Power Allocation Algorithms for ZF-THP Sum Rate Optimization in Multi-user Multi-antenna Systems

Wookbong Lee, Changick Song, Sangrim Lee, Kilbom Lee, Jin Sam Kwak*, and Inkyu Lee, *Senior Member, IEEE*

School of Electrical Engineering, Korea University, Seoul, Korea
*LG Electronics, Anyang, Korea
Email: {wookbong, generalsci, sangrim78, bachhi, inkyu}@korea.ac.kr, jinsam.kwak@lge.com

***Abstract*—In this paper, we study a power allocation technique for Tomlinson-Harashima precoding (THP) in multi-user multiple input single output (MISO) downlink systems. In contrast to previous approaches, a mutual information based method is exploited for maximizing the sum rate of zero-forcing THP systems. Then, we propose a simple power allocation algorithm which assigns proper power level for modulo operated users. Simulation results show that the proposed scheme outperforms a conventional water-filling method, and achieves near optimal performance with much reduced complexity.**

I. Introduction

A multi-user multiple input multiple output (MIMO) technique allows us to support several users in cellular communication systems [1]. A downlink transmission system is often modeled as a MIMO broadcast channel (BC) where a base station (BS) with multiple antennas transmits to geographically separated users simultaneously. In the MIMO BC scenario, how to handle multi-user interference (MUI) has been one of major problems, and various precoding techniques have been made in the literature [2]–[9].

Dirty paper coding (DPC) is known as the optimal solution in the MIMO BC [2]. However, implementing the DPC is quite challenging due to its complexity at both a transmitter and a receiver [3]. In order to reduce the operational complexity while perfectly eliminating the MUI at each receiver, zero-forcing beamforming (ZF-BF) [3], zero-forcing dirty paper coding (ZF-DPC) [4], and zero-forcing Tomlinson-Harashima precoding (ZF-THP) [5] have been proposed. Especially, the ZF-THP presubstracts residual interference utilizing the concept of decision feedback at the transmitter, and can be considered as a practical implementation method for the ZF-DPC.

Proper power allocation introduces a considerable performance improvement in the MIMO BC scenario. In [4], the authors provided a power allocation solution for the ZF-DPC where the single input single output (SISO) capacity formula for each virtual parallel channel [10] can be used for the power allocation. However, for THP based systems, there exists a THP loss [6], which incurs a difficulty in finding the optimal transmit power for each user.

Since the actual input distribution for THP schemes is not Gaussian due to the modulo operation even if input distribution before the modulo operation is Gaussian, we cannot directly adopt the BC capacity formula [11] [12] for optimization. Recently, the authors in [7] proposed a two-step design process for minimizing the total transmit power while satisfying every user's minimum data rate and maximum bit-error rate (BER) requirement. First, the BER and rate requirements are converted to "virtual rate" requirements, which account for a gap-to-capacity introduced by practical quadrature amplitude modulation (QAM) and the THP. Then, input covariance matrices are determined so that the virtual rates are provided at the minimum total transmit power. However, this approach does not give a solution for the sum rate performance maximization of the THP, since the gap-to-capacity or the THP loss is a function of the transmit power for each user.

In this paper, we propose a simple power allocation algorithm for multi-user multiple input single output (MISO) downlink systems employing the ZF-THP. Instead of the gap-to-capacity based approach, we directly optimize the sum rate of the ZF-THP by allocating proper power for each user. For a non-Gaussian input distribution, the maximum achievable rate can no longer be expressed by the capacity. Thus, we formulate the maximization problem based on mutual information of an actual input distribution according to utilized constellations and the THP. Unfortunately, the mutual information for THP systems is generally a non-convex function with respect to the power allocation. Thus, we limit the minimum power level for each modulo operated user to make the problem

