Power Allocation for Multi-Point Joint Transmission with Different Node Activeness

Jingya Li, Behrooz Makki, Tommy Svensson, and Thomas Eriksson
Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden
{jingya.li, behrooz.makki, tommy.svensson, thomase}@chalmers.se

***Abstract*—We study the power allocation problem for the downlink of a cooperative system with different node activeness, i.e., each receiving node requests for data transmission according to a certain probability. Data symbols of the active nodes are jointly transmitted from cooperative transmission points using zero-forcing precoding. The problem is cast in form of maximizing the ergodic achievable rate subject to per-transmit-point average total power constraints. The derived power allocation indicates that depending on the channel conditions and receiving nodes' activation probability, the optimal solution takes the form of either greedy power allocation or power sharing allocation.**

I. Introduction

Transmit power allocation is an effective way for increasing the sum rate of wireless communication systems. It has been proved in [1] that water-filling power allocation is optimal in the sense of maximizing the total throughput under a sum power constraint, when different data signals are transmitted through orthogonal channels without interference. Considering an interference system that allows different users to share the same channel, the optimal solution is instead achieved by assigning the total power to the user with the best channel gain, namely greedy power allocation [2], [3]. By assuming per-cell power constraints, it has been proved in [4] that the maximum sum rate for a two-cell two-user system is achieved by binary power control, i.e., each cell either transmits with full power or does not transmit.

Coordinated multi-point (CoMP) transmission has been considered as a promising technique to increase the sum rate and cell-edge performance. If both data symbols and channel state information (CSI) are shared by coordinated transmission points (TPs), multiple TPs can jointly provide data transmission to the receiving nodes (RNs) and thereby improve the received signal quality. Major setbacks of CoMP joint transmission are, however, the large CSI feedback overhead, the capacity and latency constraints of backhaul links, and the imperfect synchronization between TPs [5]. A tradeoff between the system performance and the required amount of CSI feedback and backhaul exchange has been pointed out [6], [7]. This tradeoff is one of the reasons for restricting the use of joint transmission to a limited number of TPs [6]–[10].

Part of this work has been performed in the framework of the FP7 project ICT-317669 METIS, which is partly funded by the European Union. The authors would like to acknowledge the contributions of their colleagues in METIS, although the views expressed are those of the authors and do not necessarily represent the project. This work is also supported by the Swedish Research Council VR under the project 621-2009-4555 Dynamic Multipoint Wireless Transmission.

Different power allocation schemes are proposed in, e.g., [6]–[10], with the objective of maximizing the instantaneous sum rate under a peak transmit power constraint per TP. However, to the best of our knowledge, all CoMP-based papers have studied the power allocation problem under the assumption that the users are always active requesting for new information, a point which is not valid in practice; recent studies, e.g., [11], [12], show that a large portion of the assigned spectrum is used sporadically and geographical variations in the utilization of assigned spectrum ranges from 15% to 85% with a high variance in time. Thus, it is interesting to investigate the power allocation problem in the cases where the users request new information according to some probability.

In this paper, the power allocation problem is addressed for the downlink of a two-TP two-RN cooperative system, where each RN randomly requests for data transmission according to a certain probability. Data symbols of the RNs are jointly transmitted from the TPs using zero-forcing precoding. The objective is to maximize the ergodic achievable rate subject to per-TP average total power constraints, which is normally considered for systems with limited energy resources [12]–[14]. The optimal power allocation solution is proved to fall into two categories: 1) Greedy power allocation, when two RNs are located close to one TP and the per-TP total power budget is below a threshold; 2) Power sharing solution, i.e., both RNs receive the data from the TPs, when different RNs are close to different TPs or the per-TP total power budget is above a threshold. Moreover, we show that depending on the channel condition and the RNs' activation probability, a system should switch between cooperative joint transmission and non-cooperative transmission to improve the sum rate.

II. System model

We consider the downlink of a two-TP two-RN cooperative system, where each RN has a different activation probability. The TPs and RNs are assumed to have one transmit antenna and one receive antenna respectively, while the results can be easily generalized to multiple antenna cases. Here, we focus on the nonfading channel model which, although it is a special case of the more general fading channels, is reasonable in many practical cooperative scenarios, e.g., in line-of-sight (LoS) multiple-input multiple-output (MIMO) systems applied to microwave backhaul. Also, the same problem setup is valid for the LoS wireless backhaul between macro base station and fixed relay nodes or fixed femto base stations, where the channel vary much more slowly than the change of RNs' activation status. Assume that the RNs' activation status, data symbols, and the CSI of the two RNs are perfectly known at each TP. The TPs can provide the active RNs with joint transmission at the same time using the same spectral resource.

Let $\mathbf{x} = [x_1, x_2]^T$ denote the signal vector transmitted from the two TPs, where $(\cdot)^T$ is the transpose operation. The received signal at RN $m \in \{1,2\}$ is found as

$$y_m = \mathbf{h}_m \mathbf{x} + n_m, \tag{1}$$

where $\mathbf{h}_m = [h_{m1}, h_{m2}]$ is the channel vector between RN m and both TPs, and n_m is the independent and identically distributed complex white Gaussian noise added at the m-th RN, with zero mean and covariance σ^2. By using linear precoding, the transmit signal vector $\mathbf{x}$ can be expressed as

$$\mathbf{x} = \mathbf{W}\mathbf{b}, \tag{2}$$

where $\mathbf{b} = [b_1, b_2]^T$ represents the data symbol vector of the two RNs. Also, $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2]$ is the precoding matrix used to map the data symbol vector $\mathbf{b}$ into the transmit signal vector $\mathbf{x}$, with $\mathbf{w}_m = [w_{1m}, w_{2m}]^T$ denoting the precoding vector for RN m from the two TPs. Substituting (2) into (1), the received signal of RN m can be rewritten as

$$y_m = \mathbf{h}_m \mathbf{w}_m b_m + \textstyle\sum_{i \neq m} \mathbf{h}_m \mathbf{w}_i b_i + n_m, \; m, i = 1, 2. \tag{3}$$

