ENERGY-EFFICIENT PILOT AND DATA POWER ALLOCATION IN MASSIVE MIMO COMMUNICATION SYSTEMS BASED ON MMSE CHANNEL ESTIMATION

Ye Zhang and Wei-Ping Zhu

Department of Electrical and Computer Engineering, Concordia University, Montreal, Quebec, Canada

ABSTRACT

This paper addresses the pilot and data power allocation issue in time division duplexing (TDD) massive multi-user multiple-input multiple-output (MU-MIMO) systems. By using minimum mean square error (MMSE) channel estimation along with a maximum-ratio combining (MRC) detector for the uplink transmission and a maximum-ratio transmission (MRT) precoder for the downlink transmission, a novel pilot and data power allocation scheme is proposed to minimize the total uplink and downlink transmit power under per-user signal to interference-plus-noise ratio (SINR) and power consumption constraints. The main contribution of this paper lies in formulating the original energy efficient power allocation problem and converting such a complicated optimization problem to a geometric programming one. Computer simulation shows that the proposed scheme can save up to 78% of the total power as compared to the equal power allocation among all the mobile users.

Index Terms— massive MIMO, energy efficiency, QoS constraints, green communications

1. INTRODUCTION

As green communication has become a significant trend of future wireless communication design and development, it is of great importance to design the energy efficient (EE) power allocation schemes for MU-MIMO systems [1]-[3]. One of the major focuses of EE power control in green communication is to study the power consumption trade-off between pilot and data signals [4]-[10]. In most of the previous works on energy efficient MIMO systems, all users are assumed to have the same pilot power or data power [4]-[8]. Such an equal power allocation strategy may cause "squaring effect" in low power regime [11]. On the other hand, the schemes presented in [4]-[10] considered the EE power control for the uplink and downlink transmissions separately, which has limited the practical use of MIMO systems. On contrary to most previous works, in this paper we consider a more practical scenario, where different users have different transmit pilot and data powers. Also we address the joint power control problem for both the uplink and downlink transmissions in one optimization problem, so as to achieve a minimum sum power under both per-user QoS and per-user power budget constraints.

2. SYSTEM AND DATA MODEL

Consider a time-division duplex (TDD) single-cell MU-MIMO system which consists of an M-antenna base station (BS) serving K ($K<M$) single-antenna mobiles. Let $\boldsymbol{G}$ denote the $M\times K$ channel matrix between the K users and the BS, with its element $g_{mk} \triangleq [\boldsymbol{G}]_{mk}$ representing the channel coefficient between the k-th user and the m-th antenna of the BS. By assuming flat fading channel [14], we have $\boldsymbol{G} = \boldsymbol{H}\boldsymbol{D}^{1/2}$, where the element $h_{mk} \triangleq [\boldsymbol{H}]_{mk}$ represents the small-scale coefficients and is modeled as independent and identically distributed (i.i.d.) complex Gaussian random variables (RVs). The $K\times K$ diagonal matrix $\boldsymbol{D} = diag\{\beta_1, \beta_2, \ldots, \beta_K\}$ models the large-scale fading that incorporates path-loss and shadowing effect which are assumed to be constant and known a priori.

As usual, we assume a block fading structure where the channel gains remain constant in each coherence time period. As discussed in [2][12] and [13], in pilot-assisted channel estimation, when large antenna arrays are employed at BS, it is difficult to estimate the downlink CSI at mobiles, since in this case the number of pilot symbols must be larger than or equal to the number of BS antennas. On the contrary, the uplink CSI is easy to estimate at BS as the number of uplink pilot symbols depends on the number of active mobiles rather than that of BS antennas. Under the assumption of ideal channel reciprocity, however, we can estimate the uplink CSI at BS and then use such estimated uplink CSI for both uplink and downlink data transmission.

Fig. 1 shows the transmission protocol of the pilot symbols and the uplink and downlink data symbols under the TDD operation mode, where the BS acquires downlink CSI through uplink pilot training. In the first τ ($\tau \geq K$) slots of a coherent time interval, all users synchronously transmit uplink pilot signal from mobiles to BS for the purpose of CSI estimation which is required to detect the uplink data and generate precoding matrix for downlink data transmission. After uplink pilot transmission, T_1 symbols are used for uplink data transmission followed by T_2 symbols for downlink data transmission. All transmit data signals are assumed to be stochastic in nature with zero means. It is worth mentioning that the silent slots used for BS processing as discussed in [2][13] are not included in Fig. 1.

