Article:
Peng, Tong, Wang, Yi, Burr, Alister Graham orcid.org/0000-0001-6435-3962 et al. (1 more author) (2018) Physical layer network coding in network MIMO : a new design for 5G and beyond. IEEE Transactions on Communications. 2024 - 2035. ISSN 0090-6778

https://doi.org/10.1109/TCOMM.2018.2886184

Physical Layer Network Coding in Network MIMO: A New Design for 5G and Beyond

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

***Abstract*—Physical layer network coding (PNC) has been studied to serve wireless network MIMO systems with much lower backhaul load than approaches such as Cloud Radio Access Network (Cloud-RAN) and coordinated multipoint (CoMP). In this paper, we present a design guideline of engineering applicable PNC to fulfil the request of high user densities in 5G wireless RAN infrastructure. We show that the proposed design criteria guarantee that 1) the whole system operates over binary system; 2) the PNC functions utilised at each access point overcome all singular fade states; 3) the destination can unambiguously recover all source messages while the overall backhaul load remains at the lowest level. We then develop a two-stage search algorithm to identify the optimum PNC mapping functions which greatly reduces the real-time computational complexity. The impact of estimated channel information and reduced number of singular fade states in different QAM modulation schemes are studied in this paper. Numerical results show that the proposed schemes achieve low outage probability with reduced backhaul load.**

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

I. Introduction

Over the past few years, network multiple-input multiple-output (N-MIMO) technique [1]-[3] has received significant attention due to its flexibility, power and capacity advantage over the centralized architectures in fifth generation (5G) dense cellular networks. Multiple mobile terminals (MTs) may share the same radio resources and be served by the corresponding access points (APs) where these multiple single-antenna-equipped MTs form one visual transmitter with multiple antennas, and the inter-cell interference can be effectively mitigated due to the cooperative communications in such N-MIMO systems. This was also applied in the coordinated multipoint (CoMP) approach standardized in LTE-A [4, 5]. The Cloud Radio Access Network (C-RAN) concept has been proposed in [2] with similar goals. However a potential issue in these approaches is a significant increased upload burden on the backhaul (referred to as fronthaul in C-RAN) network between APs and central processing unit (CPU) on the uplink, especially for wireless backhaul networks. Several methods have been studied in previous work in order to address this problem. The authors in [6] proposed an algorithm to achieve the optimal sum rate in backhaul network by compressing and decompressing the received message at each AP node. Iterative interference cancellation and compressive sensing algorithms are designed in [7]-[9] as alternative solutions, but the total backhaul load remains typically several times the total user data rate.

PNC is a scheme implemented at APs in which each AP attempts to infer and forward combinations of the superimposed signals transmitted from multiple sources simultaneously over an algebraic field. An important property of PNC is that the APs decode the joint messages from multiple sources to a linear function over the algebraic field rather than decode each source symbol individually. In [10, 11], novel approaches have been designed based on physical layer network coding (PNC) that keep the total backhaul load equal to the total user data rate.

The original PNC was proposed and designed in a two-way relay channel (TWRC) based on BPSK [12] and PNC contributes to network throughput improvement and motivates many research outcomes thereafter, e.g. compute-and-forward (C&F) which generalizes PNC of TWRC to multiuser relay networks by utilizing structured nested lattice codes [13], and lattice network coding [14]. However, the lattice based network coding operates over a finite field and the coset size of the quotient lattices is typically not binary-based [10]. Even a low-complexity

T. Peng was with the Department of Electronics, University of York, and now is with the Centre for Telecommunications Research, Department of Informatics, King's college London, UK (e-mail: tong.peng@york.ac.uk, tong.peng@kcl.ac.uk). Y. Wang and A. G. Burr are with the Department of Electronics, University of York, UK (e-mails: yi.wang@york.ac.uk, alister.burr@york.ac.uk). M. Shikh-Bahaei is with the Centre for Telecommunications Research, Department of Informatics, King's college London, UK (e-mail: m.sbahaei@kcl.ac.uk).

This research is funded by EPSRC NetCoM project EP/K040006/1 and partially by EPSRC IoSIRE project EP/P022723/1.

lattice based PNC is designed in [14] with channel alignment, the lattice codes have disadvantages for engineering applications as non-binary codes are required over large prime fields.

In contrast to the previous work in PNC, we focus on designing the PNC approach with conventional 2^m-ary digital modulation. When QAM modulation schemes are used at MTs, PNC has to solve the so-called singular fading problem which is typically unavoidable at the multiple access phase under some circumstances at each AP. Failure to resolve such problem results in network performance degradation. The work in [15] provides a solution for implementing PNC in binary systems with low modulation orders only. Toshiaki *et. al.* [16, 17] proposes a scheme, namely the denoise-and-forward, which employs a non-linear 5QAM PNC mapping to mitigate all singular fade states, and gives good performance. Other researches on this issue have worked on the design of linear functions over the integer finite field or ring, e.g. linear PNC (LPNC) [18]-[20] which can only be optimised for the q-ary PNC mapping where q is a prime in $\mathbb{Z}^+$. All these approaches, however, do not operate over the binary systems, and hence cannot be readily applied in the current mobile communication networks. A linear PNC designed for PAM schemes is proposed in [21] but the authors focuses on TWRC scenarios. A network coding approach combining PNC and space-time coding (STC) is proposed in [22] for uplink scenario in N-MIMO. In this design, the buffers at each AP is utilised to maximum the sum rate and minimise the errors but the redundancy of STC symbols will increase the backhaul load.

In this paper, we propose a design guideline of PNC with reduced backhaul load and unambiguous decoding for QAM modulation schemes. This work is based on our previous work in [24] and [25]. In [24], a binary PNC scheme for a uplink scenario with 2 MTs, 2 APs and 1 CPU is proposed. Different transmit power and modulation schemes are employed at each MT and we proposed a PNC scheme in a simple 5-node N-MIMO system. In [25], the algorithm in [24] is implemented and tested with OFDM scheme using multiple Universal Software Radio Peripherals (USRPs). We focused on hardware implementation and studies of how to address the synchronisation and interference issue of the proposed algorithm in OFDM systems in [25]. The main contributions are listed as follows

1) A PNC design guideline for uplink scenarios is proposed, along with a search algorithm based on this guideline to find the optimal coefficient mapping matrices, such that a) the global matrix, formed by the coefficient matrices from each AP, guarantees all source symbols to be decoded at CPU; b) the minimum number of matrices are stored at each AP and the matrices can resolve all singular fade states; c) the proposed algorithm can be applied in N-MIMO systems with multiple MTs and APs and is generalised to QAM modulation schemes of different orders.
2) The whole scheme operates over binary systems. As discussed earlier, PNC plays as a reliable role not only in TWRC but also in RAN to serve multiple MTs. In this paper, we investigate the design criteria of engineering applicable PNC over binary systems for an uplink scenario of 5G N-MIMO system and discuss how to address the singular fading problem with multiple MTs.
3) A regulated PNC approach which fulfils the low latency demand in practical networks is presented. This approach is developed based on the proposed search algorithm with a lookup table mechanism to achieve low latency in mapping selection and decoding.
4) The impact of practical issues to performance degradation of the proposed algorithm and potential corresponding solutions are studied. We also investigated the real-time computational complexity and backhaul load reduction achieved by the proposed algorithm in this paper.

The rest of this paper is organised as follows. The introduction of N-MIMO systems and definitions of PNC design criteria are given in Section II and III, respectively. The proposed binary matrix adaptive selection algorithm is derived in Section IV, followed by the discussion of methods to reduce the computational complexity of the proposed algorithm in Section V. Numerical results are given in Section VI and finally the conclusions are drawn in Section VII.

II. System Model

A two-stage uplink model of N-MIMO system is illustrated in Fig. 1. We assume MTs and APs

are all equipped with a single antenna to form N-MIMO networks. At the first stage, u MTs transmit symbols to n APs during the same period, which refers to a multi-access stage. We have studied the effect of synchronisation errors in [25] so in this paper we assume the synchronisation is perfect. Each AP receives data from all MTs and then infers and forwards a linear combination (which is referred to as the network coded symbols (NCS) in this paper) of the entire messages over a finite field or ring. The second stage is called backhaul stage where n APs forward the NCSs to CPU. In this paper, the links in multi-access stage are modelled as wireless links in order to fulfil the request of 5G systems; while the backhaul network is formed by a lossless but capacity limited 'bit-pipe'. The techniques presented in this paper are in particular suitable for wireless backhaul which is normally more cost-effective.

Each MT employs a 2^m-ary digital modulation scheme where m denotes the modulation order. Let $\mathscr{M} : \mathbb{F}_{2^m} \longrightarrow \Omega$ denotes a one-to-one mapping function, where Ω is the set of all possible complex constellation points. Hence the modulation at the ℓ^{th} MT using a 2^m-QAM scheme is to map the messages $\mathbf{w}_\ell \in \mathbb{F}_{2^m}$ to the complex symbol s_ℓ, given by $s_\ell = \mathscr{M}(\mathbf{w}_\ell)$, where $\mathbf{w}_\ell = [w_\ell^{(1)}, \cdots, w_\ell^{(m)}]$ is an m-tuple with each element $w_\ell^{(i)} \in \mathbb{F}_2$.