Let $p_m = b_m b_m^H$ be the symbol power allocated to the m-th RN across the two TPs, where $(\cdot)^H$ is the conjugate transpose operation. The signal to interference plus noise ratio (SINR) of the RN m is then given by

$$\gamma_m = \frac{\|\mathbf{h}_m \mathbf{w}_m\|^2 p_m}{\sum_{i \neq m} \|\mathbf{h}_m \mathbf{w}_i\|^2 p_i + \sigma^2}, \; m, i = 1, 2. \tag{4}$$

According to (2), the transmit power of TP n is

$$P_n = \|w_{n1}\|^2 p_1 + \|w_{n2}\|^2 p_2, \; n = 1, 2. \tag{5}$$

Using zero-forcing precoding among the coordinated TPs, the precoding matrix is obtained as

$$\mathbf{W} = \mathbf{H}^H (\mathbf{H}\mathbf{H}^H)^{-1}, \tag{6}$$

where $\mathbf{H} = [\mathbf{h}_1^T, \mathbf{h}_2^T]^T$ represents the channel matrix of the system, and $(\cdot)^{-1}$ is the matrix inversion operation. Then, (4) becomes $\gamma_m = p_m/\sigma^2$, for $m \in \{1,2\}$.

Depending on the activity of the two RNs, joint transmission can be categorized into the following three statuses: 1) Both RN 1 and RN 2 request new information; 2) Only RN 1 requests new information; 3) Only RN 2 requests new information. Let $p_m^{(k)}$ be the symbol power allocated to RN m in status k, with $m \in \{1,2\}$ and $k \in \{1,2,3\}$. Since RN 1 and RN 2 are inactive in status 3 and status 2 respectively, we have $p_1^{(3)} = p_2^{(2)} = 0$. Thus, the ergodic achievable rate of the system normalized by the bandwidth can be expressed as

$$\begin{aligned} R = {} & \alpha_1\alpha_2 \left(\log_2(1 + p_1^{(1)}/\sigma^2) + \log_2(1 + p_2^{(1)}/\sigma^2)\right) \\ & + \alpha_1(1-\alpha_2)\log_2(1 + p_1^{(2)}/\sigma^2) \\ & + \alpha_2(1-\alpha_1)\log_2(1 + p_2^{(3)}/\sigma^2), \end{aligned} \tag{7}$$

where α_1 and α_2 are activation probability for RN 1 and RN 2 respectively. Assume that the TPs have the same average total power constraint P_{tot}, which is related to an energy consumption constraint. The coordinated TPs can jointly adapt the transmit power based on the RN activeness, in order to maximize the ergodic sum rate R subject to per-TP average total power constraints. Let the vector $\mathbf{p} = [p_1^{(1)}, p_2^{(1)}, p_1^{(2)}, p_2^{(3)}] \succeq \mathbf{0}$ be the symbol power allocated to each RN in different statuses, where $\mathbf{p} \succeq \mathbf{0}$ represents componentwise inequality between $\mathbf{p}$ and $\mathbf{0}$. Then, the corresponding average total transmit power of TP 1 and TP 2, respectively denoted by $\bar{P}_{\text{TP1}}(\mathbf{p})$ and $\bar{P}_{\text{TP2}}(\mathbf{p})$, are

$$\begin{aligned} \bar{P}_{\text{TP1}}(\mathbf{p}) = {} & \alpha_1\alpha_2 \left(\Sigma_{m=1}^2 \|w_{1m}\|^2 p_m^{(1)}\right) \\ & + \alpha_1(1-\alpha_2)\|w_{11}\|^2 p_1^{(2)} \\ & + \alpha_2(1-\alpha_1)\|w_{12}\|^2 p_2^{(3)}, \end{aligned} \tag{8}$$

$$\begin{aligned} \bar{P}_{\text{TP2}}(\mathbf{p}) = {} & \alpha_1\alpha_2 \left(\Sigma_{m=1}^2 \|w_{2m}\|^2 p_m^{(1)}\right) \\ & + \alpha_1(1-\alpha_2)\|w_{21}\|^2 p_1^{(2)} \\ & + \alpha_2(1-\alpha_1)\|w_{22}\|^2 p_2^{(3)}. \end{aligned} \tag{9}$$

Hence, the optimization problem can be formulated as

$$\begin{aligned} & \max_{\mathbf{p} \succeq \mathbf{0}} R(\mathbf{p}) \\ & \text{s.t.} \; \begin{cases} \bar{P}_{\text{TP1}}(\mathbf{p}) \leq P_{\text{tot}}, \\ \bar{P}_{\text{TP2}}(\mathbf{p}) \leq P_{\text{tot}}. \end{cases} \end{aligned} \tag{10}$$

This problem is convex, since the objective function is a concave function of $\mathbf{p}$ and the constraints are affine functions. Thus, the optimal solution with respect to a given α_1, α_2 and P_{tot} can be obtained by numerical convex optimization [15].

III. Optimal Power Allocation

In this section, we solve the Karush-Kuhn-Tucker (KKT) conditions for the convex problem (10), in order to have better understanding of the optimal power allocation solution.

