

# Coordinated Tomlinson-Harashima Precoding Design Algorithms for Overloaded Multi-user MIMO Systems

Keke Zu  $\diamond\ddagger$ , Bin Song  $\diamond$ , Martin Haardt  $\diamond$  and Rodrigo C. de Lamare  $\dagger$   
 $\diamond$  Communications Research Laboratory, Ilmenau University of Technology

PO Box 100565, D-98684 Ilmenau, Germany

$\dagger$  Dept. of Electronics / CETUC, University of York / PUC-Rio

York YO10 5DD, U.K./ Marquês de S.Vicente, 225 Gávea, Rio de Janeiro, Brazil

Emails: {ke-ke.zu, bin.song, martin.haardt}@tu-ilmenau.de, rodrigo.delamare@york.ac.uk

**Abstract**—Tomlinson-Harashima precoding (THP) is a non-linear processing technique employed at the transmit side to implement the concept of dirty paper coding (DPC). The application of THP is restricted by the dimensionality constraint that the number of transmit antennas has to be greater or equal to the total number of receive antennas. In this paper, we propose an iterative coordinate THP algorithm for overloaded scenarios in which the total number of receive antennas is larger than the number of transmit antennas. The proposed algorithm is implemented on two types of THP structures, the decentralized THP (dTHP) with diagonal weighted filters at the receivers of the users, and the centralized THP (cTHP) with diagonal weighted filter at the transmitter. Simulation results show that a significantly better bit error rate (BER) and sum-rate performances can be achieved by the proposed iterative coordinate THP algorithm as compared to previously reported techniques.

**Index Terms**—Tomlinson-Harashima precoding (THP), dimensionality constraint, overloaded systems, coordinated beamforming algorithm.

## I. INTRODUCTION

In multi-user MIMO (MU-MIMO) broadcast channels with independent and identically distributed (i.i.d.) Gaussian interference, the capacity region is achieved by a combination of dirty paper coding (DPC), power allocation and scheduling [1], [2]. The concept of DPC has been originally proposed in [3] and states that the capacity of the broadcast channel is the same as if the interference was not present by setting the transmitted signal equal to the desired data minus the interference. The DPC theory, however, is not suitable for practical applications due to the requirement of infinitely long codewords [4].

Tomlinson-Harashima precoding (THP) [5], [6] is a pre-equalization technique originally proposed for channels with intersymbol interference (ISI). THP techniques were subsequently extended from temporal equalization to spatial equalization for MIMO precoding in [7], [8], where an equal number of transmit and receive antennas are assumed. Although THP still suffers a performance loss as compared to the theoretical upper bound achieved by DPC [9], it can work as a

cost-effective replacement of DPC in practice by implementing an LQ decomposition [10], [11]. In the literature, however, the system dimensionality for implementing the THP algorithms is always set as the number of transmit antennas greater than or equal to the total number of receive antennas, which is termed as the dimensionality constraint in [12].

To overcome the dimensionality constraint and allow precoding algorithms to operate in overloaded systems. Overloaded systems refer to those MIMO system scenarios in which the dimensionality constraint is violated, i.e., the total number of receive antennas exceeds the number of transmit antennas. A receive antenna selection method has been proposed in [13]. Another approach for overloaded systems is the coordinated beamforming (CBF) technique which employs iterative operations to jointly update the transmit-receive beamforming vectors [14]-[18]. The convergence behavior of the iterations is not considered in [14], and the coordinated transmission strategy in [15] only supports a single data stream to each user. A flexible coordinated beamforming (FlexCoBF) algorithm has been proposed in [16]-[18] to support the transmission of multiple data streams to each user employing linear precoding techniques [19].