The link between all MTs and the j^{th} AP forms a multiple access channel (MAC), where the j^{th} AP observes the noisy, faded and superimposed signals at a certain time slot, mathematically given by

$$y_j = \sum_{\ell=1}^{u} h_{j,\ell} s_\ell + z_j, \tag{1}$$

where z_j denotes the additive complex Gaussian noise with zero mean and variance σ^2, and $h_{j,\ell}$ represents the channel coefficient between the ℓ^{th} MT and the j^{th} AP, which is a random variable with Rayleigh distribution.

III. Design Criteria

Before presenting the proposed PNC design criteria, two main constraints for a PNC to be engineering applicable is listed as follow: 1) the PNC decoding must operate over $\mathbb{F}_2$; 2) the PNC mapping function is capable to resolve all singular fade states and achieve unambiguous recovery of all source messages at the CPU. We give details of the proposed design criteria for PNC in N-MIMO systems in this section, which relaxes above aforementioned constraints.

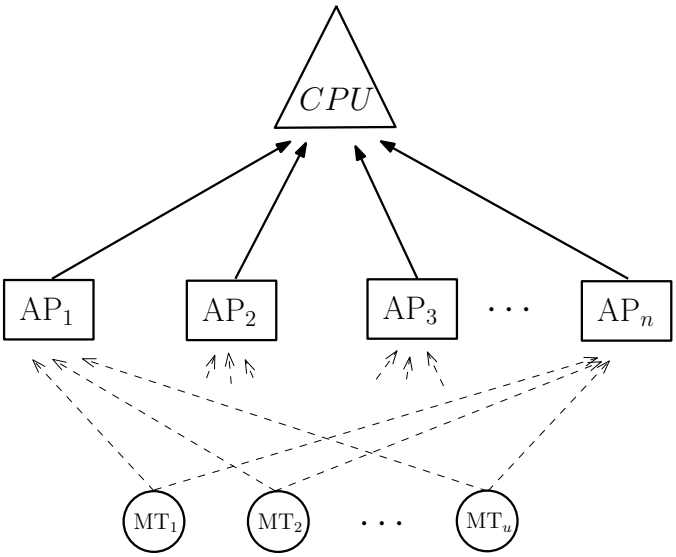

Fig. 1. The uplink system diagram.

A. Engineering Applicable PNC Function

We are primarily concerned with the MAC phase between u MTs and the j^{th} AP in the design of the PNC mapping function.Instead of using PNC approaches performing linear combinations on symbol level, such as [18], we design a method to encode PNC directly at the bit level which allows the APs to operate over a binary field for industrial application.

Definition 1: The bit-level linear PNC mapping function of the j^{th} AP for u MTs is defined as $\mathscr{N}_j(.)$ and the PNC encoding is mathematically expressed as

$$\mathbf{v}_j = \mathscr{N}_j(\mathbf{M}_j, \mathbf{w}) = \mathbf{M}_j \otimes \mathbf{w}, \tag{2}$$

where $\mathbf{w} \triangleq [\mathbf{w}_1, \cdots, \mathbf{w}_u]^T$ denotes an $mu \times 1$ joint message set with $\mathbf{w} \in \mathbb{F}_2^{mu \times 1}$, and each $\mathbf{w}_\ell$ stands for a $1 \times m$ binary data vector at the ℓ^{th} MT. $\mathbf{M}_j$ denotes a matrix with $\mathbb{F}_2^{t \times mu}$, where t stands for the size of network coded vector at the j^{th} AP and $t \geq m$, and $\otimes$ denotes the multiplication over $\mathbb{F}_2$. $\mathbf{v}_j \in \mathbb{F}_2^{t \times 1}$ is called the network coded vector (NCV) which consists of all t linear network coded bits $\mathbf{v}_j = [v_j^{(1)}, v_j^{(2)}, \cdots, v_j^{(t)}]^T$.

It is obvious that each coded bit $v_j^{(i)}$ is indeed a linear combination of all source bits over $\mathbb{F}_2$, thus,

$$v_j^{(i)} = M_j^{(i,1)} \otimes w_1^{(1)} \oplus \cdots \oplus M_j^{(i,um)} \otimes w_u^{(m)}, \tag{3}$$

where $\oplus$ denotes the addition operation over $\mathbb{F}_2$, and $M_j^{(i,1)}$ denotes the entry at the i^{th} row and the 1^{st} column of $\mathbf{M}_j$.

Definition 2: We define the constellation set which contains all possible superimposed symbols at the j^{th} AP over a given channel coefficient vector $\mathbf{h}_j \triangleq [h_{j,1}, \cdots, h_{j,u}]$ as $\mathbf{s}_{j,\triangle} \triangleq [s_{j,\triangle}^{(1)}, \cdots, s_{j,\triangle}^{(2^{mu})}]$, where

$$s_{j,\triangle}^{(\tau)} = \sum_{\ell=1}^{u} h_{j,\ell} s_\ell, \quad \forall s_\ell \in \Omega,\ \tau = 1, 2, \cdots, 2^{mu}.$$

Theorem 1: For the MAC link between u MTs and the j^{th} AP, there exists a surjective function

$$\Theta : \mathbf{s}_{j,\triangle} \longrightarrow \mathbf{v}_j, \tag{4}$$

when the size of NCV $t < mu$.

Proof: Since the QAM scheme $\mathscr{M}$ is a bijective function, we have the following relationship

$$\mathbf{v}_j \overset{\mathscr{N}_j}{\Longleftarrow} \mathbf{w} \underset{\mathscr{M}^{-1}}{\overset{\mathscr{M}}{\Longleftrightarrow}} \mathbf{s}, \tag{5}$$

where $\Leftarrow$ and $\Longleftrightarrow$ represent surjective and bijective relationships, respectively. $\mathbf{s} \triangleq [s_1, \cdots, s_u] = [\mathscr{M}(\mathbf{w}_1), \cdots, \mathscr{M}(\mathbf{w}_u)]$ stands for the set that contains the modulated symbols at all MTs. Following (4), for each element in $\mathbf{s}$, there exists a superimposed constellation point $s_{j,\triangle}$ at a given channel coefficient vector $\mathbf{h}_j$, and this proves Theorem 1. ■

We call Θ the PNC mapping function which maps a superimposed constellation point to an NCV and plays the key role in PNC encoding, where this PNC encoding performs estimation of the possible NCV outcomes $\mathbf{v}_j$ for the j^{th} AP, based on the received signals y. Let $\mathbf{V}_j$ denote the vector-based random variable with its realization $\mathbf{v}_j$. The *a posteriori* probability of the event $\mathbf{V}_j = \mathbf{v}_j$ conditioned on the MAC outputs $Y_j = y_j$ is

$$\begin{aligned}
&\Pr(\mathbf{V}_j = \mathbf{v}_j | y_j, \mathbf{h}_j) \\
=&\frac{\Pr(Y_j | \mathbf{V}_j = \mathbf{v}_j, \mathbf{h}_j)\Pr(\mathbf{V}_j = \mathbf{v}_j)}{\Pr(Y_j = y_j)} \\
=&\frac{\sum\limits_{\forall \mathbf{w}: \mathscr{N}_j(\mathbf{M}_j, \mathbf{w}) = \mathbf{v}_j} \Pr(Y_j | \mathbf{w}, \mathbf{h}_j)\Pr(\mathbf{w})}{\Pr(Y_j = y_j)} \\
=&\frac{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{v}_j} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})}{\Pr(Y_j = y_j)}.
\end{aligned} \tag{6}$$

The conditional probability density function is given by

$$\Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|y_j - \mathbf{s}_{j,\triangle}|^2}{2\sigma^2} \right). \tag{7}$$

The *a posteriori* L-value $L_{\mathbf{v}_j}$ for the event $\mathbf{V}_j = \mathbf{v}_j$ is

$$L_{\mathbf{v}_j} = \log \left(\frac{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{v}_j} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})}{\sum\limits_{\forall \mathbf{s}: \Theta(\mathbf{s}_{j,\triangle}) = \mathbf{0}} \Pr(Y_j | \mathbf{S}_{j,\triangle} = \mathbf{s}_{j,\triangle})\Pr(\mathbf{S} = \mathbf{s})} \right), \tag{8}$$

where $\mathbf{0}$ is a length-t all-zero vector over $\mathbb{F}_2^{t \times 1}$.

B. Resolving the Singular Fading

We have set up the PNC mapping approach in binary systems in previous subsection, then the next upcoming problem lies in how to resolve the singular fading in the multiple access stage. In this subsection, we demonstrate that the PNC mapping function Θ proposed above is capable of resolving all singular fade states with a simple design approach. We first define the singular fade states as follows

Definition 3: The singular fade state (SFS) at the j^{th} AP is defined as the channel fading coefficients $\mathbf{h}_j$ which makes $s_{j,\triangle}^{(\tau)} = s_{j,\triangle}^{(\tau')}$ when $\tau \neq \tau'$. □

In other words, for a given channel coefficient vector $\mathbf{h}_j$, if two or more elements in the set $\mathbf{s}_{j,\triangle}$ are the same, $\mathbf{h}_j$ results an SFS. Normally, singular fading is unavoidable at MAC stage, and multiuser detection is in principle infeasible if the j^{th} AP expects to decode all source messages. PNC is capable to overcome SFS problem when the coincident superimposed constellation points are well labelled by the NCV $\mathbf{v}_j$, which helps CPU to recover all source messages.