The Lagrangian function of (10) can be written as

$$L(\mathbf{p}, \boldsymbol{\lambda}) = -R(\mathbf{p}) + \lambda_1(\bar{P}_{\text{TP1}}(\mathbf{p}) - P_{\text{tot}}) + \lambda_2(\bar{P}_{\text{TP2}}(\mathbf{p}) - P_{\text{tot}}), \tag{11}$$

where $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are Lagrangian multipliers associated with the two power constraints. Substituting (7)-(9) into (11) and setting $\nabla_{\mathbf{p}} L(\mathbf{p}, \boldsymbol{\lambda}) = 0$, we have

$$p_1 = p_1^{(1)} = p_1^{(2)} = \left[\frac{1}{\ln 2(\lambda_1 \|w_{11}\|^2 + \lambda_2 \|w_{21}\|^2)} - \sigma^2\right]^+, \tag{12}$$

$$p_2 = p_2^{(1)} = p_2^{(3)} = \left[\frac{1}{\ln 2(\lambda_1 \|w_{12}\|^2 + \lambda_2 \|w_{22}\|^2)} - \sigma^2\right]^+, \tag{13}$$

where $[a]^+ \doteq \max(a, 0)$. From (12) and (13), we observe that the symbol power allocated to RN m will be the same for all statuses if it is active. This result is intuitive because from the RNs' perspective each RN does not receive any interference regardless of whether the other RN is active or not, since the interference is completely mitigated by performing zero-forcing joint precoding at the TPs. However, it needs to be pointed out that from the TP's perspective, the transmit power to the RN m from TP 1 and TP 2, i.e., $\|w_{1m}\|^2 p_m$ and $\|w_{2m}\|^2 p_m$, are different.

Remark 1: Equations (12) and (13) have an intuitive consequence; Depending on the channel condition ($\mathbf{H}$) and the value of the average total power budget for each TP (P_{tot}), the optimal power allocation that maximizes the ergodic sum rate of the system falls into one of the following three cases:

Case 1. Two TPs jointly transmit data to RN 1 when it is active, however, no power will be allocated to RN 2, even when RN 2 requests new information;

Case 2. Two TPs jointly transmit data to RN 2 when it is active, however, no power will be allocated to RN 1, even when RN 1 requests new information;

Case 3. Two TPs provide joint transmission to the active RNs. In other words, Remark 1 indicates that depending on the channel condition there are cases where greedy power allocation, which is practically used in interference limited systems [2], provides the optimal solution (Case 1 and Case 2). On the other hand, Case 3 is the power sharing solution, where both RNs receive the data from TP 1 and TP 2. In the following, we show the channel condition and the corresponding threshold of P_{tot}, under which different cases are the optimal solution.

Lemma 1. Let $\mathbf{p}^*$ be the optimal transmit power allocation vector of (10). Then, there is at least one TP with index $n \in \{1, 2\}$, such that the average total power constraint for this TP is satisfied with equality.

Proof: The main idea of the proof is that, for any interior point of the feasible set of (10), it is always possible to find a factor $\theta > 1$, such that at least one of the power constraints is satisfied with equality and $R(\theta\mathbf{p}) > R(\mathbf{p})$. The proof is presented in Section A of the Appendix. ∎

Theorem 1. Consider the power allocation problem of (10) and define $G_{mn} = \|h_{mn}\|^2$. The optimal solution falls into one of the following cases:

Case 1. $\left\{p_1^* = \frac{P_{\text{tot}}}{\alpha_1 \max\{\|w_{11}\|^2, \|w_{21}\|^2\}} > 0, p_2^* = 0\right\}$, if

- $G_{22} < G_{21} < G_{11}$ and $P_{\text{tot}} < \sigma^2\alpha_1(\|w_{22}\|^2 - \|w_{21}\|^2)$, or
- $G_{21} < G_{22} < G_{12}$ and $P_{\text{tot}} < \sigma^2\alpha_1(\|w_{12}\|^2 - \|w_{11}\|^2)$, or
- $G_{21} = G_{22} < \max\{G_{12}, G_{11}\}$ and $P_{\text{tot}} < \sigma^2\alpha_1(\max\{\|w_{12}\|^2, \|w_{22}\|^2\} - \|w_{11}\|^2)$.

Case 2. $\left\{p_1^* = 0, p_2^* = \frac{P_{\text{tot}}}{\alpha_2 \max\{\|w_{12}\|^2, \|w_{22}\|^2\}} > 0\right\}$, if

- $G_{11} < G_{12} < G_{22}$ and $P_{\text{tot}} < \sigma^2\alpha_2(\|w_{11}\|^2 - \|w_{12}\|^2)$, or
- $G_{12} < G_{11} < G_{21}$ and $P_{\text{tot}} < \sigma^2\alpha_2(\|w_{21}\|^2 - \|w_{22}\|^2)$, or
- $G_{11} = G_{12} < \max\{G_{22}, G_{21}\}$ and $P_{\text{tot}} < \sigma^2\alpha_2(\max\{\|w_{11}\|^2, \|w_{21}\|^2\} - \|w_{12}\|^2)$.

Case 3. $\{p_1^* > 0, p_2^* > 0\}$, otherwise.

Proof: The proof is based on the complementary slackness condition associated with the two power constraints in (10), that is, for each $n = 1, 2$, either $\lambda_n = 0$ or the power constraint associated with TP n is satisfied with equality. The details of the proof are provided in Appendix B. ∎

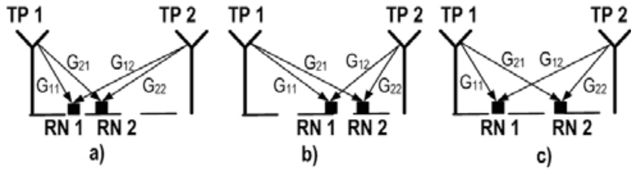

Figure 1. The downlink joint transmission of a two-TP two-RN cooperative system with different RN activities. Scenario a) both RN 1 and RN 2 are close to TP 1; b) both RN 1 and RN 2 are close to TP 2, 3) RN 1 is close to TP 1 and RN 2 is close to TP 2.