Based on our previous work on FlexCoBF, an iterative coordinate nonlinear THP algorithm is proposed in this work. We consider two THP structures according to the positions of the diagonal weighted filter, decentralized filters located at the receivers or centralized filters deployed at the transmitter, which are denoted as dTHP or cTHP, respectively [20], [21]. Most of the previous research works on THP, however, have only focused on one of the structures. In this work, we develop iterative coordinate THP schemes for both THP structures. We devise a coordinated beamforming strategy in which the transmit and the receive filters of THP structures are jointly optimized through iterations, resulting in precoding techniques with improved performance that are suitable for overloaded MU-MIMO systems. In particular, we study the proposed algorithms in relevant scenarios via simulations and compare them with the previously reported techniques. The main contributions of the work can be summarized as

- 1) An iterative coordinated strategy is developed to solve

$\ddagger$  Keke Zu now is with Ericsson Research, Sweden

the dimensionality constraint for the nonlinear THP algorithms.

- 2) The proposed iterative coordinate THP algorithm supports the transmission of multiple data streams.
- 3) The proposed iterative coordinate THP algorithm is developed for both cTHP and dTHP structures.
- 4) We investigate in a simulation study the sum-rate and the BER performances of the proposed and existing precoding algorithms.

This paper is organized as follows. The system model and the basics of THP techniques are described in Section II. The proposed iterative coordinated THP algorithm is described in detail in Section III. Simulation results and conclusions are presented in Section IV and Section V, respectively.

## II. SYSTEM MODEL AND THP ALGORITHMS

In this section, we describe the system model of a MU-MIMO broadcast system equipped with THP precoding algorithms and access to channel state information.

### A. MU-MIMO System Model

We consider an uncoded MU-MIMO broadcast system illustrated in Fig. 1, equipped with  $M_t$  transmit antennas at the base station (BS),  $K$  users in the system each equipped with  $M_k$  receive antennas, and the total number of receive antennas is  $M_r = \sum_{k=1}^K M_k$ . The combined transmit data streams are denoted as  $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_K^T]^T \in \mathbb{C}^{r \times 1}$  with  $\mathbf{s}_k \in \mathbb{C}^{r_k \times 1}$ , where  $r$  is the total number of transmit data streams and  $r_k$  is the number of data streams for user  $k$ . The combined channel matrix is denoted as  $\mathbf{H} = [\mathbf{H}_1^T, \mathbf{H}_2^T, \dots, \mathbf{H}_K^T]^T \in \mathbb{C}^{M_r \times M_t}$  and  $\mathbf{H}_k \in \mathbb{C}^{M_k \times M_t}$  is the  $k$ th user's channel matrix.

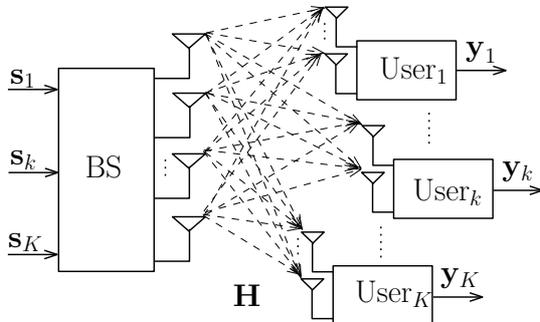


Fig. 1: The MU-MIMO System Model

Note that power-loading schemes [8] could be used to determine the number of data streams or allocate more power to a weaker user to improve the overall performance. However, for simplicity, we assume that equal power allocation between users and transmitted streams is performed since the power allocation is not the focus of this paper.

### B. Two Basic THP Structures

We assume the channel state information (CSI) is perfectly known at the transmit side, then the interference of the parallel streams of a MU-MIMO system with spatial multiplexing can

be subtracted from the current stream. This successive inter-user interference cancellation technique at the transmit side is known as THP. Generally, there are three filters to implement the THP algorithms: the feedback filter  $\mathbf{B}$ , the feedforward filter  $\mathbf{F}$ , and the scaling filter  $\mathbf{G}$ . According to the position of  $\mathbf{G}$ , there are two THP structures, which are illustrated in Fig. 2. The decentralized THP (dTHP) employs  $\mathbf{G}$  (or sub-matrices of it) at the receivers, whereas the centralized THP (cTHP) uses  $\mathbf{G}$  at the transmitter.

