UNIVERSITY *of York*

This is a repository copy of *An Adaptive Optimal Mapping Selection Algorithm for PNC using Variable QAM Modulation*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/137812/

Version: Accepted Version

White Rose
university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

An Adaptive Optimal Mapping Selection Algorithm for PNC using Variable QAM Modulation

Tong Peng, Yi Wang, Alister G. Burr and Mohammad Shikh-Bahaei

Abstract—Fifth generation (5G) wireless networks will need to serve much higher user densities than existing 4G networks, and will therefore require an enhanced radio access network (RAN) infrastructure. Physical layer network coding (PNC) has been shown to enable such high densities with much lower backhaul load than approaches such as Cloud-RAN and coordinated multipoint (CoMP). In this letter, we present an engineering applicable PNC scheme which allows different cooperating users to use different modulation schemes, according to the relative strength of their channels to a given access point. This is in contrast with compute-and-forward and previous PNC schemes which are designed for the two-way relay channel. A two-stage search algorithm to identify the optimum PNC mappings for given channel state information and modulation is proposed in this letter. Numerical results show that the proposed scheme achieves low bit error rate with reduced backhaul load.

Index Terms—adaptive PNC, industrial applicable, backhaul load, unambiguous detection.

I. Introduction

The concept of network multiple input, multiple output (N-MIMO) [1] has been known for some time as a means to overcome the inter-cell interference in fifth generation (5G) dense cellular networks, by allowing multiple access points (APs) to cooperate to serve multiple mobile terminals (MTs). This was implemented in the coordinated multipoint (CoMP) approach standardized in Long Term Evolution (LTE)-A, and the Cloud Radio Access Network (C-RAN) concept has been proposed to achieve similar goals [2]. However these approaches result in large loads on the backhaul network (also referred to as fronthaul in C-RAN) between APs and the central processing unit (CPU), many times the total user data rate.

While there has been previous work addressing backhaul load reduction in CoMP and C-RAN, using, for example, Wyner-Ziv compression [3] or compressive sensing [4], the resulting total backhaul load remains typically several times the total users' data rate. A novel approach was introduced in [5, 6], based on physical layer network coding (PNC) and its more general form called lattice network coding [7], which reduces the total backhaul load to a level in the range of the total users' data rate. However the research on PNC has mostly focused on two-way relay channel (TWRC) application [8, 9] or lattice code-based PNC design [10], which have significant disadvantages in terms of engineering applicability in N-MIMO networks. Moreover all these schemes have assumed that the same modulation order is used by all MTs sharing an
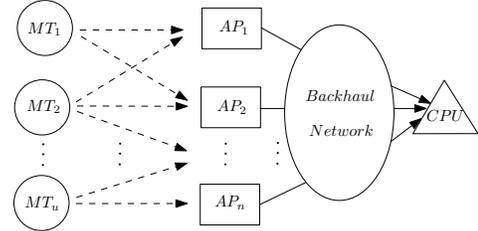
Fig. 1. The uplink system diagram.

AP, and that the signal strength at an AP is the same from all the MTs.

Different from our previous work in [12], in which a PNC scheme for MTs using a quadrature amplitude modulation (QAM) with the same modulation order and equal transmission power is presented, we consider a more practical scenario in which the MTs may transmit with different powers in this letter. In this case, different modulation schemes are employed at each MT to maintain the required bit error rate (BER) performance. In the simulations, a comparison of the proposed algorithm with the CoMP approaches is given, along with the impact of the estimated channel coefficients on the accuracy of optimal matrix selection. The primary contributions of this letter are summarized as follows: (1) an adaptive mapping selection algorithm when a different QAM scheme is employed at each MT; (2) binary mapping matrices are employed so that the proposed algorithm can be implemented in current practical systems to achieve engineering applicability; (3) the dimension of mapping matrices is minimised to achieve lower computational complexity and higher degree of freedom.

II. System Model

A two-stage uplink system model for N-MIMO is illustrated in Fig.1, where u MTs are served by n APs, and the APs are connected to a CPU via a backhaul network. A single antenna is provided at all MTs and APs. The first stage in the uplink is called the multi-access stage, in which all MTs broadcast symbols to APs at the same time. Each AP maps the superimposed signal to a PNC codeword vector and forwards it to the CPU via a backhaul network. The multi-access stage is assumed to use wireless communications, whilst the backhaul link is a lossless but capacity-limited 'bit-pipe'. Note that the backhaul link can be wireless or fixed and the proposed scheme can be implemented in both cases.