Example 1: In Fig. 1, we consider three different pairs of RN locations, corresponding to the following three scenarios:

a) two RNs are closer to TP 1;
b) two RNs are closer to TP 2;
c) RN 1 is closer to TP 1 and RN 2 is closer to TP 2.

For scenario a), we have $G_{22} < G_{21} < G_{11}$. Hence, according to Theorem 1, there is a threshold

$$P_{\text{thr}} = \sigma^2\alpha_1(\|w_{22}\|^2 - \|w_{21}\|^2) \tag{14}$$

that if $P_{\text{tot}} < P_{\text{thr}}$, the optimal power allocation falls into Case 1, i.e., only RN 1 will be served. Moreover, the threshold grows linearly with the activation probability. In scenario c), none of the conditions for Case 1 and Case 2 of Theorem 1 is satisfied, thus, joint data transmission will be provided to both RNs when they are active.

IV. Numerical Results

A. On Power Allocation Thresholds

Fig. 2 verifies the effect of the channel condition ($\mathbf{H}$), the average total power budget for each TP (P_{tot}), and the RNs' activation probability (α_1, α_2) on the optimal power allocation solution. The symbol power allocated for each RN p_m is plotted with different P_{tot} for $(\alpha_1, \alpha_2) = (0.2, 0.7), (0.5, 0.5), (0.8, 0.3)$ respectively[1]. Corresponding to the scenarios a), b) and c) in Example 1, three pairs of RN locations are considered, with channel matrices chosen as:

$$H_{\text{a}} = \begin{bmatrix} 0.61 + j0.73 & 0.21 + j0.25 \\ 0.46 + j0.45 & 0.37 + j0.43 \end{bmatrix},$$

$$H_{\text{b}} = \begin{bmatrix} 0.28 + j0.47 & 0.61 + j0.36 \\ 0.29 + j0.17 & 0.76 + j0.64 \end{bmatrix},$$

$$H_{\text{c}} = \begin{bmatrix} 0.61 + j0.73 & 0.21 + j0.25 \\ 0.29 + j0.17 & 0.76 + j0.64 \end{bmatrix}.$$

The noise power σ^2 is set to 1. We see that for scenario a) and b), there is a threshold under which only one RN is served, while for scenario c) all the RNs are always served by the TPs when they are active, in agreement with Theorem 1. Moreover, the threshold for scenario a) and scenario b) increases as the value of α_1 and α_2 increases, respectively.

B. Comparison with Non-Cooperative Scheme

Consider an interference limited system setup, with $P_{\text{tot}} = 1$ and $\sigma^2 = 0.01$. Due to symmetry, only scenarios a) and c)

[1]The transmit power to RN m from TP 1 and TP 2 can then be obtained by $\|w_{1m}\|^2 p_m$ and $\|w_{2m}\|^2 p_m$, respectively.

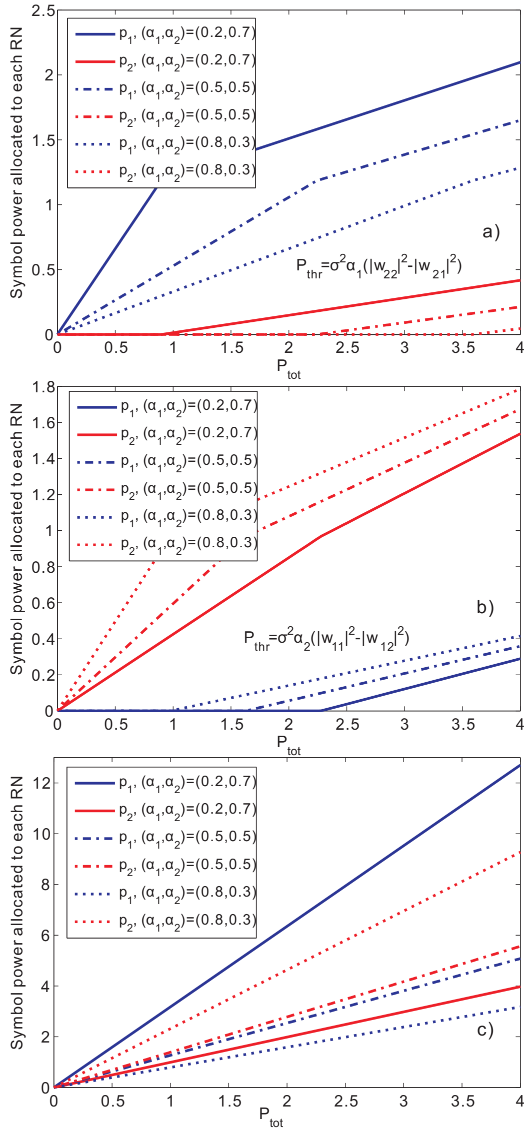

Figure 2. Symbol power allocated to each RN (p_1, p_2) vs. per-TP average total power constraint (P_{tot}) for scenarios: a) H_a, b) H_b and c) H_c.

are studied. The ergodic sum rate of the cooperative power allocation scheme in (10) (R_{Co}) is compared with the sum rate achieved by a non-cooperative scheme ($R_{\text{Non-Co}}$), where TP n always transmits to RN n with P_{tot}/α_n when RN n requests new data, $n = 1, 2$. In this case, $R_{\text{Non-Co}}$ is given as

$$\begin{aligned} R_{\text{Non-Co}} &= \alpha_1\alpha_2 \log_2(1 + \frac{G_{11}P_{\text{tot}}/\alpha_1}{\sigma^2 + G_{12}P_{\text{tot}}/\alpha_2}) \\ &+ \alpha_1\alpha_2 \log_2(1 + \frac{G_{22}P_{\text{tot}}/\alpha_2}{\sigma^2 + G_{21}P_{\text{tot}}/\alpha_1}) \\ &+ \alpha_1(1-\alpha_2)\log_2(1 + G_{11}P_{\text{tot}}/\alpha_1/\sigma^2) \\ &+ \alpha_2(1-\alpha_1)\log_2(1 + G_{22}P_{\text{tot}}/\alpha_2/\sigma^2). \end{aligned} \tag{15}$$

