) as follows:

$$\underline{\epsilon} = \arg\min_{\underline{\epsilon}} \max_n \max \left(|\mathcal{R}((\mathbf{Ps})_n + \sum_{k=1}^{K} \mathbf{P}_{n,k}\underline{\epsilon}_k)|^2, |\mathcal{I}((\mathbf{Ps})_n + \sum_{k=1}^{K} \mathbf{P}_{n,k}\underline{\epsilon}_k)|^2 \right) \tag{7}$$

This results in changing the projection on a sphere as above, into a projection onto a box. This changes Step 2)(a) of the algorithm as:

$$e_n^{(m)} = \begin{cases} x_n^{(m)} - a\left(\frac{\mathcal{R}(x_n^{(m)})}{\sqrt{2}|\mathcal{R}(x_n^{(m)})|} + j\frac{\mathcal{I}(x_n^{(m)})}{\sqrt{2}|\mathcal{I}(x_n^{(m)})|}\right) \\ \qquad \text{when } |\mathcal{R}(x_n^{(m)})| > a \ , \ |\mathcal{I}(x_n^{(m)})| > a \\ j\mathcal{I}(x_n^{(m)}) - ja\frac{\mathcal{I}(x_n^{(m)})}{\sqrt{2}|\mathcal{I}(x_n^{(m)})|} \\ \qquad \text{when } |\mathcal{R}(x_n^{(m)})| \leqslant a \ , \ |\mathcal{I}(x_n^{(m)})| > a \\ \mathcal{R}(x_n^{(m)}) - a\frac{\mathcal{R}(x_n^{(m)})}{\sqrt{2}|\mathcal{R}(x_n^{(m)})|} \\ \qquad \text{when } |\mathcal{R}(x_n^{(m)})| > a \ , \ |\mathcal{I}(x_n^{(m)})| \leqslant a \\ 0, \ \text{otherwise.} \end{cases} \tag{8}$$

We refer to the resulting algorithm as "anti-ACE box." In Fig. 5, we plot its performance together with that of the previously described algorithms. We observe that the error floor has disappeared, at least for the chosen range of SNR values. Also, more surprisingly, the anti-ACE box algorithm outperforms the greedy vector symbol perturbation method of [10] (Perturbed QZF) that operates directly on the M-dimensional transmit vector. The performance of the anti-ACE box approach is within about 2-3 dB of the unquantized ZF precoder, which is similar to the performance obtained by the more computationally complex algorithms referenced earlier. This performance is achieved with a significantly reduced computational cost.

The simulations of Fig. 5 were obtained using the best choice for the sphere radius and half box side length, a. Fig. 6 shows that the value of a must be carefully chosen for both anti-ACE and anti-ACE box in order to achieve the best possible performance. This is a key facet of the algorithm that requires further study.

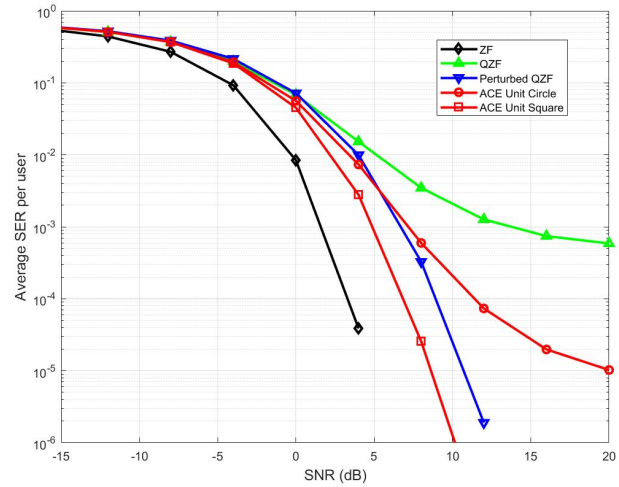

Fig. 5: Average SER vs. SNR, $K = 16$, $M = 128$, average over 50000 realizations of i.i.d channel and noise.