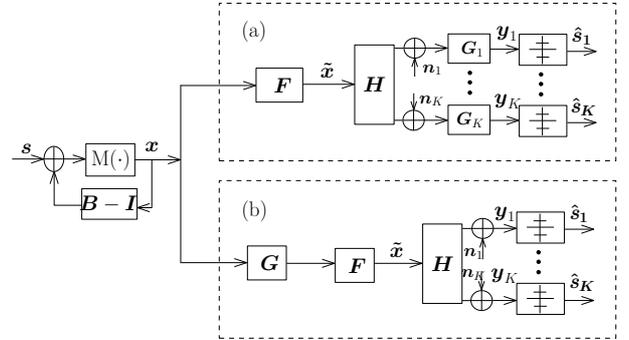


Fig. 2: The two basic THP structures

(a) Decentralized THP: the scaling matrix  $\mathbf{G}$  is separately placed at the receivers.

(b) Centralized THP: the scaling matrix  $\mathbf{G}$  is placed at the transmitter.

The feedback filter  $\mathbf{B}$  is used to successively cancel the interference caused by the previous streams from the current stream. Therefore, the feedback filter  $\mathbf{B}$  should be a lower triangular matrix with ones on the main diagonal [8]. The feedforward filter  $\mathbf{F}$  is used to enforce the spatial causality and has to be implemented at the transmit side for MU-MIMO systems because the physically distributed users cannot be processed jointly. The scaling matrix  $\mathbf{G}$  contains the corresponding positive weighting coefficient for each stream, and thus it should have a diagonal structure. The quantity  $\mathbf{x}$  is the combined transmit signal vector after the feedback operation and  $\tilde{\mathbf{x}}$  is the combined transmit signal vector after the precoding described by

$$\tilde{\mathbf{x}}^{(\text{dTHP})} = \mathbf{F}\mathbf{x}, \quad (1)$$

$$\tilde{\mathbf{x}}^{(\text{cTHP})} = \mathbf{F}\mathbf{G}\mathbf{x}. \quad (2)$$

The filtering matrices of THP can be effectively obtained by implementing an LQ decomposition [21] of the combined channel matrix  $\mathbf{H}$ , i.e.,

$$\mathbf{H} = \mathbf{L}\mathbf{Q}, \quad (3)$$

where  $\mathbf{L} \in \mathbb{C}^{M_r \times M_t}$  is a lower triangular matrix and  $\mathbf{Q} \in \mathbb{C}^{M_t \times M_t}$  is a unitary matrix. Therefore, the filters for the THP algorithms can be obtained as

$$\mathbf{F} = \mathbf{Q}^H, \quad (4)$$

$$\mathbf{G} = \text{diag}(\mathbf{L})^{-1}, \quad (5)$$

$$\mathbf{B}^{(\text{dTHP})} = \mathbf{G}\mathbf{L}, \quad (6)$$

$$\mathbf{B}^{(\text{cTHP})} = \mathbf{L}\mathbf{G}, \quad (7)$$

where  $\text{diag}(\mathbf{A})$  creates a diagonal matrix with the diagonal elements of the matrix  $\mathbf{A}$  on the main diagonal.

From Fig. 2, the transmitted symbols  $x_i$  are successively generated as

$$x_i = \text{M}\left(s_i - \sum_{j=1}^{i-1} b_{i,j} x_j\right), \quad i = 1, \dots, r, \quad (8)$$

where  $s_i$  is the  $i$ th transmit data stream with variance  $\sigma_s^2$  and  $b_{i,j}$  are the elements of  $\mathbf{B}$  in row  $i$  and column  $j$ . The transmit power is significantly increased as the amplitude of  $x_i$  exceeds the modulation boundary by implementing the successive cancellation. A modulo operation  $\text{M}(\cdot)$ , which is defined element-wise, is employed to reduce the amplitude of the channel symbol  $x_i$  to the boundary of the modulation alphabet [23]