A 2^m-QAM scheme is employed at each MT with a modulation function of $\mathcal{M} : \mathbb{F}_{2^m} \longrightarrow \Omega$, where m stands for a modulation order, in bits per symbol, and Ω stands for the set of all possible modulated symbols. In the multi-access stage, u_j MTs are served by an AP where the MTs transmit

their signals simultaneously, yielding a received signal at the jth AP of

$$r_j = \sum_{i=1}^{u_j} h_{j,i} s_i + z_j, \quad \text{for } j = 1, 2, \cdots, n, \qquad (1)$$

where $h_{j,i}$ denotes the complex Gaussian distributed random channel coefficient between the jth AP and the ith MT, and $s_i = \mathscr{M}(\mathbf{b}_i)$ is the 2^m-QAM signal, where $\mathbf{b}_i$ denotes the binary data vector with a dimension of $m \times 1$. z_j denotes the additive white Gaussian noise with zero mean and variance σ^2.

For the purpose of industrial applicability, binary mapping matrices are employed in the proposed PNC encoder and decoder. Define a mapping function $\mathbf{G}_j$ at the jth AP, the PNC encoding is then given by

$$\mathbf{x}_j = \mathbf{G}_j \otimes \mathbf{b}, \qquad (2)$$

where $\mathbf{b} \triangleq [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_{u_j}]^T$ denotes the $m_s \times 1$ binary joint message vector with $\mathbf{b} \in \mathbb{F}_2^{m_s \times 1}$ and $m_s = \sum_{i=1}^{u_j} m_i$, where m_i denotes the modulation order at the ith MT. $\mathbf{G}_j$ denotes a binary matrix with dimensions of $l_j \times m_s$, where l_j denotes the number of rows of the mapping matrix used at the jth AP, and $\otimes$ denotes matrix multiplication over $\mathbb{F}_2$. $\mathbf{x}_j \in \mathbb{F}_2^{l_j \times 1}$ is the network codeword vector (NCV) detected by the AP which consists of l_j linear combinations of the original binary data.

We define $\mathbf{s}_{c_j}$ as a set which contains all possible combinations of the modulated signals from the MTs, where u_j denotes the number of MTs served by the jth AP. Given a $1 \times u_j$ channel coefficient vector $\mathbf{h}_j \triangleq [h_{j,1}, h_{j,2}, \cdots, h_{j,u_j}]$ at the jth AP, the vector containing all $u_j^{m_s}$ possible superimposed signals can be calculated by

$$\mathbf{s}_{j,\text{sc}} \triangleq [s_{j,\text{sc}}^{(1)}, s_{j,\text{sc}}^{(2)}, \cdots, s_{j,\text{sc}}^{(u_j^{m_s})}] = \mathbf{h}_j \mathbf{s}_{c_j}, \quad \text{for } j = 1, 2, \cdots, n. \qquad (3)$$

According to the definition, the modulation function $\mathscr{M}(\cdot)$ is a one-to-one bijective mapping function which maps all possible combinations of binary data to the complex symbols $s \in \Omega$. In (2), the joint message vector $\mathbf{b}$ can be mapped to an NCV $\mathbf{x}$ by a binary matrix $\mathbf{G}$. Then with the expression in (3), we can always find a surjective PNC mapping function $\mathscr{N}_j$ that maps the superimposed constellation point $s_{j,\text{sc}}^{(k)}$ to an NCV, mathematically given by

$$\mathscr{N}_j(s_{j,\text{sc}}^{(k)}) = \mathbf{x}_j, \quad \text{for } k = 1, 2, \cdots, u_j^{m_s}. \qquad (4)$$

At each AP, an estimator calculates the conditional probability of each possible NCV given the mapping function $\mathscr{N}_j$ and the channel coefficients. The estimator returns the log-likelihood ratio (LLR) of each bit of $\mathbf{x}_j$ which is then applied to a soft decision decoder. In the backhaul stage, the NCVs will be forwarded to the CPU from each AP, and then by concatenating them and multiplying by the inverse of the binary PNC mapping matrix, the original data from each MT will be recovered. Note that the global mapping matrix $\mathbf{G} \triangleq [\mathbf{G}_1 \ \mathbf{G}_2 \ \cdots \ \mathbf{G}_{u_j}]^T$ must be non-singular in order to unambiguously recover the source data [12]. Maximum

likelihood (ML) detection can also be used for PNC decoding.