Fig. 3 plots the ergodic sum rate gain, $\frac{(R_{\text{Co}} - R_{\text{Non-Co}})}{R_{\text{Non-Co}}}$, versus different RNs' activation probabilities (α_1, α_2). For both scenarios a) and c), the sum rate gain increases as the RNs' activation probability increases. This is intuitive because if the RNs' activation probability is low, the probability that both RNs request data transmission will be very low. Then, inter-RN interference will not be the main problem for data transmission, i.e., the contribution of interference mitigation via zero-forcing joint transmission is limited.

We can also see that, for scenario c) where RN 1 is closer to TP 1 and RN 2 is closer to TP 2, cooperative joint transmission always outperforms non-cooperative transmission. The sum rate gain is about 31.1% and 92.7% when $(\alpha_1, \alpha_2) = (0.5, 0.5)$ and $(1, 1)$ respectively. However, for scenario a), cooperative joint transmission only outperforms the non-cooperative scheme when the RNs' activation probability is high. That is because, when RNs are closer to one TP, the transmit power of this TP will be limited by using zero-forcing precoding. While the other TP, which is far away from the RNs, needs to consume all its transmit power in order to contribute for joint transmission. Hence, in scenario a), inter-RN interference mitigation is performed at the cost of limiting the transmit power of the closer TP. Thus, cooperative zero-forcing joint transmission only provides sum rate gain when the RNs' activation probability is high, i.e., the probability that both RNs request for data transmission is high. Therefore, the system should switch to non-cooperative transmission when the RNs are located closer to one TP and the RNs' activation probability is low.

Note that the above comparison is obtained under the same average total power budget constraint (P_{tot}), without considering the peak transmit power constraint posed by the hardware limitations. Assume that both TPs are equipped with the same hardware, i.e., subject to the same peak power constraint. For any given pair of activation probabilities (α_1, α_2), let $P_{\text{Co}}^{\max}$ and $P_{\text{Non-Co}}^{\max}$ be the associated maximum instantaneous transmit power for the considered cooperative and non-cooperative transmission schemes, respectively. Then, we have

$$P_{\text{Co}}^{\max} = \max\left\{\Sigma_{m=1}^{2} \|w_{nm}\|^2 p_m^{(k)}\right\}, n \in \{1,2\}, k \in \{1,2,3\},$$

$$P_{\text{Non-Co}}^{\max} = \max\{P_{\text{tot}}/\alpha_n\}, n \in \{1,2\}.$$

In Fig. 4, the maximum instantaneous transmit power ratio, $P_{\text{Co}}^{\max}/P_{\text{Non-Co}}^{\max}$, is plotted as a function of the RNs' activation probabilities (α_1, α_2) for scenario a), with $P_{\text{tot}} = 1$. It can be seen that $P_{\text{Co}}^{\max}/P_{\text{Non-Co}}^{\max} \leq 1$ for all the values of (α_1, α_2). Therefore, under the same average total power budget and the same RNs' activation probabilities, the non-cooperative transmission scheme poses a higher peak power requirement on the hardware design of TPs.

V. Conclusions

We address the downlink power allocation problem for a cooperative joint transmission system with two transmission points and two receiving nodes. Each receiving node randomly requests for data transmission according to a different probability. Theoretical and numerical results indicate that there are cases where greedy power allocation provides the optimal power allocation (when the receiving nodes are closer to one transmission point and the power budget of the transmission point is below a threshold). In other cases, power sharing

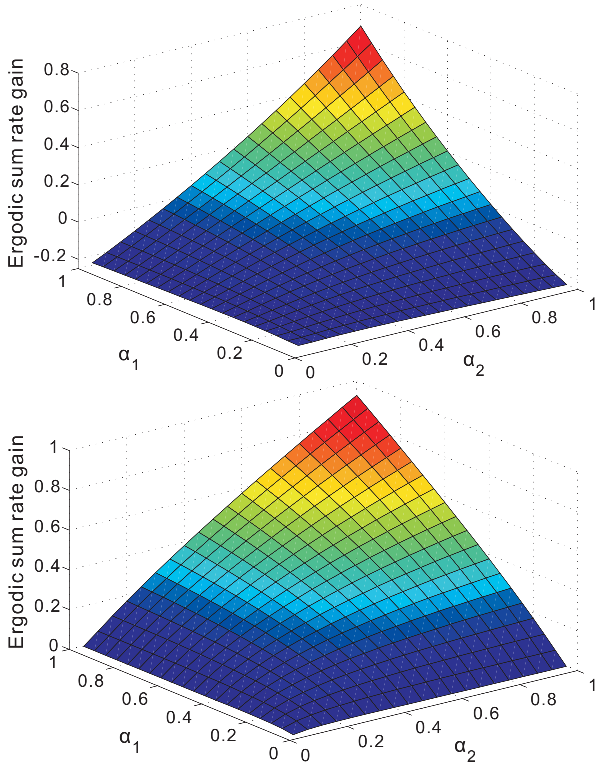

Figure 3. Ergodic sum rate gain ($(R_{\text{Co}} - R_{\text{Non-Co}})/R_{\text{Non-Co}}$) vs. RNs' activation probability (α_1, α_2) for: scenario a), H_{a} (upper plot) and scenario c), H_{c} (lower plot).