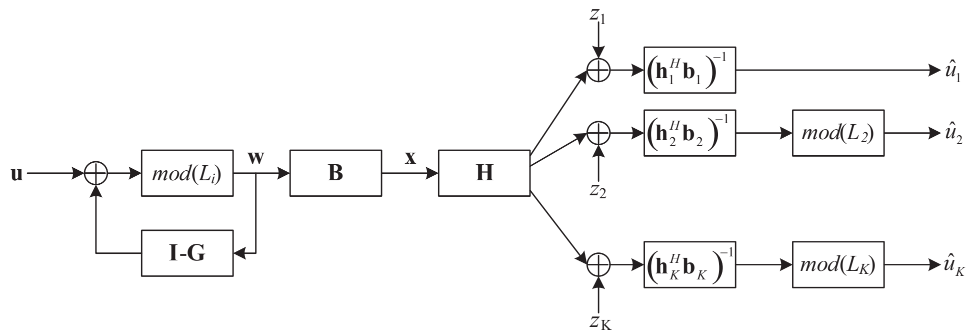

Fig. 1. System description of the ZF-THP system

tractable. Simulation results show that the sum rate provided by the proposed power allocation scheme outperforms the conventional ZF-DPC based power allocation scheme in [4]. Also, we confirm that the proposed scheme achieves near optimal performance with much reduced complexity.

The following notations are used throughout the paper. We employ uppercase boldface letters for matrices and lowercase boldface for vectors. For any general matrix $\mathbf{A}$, $\mathbf{A}^T$ and $\mathbf{A}^H$ denote transpose and Hermitian, respectively. For any complex scalar variable a, $\Re(a)$ and $\Im(a)$ represent the real and imaginary part of a, respectively.

The rest of the paper is organized as follows: In Section II, we briefly introduce the conventional THP systems. Then we formulate the optimization problem of interest in Section III, and its performance is presented in Section IV. Section V concludes the paper.

II. System model

In this section, we first review the SISO THP system. The main idea of the THP system is presubtracting known interference at the transmitter using the concept of decision feedback. Since the presubtracted signal may have very large signal power, a modulo operation is employed. Letting x and s be the transmit and the interference signal, respectively, the received signal y in the THP system with size L modulo operation is obtained by

$$\begin{aligned} y &= x + s + z \\ &= (u - s)_{mod(L)} + s + z \end{aligned}$$

where z denotes the complex additive white Gaussian noise (AWGN) with variance σ_z^2, u is the user data chosen from the M-ary constellation χ_M, and $(a)_{mod(L)}$ indicates the modulo operation whose output lies within the interval of $[-L, L)$ for each dimension (real and imaginary). Here we have

$$(a)_{mod(L)} = a + 2L\left(\left\lfloor \frac{1 - \Re(a)/L}{2} \right\rfloor + j \left\lfloor \frac{1 - \Im(a)/L}{2} \right\rfloor\right)$$

where $\lfloor b \rfloor$ represents the maximum integer value which does not exceed a real variable b. The choice of L can be determined to maintain the minimum distance of the system.

After applying the modulo operation again at the receiver, the final observation $\hat{u}$ can be expressed by

$$\begin{aligned} \hat{u} &= (y)_{mod(L)} = \left((u - s)_{mod(L)} + s + z\right)_{mod(L)} \\ &= u + z + 2L(\alpha + \beta j) \end{aligned} \tag{1}$$

where α and β are integers determined by u, s, and z so that the range of $\hat{u}$ lies within $[-L, L)$ for each dimension.

Now we review the ZF-THP system which applies THP using a ZF beamformer in MISO systems. A structure of ZF-THP system [5] is shown in Fig. 1 for the MISO BC channel. Let us consider a BS equipped with N_t antennas which tries to transmit independent information simultaneously to K users with a single antenna. Note that the performance of ZF-THP depends on the user ordering [5]. Without loss of generality we assume that the encoding order is set from the 1-st user to the K-th user stream. Then for flat fading channels, the received signal at the i-th user ($i = 1, \ldots, K$) can be written as

$$\begin{aligned} y_i &= \mathbf{h}_i^H \mathbf{x} + z_i \\ &= \sqrt{P_i}\mathbf{h}_i^H \mathbf{b}_i w_i + \mathbf{h}_i^H \sum_{i \neq j} \sqrt{P_j}\mathbf{b}_j w_j + z_i \end{aligned}$$