III. DESIGN CRITERION FOR BINARY PNC

Singular fading in the multi-access stage is a serious problem which affects detection performance at the CPU. We give a simple example here with 2 MTs to illustrate this problem and the singular fading is defined as a situation in which different pairs of transmitted signals cannot be distinguished at the receiver, mathematically given by

$$h_{j,1} s_1 + h_{j,2} s_2 = h_{j,1} s_1' + h_{j,2} s_2', \qquad (5)$$

where s_i and s_i' stand for the QAM signals at the ith MT, and $s_i \neq s_i'$. The special channel coefficient vector $\mathbf{h}_{\text{sf}} \triangleq [h_{j,1}, h_{j,2}]$ is defined as a singular fade state (SFS). Note that the solution of (5) is not unique, since there is more than one SFS for each QAM scheme. (5) implies that superimposed constellation points corresponding to two different MT data combinations coincide, and hence this data cannot be unambiguously decoded at this AP.

Besides the coincident symbols, there is a set of superimposed symbols that would map to the same NCV; this set is defined as a *cluster* and denoted by $\mathbf{s}_{\text{cl}} = [s_{j,\text{sc}}^{(1)}, s_{j,\text{sc}}^{(2)}, \cdots]$. We then define the minimum distance between different clusters as

$$d_{\min} = \min_{\mathscr{N}_j(s_{j,\text{sc}}^{(\tau_i)}) \neq \mathscr{N}_j(s_{j,\text{sc}}^{(\tau_k')})} |s_{j,\text{sc}}^{(\tau_i)} - s_{j,\text{sc}}^{(\tau_k')}|^2, \qquad (6)$$

$$\forall s_{j,\text{sc}}^{(\tau_i)} \in \mathbf{s}_{\text{cl}}^{(\tau)}, \ \forall s_{j,\text{sc}}^{(\tau_k')} \in \mathbf{s}_{\text{cl}}^{(\tau')}, \ \text{for } i = 1, 2, ..., \ k = 1, 2, ...$$

In the coincident symbols cases $d_{\min} = 0$. The design criterion is to employ a mapping function that labels the constellation points within a clash to the same NCV with maximised $d_{\min}$ in order to achieve unambiguous decoding at the CPU. A detailed mathematical proof is derived in [12] and the design criterion holds in multiple-MT case.

IV. ADAPTIVE MAPPING FUNCTION SELECTION ALGORITHM

In this section, we describe an adaptive PNC mapping matrix selection algorithm based on the design criterion introduced in the previous section. According to the criterion, the separate mapping matrices used at each AP should encode the constellation points within one cluster to the same NCV and additionally, the global mapping matrix $\mathbf{G}$ formed by the concatenation of the selected mapping matrices at each AP should be invertible for unambiguous decoding at the CPU. We define an $L \times u_j$ matrix $\mathbf{H}_{\text{sfs}} = [\mathbf{h}_{\text{sfs}_1}, \cdots, \mathbf{h}_{\text{sfs}_L}]^T$ whose rows contain all special channel vectors that cause different SFSs for u_j modulation scheme combinations. Like the work in [6], our proposed algorithm comprises two procedures, the first of which is an Off-line search and the second is an On-line search. The proposed Off-line and On-line algorithms are described in Algorithm 1.

The main differences between the proposed design and the work in [6] and [12] are as follows. In [6], MTs employ the same modulation scheme and the channels have the same average path loss, and the dimension of the mapping matrices stored at each AP is equal to $m_s/2$ so that the sizes of the NCVs at each AP are the same. In [12], a design guideline

Algorithm 1 Binary Matrices Selection (Off/On-line Search Algorithm)