Definition 4: If a set of constellation points received at APs are mapped to the same NCV, we call this set a *cluster*, denoted as

$$\mathbf{c}^{(\tau)} \triangleq [s_{j,\triangle}^{(\tau_1)}, s_{j,\triangle}^{(\tau_2)}, \cdots], \tag{9}$$

where $s_{j,\triangle}^{(\tau_i)}$ denotes the i^{th} cluster members. In a singular fading, if the values of cluster members are the same then this cluster is called a *clash*. □

Definition 5: Given two clusters $\mathbf{c}^{(\tau)}$ and $\mathbf{c}^{(\tau')}$ that the constellation points in which are mapped

to different NCVs, then the minimum inter-cluster distance, also known as the minimum distance between these two different NCVs, is defined as

$$d_{\min} = \min_{\Theta(s_{j,\triangle}^{(\tau_i)}) \neq \Theta(s_{j,\triangle}^{(\tau_k')})} |s_{j,\triangle}^{(\tau_i)} - s_{j,\triangle}^{(\tau_k')}|^2, \quad (10)$$

$$\forall s_{j,\triangle}^{(\tau_i)} \in \mathbf{c}^{(\tau)},\ \forall s_{j,\triangle}^{(\tau_k')} \in \mathbf{c}^{(\tau')},\ i = 1, 2, ...,\ k = 1, 2, ...$$

Theorem 2: The PNC mapping function Θ cannot resolve singular fading if the minimum inter-cluster distance $d_{\min} = 0$.

Proof: When $d_{\min} = 0$, the posterior probability of some outcomes of $\mathbf{V}_j$ will be very similar (in terms of (6)). This definitely introduces the ambiguities in estimating the real NCV $\mathbf{v}_j$, especially when a superimposed constellation point labelled by one NCV is close to another point that is labelled by another NCV. Hence the singular PNC mapping function is in principle not capable of decoding the NCV reliably. ■

Normally the dimension of NCV $\mathbf{v}_j$ is $t \times 1$ at the j^{th} AP. When the number of source increases (a large MAC), the singular fading problem becomes more severe and the method of SFS values calculation is different. Utilisation of an $\mathbf{M}_j$ with a larger dimension provide a potential solution to this problem. In [16, 17], the authors enlarged the size of encoding matrices to introduce additional codeword cardinalities in order to resolve some special SFS cases in TWRC networks. With mapping the superimposed symbols at the relay node to more codewords, more information will be obtained at each end user for data recovery. In this case, by increasing the size of mapping matrices, there definitely exists non-singular PNC function which is capable of resolving a kind of SFS.

Remark 1: We can obtain non-singular PNC mapping function Θ_j for the j^{th} AP if the cardinality t of the PNC encoding outcomes are determined in terms of the following criterion

$$t = \arg\min_{m \leq t < mu} \{d_{\min} - d_\alpha \geq 0\}, \quad (11)$$

where $d_\alpha > 0$ is a distance threshold.

Remark 1 reveals the second design criterion for PNC mapping function Θ_j over a u-MT and 2^m-ary digital modulation MAC, which guarantees the reliable PNC encoding with the minimum possible cardinality expansion.

C. Algebraic Work for Unambiguous Decodability

The last design criterion is that the CPU can guarantee all source messages to be unambiguously recovered. We need to carefully design each $\mathbf{M}_j$ so that $\mathbf{M} = [\mathbf{M}_1, \cdots, \mathbf{M}_n]^T$ includes a number of row coefficients which forms the following theorem:

Theorem 3: Assume $\mathbf{M} = M_{n\times n}(R)$, where the coefficients are from a commutative ring R. Source messages are drawn from a subset of R and all source messages can be unambiguously decoded at the destination if and only if the determinant of the transfer matrix is a unit in R,

$$\det(\mathbf{M}) = \mathcal{U}(R). \quad (12)$$

Proof: The proof can be found from textbooks [26], but we still list the proof with slight modifications in case that some readers found it easy to understand the research outcomes later. We first prove that (12) gives the sufficient and necessary conditions that make a matrix $\mathbf{M}$ invertible in N-MIMO networks. Suppose $\mathbf{M}$ is invertible, then there exists a matrix $\mathbf{C} \in M_{n\times n}(R)$ such that $\mathbf{MC} = \mathbf{CM} = \mathbf{I}_n$. This implies $1 = \det(\mathbf{I}_n) = \det(\mathbf{MC}) = \det(\mathbf{M})\det(\mathbf{C})$. According to the definition of a unit, we say $\det(\mathbf{M}) \in U(R)$.

We know $\mathbf{M}\cdot\text{adj}(\mathbf{M}) = \text{adj}(\mathbf{M})\cdot\mathbf{M} = \det(\mathbf{M})\mathbf{I}_n$. If $\det(\mathbf{B}) \in U(R)$, we have

$$\begin{aligned}\mathbf{M}\cdot(\det\mathbf{M})^{-1}\text{adj}(\mathbf{M})) &= ((\det(\mathbf{M})^{-1}\text{adj}(\mathbf{M}))\mathbf{M} \\ &= \det(\mathbf{M})^{-1}\det(\mathbf{M}) = \mathrm{I}_n.\end{aligned} \quad (13)$$

Hence, $\mathbf{C} = (\det(\mathbf{M})^{-1}\text{adj}(\mathbf{M}))$ is the inverse of $\mathbf{M}$ since $\mathbf{MC} = \mathbf{CM} = \mathbf{I}_n$.

If $\mathbf{M}$ is invertible, then its inverse $\mathbf{M}^{-1}$ is uniquely determined. Assuming $\mathbf{M}$ has two inverses, say, $\mathbf{C}$ and $\mathbf{C}'$, then

$$\mathbf{M}\cdot\mathbf{C} = \mathbf{C}\cdot\mathbf{M} = \mathbf{I}_n, \quad (14)$$

$$\mathbf{M}\cdot\mathbf{C}' = \mathbf{C}'\cdot\mathbf{M} = \mathbf{I}_n, \quad (15)$$

hence we have

$$\mathbf{C} = \mathbf{C}\cdot\mathbf{I}_n = \mathbf{C}\cdot\mathbf{M}\cdot\mathbf{C}' = \mathbf{I}_n\cdot\mathbf{C}' = \mathbf{C}'. \quad (16)$$

It proves the uniqueness of the invertible matrix $\mathbf{M}$ over R.

Assume $\mathbf{a} \neq \mathbf{a}'$, $\mathbf{M}\cdot\mathbf{a} = \mathbf{F}$, $\mathbf{M}\cdot\mathbf{a}' = \mathbf{F}'$, and $\mathbf{F} = \mathbf{F}'$. This means

$$\mathbf{a} = \mathbf{M}^{-1}\cdot\mathbf{F} = \mathbf{M}^{-1}\cdot\mathbf{F}' = \mathbf{a}'. \quad (17)$$

This contradicts $\mathbf{a} \neq \mathbf{a}'$. Hence it ensures unambiguous decodability:

$$\mathbf{M} \cdot \mathbf{a} \neq \mathbf{M} \cdot \mathbf{a}', \ \forall \mathbf{a} \neq \mathbf{a}'. \tag{18}$$

■

Definition 6: The ideal in R generated by all $\nu \times \nu$ minors of $M_{m\times n}(R)$ is denoted by $I_\nu(M_{m\times n}(R))$, where $\nu = 1, 2, \cdots, r = \min\{m, n\}$. □

A $\nu \times \nu$ minor of $M_{m\times n}(R)$ is the determinant of a $\nu \times \nu$ matrix obtained by deleting $m-\nu$ rows and $n-\nu$ columns. Hence there are $\binom{m}{\nu}\binom{n}{\nu}$ minors of size $\nu \times \nu$. $I_\nu(M_{m\times n}(R))$ is the ideal of R generated by all these minors.

Design Criterion: The destination is able to unambiguously decode u source messages if:

1) $u \geq \max\{\nu \mid \mathrm{Ann}_R(I_\nu(\mathbf{M}_j)) = \langle 0\rangle\}, \ \forall j = 1, 2, \cdots, n,$
2) $\mathbf{M}_j = \arg\max_{\mathbf{M}_j} \{d_{min}\}, \quad d_{min} > 0.$

where $\langle x\rangle$ denotes the ideal generated by x.