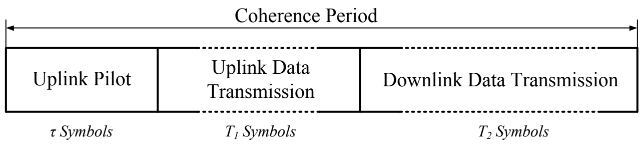

Fig. 1. Frame structure of TDD system

2.1. Channel Estimation

By using the MMSE channel estimation as discussed in [14], the estimated channel matrix can be expressed as

$$\hat{\boldsymbol{G}} = \boldsymbol{Y}_p \boldsymbol{S}_p^H (\boldsymbol{D}^{-1} + \boldsymbol{S}_p \boldsymbol{S}_p^H)^{-1} \tag{1}$$

where $\boldsymbol{S}_P$ denotes the $K \times N_p$ transmit pilot symbol matrix and $\boldsymbol{Y}_P$ is the $M \times N_p$ received pilot signal matrix. In this paper, an orthogonal pilot is used [5], which means $\boldsymbol{S}_P \boldsymbol{S}_P^H = \tau diag(p_{p,1}, p_{p,2}, \dots, p_{p,K})$, where $p_{p,k}$ $(k = 1, 2, \dots, K)$ represents the pilot power of the k-th user.

From the property of MMSE channel estimation [14], both the estimated channel matrix $\hat{\boldsymbol{G}}$ and the estimation error matrix $\Delta \boldsymbol{G} = \boldsymbol{G} - \hat{\boldsymbol{G}}$ have i.i.d. Gaussian RVs with zero mean. Let $M \times 1$ vectors $\hat{\boldsymbol{g}}_k$ and $\Delta \boldsymbol{g}_k$ denote the k-the column of $\hat{\boldsymbol{G}}$ and that of $\Delta \boldsymbol{G}$, respectively. The elements of $\hat{\boldsymbol{G}}$ are independent of that of $\Delta \boldsymbol{G}$ and the variance of the elements of $\hat{\boldsymbol{g}}_k$ and $\Delta \boldsymbol{g}_k$ can be expressed as

$$\sigma_k^2 = \frac{\beta_k^2 \tau p_{p,k}}{1 + \beta_k \tau p_{p,k}}, \;\; \varepsilon_k^2 = \frac{\beta_k}{1 + \beta_k \tau p_{p,k}} \tag{2}$$

2.2. Uplink Data Transmission

By employing the maximum-ratio combining (MRC) detector at the BS, the received uplink SINR of the user k after applying the $M \times K$ receive beamforming matrix $\hat{\boldsymbol{G}}$ can be represented as

$$\gamma_k = \frac{p_{d,k} \left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_k \right|^2}{\sum\limits_{i=1, i \neq k}^{K} p_{d,i} \left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|^2 + \sum\limits_{j=1}^{K} p_{d,j} \left| \hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_i \right|^2 + \left\| \hat{\boldsymbol{g}}_k^H \right\|^2} \tag{3}$$

where $p_{d,k}$ $(k = 1, 2, \dots, K)$ represents the uplink data transmit power for the k-th user. In this paper, the normalized white Gaussian noise with zero mean and unit variance is assumed .

2.3. Downlink Data Transmission

For the downlink transmission, the BS uses a normalized precoding vector $\hat{\boldsymbol{g}}_k / \|\hat{\boldsymbol{g}}_k\|$, and the received downlink SINR of the user k is then given by

$$\tilde{\gamma}_k = \frac{\tilde{p}_{d,k} \|\hat{\boldsymbol{g}}_k\|^2}{\sum\limits_{i=1, i \neq k}^{K} \tilde{p}_{d,i} \frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|^2}{\|\hat{\boldsymbol{g}}_i\|^2} + \sum\limits_{j=1}^{K} \tilde{p}_{d,j} \frac{\left| \Delta \boldsymbol{g}_j^H \hat{\boldsymbol{g}}_j \right|^2}{\|\hat{\boldsymbol{g}}_j\|^2} + 1} \tag{4}$$

where $\tilde{p}_{d,k}$ $(k = 1, 2, \dots, K)$ represents the downlink data power for the k-th user.

