

# **Constant Envelope Precoding for Power-Efficient Downlink Wireless Communication in Multi-User MIMO Systems Using Large Antenna Arrays**

Saif Khan Mohammed and Erik G Larsson

**Linköping University Post Print**

N.B.: When citing this work, cite the original article.

©2012 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Saif Khan Mohammed and Erik G Larsson, Constant Envelope Precoding for Power-Efficient Downlink Wireless Communication in Multi-User MIMO Systems Using Large Antenna Arrays, 2012, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-78476>

# CONSTANT ENVELOPE PRECODING FOR POWER-EFFICIENT DOWNLINK WIRELESS COMMUNICATION IN MULTI-USER MIMO SYSTEMS USING LARGE ANTENNA ARRAYS

Saif Khan Mohammed and Erik G. Larsson

Communication Systems Division, Electrical Eng. (ISY), Linköping University, Sweden

## ABSTRACT

We consider downlink cellular multi-user communication between a base station (BS) having  $N$  antennas and  $M$  single-antenna users, i.e., an  $N \times M$  Gaussian Broadcast Channel (GBC). Under an average only total transmit power constraint (APC), large antenna arrays at the BS (having tens to a few hundred antennas) have been recently shown to achieve remarkable multi-user interference (MUI) suppression with simple precoding techniques. However, building large arrays in practice, would require cheap/power-efficient Radio-Frequency (RF) electronic components. The type of transmitted signal that facilitates the use of most power-efficient RF components is a *constant envelope* (CE) signal (i.e., the amplitude of the signal transmitted from each antenna is *constant* for every channel use and every channel realization). Under certain mild channel conditions (including i.i.d. fading), we analytically show that, even under the stringent per-antenna CE transmission constraint (compared to APC), MUI suppression can still be achieved with large antenna arrays. Our analysis also reveals that, with a fixed  $M$  and increasing  $N$ , the total transmitted power can be *reduced* while maintaining a constant signal-to-interference-noise-ratio (SINR) level at each user. We also propose a novel low-complexity CE precoding scheme, using which, we confirm our analytical observations for the i.i.d. Rayleigh fading channel, through Monte-Carlo simulations. Simulation of the information sum-rate under the per-antenna CE constraint, shows that, for a fixed  $M$  and a fixed desired sum-rate, the required total transmit power decreases linearly with increasing  $N$ , i.e., an  $O(N)$  array power gain. Also, in terms of the total transmit power required to achieve a fixed desired information sum-rate, despite the stringent per-antenna CE constraint, the proposed CE precoding scheme performs *close* to the GBC sum-capacity (under APC) achieving scheme.

**Index Terms**— GBC, constant envelope, per-antenna.

## 1. INTRODUCTION

We consider a Gaussian Broadcast Channel (GBC), wherein a base station (BS) having  $N$  antennas communicates with  $M$  single-antenna users in the downlink. Large antenna arrays at the BS has been of recent interest, due to their remarkable ability to suppress multi-user interference (MUI) with very simple precoding techniques. Specifically, under an average only total transmit power constraint (APC), for a fixed  $M$ , a simple matched-filter precoder has been shown to achieve total MUI suppression in the limit as  $N \rightarrow \infty$  [1]. Additionally, due to its inherent array power gain

property<sup>1</sup>, large antenna arrays are also being considered as an enabler for reducing power consumption in wireless communications, specially since the operational power consumption at BS is becoming a matter of world-wide concern [3, 4].

Despite the benefits of large antenna arrays at BS, practically building them would require cheap and power-efficient RF components like the power amplifier (PA).<sup>2</sup> With current technology, power-efficient RF components are generally non-linear. The type of transmitted signal that facilitates the use of most power-efficient/non-linear RF components, is a *constant envelope* (CE) signal. In this paper, we therefore consider a GBC, where the signal transmitted from each BS antenna has a *constant* amplitude for every channel-use and every channel realization. Since, the per-antenna CE constraint is much more restrictive than APC, we investigate as to whether MUI suppression and array power gain can still be achieved under the stringent per-antenna CE constraint ?