$$\text{M}(x_i) = x_i - \left\lfloor \frac{\text{Re}(x_i)}{\tau} + \frac{1}{2} \right\rfloor \tau - j \left\lfloor \frac{\text{Im}(x_i)}{\tau} + \frac{1}{2} \right\rfloor \tau, \quad (9)$$

where  $\tau$  is a constant for the periodic extension of the constellation. The specific value of  $\tau$  depends on the chosen modulation alphabet. Common choices for  $\tau$  are  $\tau = 2\sqrt{2}$  for QPSK symbols. The modulo processing is equivalent to adding a perturbation vector  $\mathbf{d}$  to the transmit data  $\mathbf{s}$ , such that the modified transmit data are [21], [22]

$$\mathbf{v} = \mathbf{s} + \mathbf{d}. \quad (10)$$

Thus, the initial signal constellation is extended periodically and the effective  $k$ th transmit data symbols  $\mathbf{v}_k$  are taken from the expanded set.

The received signal for dTHP and cTHP, after the feedback, feedforward and the scaling filters, is respectively given by

$$\mathbf{y}^{(\text{cTHP})} = \beta(\mathbf{H} \cdot \frac{1}{\beta} \mathbf{F} \mathbf{G} \mathbf{x} + \mathbf{n}), \quad (11)$$

$$\mathbf{y}^{(\text{dTHP})} = \mathbf{G}(\mathbf{H} \mathbf{F} \mathbf{x} + \mathbf{n}), \quad (12)$$

where the quantity  $\mathbf{n} = [\mathbf{n}_1^T, \mathbf{n}_2^T, \dots, \mathbf{n}_K^T]^T$  is the combined Gaussian noise vector with i.i.d. entries of zero mean and variance  $\sigma_n^2$ . The factor  $\beta$  is used to impose the power constraint  $\text{E}\|\tilde{\mathbf{x}}\|^2 = \xi$  with  $\xi$  being the average transmit power.

### III. PROPOSED ITERATIVE COORDINATED THP ALGORITHM

The previously discussed THP algorithms are only feasible under the dimensionality constraint that  $M_t \geq M_r$ . In this section, the coordinated strategy is developed for THP algorithms in case of overloaded systems.

For the case when  $M_r > M_t$ , the MU-MIMO system cannot support the transmission of  $M_r$  data streams. Assume that the number of actually transmitted data streams is  $r$  and it should satisfy  $r \leq M_t$ , which means that the maximum number of transmitted data streams cannot exceed the number of transmit antennas  $M_t$ . Then, a receive beamforming matrix  $\mathbf{W}_k \in \mathbb{C}^{r_k \times M_k}$  is introduced at each user to adjust

the transmit-receive filters iteratively. The equivalent channel matrix  $\mathbf{H}_e \in \mathbb{C}^{r \times M_t}$  is obtained as

$$\mathbf{H}_e = \begin{bmatrix} \mathbf{W}_1^H & 0 & \dots & 0 \\ 0 & \mathbf{W}_2^H & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \mathbf{W}_K^H \end{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_K \end{bmatrix} = \begin{bmatrix} \mathbf{W}_1^H \mathbf{H}_1 \\ \mathbf{W}_2^H \mathbf{H}_2 \\ \vdots \\ \mathbf{W}_K^H \mathbf{H}_K \end{bmatrix}. \quad (13)$$

For the design of the receive beamforming matrices  $\mathbf{W}_k$ , we investigate the feedback process in Fig. 2. Mathematically, the feedback processing is equivalent to implementing an inversion operation  $\mathbf{B}^{-1}$ . Therefore, the transmitted vector before precoding  $\mathbf{x}$  can be rewritten as

$$\mathbf{x} = \mathbf{B}^{-1} \mathbf{v} = \mathbf{B}^{-1}(\mathbf{s} + \mathbf{d}), \quad (14)$$