3. LOWER BOUNDS OF AVERAGE SINR

3.1. Uplink Transmission

Proposition 1: By using the MRC receiver at BS, a lower bound of the uplink average SINR of user k under the MMSE channel estimation is given by

$$E\{\gamma_k\} \geq \dot{\gamma}_k^{MRC,up} \triangleq \frac{\frac{M \beta_k^2 \tau p_{p,k} p_{d,k}}{1 + \beta_k \tau p_{p,k}}}{\sum\limits_{i=1, i \neq k}^{K} \beta_i p_{d,i} + p_{d,k} \frac{\beta_k}{1 + \beta_k \tau p_{p,k}} + 1} \tag{5}$$

Proof: From (3), the uplink average SINR of user k can be written as

$$E\{\gamma_k\} = E\{ \frac{p_{d,k} \|\hat{\boldsymbol{g}}_k\|^2}{\sum\limits_{i=1, i \neq k}^{K} p_{d,i} \frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2} + \sum\limits_{j=1}^{K} p_{d,j} \frac{\left| \hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_j \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2} + 1} \} \tag{6}$$

where the vectors $\hat{\boldsymbol{g}}_i$ and $\Delta \boldsymbol{g}_i$ are spherically symmetric as their elements consist of i.i.d. zero-mean Gaussian RVs. Further, from the spherically symmetric distribution [15, chapter 4], $\frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|}{\|\hat{\boldsymbol{g}}_k\|}$ and $\frac{\hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_i}{\|\hat{\boldsymbol{g}}_k\|}$ are Gaussian RVs with zero mean and variances σ_i^2 and ε_i^2, respectively. Both $\frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|}{\|\hat{\boldsymbol{g}}_k\|}$ and $\frac{\hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_i}{\|\hat{\boldsymbol{g}}_k\|}$ are independent of $\|\hat{\boldsymbol{g}}_k\|$, which means the numerator in (6) is independent of the three terms in the denominator. Noting that the function $1/x$ is convex when x is positive and the Jensens inequality, (6) can be further expressed as

$$\begin{aligned} E\{\gamma_k\} &= E\{\|\hat{\boldsymbol{g}}_k\|^2\} E\{ \frac{p_{d,k}}{\sum\limits_{i=1, i \neq k}^{K} p_{d,i} \frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2} + \sum\limits_{j=1}^{K} p_{d,j} \frac{\left| \hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_j \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2} + 1} \} \\ &\geq E\{\|\hat{\boldsymbol{g}}_k\|^2\} \frac{p_{d,k}}{\sum\limits_{i=1, i \neq k}^{K} p_{d,i} E\{\frac{\left| \hat{\boldsymbol{g}}_k^H \hat{\boldsymbol{g}}_i \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2}\} + \sum\limits_{j=1}^{K} p_{d,j} E\{\frac{\left| \hat{\boldsymbol{g}}_k^H \Delta \boldsymbol{g}_j \right|^2}{\|\hat{\boldsymbol{g}}_k\|^2}\} + 1} \\ &= E\{\|\hat{\boldsymbol{g}}_k\|^2\} \frac{p_{d,k}}{\sum\limits_{i=1, i \neq k}^{K} p_{d,i} \sigma_i^2 + \sum\limits_{j=1}^{K} p_{d,j} \varepsilon_j^2 + 1} \end{aligned} \tag{7}$$

The term $\|\hat{\boldsymbol{g}}_k\|^2$ in (7) can be treated as a 1×1 central complex Wishart matrix with M degrees of freedom. According to the property of central Wishart matrix [16], we can then get

$$\begin{aligned} E\{\|\hat{\boldsymbol{g}}_k\|^2\} &= \sigma_k^2 E\{ \left(\frac{\hat{\boldsymbol{g}}_k}{\sigma_k} \right)^H \frac{\hat{\boldsymbol{g}}_k}{\sigma_k} \} \\ &= \sigma_k^2 E\{ tr \left[\left(\frac{\hat{\boldsymbol{g}}_k}{\sigma_k} \right)^H \frac{\hat{\boldsymbol{g}}_k}{\sigma_k} \right] \} = \frac{M \beta_k^2 \tau p_{p,k}}{1 + \beta_k \tau p_{p,k}} \end{aligned} \tag{8}$$

Substituting (8) into (7), we obtain the result as in (5).