To the best of our knowledge, there is no reported work which addresses this question. Most reported work on per-antenna communication consider an average-only or a peak-only power constraint (see [5, 6] and references therein). In this paper, firstly, we derive expressions for the MUI at each user under the per-antenna CE constraint, and then propose a low-complexity CE precoding scheme with the objective of minimizing the MUI energy at each user. For a given vector of information symbols to be communicated to the users, the proposed precoding scheme chooses per-antenna CE transmit signals in such a way that the MUI energy at each user is small.

Secondly, under certain mild channel conditions (including i.i.d. fading), using a novel probabilistic approach, we analytically show that, *MUI suppression can be achieved even under the stringent per-antenna CE constraint*. Specifically, for a fixed  $M$  and fixed user information symbol alphabets, an arbitrarily low MUI energy can be guaranteed at each user, by choosing a sufficiently large  $N$ . Our analysis further reveals that, for i.i.d channels, with a fixed  $M$  and increasing  $N$ , the total transmitted power can be *reduced* while maintaining a constant SINR level at each user.

Thirdly, through simulation, we confirm our analytical observations for the i.i.d. Rayleigh fading channel. We numerically compute an achievable ergodic information sum-rate under the per-antenna CE constraint, and show that, for a fixed  $M$  and a fixed desired ergodic sum-rate, *the required total transmit power reduces linearly with increasing  $N$* . We also observe that, to achieve a given desired ergodic information sum-rate, compared to the optimal GBC sum-capacity achieving scheme under APC, the extra total transmit power required by the proposed CE precoding scheme is *small* (less than 1.7 dB for large  $N$ ).

This work was supported by the Swedish Foundation for Strategic Research (SSF) and ELLIT. E. G. Larsson is a Royal Swedish Academy of Sciences (KVA) Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

<sup>1</sup> Under an APC constraint, for a fixed  $M$  and a fixed desired information sum-rate, the required total transmit power decreases with increasing  $N$  [2].

<sup>2</sup>In conventional BS, power-inefficient PA's contribute to roughly 40-50 percent of the total operational power consumption [4].

## 2. SYSTEM MODEL

Let the complex channel gain between the  $i$ -th BS antenna and the  $k$ -th user be denoted by  $h_{k,i}$ . The vector of channel gains from the BS antennas to the  $k$ -th user is denoted by  $\mathbf{h}_k = (h_{k,1}, h_{k,2}, \dots, h_{k,N})^T$ .  $\mathbf{H} \in \mathbb{C}^{M \times N}$  is the channel gain matrix with  $h_{k,i}$  as its  $(k, i)$ -th entry. Let  $x_i$  denote the complex symbol transmitted from the  $i$ -th BS antenna. Further, let  $P_T$  denote the average total power transmitted from all the BS antennas. Under the APC constraint, we must have  $\mathbb{E}[\sum_{i=1}^N |x_i|^2] = P_T$ , whereas under the per-antenna CE constraint we have  $|x_i|^2 = P_T/N$  which is clearly a more stringent constraint compared to APC. Further, due to the per-antenna CE constraint, it is clear that  $x_i$  is of the form  $x_i = \sqrt{P_T/N} e^{j\theta_i}$ , where  $\theta_i$  is the phase of  $x_i$ . Under CE transmission, the symbol received by the users is therefore given by

$$y_k = \sqrt{\frac{P_T}{N}} \sum_{i=1}^N h_{k,i} e^{j\theta_i} + w_k, \quad k = 1, 2, \dots, M \quad (1)$$

where  $w_k \sim \mathcal{CN}(0, \sigma^2)$  is the AWGN noise at the  $k$ -th receiver. For the sake of notation, let  $\Theta = (\theta_1, \dots, \theta_N)^T$  denote the vector of transmitted phase angles. Let  $\mathbf{u} = (\sqrt{E_1}u_1, \dots, \sqrt{E_M}u_M)^T$  be the vector of scaled information symbols, with  $u_k \in \mathcal{U}_k$  denoting the information symbol to be communicated to the  $k$ -th user. Here  $\mathcal{U}_k$  denotes the unit average energy information alphabet of the  $k$ -th user.  $E_k, k = 1, 2, \dots, M$  denote the information symbol energy for each user. Also, let  $\mathcal{U} \triangleq \sqrt{E_1}\mathcal{U}_1 \times \sqrt{E_2}\mathcal{U}_2 \times \dots \times \sqrt{E_M}\mathcal{U}_M$ . Subsequently, in this paper, we would be interested in scenarios where  $M$  is fixed and  $N$  is allowed to increase. Also, throughout this paper, for a fixed  $M$ , the alphabets  $\mathcal{U}_1, \dots, \mathcal{U}_M$  are also fixed and do not change with increasing  $N$ .