where $\mathbf{h}_i \in \mathbb{C}^{N_t \times 1}$ is the complex channel response vector between the transmitter and the i-th user, $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ represents the transmit signal vector with $E[\mathbf{x}^H\mathbf{x}] = P$, $P_i = \varepsilon_{x_i}$ designates the transmit signal power for the i-th user satisfying the total transmit power constraint $\sum_{i=1}^K P_i = P$, z_i indicates the AWGN at the i-th user, w_i is defined by $w_i = \left(u_i - \sum_{j=1}^{i-1} g_{i,j} w_j\right)_{mod(L_i)}$ with $g_{i,j}$ being the (i,j)-th element of the interference cancelation matrix $\mathbf{G} \in \mathbb{C}^{K \times K}$, u_i denotes the i-th user signal drawn from the size M_i constellation point χ_{M_i} with unit variance, L_i equals the i-th user's modulo size, and $\mathbf{b}_i \in \mathbb{C}^{N_t \times 1}$ stands for the transmit precoder for the i-th user. Note that since the output signal uniformly spreads over $[-L_i, L_i)$ in each dimension [6], the variance of w_i becomes $\sigma_{w,i}^2 = 2L_i^2/3$ for $i \neq 1$ and $\sigma_{w,1}^2 = 1$.

For the ZF-THP system, the transmit precoding matrix $\mathbf{B} \in \mathbb{C}^{N_t \times K}$ can be obtained by QR decomposition of the stacked channel matrix $\mathbf{H} \triangleq [\mathbf{h}_1\mathbf{h}_2 \ldots \mathbf{h}_K]^T = \mathbf{R}^H\mathbf{Q}^H$ where $\mathbf{Q} \in \mathbb{C}^{N_t \times N_t}$ and $\mathbf{R} \in \mathbb{C}^{N_t \times K}$ are a unitary and an upper triangular matrix, respectively. Then, we have

$$\mathbf{B} = \left[\sqrt{P_1}\mathbf{b}_1 \ \sqrt{P_2}\mathbf{b}_2 \ \ldots \ \sqrt{P_K}\mathbf{b}_K\right] \tag{2}$$

where $\mathbf{b}_i = \mathbf{q}_i/\sigma_{w,i}$ with $\mathbf{q}_i$ being the i-th column vector of $\mathbf{Q}$.

For the first user, the received signal after eliminating the channel gain $\sqrt{P_1}\mathbf{h}_1^H\mathbf{b}_1$ at the receiver is now written as

$$\hat{u}_1 = u_1 + \frac{z_1}{\sqrt{P_1}\mathbf{h}_1^H\mathbf{b}_1}. \tag{3}$$

Also for the i-th receiver ($i \neq 1$), after removing the channel gain $\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i$ and applying the modulo operation, the received signal becomes

$$\begin{aligned}\hat{u}_i &= \left(w_i + \sum_{j=1}^{i-1}\frac{\sqrt{P_j}\mathbf{h}_i^H\mathbf{b}_j}{\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i}w_j + \frac{z_i}{\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i}\right)_{mod(L_i)}\\ &= \left(u_i - \sum_{j=1}^{i-1}g_{i,j}w_j + \sum_{j=1}^{i-1}\frac{\sqrt{P_j}\mathbf{h}_i^H\mathbf{b}_j}{\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i}w_j + \frac{z_i}{\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i}\right)_{mod(L_i)}\end{aligned} \tag{4}$$

Here, we set $g_{i,j}$ to be $\sqrt{P_j}\mathbf{h}_i^H\mathbf{b}_j/\sqrt{P_i}\mathbf{h}_i^H\mathbf{b}_i$ for $i > j$, and 0 for $i \leq j$ to completely eliminate other users' interference. Note that we do not need the modulo operator at the first user.