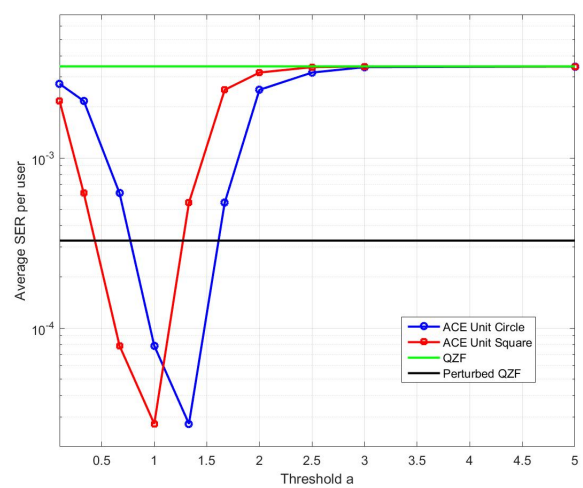

Fig. 6: Average SER vs. projection threshold a, $K = 16$, $M = 128$, for SNR= 8dB, average over 50000 realizations of i.i.d channel and noise.

V. Conclusion

We have proposed a new algorithm for one-bit precoding based on the ACE algorithm for multi-user massive MIMO. The algorithm has reduced computational complexity compared with other direct design methods based on MSE or BER reduction via constructive interference, but nonetheless

achieves similar performance. While based on ACE, the proposed algorithm actually uses an update with a sign change that actually enforces rather than minimizes the amplitude variation across the transmit vector prior to the one-bit quantization. Surprisingly, this effect appears to enhance the constructive interference effect once the one-bit quantization is applied. The method is somewhat sensitive to the type of constraint used, an observation that requires further study for a complete understanding.

ACKNOWLEDGMENT

The research was supported by the National Science Foundation under Grant ECCS-1547155, and by the Institute for Advanced Study at the Technische Universität München, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement No. 291763, and by the European Union under the Marie Curie COFUND Program, and by the French CNRS.

REFERENCES

[1] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communication Magazine*, vol. 52, no. 2, pp. 186-195, Feb 2014.

[2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742-758, Oct. 2014.

[3] D. Ha, K. Lee and J. Kang, "Energy efficiency analysis with circuit power consumption in massive MIMO systems", in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, London, United Kingdom, 2013, pp. 938-942.

[4] M. Ariaudo, I. Fijalkow, J.-L. Gautier, M. Brandon, B. Aziz and B. Milevsky, *Green radio despite "Dirty RF" front-end*, in EURASIP Journal on Wireless Communications and Networking, special issue on Green Radio, 2012:146 doi:10.1186/1687-1499-2012-146, April 2012.

[5] J.A. Nossek, and M.T. Ivrlac, "Capacity and coding for quantized MIMO systems," in *Proc. of the 2006 International Conference on Wireless Communications and Mobile Computing*, Vancouver, Canada, 2006, pp. 1387-1392

[6] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4670-4684, Nov. 2017.

[7] H. Jedda, A. Noll, and J. Nossek, "PSK precoding in multi-user MISO systems," in *21st International ITG Workshop on Smart Antennas (WSA)*, March 2017.

[8] H. Jedda, A. Mezghani, J. Nossek, and A. Swindlehurst, "Massive MIMO downlink 1-bit precoding with linear programming for PSK signaling," in *IEEE Signal Processing Workshop on Advances in Wireless Communications (SPAWC)*, July 2017.

[9] A. K. Saxena, I. Fijalkow and A. Swindlehurst, "Analysis of One-Bit Quantized Precoding for the Multiuser Massive MIMO Downlink," in *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4624-4634, Sept. 2017.

[10] A. Swindlehurst, A. Saxena, A. Mezghani, I. Fijalkow, "Minimum probability-of-error perturbation precoding for the one-bit massive MIMO downlink," in *Proc. IEEE ICASSP*, New Orleans, USA, March 2017.

[11] B.S. Krongold, D.L. Jones, "PAR Reduction in OFDM via Active Constellation Extension," IEEE Transactions on Broadcasting, Vol.49, No.3, September, 2003.

[12] C.B. Peel, B.M. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: Channel inversion and regularization," *IEEE Transactions on Communications*, vol 53, no. 1, pp: 195-202, 2005.

[13] F. Liu, C. Masouros, P. V. Amadori, and H. Sun, "An efficient manifold algorithm for constructive interference based constant envelope precoding," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1542–1546, Oct 2017.

[14] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[15] P. V. Amadori and C. Masouros, "Constant envelope precoding by interference exploitation in phase shift keying-modulated multiuser transmission," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 538–550, Jan 2017.