3.2. Downlink Transmission

Proposition 2: Assuming that an MRT precoder is employed at BS, a lower bound of the downlink average SINR of user *k* can be expressed as

$$
\begin{aligned}
&E\{\tilde{\gamma}_k\} \geq \dot{\gamma}_k^{MRC,dn} \\
&\triangleq \frac{\frac{(M-1)\beta_k^2 \tau p_{p,k}}{1+\beta_k \tau p_{p,k}} \tilde{p}_{d,k}}{\frac{(M-1)\beta_k^2 \tau p_{p,k}}{M(1+\beta_k \tau^{dn} p_{p,k})} \sum\limits_{i=1,i\neq k}^{K} \tilde{p}_{d,i} + \sum\limits_{j=1}^{K} \tilde{p}_{d,j} \frac{\beta_j}{1+\beta_j \tau p_{p,j}} + 1}
\end{aligned}
\tag{9}
$$

The proof is similar to the uplink transmission case and is omitted here due to the space limit.

4. ENERGY-EFFICIENT PILOT AND DATA POWER ALLOCATION

In this section, we develop an algorithm for power allocation between pilot and data symbols to minimize the total uplink and downlink transmit power while guaranteeing per-user QoS and power consumption constraints.

Let P be the total power for one transmission frame. In order to find the best power-consumption trade-off between the uplink and downlink transmission, a weighted sum-power minimization is considered with a positive weight parameter ξ. By denoting $\boldsymbol{p}_p \triangleq [p_{p,1}, p_{p,2}, \ldots, p_{p,K}]$, $\boldsymbol{p}_d \triangleq [p_{d,1}, p_{d,2}, \ldots, p_{d,K}]$ and $\tilde{\boldsymbol{p}}_d \triangleq [\tilde{p}_{d,1}, \tilde{p}_{d,2}, \ldots, \tilde{p}_{d,K}]$, the power allocation problem which minimizes the total transmit power based on the obtained average SINR lower bounds for MRC receiver and MRT precoder can be formulated as

$$
\begin{aligned}
\min_{\boldsymbol{p}_p, \boldsymbol{p}_d, \tilde{\boldsymbol{p}}_d} \; & P = \sum_{k=1}^{K} (\tau p_{p,k} + T_1 p_{d,k} + \zeta T_2 \tilde{p}_{d,k}) \\
\text{s.t.} \quad & \frac{\frac{M\beta_k^2 \tau p_{p,k} p_{d,k}}{1+\beta_k \tau p_{p,k}}}{\sum\limits_{i=1,i\neq k}^{K} \beta_i p_{d,i} + p_{d,k} \frac{\beta_k}{1+\beta_k \tau p_{p,k}} + 1} \geq \gamma_1 \\
& \frac{\frac{(M-1)\beta_k^2 \tau p_{p,k}}{1+\beta_k \tau p_{p,k}} \tilde{p}_{d,k}}{\frac{(M-1)\beta_k^2 \tau p_{p,k}}{M(1+\beta_k \tau p_{p,k})} \sum\limits_{i=1,i\neq k}^{K} \tilde{p}_{d,i} + \sum\limits_{j=1}^{K} \tilde{p}_{d,j} \frac{\beta_j}{1+\beta_j \tau p_{p,j}} + 1} \geq \gamma_2 \\
& \tau p_{p,k} + T_1 p_{d,k} \leq P_1 \\
& T_2 \sum_{k=1}^{K} \tilde{p}_{d,k} \leq P_2 \\
& p_{p,k} \geq 0, p_{d,k} \geq 0, \tilde{p}_{d,k} \geq 0
\end{aligned}
\tag{10}
$$

Here, the cost function is the weighted sum-power accounting for the pilot power consumption and the uplink and downlink data transmission. The first and second constraints represent the uplink and downlink QoS requirement with per-user SINR targets γ_1 and γ_2, respectively. The third and fourth constraints P_1 and P_2 are the power constraint of each mobile user and that of BS, respectively.