Substituting (14) into (11) and (12), the received signal for cTHP and dTHP can be, respectively, rewritten as

$$\mathbf{y}^{(\text{cTHP})} = \mathbf{v} + \beta \mathbf{n}, \quad (15)$$

$$\mathbf{y}^{(\text{dTHP})} = \mathbf{v} + \mathbf{G} \mathbf{n}. \quad (16)$$

From the above equations (15) and (16), the MU-MIMO broadcast channel is transformed into effective parallel single-user MIMO (SU-MIMO) channels. Therefore, the multi-user interference (MUI) is forced to zero at each user by the successive THP processing. In the proposed iterative coordinated THP algorithm, the receive beamforming matrix  $\mathbf{W}_k$  at each receiver is initialized with random matrices. Then, iterative computations are employed to update  $\mathbf{W}_k$  to enforce the zero MUI constraint at receiver. Assuming that the variable  $p$  represents the iteration index, the proposed iterative coordinated THP algorithm is performed in the following 5 steps:

- 1) Initialize the iteration index  $p$  to zero and  $\mathbf{W}_k^{(0)}$  to random matrices. Set the constant  $\epsilon$  as the threshold to iteratively enforce the zero MUI constraint at the receiver.
- 2) Set  $p = p+1$  and compute the equivalent channel matrix  $\mathbf{H}_e^{(p)}$  as

$$\mathbf{H}_e^{(p)} = \begin{bmatrix} \mathbf{W}_1^{(p-1)H} \mathbf{H}_1 \\ \mathbf{W}_2^{(p-1)H} \mathbf{H}_2 \\ \vdots \\ \mathbf{W}_K^{(p-1)H} \mathbf{H}_K \end{bmatrix}.$$

- 3) Apply the LQ decomposition on the equivalent channel matrix  $\mathbf{H}_e^{(p)} = \mathbf{L}_e^{(p)} \mathbf{Q}_e^{(p)}$  to obtain the THP filters as

$$\begin{aligned} \mathbf{F}_e^{(p)} &= \mathbf{Q}_e^{(p)H}, \\ \mathbf{G}_e^{(p)} &= \text{diag}(\mathbf{L}_e^{(p)})^{-1}, \\ \mathbf{B}_e^{(p)(\text{cTHP})} &= \mathbf{L}_e^{(p)} \mathbf{G}_e^{(p)}, \\ \mathbf{B}_e^{(p)(\text{dTHP})} &= \mathbf{G}_e^{(p)} \mathbf{L}_e^{(p)}. \end{aligned}$$

- 4) Update the  $p$ th combined receive beamforming matrix  $\mathbf{W}^{(p)}$  as the maximum ratio combining (MRC) receiver.

Alternatively, the minimum mean square error (MMSE) strategy could also be used.

$$\mathbf{W}_e^{(p)(c\text{THP})} = \mathbf{H}\mathbf{F}_e^{(p)}\mathbf{G}_e^{(p)}\left[\mathbf{B}_e^{(p)(c\text{THP})}\right]^{-1},$$

$$\mathbf{W}_e^{(p)(d\text{THP})} = \mathbf{H}\mathbf{F}_e^{(p)}\left[\mathbf{B}_e^{(p)(d\text{THP})}\right]^{-1}.$$

- 5) Track the alterations of the residual MUI after the THP processing as

$$\text{MUI}(\mathbf{H}_e^{(p+1)}\mathbf{P}_e^{(p)(c\text{THP})}) = \|\text{off}(\mathbf{H}_e^{(p+1)}\mathbf{P}_e^{(p)(c\text{THP})})\|_F^2,$$

$$\text{MUI}(\mathbf{H}_e^{(p+1)}\mathbf{P}_e^{(p)(d\text{THP})}) = \|\text{off}(\mathbf{H}_e^{(p+1)}\mathbf{P}_e^{(p)(d\text{THP})})\|_F^2,$$

where the equivalent precoding matrix

$$\mathbf{P}_e^{(p)(c\text{THP})} = \mathbf{F}_e^{(p)}\mathbf{G}_e^{(p)}\left[\mathbf{B}_e^{(p)(c\text{THP})}\right]^{-1}$$

$$\mathbf{P}_e^{(p)(d\text{THP})} = \mathbf{F}_e^{(p)}\left[\mathbf{B}_e^{(p)(d\text{THP})}\right]^{-1}.$$