III. Mutual Information Based Optimization

In this section, we propose a mutual information based simple power allocation algorithm to optimize the sum rate performance of ZF-THP. It was recognized in [6] that THP systems suffer from the following loss with respect to capacity.

- Shaping loss: When the power of the interference signal s is large, x is uniformly distributed over $[-L, L)$ in each dimension. Since the capacity is achieved with the Gaussian distributed input, this introduces a shaping loss.
- Power loss: The transmitted signal x in the THP can have higher power than the intended signal u. For example, with the input signal of unit power, the transmitted signal power ε_x equals $2L^2/3$.
- Modulo loss: After the modulo operation at the receiver, the noise is no longer AWGN and the number of nearest neighbors increases.

As a consequence of these THP loss, the gap between capacity and mutual information becomes large. To better reflect supportable rates, we introduce mutual information based optimization.

First of all, we need to calculate the mutual information between u and $\hat{u}$ in equation (1). Since the real and imaginary parts are orthogonal, the mutual information for the size M

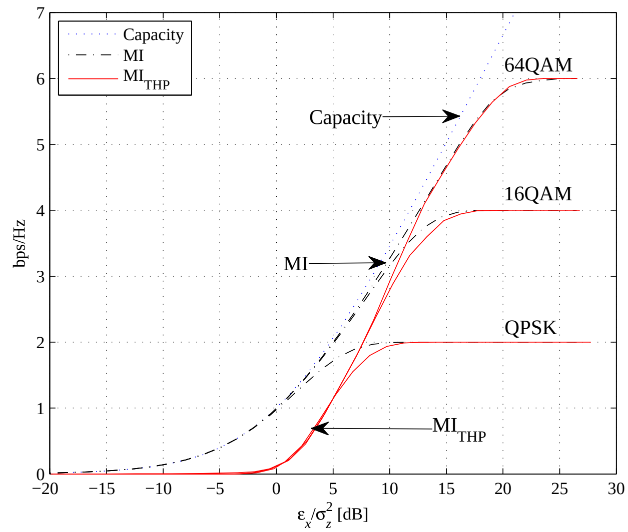

Fig. 2. Mutual information as a function of ε_x/σ_z^2

constellation can be obtained by

$$\begin{aligned}\mathrm{MI}_{\mathrm{THP},M} &= I(u;\hat{u}) = 2\left(H(\Re(\hat{u})) - H(\Re(\hat{u})|\Re(u))\right)\\ &= \frac{2}{M}\sum_{m=1}^{M}\int_{-\infty}^{\infty} p_m \log_2\left(\frac{p_m}{\frac{1}{M}\sum_{k=1}^{M}p_k}\right)d\Re(\hat{u})\end{aligned}$$

where p_m represents $p\left(\Re(\hat{u})|\Re(u) = \Re(a_m)\right)$ and a_m is the m-th element of χ_M.

Also, for an arbitrary s and the additive Gaussian noise z with variance $\sigma_z^2/2$ per dimension, p_m can be expressed as

$$p_m = \frac{1}{\sqrt{\pi\sigma_z^2}}\sum_{l=-\infty}^{\infty}\exp\left(-\frac{|\Re(\hat{u}) - \Re(a_m) + 2lL|^2}{\sigma_z^2}\right)$$

for $-L \leq \Re(\hat{u}) < L$ and 0 for elsewhere. Fig. 2 illustrates the $\mathrm{MI}_{\mathrm{THP},M}$ values for M-QAM at $E[|x|^2]/E[|z|^2] = \varepsilon_x/\sigma_z^2$ in comparison to the mutual information without the modulo operation denoted by MI_M.

Let us define $\mathrm{MI}_{\mathrm{THP}}$ and MI as the mutual information with the maximum constellation level $M_{\max}$ with and without the modulo operation, respectively. Then, from (2)-(4), the mutual information based optimization problem for ZF-THP systems can be formulated as

$$\begin{aligned}\max_{P_1 \ldots P_K} \quad & \mathrm{MI}(\gamma_1 P_1) + \textstyle\sum_{i=2}^{K}\mathrm{MI}_{\mathrm{THP}}(\gamma_i P_i)\\ s.t. \quad & \textstyle\sum_{i=1}^{K} P_i \leq P\end{aligned} \tag{5}$$

where $\gamma_i = |\mathbf{h}_i^H\mathbf{b}_i|^2/\sigma_z^2$.