It is easy to see that the optimization problem (10) is non-convex since the first and second constraint functions are non-convex. As such, it is very difficult to solve it directly. In order to simplify the optimization problem (10), we introduce a new set of variables

$$
a_k \triangleq \frac{\beta_k}{1 + \beta_k \tau p_{p,k}} \tag{11}
$$

along with the constraints $0 < \alpha_k \leq \beta_k$. By substituting (11) into (10) and after some operations, the second constraint in (10) can be expressed as

$$
\begin{aligned}
&\tfrac{M-1}{M} (p_{d,k})^{-1} \sum_{i=1,i\neq k}^{K} \tilde{p}_{d,i} + \tfrac{1}{\beta_k - a_k} (p_{d,k})^{-1} \sum_{j=1}^{K} \tilde{p}_{d,j} a_j \\
&\quad + \tfrac{1}{\beta_k - a_k} (p_{d,k})^{-1} \leq \tfrac{(M-1)}{\gamma_2}
\end{aligned}
\tag{12}
$$

By replacing the term $\frac{1}{\beta_k - a_k}(p_{d,k})^{-1} \sum_{j=1}^{K} \tilde{p}_{d,j} a_j + \frac{1}{\beta_k - a_k}(p_{d,k})^{-1}$ with a new variable b_k, we have constraint $\frac{1}{\beta_k - a_k}(p_{d,k})^{-1} \sum_{j=1}^{K} \tilde{p}_{d,j} a_j + \frac{1}{\beta_k - a_k}(p_{d,k})^{-1} \leq b_k$. Then, the minimization problem in (10) can be rewritten as

$$
\begin{aligned}
\min_{\boldsymbol{a}, \boldsymbol{p}_d, \boldsymbol{p}_d^{dn}} \; & P = \sum_{k=1}^{K} (\tfrac{1}{a_k} + T_1 p_{d,k} + \zeta T_2 \tilde{p}_{d,k}) \\
\text{s.t.} \; & \sum_{i=1,i\neq k}^{K} \beta_i p_{d,i} (p_{d,k})^{-1} + (p_{d,k})^{-1} + (\tfrac{M}{\gamma_1} + 1) a_k \leq \tfrac{M\beta_k}{\gamma_1} \\
& \tfrac{M-1}{M} (p_{d,k})^{-1} \sum_{i=1,i\neq k}^{K} \tilde{p}_{d,i} + b_k \leq \tfrac{(M-1)}{\gamma_2} \\
& (b_k p_{d,k})^{-1} \sum_{j=1}^{K} \tilde{p}_{d,j} a_j + (b_k p_{d,k})^{-1} + a_k \leq \beta_k \\
& a_k^{-1} + p_{d,k} T_1 \leq P_1 + \tfrac{1}{\beta_k} \\
& T_2 \sum_{k=1}^{K} \tilde{p}_{d,k} \leq P_2 \\
& 0 < a_k \leq \beta_k, p_{d,k} \geq 0, \tilde{p}_{d,k} \geq 0
\end{aligned}
\tag{13}
$$

Now, the optimization problem in (13) is a standard geometric programming (GP) problem, since its cost function and constraints are all posynomials [17][18]. Then, this GP problem can be easily solved by using some standard numerical optimization packages, for example ConVeX (CVX) [19]. It is worth mentioning that the above GP problem can be converted to a convex optimization problem through a logarithmic change of the variables and a logarithmic transformation of the cost and constraint functions as discussed in [17][18].

5. SIMULATION RESULTS

In this section, numerical simulations are carried out by following the parameter setting in [5]. We consider a single cell MU-MIMO system with $K = 4$ users randomly located within a circular area with a radius of 1000m. We choose the smallest amount of training $\tau = K$. The symbols for uplink and downlink data transmission are assumed to be the same in one coherent time interval as $T_1 = T_2 = 96$.The weight parameter ζ is assumed to be one. The pilot and data powers are normalized according to white Gaussian noise power.