Condition 1 can be proved as follows. According to Laplace's theorem, every $(\nu+1)\times(\nu+1)$ minor of $M_{m\times n}(R)$ must lie in $I_\nu(M_{m\times n}(R))$. This suggests an ascending chain of ideals in R:

$$\begin{aligned}\langle 0\rangle = I_{r+1}(\mathbf{M}_j) \subseteq I_r(\mathbf{M}_j) \subseteq \cdots \subseteq I_1(\mathbf{M}_j) \\ \subseteq I_0(\mathbf{M}_j) = R.\end{aligned} \tag{19}$$

Computing the annihilator of each ideal in (19) produces another ascending chain of ideals,

$$\begin{aligned}\langle 0\rangle = \mathrm{Ann}_R(R) \subseteq \mathrm{Ann}_R(I_1(\mathbf{M}_j)) \subseteq \cdots \\ \subseteq \mathrm{Ann}_R(I_r(\mathbf{M}_j)) \subseteq \mathrm{Ann}_R(\langle 0\rangle) = R.\end{aligned} \tag{20}$$

It is obvious that:

$$\begin{aligned}&\mathrm{Ann}_R(I_k(\mathbf{M}_j)) \neq \langle 0\rangle \\ \Rightarrow &\mathrm{Ann}_R(I_{k'}(\mathbf{M}_j)) \neq \langle 0\rangle, \quad \forall k \leq k'.\end{aligned} \tag{21}$$

The maximum value of ν which satisfies $\mathrm{Ann}_R(I_\nu(\mathbf{M}_j)) = \langle 0\rangle$ guarantees that $I_k(\mathbf{M}_j) \in R$, $\forall k < \nu$. Hence we define the rank of $\mathbf{M}_j$ as $\mathrm{rk}(\mathbf{M}_j) = \max\{\nu \mid \mathrm{Ann}_R(I_\nu(\mathbf{M}_j)) = \langle 0\rangle\}$. Suppose that $\mathbf{M}_k \in M_{m\times p}(R)$ and $\mathbf{M}_{k'} \in M_{p\times n}(R)$, then $\mathrm{rk}(\mathbf{M}_k\mathbf{M}_{k'}) \leq \min\{\mathrm{rk}(\mathbf{M}_k), \mathrm{rk}(\mathbf{M}_{k'})\}$, and we can easily prove that $0 \leq \mathrm{rk}(M_{m\times n}(R)) \leq \min\{m, n\}$. Thus, in order to guarantee there are at least u unambiguous linear equations available at the CPU, $\mathrm{rk}(\mathbf{M}_j)$ must be at least u, $\forall j = 1, 2, \cdots, n$.

The special case of condition 1 is that the entry of the coefficient matrix $\mathbf{M}_j \in M_{m\times n}(F)$ is from a finite field $F \in \mathbb{F}$. Then condition 1 of the above *Design Criterion* may be changed to "the maximum number of linearly independent rows (or columns)" since $\mathrm{Ann}_R(I_\nu(\mathbf{M}_j)) = \langle 0\rangle$ if and only if $I_\nu(\mathbf{M}_j) \neq 0$. In other words, the largest ν such that the $\nu \times \nu$ minor of $\mathbf{M}_j$ is a non-zero divisor represents how many reliable linear combinations the j^{th} layer may produce. Hence condition 1 is a strict definition which ensures unambiguous decodability of the u sources.

Condition 2 ensures that the selected coefficient matrix maximises the minimum inter-cluster distance d_{min}, which, in terms of Theorem 2, guarantees the unambiguious decoding. The design criteria provides the key guideline to design and explore the PNC technique applied in the large and complicated cloud network.

IV. Binary Matrix Adaptive Selection Algorithm

According to the design criteria proposed in the previous section, we can summarise that given a QAM scheme, the optimal binary PNC mapping function contains the following properties

1) it maximises the minimum distance between different NCVs ;
2) the composited global mapping matrix is invertible.

In order to achieve these properties and applicability in practical N-MIMO systems, we propose a two-stage binary matrix adaptive selection (BMAS) algorithm. The first stage is called Off-line search, in which an exhaustive search is implemented among all $t \times mu$ binary mapping matrices to find a set of candidate matrices which resolves all SFSs with the above property 1); the second stage is called Online search, in which a selection from the candidate matrices in order to obtain the invertible $mu \times mu$ global mapping matrix according to property 2). The computational complexity of the proposed BMAS algorithm is mainly caused by the Off-line search, especially in higher order modulation schemes due to the increased number of SFSs and enlarged size of matrices. However, this Off-line search algorithm only needs to be done once for each QAM scheme.

A. Off-Line Search Algorithm

An exhaustive search among $t \times mu$ binary matrices is implemented in order to obtain all the possible

solutions for each SFS. The matrices containing a high possibility to form an invertible $mu \times mu$ global mapping matrix when combining with other mapping matrices candidates resolving other SFSs will be selected and stored at the APs and CPU. In this case, the global mapping matrix is capable to resolve a combination of SFSs and recover the original message at the CPU. A brief description is given here and a complete Off-line search is presented in Algorithms 2 and 3 in Appendix A.

Define a set $\mathbf{W}_{joint} \triangleq [\mathbf{w}_{jo_1}, \mathbf{w}_{jo_2}, \cdots, \mathbf{w}_{jo_N}]$ which contains all possible binary joint message combinations with $N = 2^{um}$, so that each $\mathbf{w}_{jo_i}$ in this set stands for a $1 \times mu$ binary joint message vector from u MTs, for $i = 1, 2, \cdots, N$. By applying a QAM scheme $\mathscr{M}$ over each $\mathbf{w}_{jo_i}$ in $\mathbf{W}_{joint}$, a set $\mathbf{S}_{joint} \triangleq [\mathbf{s}_{jo_1}, \mathbf{s}_{jo_2}, \cdots, \mathbf{s}_{jo_N}]$ is obtained, where $\mathbf{s}_{jo_i} = [s_1^{(jo_i)} \ s_2^{(jo_i)} \cdots \ s_u^{(jo_i)}]$ stands for the i^{th} combination of u modulated symbols and $s_\ell^{(jo_i)} \in \Omega$ for $\ell = 1, \cdots, u$. The next step is to calculate the NCS $s_{n,\triangle}^{(q)}$ and its corresponding NCV $\mathbf{v}_{i,n}^{(q)}$ under all L SFS circumstances, mathematically given by

$$s_{n,\triangle}^{(q)} = \mathbf{h}_{\mathbf{v}SFS}^{(q)} \mathbf{s}_{jo_n}^T, \ \mathbf{v}_{i,n}^{(q)} = \mathbf{M}_i \otimes \mathbf{w}_{jo_n}^{(q)}, \quad (22)$$
$$n = 1, \cdots, N, \ q = 1, \cdots, L, \ i = 1, \cdots, N^2,$$

where $\mathbf{h}_{\mathbf{v}SFS}^{(q)}$ denotes the channel coefficient vector causes SFS. Due to SFS, the same $s_{n,\triangle}^{(q)}$ could be obtained with different $\mathbf{s}_{jo_n}$ sets in a clash. In that case, these joint symbol sets should be encoded to the same NCV according to the unambiguous decodability theorem. Then an exhaustive search among all $t \times mu$ binary matrices is implemented to obtain the matrices that is capable to map the joint message combinations in the same groups to the same NCVs. Note that the Condition 1) listed in this section is achieved by mapping the clashed superimposed messages to the same NCV.

B. On-Line Search Algorithm

The proposed Off-line search is implemented before the transmission to minimise the coding delay in real-time transmissions by reducing the number of mapping matrices. The procedure in the proposed On-line search is almost the same as that in the Off-line search, except a rank check of the global mapping matrix in order to fulfil Condition 2) presented at the beginning of this section. The proposed On-line search is applied at the CPU and the results will be sent back to the APs.

During the real-time transmission, the channel coefficients are estimated at each AP and forwarded to the CPU. The CPU calculates the fade state and implement an exhaustive search among all stored mapping matrices in order to obtain an invertible $mu \times mu$ binary matrix satisfies the *Design Criterion* presented in Section III-C. When the optimal mapping matrix is selected, the indexes of the selected mapping matrices will be sent back to each AP through the backhaul channel. A log-likelihood ratio (LLR) estimator with a soft decision decoder is deployed at each AP to estimate each bit of NCV $\mathbf{v}_j$ based on the selected mapping matrix. Note that the LLR algorithm does not require to detect individual symbols transmitted from each MT but a linear combination of the binary messages using the selected mapping matrices. Then the NCV at each AP are forwarded to the CPU and the original messages from all MTs are recovered by multiplying the inverse of the global binary PNC mapping matrix.

V. Analysis and Discussion

In this section, we discuss how to apply the proposed BMAS algorithm to a general N-MIMO network, including utilisation of reduced number of SFSs and discussion of resolving the SFS problem with more than 2 MTs. With a study of the properties of SFSs, a regulated On-line search algorithm based on lookup table mechanism with a limited performance degradation is proposed in this section in order to fulfil the request of low latency in 5G RAN.

A. Image SFSs and Principal SFSs

Since an exhaustive search is carried out among all $t \times mu$ binary matrices in the proposed Off-line search algorithm, the computational complexity increases due to a large number of SFSs as well as an increased value of um in higher order modulation schemes with large number of MTs. For example, in 2-MT and 2-AP case, the number of SFSs need to be resolved at each AP is $L = 13$ in 4QAM and $L = 389$ in 16QAM. Thus for the 4QAM case, at least 13 binary matrices should be stored at each AP for On-line search. Then at least 389 binary matrices are required in 16QAM case which results an increased coding delay.