To solve the problem (5), we apply the Lagrange multiplier as

$$J = \mathrm{MI}(\gamma_1 P_1) + \sum_{i=2}^{K}\mathrm{MI}_{THP}(\gamma_i P_i) + \lambda\left(\sum_{i=1}^{K}P_i - P\right).$$

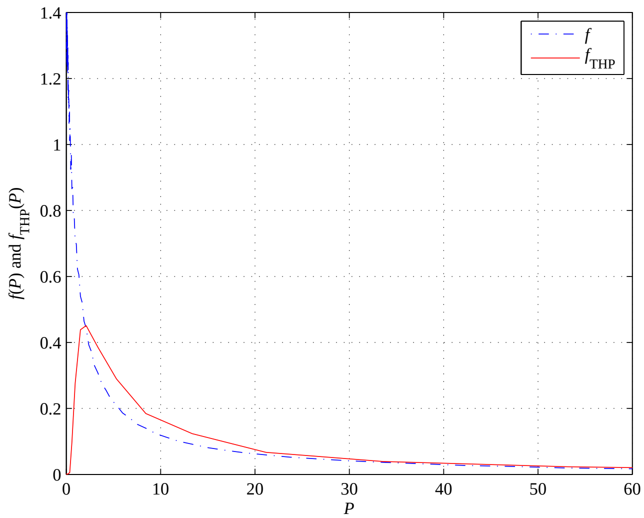

Fig. 3. Plot of derivatives of MI

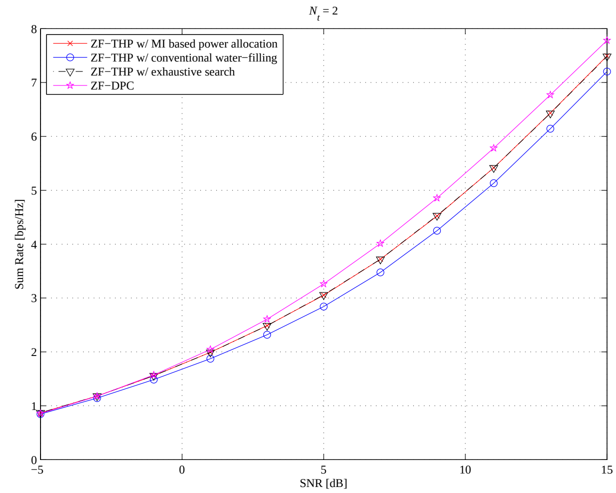

Fig. 4. Sum rate performance of two transmit antenna systems

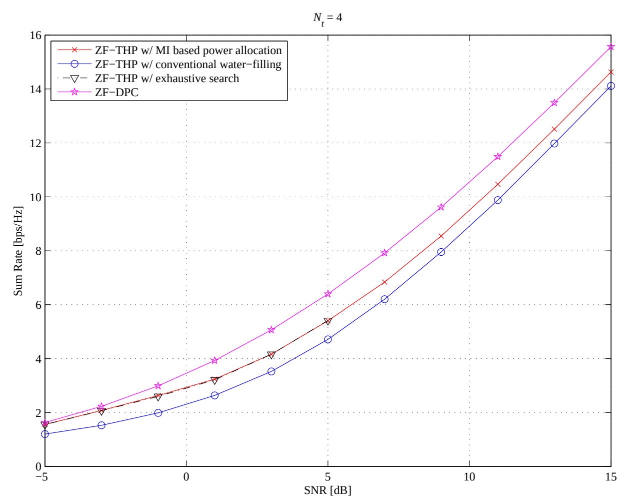

Fig. 5. Sum rate performance of four transmit antenna systems

Setting the gradient of J with respect to P_i equal to zero, we have

$$\frac{\partial J}{\partial P_1} = \gamma_1 \frac{\partial \mathrm{MI}(\gamma_1 P_1)}{\partial P_1} + \lambda = 0 \qquad \text{for } i = 1$$

and

$$\frac{\partial J}{\partial P_i} = \gamma_i \frac{\partial \mathrm{MI}_{THP}(\gamma_i P_i)}{\partial P_i} + \lambda = 0 \qquad \text{for } i \neq 1.$$