In Fig. 2, we first give the numerical results for the original SINR and the derived lower bounds for comparison. Here, we have assumed that equal pilot and data power allocation among all users is applied with $p_{p,k} = p_{d,k} = \tilde{p}_{d,k}$ for any $k \in K$ as in [5]. It is clearly seen that the derived lower bounds are tight in all cases despite the number of BS antennas.

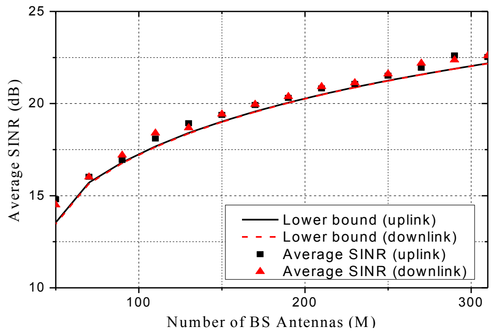

Fig. 2. Average SINR versus the number of BS antennas

Fig. 3 shows the uplink and downlink power for all the users versus the number of BS antennas with $\gamma_1 = \gamma_2 = 5dB$ and $\gamma_1 = \gamma_2 = 15dB$, respectively. The uplink power includes the power of both pilot and uplink data signal. All the powers are normalized according to the noise power. It is obvious that as M grows, both uplink and downlink powers decrease, showing that the use of massive MU-MIMO can save a great deal of transmit power. Note that when the required SINR is chosen as $15dB$, there is no solution for $M \leq 70$ because of the significant crosstalk interference.

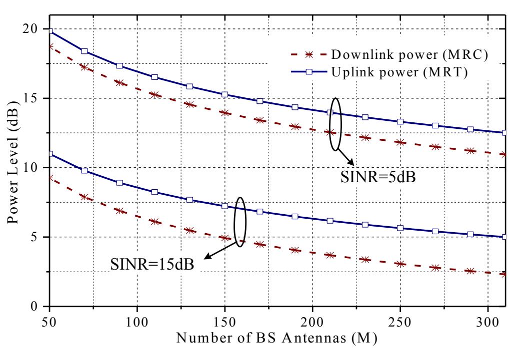

Fig. 3. Pilot-data power allocation versus number of BS antennas

In order to demonstrate the advantage of our proposed power allocation algorithm as compared with a simple equal pilot-data power allocation where the pilot and data signal have the same power p_u for all the users as in [5], we define

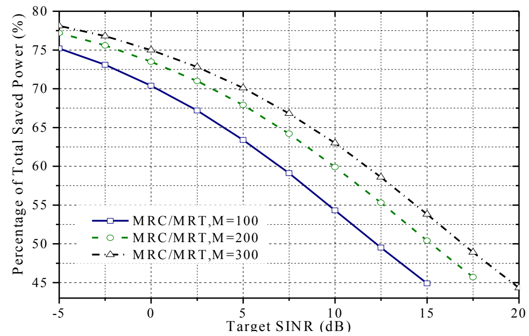

Fig. 4. Percentage of power saving versus target SINRs

the percentage of the total power saving as

$$\frac{K(\tau + T_1 + \zeta T_2)p_u - \sum_{k=1}^{K} (\tau p_{p,k} + T_1 p_{d,k} + \zeta T_2 \tilde{p}_{d,k})}{K(\tau + T_1 + \zeta T_2)p_u} \tag{14}$$

where P_u can be easily found by setting $p_{p,k} = p_{d,k} = \tilde{p}_{d,k} = p_u$ in the previous optimization problem (14). From Fig. 4, it can be seen that about 75% to 78% total power has been saved by our method in low target SINR region, depending on the number of BS antennas. The percentage of power saving decreases as the required per-user SINR increases. It should be mentioned that the benefit of deploying a large number of BS antennas tends to become marginal, since the ultimate SINR performance is limited by the interference and channel estimation error.