## 3. PROPOSED CE PRECODING SCHEME

For any given information symbol vector  $\mathbf{u}$  to be communicated, with  $\Theta$  as the transmitted phase angle vector, using (1) the received signal at the  $k$ -th user can be expressed as

$$y_k = \sqrt{P_T} \sqrt{E_k} u_k + \sqrt{P_T} s_k + w_k$$

$$s_k \triangleq \left( \frac{\sum_{i=1}^N h_{k,i} e^{j\theta_i}}{\sqrt{N}} - \sqrt{E_k} u_k \right) \quad (2)$$

where  $\sqrt{P_T} s_k$  is the MUI term at the  $k$ -th user. For reliable communication to each user, the precoder at the BS, must therefore choose a  $\Theta$  such that  $|s_k|$  is as small as possible for each  $k = 1, 2, \dots, M$ . This motivates us to consider the following non-linear least squares (NLS) problem

$$\Theta^u = (\theta_1^u, \dots, \theta_N^u) = \arg \min_{\theta_i \in [-\pi, \pi], i=1, \dots, N} g(\Theta, \mathbf{u})$$

$$g(\Theta, \mathbf{u}) \triangleq \sum_{k=1}^M \left| \frac{\sum_{i=1}^N h_{k,i} e^{j\theta_i}}{\sqrt{N}} - \sqrt{E_k} u_k \right|^2. \quad (3)$$

This NLS problem is non-convex and has multiple local minima. However, as the ratio  $N/M$  becomes large, due to the large number of extra degrees of freedom ( $N - M$ ), the value of the objective function  $g(\Theta, \mathbf{u})$  at most local minima has been observed to be small, enabling gradient descent based methods to be used. However, due to the slow convergence of gradient descent based methods, we propose a novel iterative method, which has been experimentally observed to achieve similar performance as the gradient descent based methods, but with a significantly faster convergence.

In the proposed iterative method to solve (3), we start with the  $p = 0$ -th iteration, where we initialize all the angles to 0. Each iteration consists of  $N$  sub-iterations. Let  $\Theta^{(p,q)} = (\theta_1^{(p,q)}, \dots, \theta_N^{(p,q)})^T$  denote the phase angle vector after the  $q$ -th sub-iteration ( $q = 1, 2, \dots, N$ ) of the  $p$ -th iteration (subsequently we shall refer to the  $q$ -th sub-iteration of the  $p$ -th iteration as the  $(p, q)$ -th iteration). After the  $(p, q)$ -th iteration, the algorithm moves either to the  $(p, q + 1)$ -th iteration (if  $q < N$ ), or else it moves to the  $(p + 1, 1)$ -th iteration. In general, in the  $(p, q + 1)$ -th iteration, the algorithm attempts to reduce the current value of the objective function i.e.,  $g(\Theta^{(p,q)}, \mathbf{u})$  by only modifying the  $(q + 1)$ -th phase angle while keeping the other phase angles fixed to the values from the previous iteration. Therefore, the new phase angles after the  $(p, q + 1)$ -th iteration, are given by

$$\theta_{q+1}^{(p,q+1)} = \arg \min_{\Theta = (\theta_1^{(p,q)}, \dots, \theta_q^{(p,q)}, \phi, \theta_{q+2}^{(p,q)}, \dots, \theta_N^{(p,q)})^T, \phi \in [-\pi, \pi]} g(\Theta, \mathbf{u})$$

$$= \pi + \arg \left( \sum_{k=1}^M h_{k,q+1}^* \frac{\sum_{i=1, \neq (q+1)}^N h_{k,i} e^{j\theta_i^{(p,q)}}}{\sqrt{N}} - \sqrt{E_k} u_k \right)$$

$$\theta_i^{(p,q+1)} = \theta_i^{(p,q)}, \quad i = 1, 2, \dots, N, \quad i \neq q + 1.$$

The algorithm is terminated after a pre-defined number of iterations.<sup>3</sup> We denote the phase angle vector after the last iteration by  $\hat{\Theta}^u = (\hat{\theta}_1^u, \dots, \hat{\theta}_N^u)^T$ .