In the proposed BMAS algorithm, we resolve this problem by keeping the number of useful SFSs minimum. According to (22), we found some of different SFSs generate the same clashes which can be resolved by the same binary matrices. We then define such SFSs as image SFSs (iSFSs) and keep only one in the proposed search algorithm.

However, this problem still remains when QAM schemes with higher modulation order are employed at MTs. Due to a larger constellation in a QAM scheme with high modulation order, a few SFSs cannot be resolved by a binary mapping matrix. In order to address this problem, we focused on the occurrence probability of an SFS in different QAM schemes and noticed that not all SFSs occur frequently, so that we can ignore those "nonactive" SFSs with low appearance probabilities to minimise the number of mapping matrices utilised in the proposed On-line search. We define the SFSs with high appearance probabilities as principal SFSs (pSFSs) and a trade-off between the performance degradation and the number of pSFSs used in Offline search is illustrated in the next section.

B. Calculation of Singular Fade States

We illustrate how to determine a singular fade state for a QAM modulation scheme with a simple network first with $u = 2$ MTs, and discuss the SFS calculation issue in a network with more than 2 MTs later. Following Definition 3, given a QAM modulation scheme, a singular fading results $s_{j,\triangle}^{(\tau)} = s_{j,\triangle}^{(\tau')}$ when $\tau \neq \tau'$ in the constellation, mathematically derived as

$$\begin{aligned} &s_{1,\triangle}^{(\tau)} = s_{1,\triangle}^{(\tau')},\ \mathbf{hs}^{(\tau)} = \mathbf{hs}^{(\tau')},\ \text{for } \mathbf{s}^{(\tau)} \neq \mathbf{s}^{(\tau')}, \\ &h_{1,1}s_1^{(\tau)} + h_{1,2}s_2^{(\tau)} = h_{1,1}s_1^{(\tau')} + h_{1,2}s_2^{(\tau')}, \end{aligned}$$

where $h_{j,\ell}$ denotes the channel coefficient between the ℓ^{th} MT and the j^{th} AP, and $\mathbf{s}^{(\tau)} \triangleq [s_1^{(\tau)} s_2^{(\tau)}]$ refers to a joint symbol set which contains the modulated symbols at both MTs, and $\mathbf{s}^{(\tau)} \neq \mathbf{s}^{(\tau')}$ means at least one symbol is different in $\mathbf{s}^{(\tau)}$ and $\mathbf{s}^{(\tau')}$. Then we define $\mathcal{H}_{SFS} = [h_{SFS}^{(1)}, h_{SFS}^{(2)}, \cdots, h_{SFS}^{(L)}]$ as the set contains all unique value of SFSs and calculated by

$$h_{SFS}^{(q)} = \frac{h_{1,2}}{h_{1,1}} = \frac{s_1^{(\tau)} - s_1^{(\tau')}}{s_2^{(\tau')} - s_2^{(\tau)}}, \quad \forall s_l^{(\tau)}, s_l^{(\tau')} \in \Omega. \tag{23}$$

By substituting the QAM symbols with all possible combinations to (23), we can find all SFS values for this QAM scheme when $u = 2$ MTs.

In the multiple-MT ($u > 2$) case, (23) is no longer suitable for SFS values representation due to the increased number of MTs in MAC stage. The relationship between the values of SFSs and channel coefficients is no longer able to expressed by a simple linear equation, e.g. the SFSs form different surfaces with infinite values in 3-MT case. This is still an open issue in the literature for binary PNC design and we give our potential solution here. One solution is to utilise clashes instead of calculation of the values of SFSs in the proposed algorithm. In the multiple-MT case, an SFS still causes clashes and different clashes can be always found according to Definition 5. Then a binary matrix maps the superimposed symbols within a clash to the same NCV and keeps the value of d_{min} maximised at the same time will be selected in the proposed algorithm, without SFS calculation.

Another solution is to divide the whole networks into multiple 2-MT subnetworks. One approach to achieve this is by allocating different pairs of MTs to different frequencies or time slots. In this case, the superimposed symbol at an AP is always from two MTs and then SFSs can be calculated by (23). The only issue of this approach is that multiple On-line search algorithms for different MT pairs are required to be implemented which may cause extra computational complexity and coding latency. An alternative approach is to consider two MTs with the similar channel strength and transmit power as the prime MT pair and trade the other received signals as additional noise. According to our research in [25], when a strong signal with much higher energy comapred to the rest of signals is received at an AP, it is difficult to find a matrix with unambigous recovery capability [25]. Thus the 2 MTs whose signals are allocated at the similar energy level in the multiple access stage can be paired to form a subnetwork for PNC encoding and by pairing different MTs and APs, the multiple-MT-multiple-AP case is replaced by multiple 2-MT-2-AP cases.

C. Regulated BMAS Search Algorithm

In order to fulfil the request of low latency in 5G communications, we present a regulated BMAS (R-BMAS) approach with a lookup table mechanism in

this subsection. According to the definition of clash, the superimposed symbols in a clash have an intra-cluster distance of '0' and by the calculation in (23), the clash groups in an SFS are mainly determined by the absolute value and the angle of the ratio of two channel coefficients. So following the design rules in the propose algorithm, we have:

Theorem 4: The mapping matrix which resolves an SFS can always resolve the non-singular fade states with the values close to this SFS.

Proof: When a non-singular fade state (nSFS) happens, different superimposed symbols received at an AP will not be coincided so no clash is observed. When this nSFS holds a similar absolute value and rotation angle to an SFS, the superimposed symbols, which form a clash in this SFS, will form a cluster in this nSFS with a smaller intra-cluster distances compared to inter-cluster distances to other clusters. In this case, the mapping matrices that maps the coincided superimposed symbols in a clash to an NCV achieve the maximum $d_{\min}$ in this nSFS by mapping the superimposed symbols in the cluster to the same NCV. ■

An example is given in Appendix B. According to Theorem 4, the proposed On-line search approach could be replaced by a lookup table based mechanism for optimal mapping matrices selection. A table contains all SFS combinations and the corresponding $mu \times mu$ optimal binary mapping matrices could be established in the Off-line search. During the real-time transmission, the value of the fade state v_{FS_j} is calculated by $h_{FS_j} = h_{j,2}/h_{j,1}$ and then the closest SFS to this nSFS is obtained by

$$q_{min_j} = \min d_{FS}^{(q_j)} = \min |h_{FS_j} - h_{SFS}^{(q_j)}|^2, \quad \text{for } q = 1, \cdots, L,\ j = 1, \cdots, n. \tag{24}$$

The index q_{min_j} will be forwarded to the CPU. By checking the table, the optimal mapping matrix indexes are sent back to the APs for PNC encoding. The algorithm is summarised in Algorithm 1.

As shown in Algorithm 1, most of the calculations for mapping selection have been done before the real-time transmission. Moreover, the latency is reduced by applying the regulated On-line search in Algorithm 1 instead of Algorithm 3. However, a disadvantage of the R-BMAS algorithm is the performance degradation caused by selecting sub-optimal global mapping matrices due to the FS estimation errors. In this case, accurate channel estimation (FS estimation) algorithm is required for the R-BMAS algorithm and we illustrate our study in the next Section.

Algorithm 1 Regulated Binary Matrices Adaptive Selection (R-BMAS) Algorithm

Off-line Search
for $i = 1 : L$ **do** ▷ each SFS
  Apply Algorithm 2 and 3 for $\mathbf{M}_i$
end for
for $i_1 = 1 : L$ **do** ▷ all SFS combinations
  ⋮
  for $i_n = 1 : L$ **do** ▷ all n APs
    $\mathbf{G} = \begin{bmatrix} \mathbf{M}_{i_1} \\ \vdots \\ \mathbf{M}_{i_n} \end{bmatrix}$
    $\delta \leftarrow \det(\mathbf{G})_{|\mathbb{F}_2}$ ▷ determinant over $\mathbb{F}_2$.
    if $\delta = 1$ **then**
      $\mathcal{G}_l \leftarrow \mathcal{G}_l \cup \mathbf{G}$
      Add l as the optimal mapping matrix for SFS combination $[i_1 \cdots i_n]$
    end if
  end for
  ⋮
end for
On-line Search
for $j = 1 : n$ **do** ▷ each AP
  $h_{FS_j} = h_{2,j}/h_{1,j}, \quad j \in [1, 2]$ ▷ fade state
  for $i = 1 : L$ **do**
    $\mathbf{d}_{SF}^{(i_j)} = |h_{FS_j} - h_{SFS}^{(i)}|^2$
  end for
  $[d^{(k_j)}, k_j] = \min \mathbf{d}_{SF}^{(i_j)}$ ▷ SFS index
end for
Forward $[k_1 \cdots k_n]$ to CPU
Look up the table to obtain the optimal mapping matrix index l
Send the index back to APs