The operation  $\text{off}(\mathbf{A})$  denotes the off-diagonal elements of the matrix  $\mathbf{A}$ . If the residual MUI is above the threshold  $\epsilon$ , go back to step 2. Otherwise, convergence is achieved and the iterative procedure can be ended. The maximum number of iteration is restricted to enforce convergence according to the specific system design requirement.

For setting the value of the threshold  $\epsilon$ , we usually apply  $\epsilon = 10^{-5}$  to perform the proposed coordinated THP algorithms. It is worth noting that the receive beamforming matrix  $\mathbf{W}_e^{(p)}$  is calculated centrally and needs to be informed to each receiver, which may bring extra control overhead or computational complexity.

#### IV. SIMULATION RESULTS

In this section, we assess the performance of the proposed iterative coordinated THP algorithms. A system with  $M_t = 8$  transmit antennas and  $K = 4$  users each equipped with  $M_k = 3$  receive antennas is considered; this scenario is denoted as the  $(3, 3, 3, 3) \times 8$  case. The quantity  $E_b/N_0$  is defined as  $E_b/N_0 = \frac{M_r E_s}{M_t N \sigma_n^2}$  with  $N$  being the number of information bits transmitted per channel symbol. An uncoded QPSK modulation scheme is employed in the simulations. The threshold  $\epsilon$  is set to  $10^{-5}$ , and the maximum number of iteration is restricted to 50. The channel matrix  $\mathbf{H}$  is assumed to be a complex i.i.d. Gaussian matrix with zero mean and unit variance.

For the  $(3, 3, 3, 3) \times 8$  MU-MIMO broadcast channel, the proposed iterative coordinated THP is implemented to support the maximum number of data streams at each user, i.e.,  $r_k = 2$ . The BER performance is illustrated in Fig. 3. The proposed iterative coordinated dTHP achieves a gain of more than 8 dB at a BER of  $10^{-2}$  as compared to our previously proposed linear iterative coordinated ZF precoding [18]. The iterative coordinated cTHP can still work well in this overloaded system

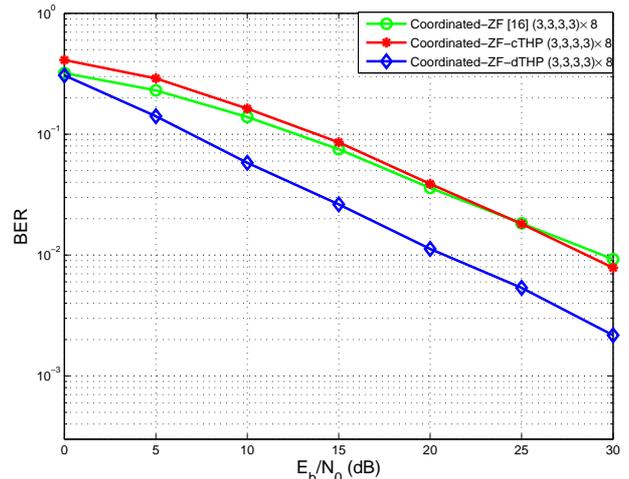


Fig. 3: BER performance with QPSK

but its performance is not as good as that of the linear one at low  $E_b/N_0$ s.

The sum-rate performance is displayed in Fig. 4. The DPC upper bound for 8 data streams is shown by the black curve. A much better sum-rate performance is achieved by both of the proposed iterative coordinated cTHP and dTHP algorithms than our previously proposed coordinated linear ZF precoding, and the sum-rate performance of the iterative coordinated dTHP algorithm is very close to that of the DPC. For the iterative coordinated cTHP, its sum-rate performance is not so sensitive as its BER performance compared to the original cTHP.