With  $\hat{\Theta}^u$  as the transmitted phase angle vector, the received signal-to-noise-and-interference-ratio (SINR) at the  $k$ -th user is given by

$$\gamma_k(\mathbf{H}, E) = \frac{E_k}{\mathbb{E}_{u_1, \dots, u_M} [|\hat{s}_k|^2] + \frac{\sigma^2}{P_T}}$$

$$\hat{s}_k \triangleq \left( \frac{\sum_{i=1}^N h_{k,i} e^{j\hat{\theta}_i^u}}{\sqrt{N}} - \sqrt{E_k} u_k \right) \quad (4)$$

where  $E \triangleq (E_1, E_2, \dots, E_M)^T$  is the vector of information symbol energy. For each user, we would be ideally interested to have a low value of the MUI energy  $\mathbb{E}[|\hat{s}_k|^2]$ , since this would imply a larger SINR.

## 4. MUI ANALYSIS

In this Section, for any general CE precoding scheme (without restricting to the proposed CE precoding algorithm in Section 3), through analysis, we aim to get a better understanding of the MUI energy level at each user. Towards this end, we firstly study the dynamic range of values taken by the noise-free received signal at the users, which is given by the set

$$\mathcal{M}(\mathbf{H}) \triangleq \left\{ \mathbf{v} = (v_1, \dots, v_M) \mid v_k = \frac{\sum_{i=1}^N h_{k,i} e^{j\theta_i}}{\sqrt{N}}, \theta_i \in [-\pi, \pi] \right\} \quad (5)$$

For any vector  $\mathbf{v} \in \mathcal{M}(\mathbf{H})$ , from (5) it follows that there exists a  $\Theta^v = (\theta_1^v, \dots, \theta_N^v)^T$  such that  $v_k = \frac{\sum_{i=1}^N h_{k,i} e^{j\theta_i^v}}{\sqrt{N}}$ . This sum can

<sup>3</sup> Experimentally, we have observed that, for the i.i.d. Rayleigh fading channel, with a sufficiently large  $N/M$  ratio, beyond the  $p = L$ -th iteration (where  $L$  is some constant integer), the incremental reduction in the value of the objective function is minimal. Therefore, we terminate at the  $L$ -th iteration. Since there are totally  $LN$  sub-iterations, from the phase angle update equation above, it follows that the complexity of this algorithm is  $O(MN)$ .

now be expressed as a sum of  $N/M$  terms (without loss of generality let us assume that  $N/M$  is integral only for the argument presented here)

$$v_k = \sum_{q=1}^{N/M} v_k^q, \quad v_k^q \triangleq \frac{\sum_{r=(q-1)M+1}^{qM} h_{k,r} e^{j\theta_r^v}}{\sqrt{N}},$$

$$q = 1, \dots, \frac{N}{M}. \quad (6)$$

From (6) it immediately follows that  $\mathcal{M}(\mathbf{H})$  can be expressed as a direct-sum of  $N/M$  sets, i.e.

$$\begin{aligned} \mathcal{M}(\mathbf{H}) &= \mathcal{M}(\mathbf{H}^1) \oplus \mathcal{M}(\mathbf{H}^2) \oplus \dots \oplus \mathcal{M}(\mathbf{H}^{N/M}) \\ \mathcal{M}(\mathbf{H}^q) &\triangleq \left\{ \mathbf{v} = (v_1, \dots, v_M) \mid \right. \\ &\left. v_k = \frac{\sum_{i=1}^M h_{k,(q-1)M+i} e^{j\theta_i}}{\sqrt{N}}, \theta_i \in [-\pi, \pi) \right\} \end{aligned} \quad (7)$$