Define $f(P) \triangleq \frac{\partial \mathrm{MI}(P)}{\partial P}$, $f_{\mathrm{THP}}(P) \triangleq \frac{\partial \mathrm{MI}_{\mathrm{THP}}(P)}{\partial P}$, and its inverse as f^{-1} and f_{THP}^{-1}, respectively. Then, the solution is easily computed for $i = 1$ as

$$P_1^* = \begin{cases} \frac{1}{\gamma_1} f^{-1}\left(-\frac{\lambda}{\gamma_1}\right) & \text{for } -\frac{\lambda}{\gamma_1} \leq \max(f(P)) \\ 0 & \text{else} \end{cases} \tag{6}$$

and for $i \neq 1$ as

$$P_i^* = \begin{cases} \frac{1}{\gamma_i} f_{\mathrm{THP}}^{-1}\left(-\frac{\lambda}{\gamma_i}\right) & \text{for } -\frac{\lambda}{\gamma_i} \leq \max(f_{\mathrm{THP}}(P)) \\ 0 & \text{else.} \end{cases} \tag{7}$$

Here, each solution must satisfy the total power constraint,

$$\left(\sum_{k=1}^{K} P_k(\lambda)\right) - P = 0 \tag{8}$$

where $P_k(\lambda)$ represents P_k^* for the corresponding λ. Since $\mathrm{MI}_{\mathrm{THP}}$ is not convex with repect to P as shown in Fig. 2, a solution of equation (5) may not be globally optimal.

In equations (6) and (7), $f(P)$ and $f_{\mathrm{THP}}(P)$ should be found numerically from MI and $\mathrm{MI}_{\mathrm{THP}}$, respectively, as illustrated in Fig. 3, because there is no closed form solution for the mutual information and its derivative. Thus, we need to use, for example, a look-up table to obtain the final solution. Also, f_{THP} is not a monotonic function, and thus f_{THP}^{-1} does not have a unique solution. To overcome this issue, we ignore the low power region below the peak for f_{THP}^{-1} by limiting the minimum power level for modulo operated users. Then, f_{THP}^{-1} becomes a monotonically decreasing function, and a simple bisection method can be applied to find a solution in (8). We will show in the following section that our approach does not degrade the performance while almost the same complexity is maintained as the conventional water-filling algorithm.

IV. Simulation Results

In this section, we demonstrate the efficiency of our algorithm using Monte-Carlo simulations compared to the conventional water-filling optimization method in [4]. The sum rate of ZF-THP systems is calculated with $M_{\max} = 64$. We include the performance of an exhaustive search method

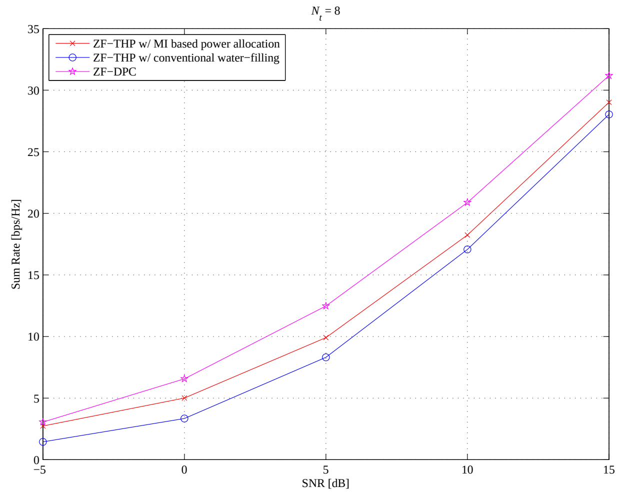

Fig. 6. Sum rate performance of eight transmit antenna systems

that serves as a benchmark to validate our algorithm. Here the exhaustive search method refers collecting best performance among all possible power level combinations with 0.1dB step. Also, we include ZF-DPC results, which are upper bounds for the ZF-THP system.