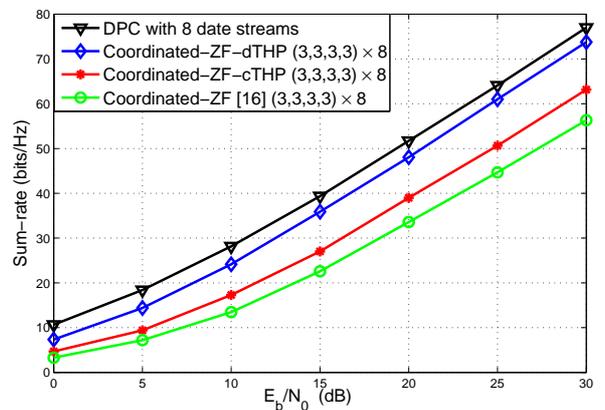


Fig. 4: Sum-rate performance of analyzed precoders

The complementary cumulative distribution function (CCDF) of the required number of iterations is illustrated in Fig. 5. The coordinated dTHP requires more iterations compared to the coordinated cTHP and coordinated ZF precoding. For most of the cases, the proposed coordinated THP algorithms converge in less than 20 iterations.

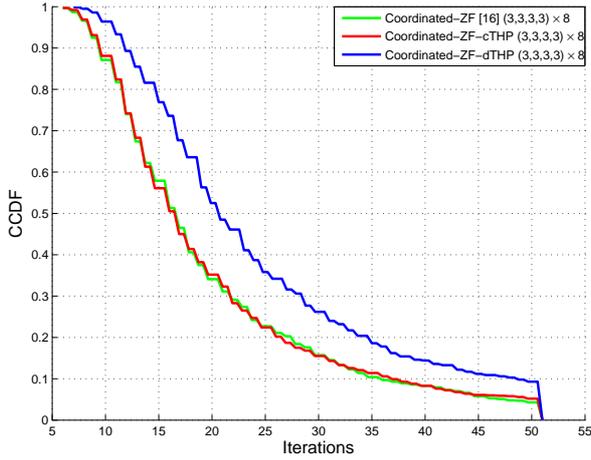


Fig. 5: CCDF of required iterations

Therefore, we have found that better BER and sum-rate performances can be achieved by the iterative coordinated dTHP as compared to the iterative coordinated cTHP. The BER performance of the iterative coordinated cTHP is more sensitive to the receive beamforming matrices  $\mathbf{W}_k$  but its sum-rate performance is not affected by overloading the MU-MIMO system. For the fixed threshold  $\epsilon = 10^{-5}$ , the coordinated dTHP requires more iterations compared to the coordinated cTHP.

## V. CONCLUSION

In this paper, an iterative coordinate THP algorithm has been proposed to relax the dimensionality constraints suffered by the original THP algorithms. Therefore, we consider an overloaded scenario where  $M_r > M_t$ . This condition is fulfilled in many situations that have been studied recently. For example, the users across cell borders have to be considered jointly by base stations (BSs) for coordinated multi-point (CoMP) transmission. Furthermore, when each user is equipped with multiple antennas, the BS simultaneously serves as many users as possible, which corresponds to a large number of receive antennas. The performance improvements of our proposed algorithm have been demonstrated by the simulation results compared to the previous CBF with linear precoding. The proposed coordinated dTHP illustrates much better BER and sum-rate performance than that of the cTHP. The improved performance of dTHP is at the cost of transmitting  $\mathbf{G}$  to each receiver.

## REFERENCES

[1] Q. Spencer, C. Peel, A. Swindlehurst and M. Haardt, "An introduction to the multi-user MIMO downlink," *IEEE Commun. Mag.*, vol. 42, issue: 10, pp. 60-67, Oct. 2004.

[2] H. Weingarten, Y. Steinberg and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 3936-3964, Sept. 2006.

[3] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, no. 3, pp. 439-441, May 1983.