where  $\mathcal{M}(\mathbf{H}^q) \subset \mathbb{C}^M$  is the dynamic range of the received noise-free signals when only the  $M$  BS antennas numbered  $(q-1)M+1, (q-1)M+2, \dots, qM$  are used and the remaining  $N-M$  antennas are inactive. If the statistical distribution of the channel gain vector from a BS antenna to all the users is identical for all the BS antennas, then, on an average the sets  $\mathcal{M}(\mathbf{H}^q)$  would have similar topological properties. Since,  $\mathcal{M}(\mathbf{H})$  is a direct-sum of  $N/M$  topologically similar sets, it is expected that for a fixed  $M$ , on an average the region  $\mathcal{M}(\mathbf{H})$  expands/enlarges with increasing  $N$ . Based on this discussion, for i.i.d. channels, we have the following two important remarks in Section 4.1 and 4.2.

#### 4.1. Diminishing MUI with increasing $N$ , for fixed $M$ and $E_k$

**Theorem 1** For a fixed  $M$  and increasing  $N$ , consider a sequence of channel gain matrices  $\{\mathbf{H}_N\}_{N=M+1}^{\infty}$  satisfying the mild conditions

$$\lim_{N \rightarrow \infty} \frac{|\mathbf{h}_k^{(N)H} \mathbf{h}_l^{(N)}|}{N} = 0, \quad \forall k \neq l \quad (cnd.1)$$

$$\lim_{N \rightarrow \infty} \frac{\sum_{i=1}^N |h_{k_1,i}| |h_{l_1,i}| |h_{k_2,i}| |h_{l_2,i}|}{N^2} = 0, \quad \forall k_1, l_1, k_2, l_2 \in (1, 2, \dots, M) \quad (cnd.2)$$

$$\lim_{N \rightarrow \infty} \frac{\|\mathbf{h}_k^{(N)}\|^2}{N} = c_k, \quad k = 1, 2, \dots, M \quad (cnd.3) \quad (8)$$

where  $c_k$  are positive constants and  $\mathbf{h}_k^{(N)}$  denotes the  $k$ -th row of  $\mathbf{H}_N$ . (From the law of large numbers, it follows that i.i.d. channels satisfy these conditions with probability 1.)

For any given fixed finite alphabet  $\mathcal{U}$  (fixed  $E_k, k = 1, \dots, M$ ) and any given  $\Delta > 0$ , there exists a corresponding integer  $N(\{\mathbf{H}_N\}, \mathcal{U}, \Delta)$  such that with  $N \geq N(\{\mathbf{H}_N\}, \mathcal{U}, \Delta)$  and  $\mathbf{H}_N$  as the channel gain matrix, for any  $\mathbf{u} \in \mathcal{U}$  to be communicated, there exist a phase angle vector  $\Theta_N^u(\Delta) = (\theta_1^u(\Delta), \dots, \theta_N^u(\Delta))^T$  which when transmitted, results in the MUI energy at each user being upper bounded by  $2\Delta^2$ , i.e.

$$\left| \frac{\sum_{i=1}^N h_{k,i}^{(N)} e^{j\theta_i^u(\Delta)}}{\sqrt{N}} - \sqrt{E_k} u_k \right|^2 \leq 2\Delta^2, \quad k = 1, \dots, M \quad (9)$$

where  $h_{k,i}^{(N)}$  denotes the  $i$ -th component of  $\mathbf{h}_k^{(N)}$ .

Due to limited space, we present a sketch of the proof of Theorem 1 in Appendix A. In Theorem 1,  $\Delta$  can be chosen to be arbitrarily

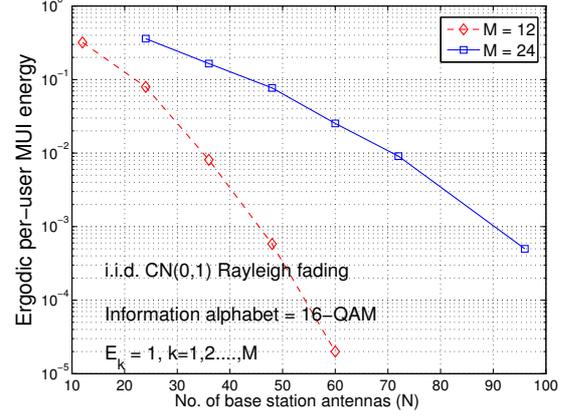


Fig. 1. Reduction in MUI with increasing  $N$ . Fixed  $E_k$ .