To determine the encoding order of the ZF-THP system, we evaluate all possible $K!$ combinations to identify the best one. For all simulations, we assume that the transmitter supports at most N_t users. Since power allocation algorithms may provide zero power to some users, the number of effective number of users can be less than N_t.

Figures 4, 5, and 6 show simulation results for the case of $N_t = 2$, 4, and 8, respectively. Note that we do not include an exhaustive search method for $N_t = 8$ case due to its extremely high complexity. From these plots, we confirm that the proposed scheme outperforms the ZF-THP system with the conventional water-filling optimization method. Also, we can see that the MI based optimization exhibits performance identical to the exhaustive scheme, which shows that limiting the minimum power level for modulo operated users does not have any impact on the performance. One interesting point here is that our scheme shows high performance gains over the conventional water-filling optimization for the moderate SNR range, which is important especially for cellular networks where other cell interference exists [13]. For N_t =4 and 8, 5∼30 % and 7∼50 % gains at 0∼10dB can be obtained, respectively, while the performance gap is reduced as SNR grows. This is due to the fact that the performance difference is large when a low constellation level is selected, since a THP loss is larger for lower constellation levels.

V. Conclusion

In this paper, we have proposed a simple power allocation method for ZF-THP systems over MISO broadcast channels. We have first formulated the optimization problem based on the mutual information. Then we have limited the minimum power level for modulo operated users to simplify the algorithm. Simulation results show that the proposed scheme outperforms the ZF-THP system with the conventional water-filling optimization method, and also achieves near optimal performance with much reduced complexity.

Acknowledgment

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2010- 0017909).

References

[1] Q. Li, G. Li, W. Lee, M. Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO techniques in WiMAX and LTE: a feature overview," *IEEE Communication Magazine*, vol. 48, pp. 86–92, May 2010.
[2] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 2658–2668, October 2003.
[3] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 52, pp. 461–471, February 2004.
[4] G. Caire and S. Shamai, "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, pp. 1691–1706, July 2003.
[5] C. Windpassinger, R. F. Fisher, T. Vencel, and J. B. Huber, "Precoding in Multiantenna and Multiuser Communications," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1305–1316, July 2004.
[6] W. Yu, D. P. Varodayan, and J. M. Cioffi, "Trellis and Convolutional Precoding for Transmitter-Based Interference Presubtraction," *IEEE Transactions on Communication*, vol. 53, pp. 1220–1230, July 2005.
[7] C.-H. F. Fung, W. Yu, and T. J. Lim, "Precoding for the Multiantenna Downlink: Multiuser SNR Gap and Optimal User Ordering," *IEEE Transactions on Communication*, vol. 55, pp. 188–197, January 2007.
[8] H.-S. Han, S.-H. Park, S. Lee, and I. Lee, "Modulo Loss Reduction for Vector Perturbation Systems," *IEEE Transactions on Communications*, vol. 58, pp. 3392–3397, December 2010.
[9] S.-H. Park, H.-S. Han, and I. Lee, "A Decoupling Approach for Low-Complexity Vector Perturbation in Multiuser Downlink Systems," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 1697–1701, June 2011.
[10] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, 2004.
[11] W. Yu and J. M. Cioffi, "Sum Capacity of Gaussian Vector Broadcast Channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 1875–1892, September 2004.
[12] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, Achievable Rates, and Sum-Rate Capacity of Gaussian MIMO Breadcast Channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 2658–2668, October 2003.
[13] J. G. Andrews, W. Choi, and Robert W. Heath Jr., "Overcoming Interference in Spatial Multiplexing MIMO Cellular Networks," *IEEE Transactions on Information Theory*, vol. 14, pp. 95–104, December 2007.