D. Computational Complexity and Backhaul Load

We investigate the computational complexity of the ideal CoMP, non-ideal CoMP and the proposed algorithms in this subsection. In ideal CoMP, the bandwidth of backhaul network is assumed unlimited so that the received signal y_j in (1) at each AP will be forwarded to the CPU for joint multiuser ML detection, given by

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s} \in \Omega^u} \|\mathbf{y} - \mathbf{H}\mathbf{s}\|, \tag{25}$$

where $\mathbf{s}$ denotes the $u \times 1$ symbol vector at the MTs and $\hat{\mathbf{s}}$ is the estimated version, and $\mathbf{H}$ stands for the $n \times u$ channel matrix. In non-ideal CoMP, the backhaul network is bandwidth-limited and each AP employs an LLR based multiuser detection algorithm to estimate the transmitted symbols from each MT, mathematically given by

$$\chi_{i,\ell} = \frac{\sum\limits_{w_{i,\ell}=0} P(s_\ell)p(y_j \mid s_\ell)}{\sum\limits_{w_{i,\ell}=1} P(s_\ell)p(y_j \mid s_\ell)}, \ i = 1, 2, \cdots, m, \tag{26}$$

where $\chi_{i,\ell}$ stands for the LLR corresponding to the i^{th} bit of the binary message vector $\mathbf{w}_\ell$. A scalar quantizer which quantizes $\chi_{i,\ell}$ into binary bits is employed after the estimation. The quantised bits are sent to the CPU via the backhaul network. A detailed computation of (26) is given in [23].

In the proposed BMAS algorithm, each AP estimates the linear combination of the messages from MTs based on the ML rule and mathematically expressed in (6) - (8). The proposed On-line search is implemented during the real-time transmission with a reduced number of mapping matrix candidates, e.g. $K = 5$ matrices at each AP for 4QAM modulation scheme in Algorithm 3. While in the proposed R-BMAS algorithm, calculations in the mapping selection are replaced by a lookup table mechanism and the computation complexity is reflected in distance comparison in (24). Then an LLR estimation of each bit in the NCV is applied which is the same as in the BMAS algorithm.

Consider a quantisation scheme of 2 bits is employed at each AP and 4QAM is utilised at 2 MTs in a simple 5-node system, then 4 LLRs are calculated by (26) at each AP and a total 16 bits are sent via the backhaul network in non-ideal CoMP. In order to achieve a good performance in terms of error rate and outage probability, a quantization scheme with a larger number of quantized bits is required. A binary mapping matrix with the minimum size of 2×4 is utilised in the proposed BMAS algorithm, which results in a total 4 bits backhaul load. Note an AP could employ a mapping matrix with the maximum size of $mu \times mu$ to generate the NCV and in this case, the other APs will not participate in the PNC encoding because they fail to receive any useful signals and the total backhaul load is still equal to mu.

VI. Numerical Results

In this section, we illustrate the outage probability performances of the proposed BMAS algorithm, R-BMAS algorithm and CoMP in a 5-node system which includes 2 MTs, 2 APs and 1 CPU. The 5-node network is the smallest N-MIMO network and used as a baseline to illustrate the advantage of the proposed algorithms. Each node contains 1 antenna for transmitting and receiving and 4QAM/16QAM modulation schemes is employed at both MTs. Convolutional code is deployed in the simulation. We assume the multi-access links are wireless and the backhaul is wired which allows only binary data to be transmitted. Then a quantizer is deployed at each AP in ideal and non-ideal CoMP but the backhaul capacity is limitless in ideal CoMP and different quantization bits (2 bits and 4 bits) are used in non-ideal CoMP.

Fig. 2 illustrates the outage probabilities of the proposed BMAS algorithm in 4QAM and 16QAM schemes. As we can see from the figure, the outage probability curves achieve the same diversity order. The ideal CoMP achieves the optimal performance in both 4QAM and 16QAM cases due to the unlimited backhaul capacity and joint detection. When the backhaul load is capacity limited, non-ideal CoMP occupies 8 bits and 16 bits backhaul load in 4QAM case with 2-bit and 4-bit quantizer, respectively. Compared to the ideal CoMP, a 8dB and 13dB performance degradation in the non-ideal CoMP can be observed; whilst the degradation is limited to only approximately 3dB by deploying the proposed algorithm with 4-bit backhaul load. Instead of obtaining all SFSs-resolvable mapping matrices for 16QAM schemes, we consider only 4, 12 and 50 pSFSs in the proposed BMAS algorithm for computational complexity reduction. As illustrated in the figure, there is approximately 10dB degradation in outage performance when using only 4 pSFSs in the BMAS algorithm. When 12 and 50 pSFSs are considered, the degradation is reduced to 7dB and 5dB respectively and the gap will reduce when more pSFSs are considered in the proposed BMAS algorithm.

In Fig. 3, outage probability comparison between the proposed BMAS and R-BMAS algorithm is given. When 4QAM is used at both MTs, outage probability performance of BMAS algorithm is 1dB better than that of R-BMAS algorithm due to SFS

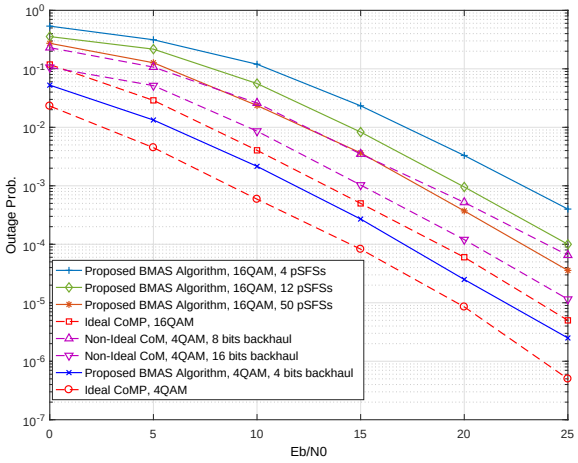

Fig. 2. Outage Probability of the Proposed BMAS Algorithm in 4QAM and 16QAM.

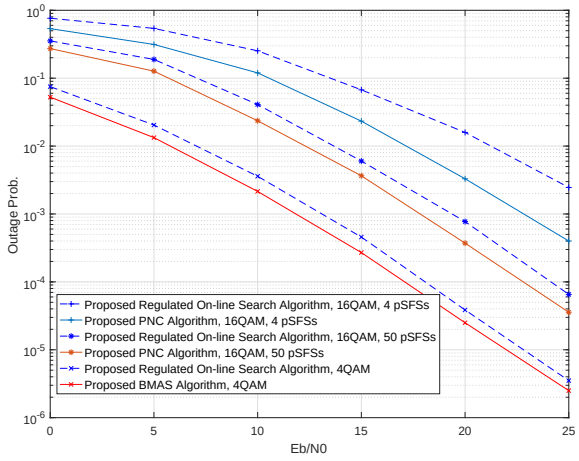

Fig. 3. Outage Probability of the Proposed BMAS Algorithm vs Regulated On-line Search Algorithm.

estimation errors used in the R-BMAS algorithm. In term of reduced computational complexity in the R-BMAS algorithm, an index searching among 25 binary 4 mapping matrices are implemented to resolve all possible SFS combinations (5 pSFSs at each AP). When 16QAM is employed at both MTs, the performance degradation between the proposed BMAS and the R-BMAS algorithm depends on the number of pSFSs utilised in the algorithm. When only 4 pSFSs are used for optimal mapping matrix selection in both algorithms, BMAS achieves about 5dB gains compared to R-BMAS. A big gap in 4 pSFSs case is caused by inefficient mapping matrix candidates in the R-BMAS algorithm due to SFS estimation errors. With the number of pSFSs increasing to 50, more pSFS candidates are used in the R-BMAS algorithm which improves the outage performance about 7dB and reduces the gap to 1dB only.

We investigate how estimated channel state information (CSI) affects the network performance and the results are in Fig.s 4 and 5. The values of FSs are calculated by the first equation in (23) so that during the transmission, the accuracy of the CSI in access link is important because it determines if the optimal mapping matrix can be selected. In Fig. 4, we illustrate the impact of estimated CSI with different pilot length to the optimal mapping selection. The higher value of "mis-mapping " means the less probability that the optimal mapping matrix will be selected. As we can see from the figure, the mis-mapping percentage decreases in any pilot length circumstances with increase of E_b/N_0. By using a short-length pilot sequence, the mis-mapping percentages are quite high which is caused by the fact that inaccurate fade states are calculated by using the estimated CSI. The low mis-mapping percentage shown in Fig. 4 leads to better outage probability performance in Fig. 5. For example, at 15dB, the outage probability using only 1 pilot symbol is 10^{-2} which refers to a mis-mapping percentage of 22%; while using 10 pilot symbols, the outage probability reduces to 3×10^{-3} and a mis-mapping percentage of only 8% is achieved. By comparing the outage performances in Fig. 5 with perfect and estimated CSI, we can conclude that pilot sequence with length of 10 is good enough for the proposed BMAS algorithm.