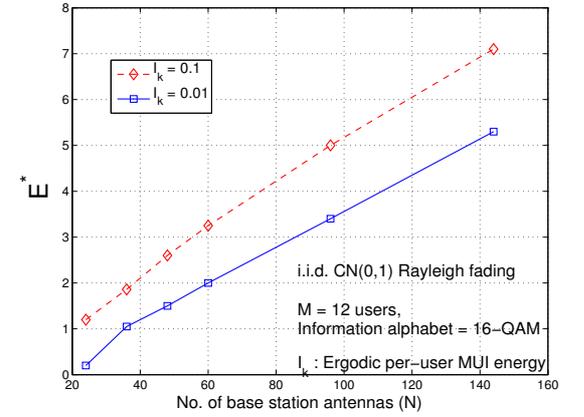


Fig. 2.  $E^*$  vs.  $N$ . Fixed MUI energy (same for each user).

small, and therefore, the MUI energy at each user can be guaranteed to be arbitrarily small, by choosing a sufficiently large  $N$ . In Fig. 1, for the i.i.d.  $\mathcal{CN}(0, 1)$  Rayleigh fading channel, with fixed information alphabets  $\mathcal{U}_1 = \mathcal{U}_2 = \dots = \mathcal{U}_M = 16$ -QAM and fixed information symbol energy  $E_k = 1, k = 1, \dots, M$ , we plot the ergodic (averaged w.r.t. channel statistics) MUI energy  $\mathbb{E}_{\mathbf{H}}[\|\hat{s}_k\|^2]$  (observed to be same for each user) as a function of increasing  $N$  ( $\hat{s}_k$  is given by (4)). It is observed that, for a fixed  $M$ , fixed information alphabets and fixed information symbol energy, the ergodic per-user MUI energy decreases with increasing  $N$ .

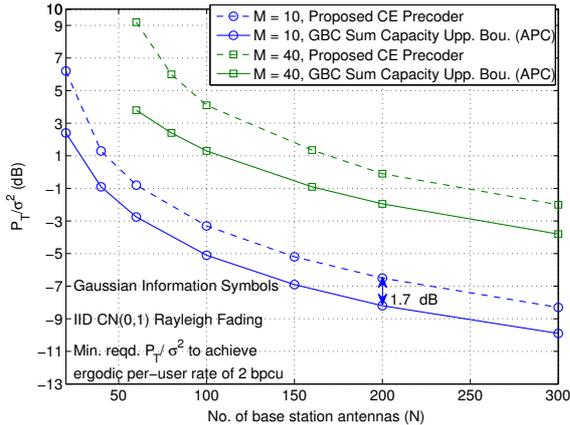
#### 4.2. Increasing $E_k$ with increasing $N$ , for a fixed MUI

From (4), it is clear that, for a fixed  $M$  and  $N$ , increasing  $E_k, k = 1, \dots, M$  would enlarge  $\mathcal{U}$  which could then increase MUI energy level at each user. However, since increase in  $N$  results in reduction of MUI (Theorem 1), it can be argued that, with increasing  $N$  the information symbol energy of each user can be increased while still maintaining a fixed MUI energy level at each user. We illustrate this through the following example. Let the fixed desired ergodic MUI energy level for the  $k$ -th user be denoted by  $I_k, k = 1, 2, \dots, M$ . For the sake of simplicity we consider  $\mathcal{U}_1 = \mathcal{U}_2 = \dots = \mathcal{U}_M$ . Consider the following optimization

$$E^* \triangleq \arg \max_{p > 0} p \quad (10)$$

$$p > 0 \mid E_k = p, I_k = \mathbb{E}_{\mathbf{H}} \left[ \mathbb{E}_{u_1, \dots, u_M} [|\hat{s}_k|^2] \right], k = 1, \dots, M$$

which finds the highest possible equal energy of the information symbols under the constraint that the ergodic MUI energy level is