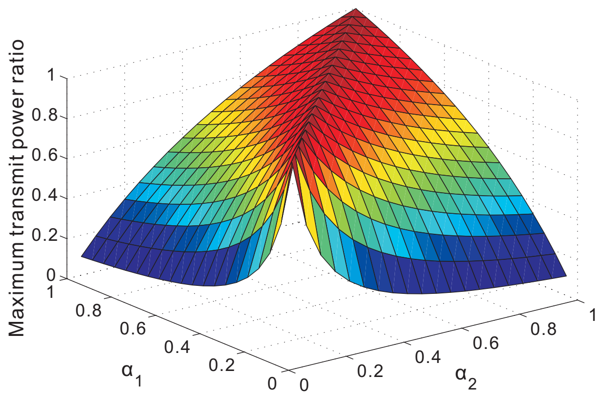

Figure 4. Maximum instantaneous transmit power ratio ($P^{\max}_{\text{Co}}/P^{\max}_{\text{Non-Co}}$) vs. RNs' activation probability (α_1, α_2) for scenario a, H_{a}.

is optimal. Comparing with the non-cooperative transmission scheme, we show that cooperative zero-forcing joint transmission can provide large sum rate gain when different receiving nodes are close to different transmission points. Moreover, depending on the channel condition and the nodes' activation probability, a system should switch between cooperative joint transmission and non-cooperative transmission to improve the ergodic sum rate.

The cooperative transmission scheme in this paper are designed with zero-forcing precoding, subject to per-transmit-point average total power constraints. In future work, different cooperative transmission schemes will be considered, and peak transmit power constraints will also be taken into account.

VI. Appendix

A. Proof of Lemma 1

Here, we prove that the optimal transmit power allocation vector of (10) will result in at least one per-TP average total power constraint satisfied with equality.

Proof: Let $\mathcal{F}$ be the feasible set of (10). Consider an interior point $\mathbf{p} = [p_1^{(1)}, p_2^{(1)}, p_1^{(2)}, p_2^{(3)}] \in \mathcal{F}$, with $\bar{P}_{\text{TP1}}(\mathbf{p}) < P_{\text{tot}}$ and $\bar{P}_{\text{TP2}}(\mathbf{p}) < P_{\text{tot}}$. Note that $\bar{P}_{\text{TP1}}(\mathbf{p})$ and $\bar{P}_{\text{TP2}}(\mathbf{p})$ are affine functions of $\mathbf{p}$. Hence, it is possible to find a factor θ, with

$$\theta = \frac{P_{\text{tot}}}{\max\{\bar{P}_{\text{TP1}}(\mathbf{p}), \bar{P}_{\text{TP2}}(\mathbf{p})\}} > 1 \tag{16}$$

such that at least one of the per-cell power constraints is satisfied with equality, i.e., $\max\{\bar{P}_{\text{TP1}}(\theta\mathbf{p}), \bar{P}_{\text{TP2}}(\theta\mathbf{p})\} = P_{\text{tot}}$. According to (7), $R(\mathbf{p})$ is an increasing function of $\mathbf{p}$, hence,

$$R(\theta\mathbf{p}) > R(\mathbf{p}). \tag{17}$$

Thus, for any feasible power allocation vector $\mathbf{p}$ satisfying $\bar{P}_{\text{TP1}}(\mathbf{p}) < P_{\text{tot}}$ and $\bar{P}_{\text{TP2}}(\mathbf{p}) < P_{\text{tot}}$, it is always possible to find a power allocation vector $\theta\mathbf{p}$ that achieves a larger sum rate R, when one of the per-TP power constraints is satisfied with equality. ■

B. Proof of Theorem 1

In this appendix, we prove that depending on the channel conditions H and the per-TP average total power budget P_{tot} the optimal power allocation solution of (10) falls into on of the three cases listed in Theorem 1.

Let $\mathbf{p}^* = [p_1^{(1)*}, p_2^{(1)*}, p_1^{(2)*}, p_2^{(3)*}]$ be the optimal power allocation vector, and $\boldsymbol{\lambda}^*$ be the dual optimal vector associated with the two power constraints. Note that the primal optimization problem (10) is a convex problem. Hence, according to (12) and (13), we have $p_1^* = p_1^{(1)*} = p_1^{(2)*}$ and $p_2^* = p_2^{(1)*} = p_2^{(3)*}$. Based on the complementary slackness KKT condition, we have $\lambda_1^*(\bar{P}_{\text{TP1}}(\mathbf{p}^*) - P_{\text{tot}}) = 0$ and $\lambda_2^*(\bar{P}_{\text{TP2}}(\mathbf{p}^*) - P_{\text{tot}}) = 0$, where $\bar{P}_{\text{TP1}}(\mathbf{p}^*)$ and $\bar{P}_{\text{TP2}}(\mathbf{p}^*)$ are derived according to (8) and (9), respectively. According to Lemma 1, at least one TP power constraint is satisfied with equality, that is, $\bar{P}_{\text{TP1}}(\mathbf{p}^*) = P_{\text{tot}}$ or/and $\bar{P}_{\text{TP2}}(\mathbf{p}^*) = P_{\text{tot}}$. The proof of the Theorem 1 can be summarized over the following propositions:

Proposition 1. If only the power constraint of TP 1 is satisfied with equality, then, the optimal solution falls into Case 1, iff $\|w_{21}\|^2 < \|w_{11}\|^2 < \|w_{12}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{12}\|^2 - \|w_{11}\|^2)$; the optimal solution falls into Case 2, iff $\|w_{22}\|^2 < \|w_{12}\|^2 < \|w_{11}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_2(\|w_{11}\|^2 - \|w_{12}\|^2)$; otherwise, the optimal solution falls into Case 3.