VII. Conclusion

In this paper we present a design guideline of engineering applicable physical layer network coding in the uplink of N-MIMO networks. The proposed design criteria guarantee unambiguous recovery of all messages and the traffic in the backhaul network is reduced to the level of total user data rate at the same time. We then propose a two-step optimal mapping matrix selection algorithm based on the design criteria in order to reduce the real-time computational complexity. The proposed algorithm is not only designed for a simple 5-node network but for a general N-MIMO network serves multiple MTs. In addition, a regulated On-line search algorithm based on lookup table mecha-

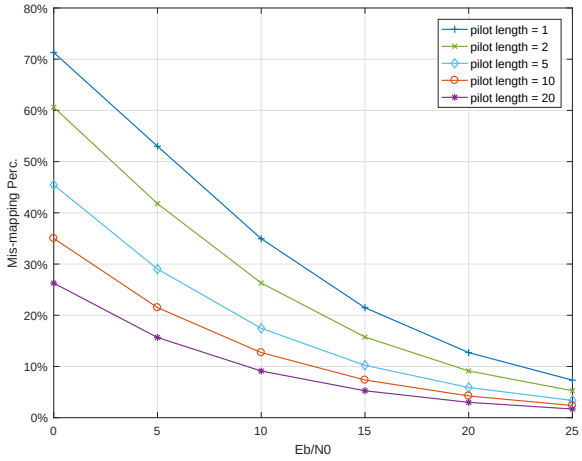

Fig. 4. Miss-Mapping Probabilities with Different Pilot Lengths.

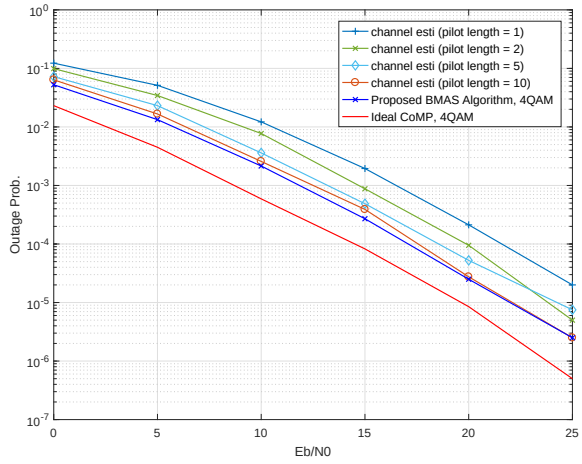

Fig. 5. Outage Probability with Different Pilot Lengths.

nism is also presented in order to further reduce the computational complexity and latency without much performance degradation. With reduced backhaul load, the proposed algorithms achieve higher outage probability performance compared to the practical non-ideal CoMP approaches.

Appendix A

We illustrate a detailed Off-line search algorithm in Algorithms 2 and 3. Part I indicates how to calculate SFSs and how to remove the image SFS in order to reduce the computational complexity, while Part II focuses on optimal mapping matrix selection. The steps in the proposed On-line search is the same as that in the Off-line search but with less number of matrix candidates so we will not show the repeated work here. $\mathcal{Q}_{\rm d}$ in Algorithm 2 is defined as a vector contains all the d_{min} between different NCSs for every binary mapping matrices in each SFS.

Algorithm 2 SFS Calculation and Image SFS Remove (Off-line Search Algorithm. Part I)

1: **for** $i = 1 : L$ **do** ▷ each singular fade state
2: $h = \mathcal{S}(i)$ ▷ h is a $1 \times m$ vector.
3: **for** $j = 1 : K$ **do** ▷ each binary matrix
4: $[\xi, T_\xi] = N(\mathcal{M}(j))$
5: $\xi_{\rm f} \leftarrow \mathscr{F}(T_\xi, h)$ ▷ $\mathscr{F}(\cdot)$ produces all faded NCSs.
6: $d_{\rm min} \leftarrow \mathscr{D}(\xi_{\rm f})$ ▷ $\mathscr{D}(\cdot)$ calculates the minimum distance of all NCSs.
7: $\mathcal{Q}_{\rm d} \leftarrow \mathcal{Q}_{\rm d} \cup d_{\rm min}$ ▷ store all $d_{\rm min}$ in $\mathcal{Q}_{\rm d}$.
8: **end for**
9: $[\boldsymbol{\beta}(i), \boldsymbol{\alpha}(i)] \leftarrow \mathscr{C}(\mathcal{Q}_{\rm d})$ ▷ $\mathscr{C}(\cdot)$ sorts $\mathcal{Q}_{\rm d}$ in descending order stored in $\boldsymbol{\beta}(i)$ and outputs the rearranged index vector $\boldsymbol{\alpha}(i)$.
10: **end for**
11: $\mathcal{S}' \leftarrow \mathscr{I}(\mathcal{S}, \boldsymbol{\alpha})$ ▷ delete all image singular fade states and $\mathcal{S}'$ has L' singular states, $L' < L$.
12: $\boldsymbol{\alpha} \leftarrow \boldsymbol{\alpha} \setminus \boldsymbol{\alpha}(\boldsymbol{\beta} = 0)$ ▷ delete the index element of $\boldsymbol{\beta} = 0$.
13: $\boldsymbol{\alpha}' \leftarrow \boldsymbol{\alpha}(i|\mathcal{S}')$ ▷ $\boldsymbol{\alpha}'$ corresponds to only $\mathcal{S}'$.

Appendix B

We illustrate an example of Theorem 4 here. In Fig. 6, a received constellation of an SFS with $h_{SFS_1} = i$ is illustrated. The number of MTs is $u = 2$ and 4QAM modulation is employed. We can observe the clashes clearly from the figure and their values can be calculated according to (23), e.g. 4 constellation points are superimposed at $(0, 0)$ and 2 are at $(0, 2)$. Then the optimal binary mapping matrix will encode the superimposed constellation points in a clash to the same NCV and maximise the distance between different NCVs at the same time according to the design criteria.

Fig. 7 illustrates the received constellation of all possible superimposed symbols of another SFS with $h_{SFS_2} = 1/2 + 1/2i$. In this case, 2 constellation points are superimposed at $(0, 1)$, $(0, -1)$, $(1, 0)$ and $(-1, 0)$, respectively. The optimal mapping matrices for SFS_1 and SFS_2 are different due to the differences between the clashed constellation points.

Algorithm 3 Binary Matrix candidates Selection for Each AP (Off-line Search Algorithm. Part II)

for $l_{L'} = 1 : L'$ **do**
  $\mathcal{S}^{\dagger}_{L'-1} \leftarrow \mathcal{S}' \setminus \mathcal{S}'(l_{L'})$
  $\theta_{L'-1} \leftarrow \mathscr{E}(\mathcal{S}^{\dagger}_{L'-1})$ ▷ Index set of $\mathcal{S}'$ excluding the l^{th} element.
  for $l_{L'-1} = \theta_{L'-1}$ **do**
    $\mathcal{S}^{\dagger}_{L'-2} \leftarrow \mathcal{S}^{\dagger}_{L'-1} \setminus \mathcal{S}^{\dagger}_{L'-1}(l_{L'-1})$
    $\theta_{L'-2} \leftarrow \mathscr{E}(\mathcal{S}^{\dagger}_{L'-2})$
    $\vdots$
    for $l_{L'-n+1} = \theta_{L'-n+1}$ **do**
      for $i_1 = 1 : K$ **do**
        $\vdots$
        for $i_n = 1 : K$ **do**
          $\mathbf{M} = \begin{bmatrix} \mathcal{M}[\boldsymbol{\alpha}(l_{L'}, i_1)] \\ \vdots \\ \mathcal{M}[\boldsymbol{\alpha}(l_{L'-n+1}, i_n)] \end{bmatrix}$
          $\delta \leftarrow \det(\mathbf{M})_{\mid \mathbb{F}_2}$ ▷ determinant over $\mathbb{F}_2$.
          if $\delta = 1$ **then**
            $\mathscr{R} \leftarrow \mathscr{R} \cup (l_{L'} \cdots l_{L'-n+1}; i_1 \cdots i_n)$
            $\mathbf{G} \leftarrow \mathbf{G} \cup \mathbf{M}$ ▷ $\mathbf{M} \leftrightarrow \mathbf{G}_{A(k)}$ in $\mathbf{G}$ has unique address $A(k) = (l^{(k)}_{L'} \cdots l^{(k)}_{L'-n+1}; i^{(k)}_1 \cdots i^{(k)}_n)$, $k = 1, \cdots \frac{L'!}{(L'-n)!}$.
            return (21)
          end if
        end for
      end for
    end for
  end for
end for

$[\mathbf{G}_{A(k_1)} \cdots \mathbf{G}_{A(k_n)}] \leftarrow \mathscr{X}(\mathbf{G})$ ▷ find n $\mathbf{M}$ from $\mathbf{G}$ satisfying bijection relations $(l^{(k_e)}_{L'} \cdots l^{(k_e)}_{L'-n+1}) \Leftrightarrow \mathcal{S}'$ for $k_e = k_1 \cdots k_n$.
for $i = 1 : n$ **do**
  $\mathcal{Q}_i \leftarrow [\mathbf{G}^i_{A(k_1)} \cdots \mathbf{G}^i_{A(k_n)}]$ ▷ $\mathbf{G}^i_{A(k_i)} = \mathcal{M}[\boldsymbol{\alpha}(l^{(k_i)}_{L'-i+1}, i^{(k_i)}_i)]$
end for
Output: n stacks $\mathcal{Q}_i$ with each including L' binary matrices.