Off-line Search
1: Define $\mathbf{b}_s = [\mathbf{b}_{s_1} \cdots \mathbf{b}_{s_u}]$ as a $u^{m_s} \times m_s$ matrix containing all possible combinations of binary data
2: Define $\mathbf{G}_{\text{ofl}}$ as a set that contains all $l \times m_s$ binary matrices for $l \in [1, m_s]$, where the number of matrices is N_{ofl}
3: $\mathbf{s}_c = [M_1(\mathbf{b}_{s_1}), \cdots, M_2(\mathbf{b}_{s_u})]^T$
4: **for** $i = 1 : L$ **do**   ▷ each SFS
5:     $\mathbf{s}_{i,\text{sfs}} \triangleq [s_{i,\text{sfs}}^{(1)}, \cdots, s_{i,\text{sfs}}^{(u^{m_s})}] = \mathbf{h}_{\text{sfs}_i} \mathbf{s}_c$
6:     $d = |s_{i,\text{sfs}}^{(n_1)} - s_{i,\text{sfs}}^{(n_2)}|^2, \forall n_1, n_2, n_1 \neq n_2$
7:     Find N_{cl} different clashes $\mathbf{s}_{\text{cl}}$ containing the $s_{i,\text{sfs}}$ with $d = 0$ and the corresponding binary data combinations $\mathbf{b}_{\text{cl}}$
8:     **for** $n_{\text{ofl}} = 1 : N_{\text{ofl}}$ **do**   ▷ each binary matrix
9:         **for** $n_{\text{cl}} = 1 : N_{\text{cl}}$ **do**   ▷ each clash
10:            $\mathbf{x}_{\text{cl}}^{(n_{\text{cl}})} = \mathbf{G}_{\text{ofl}}^{(n_{\text{ofl}})} \otimes \mathbf{b}_{\text{cl}}^{(n_{\text{cl}})}$ ▷ NCV calculation
11:        **end for**
12:        **if** all $\mathbf{x}_{\text{cl}}^{(n_{\text{cl}})}$ are the same **then**
13:            Store $\mathbf{G}_{\text{ofl}}^{(n_{\text{ofl}})}$ as a matrix candidate
14:        **end if**
15:     **end for**
16: **end for**
On-line Search
17: Define $\mathbf{G}_{\text{onl}}$ as the set that contains the N_{sel} selected mapping matrix candidates
18: **for** $j = 1 : J$ **do**   ▷ each AP
19:     $\mathbf{s}_{j,\text{sc}} = \mathbf{h}_j \mathbf{s}_c$ ▷ All possible superimposed signals
20:     **for** $n_{\text{sel}} = 1 : N_{\text{sel}}$ **do**   ▷ each selected matrix
21:         $\mathbf{x}_{\text{sel}}^{(n_{\text{sel}})} = \mathbf{G}_{\text{onl}}^{(n_{\text{sel}})} \otimes \mathbf{b}_s$ ▷ NCV calculation
22:         Find the set of clusters from the positions of repeated columns in $\mathbf{x}_{\text{sel}}^{(n_{\text{sel}})}$, and among all pairs of clusters:
23:         $d_{\min} = \min_{\forall i,k,i\neq k} |s_{j,\text{sc}}^{(\text{cl}_i)} - s_{j,\text{sc}}^{(\text{cl}_k)}|^2,$
         where cl_i and cl_k denote the cluster indexes
24:     **end for**
25: **end for**
26: The non-singular global mapping matrix selection: $\mathbf{G} = [\mathbf{G}_1 \cdots \mathbf{G}_J]^T$ with maximum $d_{\min}$

Fig. 2. BER of the proposed algorithm with 3dB difference in channel quality.

for an arbitrary number of cooperating users employing the same QAM scheme and transmission power is presented. The approach presented in this paper focuses on the situation in which the modulated signals from MTs are received at an AP with different powers, then in order to maintain a good performance in terms of error probability, different modulation schemes should be used. In order to achieve a higher degree of freedom, mapping matrices with different dimensions may be utilised. The non-singular global mapping matrix used for PNC decoding at the CPU is formed by concatenating the mapping matrices used at each AP with the maximum $d_{\min}$. When the channel qualities between the MTs and the jth AP are poor, the value of $d_{\min}$ may be much smaller than that at the other APs, and this leads to performance degradation. Our solution is to employ $l_j \times m_s$ mapping matrices at the jth AP with $l_j \in [1, m_s]$. An $l \times m_s$ matrix with $l > m_s/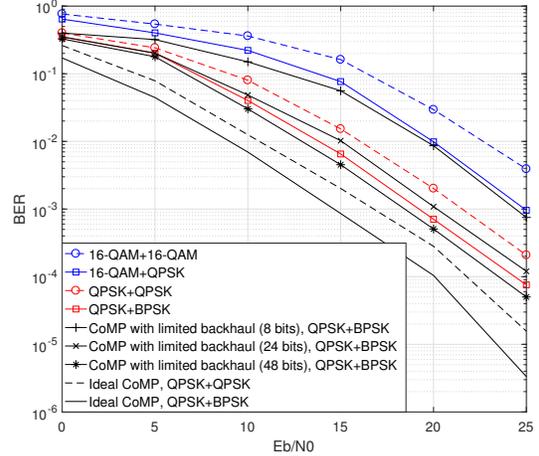2$ at one AP and $l < m_s/2$ at the other may give a greater overall minimum inter-cluster distance $d_{\min}$ over both APs than choosing $l = m_s/2$ at both.