Proof: If only the power constraint of TP 1 is satisfied with equality, $\bar{P}_{\text{TP1}}(\mathbf{p}^*) = P_{\text{tot}}$ and $\bar{P}_{\text{TP2}}(\mathbf{p}^*) < P_{\text{tot}}$. Based on the complementary slackness KKT condition, $\lambda_2^* = 0$. Plugging this into (12) and (13), the optimal solution is therefore $p_1^* = \left[\frac{1}{\ln 2\lambda_1^*\|w_{11}\|^2} - \sigma^2\right]^+$, $p_2^* = \left[\frac{1}{\ln 2\lambda_1^*\|w_{12}\|^2} - \sigma^2\right]^+$. If the optimal solution falls into Case 1, then,

$$p_1^* = \frac{1}{\ln 2\lambda_1^*\|w_{11}\|^2} - \sigma^2 > 0, \tag{18}$$

$$p_2^* = 0, \tag{19}$$

$$\frac{1}{\ln 2\sigma^2 \|w_{12}\|^2} \le \lambda_1^* < \frac{1}{\ln 2\sigma^2 \|w_{11}\|^2}. \tag{20}$$

Plugging (18) and (20) into $\bar{P}_{\text{TP1}}(\mathbf{p}^*) = P_{\text{tot}}$, it is found that

$$\lambda_1^* = \frac{\alpha_1}{\ln 2(P_{\text{tot}} + \sigma^2\alpha_1 \|w_{11}\|^2)}, \tag{21}$$

$$p_1^* = \frac{P_{\text{tot}}}{\alpha_1 \|w_{11}\|^2}. \tag{22}$$

Based on (20) and (21), we have $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{12}\|^2 - \|w_{11}\|^2)$ and $\|w_{11}\|^2 < \|w_{12}\|^2$. Plugging (20) and (22) into $\bar{P}_{\text{TP2}}(\mathbf{p}^*) < P_{\text{tot}}$, it is obtained that $\|w_{21}\|^2 < \|w_{11}\|^2$. Thus, if only the power constraint of TP 1 is satisfied with equality, then the optimal solution falls into Case 1, iff $\|w_{21}\|^2 < \|w_{11}\|^2 < \|w_{12}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{12}\|^2 - \|w_{11}\|^2)$. In this case, $p_1^* = P_{\text{tot}}/(\alpha_1 \|w_{11}\|^2) > 0$ and $p_2^* = 0$.

Similarly, it can be proved that, if only the power constraint of TP 1 is satisfied with equality, then, the optimal solution falls into Case 2, iff $\|w_{22}\|^2 < \|w_{12}\|^2 < \|w_{11}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_2(\|w_{11}\|^2 - \|w_{12}\|^2)$. In this case, $p_2^* = P_{\text{tot}}/(\alpha_2 \|w_{12}\|^2) > 0$ and $p_1^* = 0$. ■

Similar to the proof for Proposition 1, we have Proposition 2 given below.

Proposition 2. If only the power constraint of TP 2 is satisfied with equality, then, the optimal solution falls into Case 1 with $\{p_1^* = P_{\text{tot}}/(\alpha_1 \|w_{21}\|^2) > 0$ and $p_2^* = 0\}$, iff $\|w_{11}\|^2 < \|w_{21}\|^2 < \|w_{22}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{22}\|^2 - \|w_{21}\|^2)$; the optimal solution falls into Case 2 with $\{p_2^* = P_{\text{tot}}/(\alpha_2 \|w_{22}\|^2) > 0$ and $p_1^* = 0\}$, iff $\|w_{12}\|^2 < \|w_{22}\|^2 < \|w_{21}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_2(\|w_{21}\|^2 - \|w_{22}\|^2)$; otherwise, the optimal solution falls into Case 3.

Proposition 3. If both power constraints are satisfied with equality, then the optimal solution falls into Case 1, iff $\|w_{11}\|^2 = \|w_{21}\|^2$ and $P_{\text{tot}} < \sigma^2\alpha_1(\max\{\|w_{12}\|^2, \|w_{22}\|^2\} - \|w_{11}\|^2)$; the optimal solution falls into Case 2, iff $\|w_{22}\|^2 = \|w_{12}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_2(\max\{\|w_{11}\|^2, \|w_{21}\|^2\} - \|w_{22}\|^2)$; otherwise, the optimal solution falls into Case 3.

Proof: If both power constraints are satisfied with equality, then $\bar{P}_{\text{TP1}}(\mathbf{p}^*) = P_{\text{tot}}$ and $\bar{P}_{\text{TP2}}(\mathbf{p}^*) = P_{\text{tot}}$. Thus,

$$p_1^* = \frac{(\|w_{22}\|^2 - \|w_{12}\|^2)P_{\text{tot}}}{\alpha_1(\|w_{11}\|^2 \|w_{22}\|^2 - \|w_{12}\|^2 \|w_{21}\|^2)}, \tag{23}$$

$$p_2^* = \frac{(\|w_{11}\|^2 - \|w_{21}\|^2)P_{\text{tot}}}{\alpha_2(\|w_{11}\|^2 \|w_{22}\|^2 - \|w_{12}\|^2 \|w_{21}\|^2)}. \tag{24}$$

If the optimal solution falls into Case 1, then $p_2^* = 0$, $p_1^* = \frac{P_{\text{tot}}}{\alpha_1\|w_{11}\|^2}$, $\|w_{11}\|^2 = \|w_{21}\|^2$. Plugging this into (12) and (13), it is obtained that

$$\lambda_1^*(\|w_{12}\|^2 - \|w_{22}\|^2) = \frac{1}{\ln 2\sigma^2} - \frac{\alpha_1 \|w_{22}\|^2}{\ln 2(P_{\text{tot}} + \sigma^2\alpha_1 \|w_{11}\|^2)}, \tag{25}$$

$$\lambda_2^*(\|w_{22}\|^2 - \|w_{12}\|^2) = \frac{1}{\ln 2\sigma^2} - \frac{\alpha_1 \|w_{12}\|^2}{\ln 2(P_{\text{tot}} + \sigma^2\alpha_1 \|w_{11}\|^2)}. \tag{26}$$