**Fig. 3.** Req'd.  $P_T/\sigma^2$  vs.  $N$ . Fixed ergodic per-user rate = 2 bpcu.

fixed at  $I_k$ ,  $k = 1, 2, \dots, M$ . In (10),  $\hat{s}_k$  is given by (4). In Fig. 2, for the i.i.d. Rayleigh fading channel, for a fixed  $M = 12$  and a fixed  $\mathcal{U}_1 = \dots = \mathcal{U}_M = 16$ -QAM, we plot  $E^*$  as a function of increasing  $N$ , for two different fixed desired MUI energy levels,  $I_k = 0.1$  and  $I_k = 0.01$  (same  $I_k$  for each user). (Due to same channel distribution and information alphabet for each user, it is observed that the ergodic MUI energy level is also same if the users have equal information symbol energy.) From Fig. 2, it can be observed that for a fixed  $M$  and fixed  $\mathcal{U}_1, \dots, \mathcal{U}_M$ , indeed,  $E^*$  increases linearly with increasing  $N$ , while still maintaining a fixed MUI energy level at each user. At low MUI energy levels, from (4) it follows that  $\gamma_k \approx P_T E_k / \sigma^2$ . Since  $E_k$  ( $k = 1, 2, \dots, M$ ) can be increased linearly with  $N$  (while still maintaining low MUI level), it can be argued that a desired fixed SINR level can be maintained at each user by simply reducing  $P_T$  linearly with increasing  $N$ . This suggests the achievability of an  $O(N)$  array power gain.

## 5. ACHIEVABLE ARRAY POWER GAIN

For the proposed CE precoding scheme in Section 3, with the same Gaussian information alphabet for each user, an achievable ergodic information sum-rate for GBC under the per-antenna CE constraint, can be shown to be given by  $R_{CE}(E) = \sum_{k=1}^M \mathbb{E}_{\mathbf{H}} \log_2(\gamma_k(\mathbf{H}, E))$  (Here we have used the fact that, with Gaussian alphabet, Gaussian noise is the worst noise in terms of achievable mutual information). We numerically optimize  $R_{CE}(E)$  subject to the constraint  $E_1 = \dots = E_M$ . Based on this optimized ergodic sum-rate, in Fig. 3, for the i.i.d.  $CN(0, 1)$  Rayleigh fading channel, we plot the required  $P_T/\sigma^2$  to achieve an ergodic per-user information rate of 2 bits-per-channel-use (bpcu) (We have observed that the ergodic information rate achieved by each user is  $1/M$  of the ergodic sum-rate). It is observed that, for a fixed  $M$ , at sufficiently large  $N$ , the required  $P_T/\sigma^2$  reduces by roughly 3 dB for every doubling in  $N$  (i.e., the required  $P_T/\sigma^2$  reduces linearly with increasing  $N$ ). This shows that, for a fixed  $M$ , an array power gain of  $O(N)$  can indeed be achieved even under the stringent per-antenna CE constraint. For the sake of comparison, we have also plotted the minimum  $P_T/\sigma^2$  required under the APC constraint (we have used the co-operative upper bound on the GBC sum-capacity [7]). We observe that, for large  $N$  and a fixed per-user desired ergodic information rate of 2 bpcu, compared to the APC only constrained GBC, the extra total transmit power required under the more stringent per-antenna CE constraint is *small* (only 1.7 dB).

## 6. REFERENCES

- [1] T. L. Marzetta, "Non-cooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. on Wireless Communications*, vol. 9, pp. 3590–3600, Nov. 2010.
- [2] D. N. C. Tse and P. Vishwanath, "Fundamentals of wireless communications," *Cambridge University Press*, 2005.
- [3] "http://www.eweekurope.co.uk/news/greentouch-shows-low-power-wireless-19719," GreenTouch Consortium.
- [4] V. Mancuso and S. Alouf, "Reducing costs and pollution in cellular networks," *IEEE Communications Magazine*, pp. 63–71, Aug. 2011.
- [5] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per antenna power constraints," *IEEE Trans. on Signal Processing*, vol. 55, pp. 2646–2660, June 2007.
- [6] K. Kemai, R. Yates, G. Foschini, and R. Valenzuela, "Optimum zero-forcing beamforming with per-antenna power constraints," in *IEEE International Symposium on Information Theory (ISIT'07)*, 2007, pp. 101–105.
- [7] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. on Information Theory*, vol. 49, pp. 2658–2668, Oct. 2003.
- [8] P. Billingsley, "Probability and measure," *John Wiley and Sons*.