Our proposed algorithm provides a solution to reduce backhaul load in N-MIMO networks with a high user density by employing PNC technique. However, the calculation of singular fading with multiple MTs in PNC is still an open question in binary systems. In order to apply the proposed algorithm in practical scenarios, orthogonal frequency division multiplexing (OFDM) [15] could be utilised by allocating a pair of MTs to the same frequency in multiple access stage. In [13], the proposed algorithm in an OFDM scheme using multiple Universal Software Radio Peripherals (USRPs) is presented. We focused on network design and hardware implementation, such as work in [11], and studies of how to address the synchronisation, channel estimation and interference issues in the perspective of hardware design.

According to [12], the number of SFSs increases exponentially with increasing modulation order index m which leads to an increase in the number of matrices returned by the Off-line search and makes the On-line search algorithm impractical. However, the 2^m-QAM schemes are all defined in $\mathbb{F}_{2^m}$ and the same constellation points can be found in different constellation books. This property allows us to use the same mapping matrix to resolve different SFSs, and the number of mapping matrices being stored reduces dramatically. We term SFSs which can be resolved by the same matrix as another SFS, *image SFSs*. Then the original number of SFSs in quadrature phase shift keying (QPSK) is 13 and that in 16-QAM is 389 but the number is reduced to 5 in QPSK and 169 in 16-QAM after the image SFSs are removed. A detailed study of the impact to the network performance with fewer SFSs is derived in [12] with computational complexity analysis.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we illustrate the simulation results of the proposed algorithm in the 5-node system shown in Fig. 1. We assume the multi-access links are wireless and the backhaul
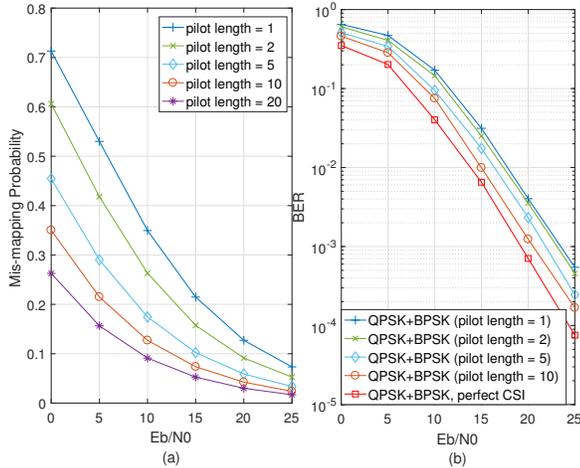
Fig. 3. Non-optimal mapping matrices selection probability with estimated channel.

is a capacity-limited wired link. For simulation simplicity, two MTs are served by two APs and we assume a 3dB difference in average path loss between each AP and the two MTs. The MT with better channel quality employs a higher order modulation scheme, corresponding to a higher rate. A convolutional code is used in the simulation while other channel codes, such as low-density parity-check (LDPC) code [14], may also be employed. As benchmarks we use ideal CoMP with an unrestricted number of bits transmitted in the backhaul network and non-ideal CoMP in which the soft information on the backhaul is quantized to a total of 12, 6 and 2 bits per symbol, resulting in the total backhaul load of 48, 24 and 8 respectively (since 4 LLRs are calculated at each AP).

In Fig. 2, the BER performances of different approaches are illustrated. We assume block fading in the multi-access stage and that perfect channel information can be obtained to obtain accurate SFSs. As shown in the figure, a 2.5dB and 3dB improvement can be observed when employing the proposed algorithm with QPSK+BPSK and 16-QAM+QPSK compared to employing the algorithm in [12] using QPSK+QPSK and 16-QAM+16-QAM, respectively. The ideal CoMP is the best among the approaches due to the availability of unrestricted bits in the backhaul network; while quantization with limited precision results in a performance degradation in non-ideal CoMP. In the QPSK+BPSK case, the proposed algorithm outperforms CoMP with 24 bits, while requiring a total of only 9 backhaul bits.