Note that $\|w_{11}\|^2 = \|w_{21}\|^2$. Therefore, if $\|w_{12}\|^2 > \|w_{22}\|^2$, the two power constraints of the original problem (10) are equivalent to $\bar{P}_{\text{TP1}}(\mathbf{p}) = P_{\text{tot}}$, $\bar{P}_{\text{TP2}}(\mathbf{p}) \le P_{\text{tot}}$. According to KKT conditions, the dual constraint $\lambda_2^* \ge 0$ must be satisfied. From (26), we obtain $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{12}\|^2 - \|w_{11}\|^2)$. Similarly, if $\|w_{12}\|^2 < \|w_{22}\|^2$, we have $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{22}\|^2 - \|w_{11}\|^2)$. Else if $\|w_{12}\|^2 = \|w_{22}\|^2$, the two power constraints of (10) are reduced to one power constraint, where the optimal solution is obtained via water-filling. In this case, $p_2^* = 0$ and $p_1^* = \frac{P_{\text{tot}}}{\alpha_1\|w_{11}\|^2}$ iff $P_{\text{tot}} \le \sigma^2\alpha_1(\|w_{22}\|^2 - \|w_{11}\|^2)$. Hence, it is concluded that if both power constraints are satisfied with equality, then, the optimal solution falls into Case 1 with $p_1^* = \frac{P_{\text{tot}}}{\alpha_1\|w_{11}\|^2}$ and $p_2^* = 0$, iff $\|w_{11}\|^2 = \|w_{21}\|^2$ and $P_{\text{tot}} \le \sigma^2\alpha_1(\max\{\|w_{12}\|^2, \|w_{22}\|^2\} - \|w_{11}\|^2)$.

Similarly, it can be proved that, when both power constraints are satisfied with equality, the optimal solution falls into Case 2 with $p_2^* = P_{\text{tot}}/(\alpha_2 \|w_{12}\|^2)$ and $p_1^* = 0$, iff $\|w_{22}\|^2 = \|w_{12}\|^2$ and $P_{tot} \le \sigma^2\alpha_2(\max\{\|w_{11}\|^2, \|w_{21}\|^2\} - \|w_{22}\|^2)$. Otherwise, $p_1^* > 0$ and $p_2^* > 0$, i.e., the optimal solution falls into Case 3. ■

Based on the above Propositions, and the fact that $\mathbf{W}$ is obtained as the pseudo-inverse of the channel matrix $\mathbf{H}$, we conclude that the optimal solution falls into three cases listed in Theorem 1.

REFERENCES

[1] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, no. 1, pp. 10-21, Jan. 1949.

[2] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems", *IEEE J. Select. Areas Commun.*, vol. 21, pp. 171-178, 2003.

[3] David N.C. Tse, "Optimal power adaptation over parallel Gaussian Broadcast Channels", *in Proc. IEEE ISIT'97*, vol. 21, pp. 27, 1997.

[4] A. Gjendemsjø, et al., "Optimal power allocation and scheduling for two-cell capacity maximization," *in Proc. WiOpt'06*, 2006.

[5] D. Gesbert, S. Hanly, H. Huang, et al., "Multi-cell MIMO cooperative networks: a new look at interference," *IEEE J. Select. Areas Commun.*, vol. 28, pp. 1380-1408, 2010.

[6] F. Yuan and C. Yang, "Bit allocation between per-cell codebook and phase ambiguity quantization for limited feedback coordinated multi-point transmission systems," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2546-2559, 2012.

[7] A. Papadogiannis, H. J. Bang, D. Gesbert and E. Hardouin, "Efficient selective feedback design for multicell cooperative networks," *IEEE Trans.Vehicular Techn.*, vol. 60, no. 1, pp. 196-205, Jan. 2011.

[8] B. Luo, Q. Cui, H. Wang, et al., "Optimal joint water-filling for coordinated transmission over frequency-selective fading channels," *IEEE Commun. Letters*, vol. 15, no. 2, pp. 190-192, Feb. 2011.

[9] J. Li, T. Svensson, C. Botella, T. Eriksson, X. Xu, and X. Chen, "Joint scheduling and power control in coordinated multi-point clusters," *in Proc. IEEE VTC'11*, Sept. 2011.

[10] J. Li, T. Eriksson, T. Svensson, and C. Botella, "Power allocation for two-cell two-user joint transmission," *IEEE Commun. Letters*, vol. 16, no. 9, pp. 1474-1477, 2012.

[11] F. Berggren, et. al., Dynamic spectrum access, phase 1: scenarios and research challenges, Sept. 2004, available: http://www.queseth.se/olav/pubs/DSAReportPhase1.pdf.

[12] B. Makki and T. Eriksson, "On the ergodic achievable rates of spectrum sharing networks with finite backlogged primary users and an interference indicator signal," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3079-3089, 2012.

[13] B. Makki, N. Seifi, and T. Eriksson, "Multi-user diversity with two-step channel state information feedback," *IET Commun.*, vol. 6, no. 9, pp. 1119-1125, 2012.

[14] B. Makki and T. Eriksson, "On hybrid ARQ and quantized CSI geedback schemes in quasi-static fading channels," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 986-997, April 2012.

[15] S. Boyd and L. Vandenberghe, "Convex optimization", Cambridge Univesity Press, NY, USA, 2004.