6. CONCLUSION

In this paper, we have proposed a pilot-data power allocation for EE communications in single-cell MU-MIMO systems with an objective of minimizing the total uplink and downlink transmit power. We have first analysed the uplink and downlink SINRs and then derived their lower bounds, based on which an EE power allocation optimization problem is formulated under the per-user SINR and power consumption constraints. The non-convex optimization problem is then converted to a standard GP problem to facilitate its solution. Numerical simulation results have confirmed the advantage of the proposed power allocation scheme.

7. ACKNOWLEDGEMENT

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Regroupement Stratgique en Microlectronique du Qubec (ReSMiQ).

8. REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of BS antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590-3600, Nov. 2010.

[2] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin and R. Zhang, An overview of massive MIMO: benefits and challenges IEEE J. Sel. Topics Signal Process., vol. 8, no. 5, pp. 742-758, Oct. 2014.

[3] G. Miao, N. Himayat, G. Y. Li, and A. Swami, Cross-layer optimization for energy-efficient wireless communications: a survey, Wireless Commun. and Mobile Comput. vol. 9, no. 4. pp. 529542, 2009.

[4] M. Biguesh and A. B. Gershman, Training-based MIMO channel estimation: a study of estimator tradeoffs and optimal training signals, IEEE Trans. Signal Process., vol. 54, no. 3, pp. 884893, Mar. 2006.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, Energy and spectral efficiency of very large multiuser MIMO systems, IEEE Trans. Communun., vol. 61, no. 4, pp. 14361449, Apr. 2013.

[6] J. Wang, O. Wen and S. Li, Near-optimum pilot and data symbols power allocation for MIMO spatial multiplexing system with Zero-Forcing receiver, IEEE Sign. Proces. Lett., vol. 16, May 2009.

[7] L. Chen, G. Wei, Energy-efficient power allocation for training-based multiple-input multiple-output system with and without feedback, IET Commun., vol. 7, iss. 15, pp. 16971707, 2013.

[8] X. Zhou, T. A. Lamahewa, P. Sadeghi, and S. Durrani, Two-way training: optimal power allocation for pilot and data transmission, IEEE Trans. Wireless Commun., vol. 9, no. 2, pp. 564 - 569, Feb. 2010.

[9] Y. Kim, G. Miao and T. Hwang, Energy Efficient Pilot and Link Adaptation for Mobile Users in TDD Multi-User MIMO Systems, IEEE Trans. Wireless Commun., vol. 13, no. 1, pp. 382-393, Jan. 2014.

[10] K. Guo and G. Ascheid, Performance analysis of multi-cell MMSE based receivers in MU-MIMO systems with very large antenna arrays, in Proc. IEEE Wireless Communications and Networking Conference (WCNC), pp. 31753179, Apr. 2013.

[11] 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.

[12] E. Bjrnson, J. Hoydis, M. Kountouris, and M. Debbah, Massive MIMO systems with non-ideal hardware: energy efficiency, estimation, and capacity limits, IEEE Trans. Inf. Theory, vol. 60, no. 11, pp. 1182, Nov. 2014.

[13] J. Shen, J. Zhang and K.d B. Letaief, Downlink user capacity of massive MIMO under pilot contamination, IEEE Trans. Wireless Commun., vol. 14, no. 6, pp. 31833193, Jun. 2015.

[14] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[15] W. Bryc, Rotation invariant distributions, Lecture Notes in Statist., vol. 100, pp. 51-69, 1995.

[16] A. M. Tulino and S. Verd, Random matrix theory and wireless communications, Foundations Trends Commun. Inf. Theory, vol. 1, no. 1, pp. 1182, June 2004.

[17] M. Chiang, Geometric programming for communication systems, Foundations Trends Commun. Inf. Theory, vol. 2: no. 12, pp 1-154, Jul. 2005.

[18] S. Boyd, S. J. Kim, L. Vandenberghe, and A. Hassibi, A tutorial on geometric programming, Optimization and Engineering, vol. 8, no. 1, pp. 67127, 2007.

[19] M. Grant, S. Boyd, and Y. Ye, CVX: Matlab Software for Disciplined Convex Programming, ver. 1.1, Nov. 2007. Online Available: www.stanford.edu./ boyd/cvx