The channel coefficients obtained at each AP are important to the proposed algorithm because its performance relies on the accurate identification of the closest SFS. In practical communication networks, the trade-off between the length of information symbols and that of pilot symbols for channel estimation has been discussed in [15]. We have illustrated the effect of channel estimation on the proposed algorithm in terms of the probability of selecting a non-optimal matrix, which is shown in Fig. 3. In the simulation, different numbers of pilot symbols have been employed to estimate the

channel, which affects the accuracy of identification of the SFS calculated at each AP. As illustrated in Fig. 3(a), when a short pilot sequence is employed, there is a maximum of 70% probability that the optimal mapping matrix will not be selected in a low Eb/N0 scenario which corresponds to a 5dB degradation compared to that when perfect channel knowledge is employed. With a pilot sequence length of 5, and at Eb/N0s sufficient to ensure a low BER, the mis-mapping probability drops below 10% and the performance degradation reduces to 3dB. According to Fig. 3, a pilot sequence with a length between $5-10$ is enough for the proposed adaptive selection algorithm.

VI. Conclusion

In this letter, an engineering applicable approach is presented to implement adaptive PNC in binary N-MIMO systems when different MTs transmit with different powers. Unlike our work in [12] and [13], we focus on applying the proposed algorithm in practical scenarios in which multiple MTs could employ different modulation schemes. We also illustrate how the estimated channel affects the optimal matrix selection accuracy. The simulation results illustrate the benefits of the proposed PNC approach in terms of backhaul load reduction and BER improvement compared to the non-ideal CoMP with quantized bits in the backhaul channel.

References

[1] M. V. Clark, *et al.*, "Distributed versus centralized antenna arrays in broadband wireless networks,"*Proc. 2001 Spring IEEE Vehicular Technology Conference.*, pp. 33-37, Rhodes Island, Greece, May 2001.
[2] D. Lee, *et al.*, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges, "*IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148-155, February 2012.
[3] L. Zhou and W. Yu, "Uplink Multicell Processing with Limited Backhaul via Per-Base-Station Successive Interference Cancellation,"*IEEE JSAC*, vol. 31, no. 10, pp. 1981-1993, Oct. 2013.
[4] Y. Wang, *et al.*, "Compressive sensing for uplink cloud radio access network with limited backhaul capacity, "*2015 4th ICCSNT*, pp. 898-902, Harbin, China, Dec. 2015.
[5] Q. T. Sun, *et al.*, "Lattice network codes based on Eisenstien integers, "*IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2713-2725, Jul. 2013.
[6] D. Fang and A. G. Burr, "Linear Physical-Layer Network Coding for 5G Radio Access Networks , "*2014 1st International Conference on 5GU*, pp. 116-221, Akaslompolo, Finland, Nov. 2014.
[7] Y. Wang, *et al.*, "A Multilevel Framework to Lattice Network Coding, "*IEEE Infor. Theory*, under review. Available at arXiv: https://arxiv.org/abs/1511.03297.
[8] T. Koike-Akino, *et al.*, "Optimized constellations for two-way wireless relaying with physical network coding, "*IEEE J. Sel. Areas Commun*, vol. 27, no. 5, pp. 773-787, June 2009.
[9] D. Fang and A. G. Burr, "Uplink of Distributed MIMO: Wireless Network Coding versus Coordinated Multipoint, "*IEEE Comm. Letters*, vol. 19, no. 7, pp. 1229-1232, Jul. 2015.
[10] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes, "*IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6463-6486, Oct. 2011.
[11] G. Chiurco, *et al.*, "FPGA Design and Performance Evaluation of a Pulse-Based Echo Canceller for DVB-T/H, "*IEEE Transactions on Broadcasting*, vol. 58, no. 4, pp. 660-668, Dec. 2012.
[12] T. Peng, *et al.*, "Adaptive Wireless Network Coding in Network MIMO: A New Design for 5G and After, "*IEEE Trans. on Commun.*, under review. Available at arXiv: https://arxiv.org/abs/1801.07061.
[13] Y. Chu, *et al.*, "Implementation of Uplink Network Coded Modulation for Two-Hop Networks, "*IEEE Access*, under review. Available at arXiv: https://arxiv.org/abs/1808.07354.
[14] A. G. Burr, "*Modulation and Coding for Wireless Communications*,"Prentice-Hall, 2000.
[15] H. D. Taun, *et al.*, "Optimized Training Sequences for Spatially Correlated MIMO-OFDM,"*IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2768-2778, Sept. 2011.