### A. PROOF OF THEOREM 1 (SKETCH)

Let us consider a probability space with the transmitted phase angles  $\theta_i, i = 1, 2, \dots, N$  being i.i.d. r.v's uniformly distributed in  $[-\pi, \pi)$ . For a given sequence of channel matrices  $\{\mathbf{H}_N\}$ , we define a corresponding sequence of r.v's  $\{\mathbf{z}_N\}$ , with  $\mathbf{z}_N \triangleq (z_1^{I(N)}, z_1^{Q(N)}, \dots, z_M^{I(N)}, z_M^{Q(N)}) \in \mathbb{R}^{2M}$ , where we have  $z_k^{I(N)} \triangleq \text{Re}\left(\frac{\sum_{i=1}^N h_{k,i}^{(N)} e^{j\theta_i}}{\sqrt{N}}\right)$ ,  $z_k^{Q(N)} \triangleq \text{Im}\left(\frac{\sum_{i=1}^N h_{k,i}^{(N)} e^{j\theta_i}}{\sqrt{N}}\right)$ ,  $k = 1, \dots, M$ . Using the Lyapunov Central Limit Theorem (CLT) [8], it can be shown that, for any channel sequence  $\{\mathbf{H}_N\}$  satisfying the conditions in (8), as  $N \rightarrow \infty$ , the corresponding sequence of r.v's  $\{\mathbf{z}_N\}$  converges in distribution to a  $2M$ -dimensional real Gaussian random vector  $X = (X_1^I, X_1^Q, \dots, X_M^I, X_M^Q)^T$  with independent zero-mean components and  $\text{var}(X_k^I) = \text{var}(X_k^Q) = c_k/2$ .

For a given  $\mathbf{u} \in \mathcal{U}$ , and  $\Delta > 0$ , we next consider the box

$$\mathcal{B}_\Delta(\mathbf{u}) \triangleq \left\{ \mathbf{b} = (b^I, b_1^Q, \dots, b_M^I, b_M^Q)^T \in \mathbb{R}^{2M} \mid \begin{aligned} |b_k^I - \sqrt{E_k} u_k^I| &\leq \Delta, |b_k^Q - \sqrt{E_k} u_k^Q| \leq \Delta, k = 1, 2, \dots, M \\ u_k^I &\triangleq \text{Re}(u_k), u_k^Q \triangleq \text{Im}(u_k) \end{aligned} \right\} \quad (11)$$

The box  $\mathcal{B}_\Delta(\mathbf{u})$  contains all those vectors in  $\mathbb{R}^{2M}$  whose component-wise displacement from  $\mathbf{u}$  is upper bounded by  $\Delta$ . Using the fact that  $\mathbf{z}_N$  converges in distribution to a Gaussian r.v. with  $\mathbb{R}^{2M}$  as its range space, continuity arguments show that, for any  $\Delta > 0$ , there exist an integer  $N(\{\mathbf{H}_N\}, \mathcal{U}, \Delta)$ , such that for all  $N \geq N(\{\mathbf{H}_N\}, \mathcal{U}, \Delta)$

$$\text{Prob}(\mathbf{z}_N \in \mathcal{B}_\Delta(\mathbf{u})) > 0, \forall \mathbf{u} \in \mathcal{U}. \quad (12)$$

Since, the probability that  $\mathbf{z}_N$  lies in the box  $\mathcal{B}_\Delta(\mathbf{u})$ , is *strictly positive*, it follows that there exist a phase angle vector  $\Theta_N^u(\Delta) = (\theta_1^u(\Delta), \dots, \theta_N^u(\Delta))^T$  which satisfies (9).