[4] A. Khina and U. Erez, "On the Robustness of Dirty Paper Coding," *IEEE Trans. Commun.*, vol. 58, no. 5, May 2010.

[5] M. Tomlinson, "New automatic equaliser employing modulo arithmetic," *Electronic Letters*, vol. 7, Mar. 1971.

[6] H. Harashima and H. Miyakawa, "Matched-transmission technique for channels with intersymbol interference," *IEEE Trans. Commun.*, vol. 20, Aug. 1972.

[7] R. Fischer, C. Windpassinger, A. Lampe and J. Huber, "Space-time transmission using Tomlinson-Harashima precoding," in *Proc. ITG Conf. on Source and Channel Coding (SCC)*, Berlin, Jan. 2002, pp. 139-147.

[8] C. Windpassinger, R. Fischer, T. Vencel and J. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, Jul. 2004.

[9] W. Y. D. Varodayan and J. M. Cioffi, "Trellis and convolutional precoding for transmitter based interference presubstraction," *IEEE Trans. Commun.*, pp. 1220-1230, Jul. 2005.

[10] U. Erez, S. Shamai and R. Zamir, "Capacity and lattice strategies for canceling known interference," *IEEE Trans. Inf. Theory*, pp. 3820-3833, Nov. 2005.

[11] K. Zu, R. C. de Lamare and M. Haardt, "Multi-Branch Tomlinson-Harashima precoding for single-user MIMO systems," in *Proc. ITG/IEEE Workshop on Smart Antennas*, Dresden, Germany, Mar. 2012.

[12] Q. Spencer, A. Swindlehurst and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461-471, Feb. 2004.

[13] Z. Shen, R. Chen, J. Andrews, R. Health and B. Evans, "Sum capacity of multiuser MIMO broadcast channels with block diagonalization," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2040-2045, Jun. 2007.

[14] Z. Zhou, W. Hardjawana and B. Vucetic, "Iterative multiple beamforming algorithm for MIMO broadcast channels," *IEEE Commun. Lett.*, vol. 12, no. 10, pp. 743-745, Oct. 2008.

[15] C. Chae, D. Marzazese, T. Inoue and R. Health, "Coordinated beamforming for the multiuser MIMO broadcast channel with limited feed-forward," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6044-6056, Dec. 2008.

[16] B. Song and M. Haardt, "Closed-form coordinated beamforming for multi-user MIMO broadcast channels," in *Proc. Wireless World Research Forum (WWRF)*, Beijing, China, 2009.

[17] B. Song, F. Roemer and M. Haardt, "Using a new structured joint congruence (STJOCO) transformation of Hermitian matrices for precoding in multi-user MIMO systems," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, TX, USA, Mar. 2010.

[18] B. Song, F. Roemer and M. Haardt, "Flexible coordinated beamforming (FlexCoBF) for the downlink of multi-user MIMO systems in single and clustered multiple cells," *Signal Processing*, vol. 93, no. 9, pp. 2462-2473, Sept. 2013.

[19] M. Joham, W. Utschick and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53 no. 8, pp. 2700-2712, Aug. 2005.

[20] M. Huang, S. Zhou and J. Wang, "Analysis of Tomlinson-Harashima precoding in multiuser MIMO systems with imperfect channel state information," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, Sep. 2008.

[21] K. Zu, R. C. de Lamare and M. Haardt, "Multi-branch Tomlinson-Harashima precoding design for MU-MIMO systems: Theory and algorithm," *IEEE Trans. Commun. (2nd round review)*, 2013.

[22] M. Joham and W. Utschick, "Ordered spatial Tomlinson-Harashima precoding," in *Smart Antennas-State-of-the-Art*, ser. EURASIP Book Series on Signal Processing and Communications. New York: Hindawi Publishing Corporation, 2005.

[23] F. Dietrich, P. Breun and W. Utschick, "Robust Tomlinson-Harashima precoding for the wireless broadcast channel," in *IEEE Trans. Signal*