Then we consider a received constellation with a non-singular fading with $h_{nSFS} = 7/10 + 7/10i$

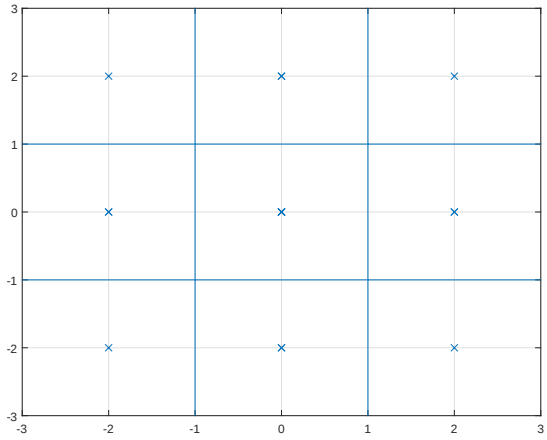

Fig. 6. Constellation of the Received Signals at AP, $h_{SFS_1} = i$.

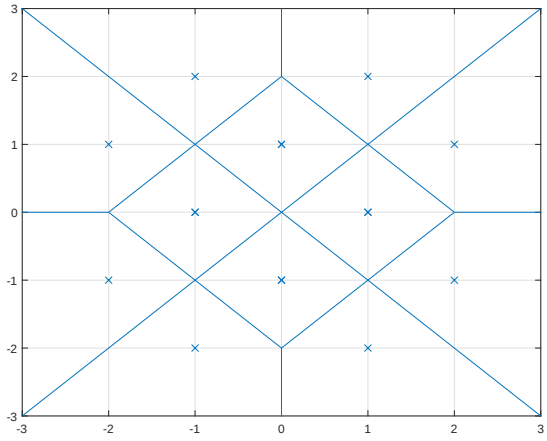

Fig. 7. Constellation of the Received Signals at AP, $h_{SFS_1} = 1/2 + 1/2i$.

illustrated in Fig. 8. Clearly, none of the constellation points are superimposed but we can easily indicate different clusters by the distances between constellation points. Also we can find that the distance, which refers the absolute value and rotation angle in (11), between h_{SFS_1} and h_{nSFS} is smaller than that between h_{SFS_2} and h_{nSFS}. Then according to the unambiguous detection theorem, the 4 points around $(0, 0)$ in Fig. 8 should be mapped to the same NCV to maximise the inter-cluster distance. The same criteria should be satisfied by the 8 points near $(0, 2)$, $(-2, 0)$, $(0, -2)$ and $(2, 0)$. Then the initial cluster groups in Fig. 6 and Fig. 8 are the same which leads to the same optimal mapping matrices could be used in both circumstances.

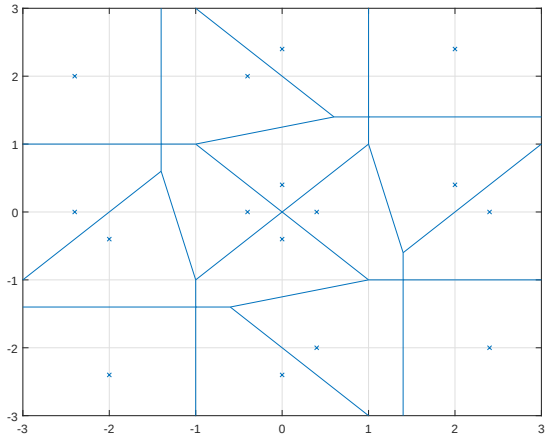

Fig. 8. Constellation of the Received Signals at AP, $h_{nSFS} = 7/10 + 7/10i$.

REFERENCES

[1] M. V. Clark, T. M. III Willis, L. J. Greenstein, A. J. Rustako, V. Erceg and R. S. Roman, "Distributed versus centralized antenna arrays in broadband wireless networks,"*Proc. 2001 Spring IEEE Vehicular Technology Conference.*, Rhodes Island, Greece, May 2001, pp. 33-37.

[2] A. Checko, *et al.*, "Cloud RAN for Mobile — Networks A Technology Overview, "*IEEE Communications Surveys & Tutorials*, vol. 17, no. 1, pp. 405-426, Firstquarter 2015.

[3] H. Huang, H. Barani and H. Al-Azzawi, "Network Multiple-Input and Multiple-Output for Wireless Local Area Networks, "available on: https://arxiv.org/abs/1402.1882

[4] S. Kanchi and S. Sandilya and D. Bhosale and A. Pitkar and M. Gondhalekar, "Overview of LTE-A technology, "*2013 IEEE Global High Tech Congress on Electronics (GHTCE)*, Shenzhen, China, Nov. 2013, pp. 195-200.

[5] J. Zhang, *et al.*, "Reconfigurable optical mobile fronthaul networks for coordinated multipoint transmission and reception in 5G, "*IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 6, pp. 489-497, June 2017.

[6] S.-H. Park, O. Simeone, O. Sahin and S. Shamai, "Multihop Backhaul Compression for the Uplink of Cloud Radio Access Networks,"*IEEE Trans. on Vehicular Tech.,* vol. 65, no. 5, pp. 3185-3199, MAY 2016.

[7] T. R. Lakshmana, *et al.*, "Scheduling for Backhaul Load Reduction in CoMP, "*2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 227-232, Shanghai, China, Apr. 2013.

[8] F. Zhang, *et al.*, "Coordinated beamforming for heterogeneous small-cell networks with a non-ideal backhaul, "*IET Communications*, vol. 12, no. 5, March 2018.

[9] Y. Wang, Z. Chen and M. Shen, "Compressive sensing for uplink cloud radio access network with limited backhaul capacity, "*2015 4th International Conference on Computer Science and Network Technology (ICCSNT)*, pp. 898-902, Harbin, China, Dec. 2015.

[10] Q. T. Sun, J. Yuan, T. Huang, and W. K. Shum, "Lattice network codes based on Eisenstien integers, "*IEEE Trans. Commun.*, vol. 61, no. 7, pp. 2713-2725, Jul. 2013.

[11] D. Fang and A. G. Burr, "Linear Physical-Layer Network Coding for 5G Radio Access Networks , "*2014 1st International Conference on 5G for Ubiquitous Connectivity (5GU)*, pp. 116-221, Akaslompolo, Finland, Nov. 2014.

[12] S. L. Zhang, S. C. Liew and P. P. Lam, "Hot Topic: Physical-layer Network Coding, "*Proceedings of the 12th Annual International Conference on Mobile Computing and Networking, MOBICOM 2006*, Los Angeles, CA, USA, Sept. 2006, pp. 358-365.

[13] M. M. Molu, K. Cumanan and A. G. Burr, "Low-Complexity Compute-and-Forward Techniques for Multisource Multirelay Networks, "*IEEE Comm. Letters*, vol. 20, no. 5, pp. 926-929, May 2016.

[14] Y. Tan, S.C. Liew and T. Huang, "Mobile Lattice-Coded Physical-Layer Network Coding with Practical Channel Alignment, "*IEEE Trans. on Mobile Computing*, vol. 17, no. 8, pp. 1908-1923, August 2018.

[15] 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.

[16] T. Koike-Akino and P. Popovski and V. Tarokh, "Denoising Maps and Constellations for Wireless Network Coding in Two-Way Relaying Systems, "*2008 IEEE Global Telecommunications Conference (GLOBECOM)*, New Orleans, LO, USA, Nov. 2008, pp. 1-5.

[17] T. Koike-Akino, P. Popovski, and V. Tarokh, "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.

[18] L. Shi and S. C. Liew, "Complex Linear Physical-Layer Network Coding, "*IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 4949-4981, April 2017.

[19] A. G. Burr and D. Fang, "Linear physical layer network coding based on rings, "*2014 IEEE Wireless Communications and Networking Conference (WCNC)*, Istanbul, Turkey, April 2014, pp. 370-375.

[20] P. Prochazka, "Nonbinary Channel Coded Physical Layer Network Coding Over Modulo-Sum Algebraic Ring Structures, "*IEEE Communications Letters*, vol. 20, no. 3, pp. 538-541, March 2016.

[21] L. Shi, S. C. Liew and L.Lu, "On the Subtleties of q-PAM Linear Physical-Layer Network Coding, "*IEEE Trans. Infor. Theory*, vol. 62, no. 5, pp. 2520-2544, May 2016.

[22] J. Gu, R. C. de Lamare and M. Huemer, "Buffer-Aided Physical-Layer Network Coding With Optimal Linear Code Designs for Cooperative Networks, "*IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2560-2575, June 2016.

[23] S. Verdu, "*Multiuser Detection*,"Cambridge University Press, 1998.

[24] T. Peng, Y. Wang, A. G. Burr and M. Shikh-Bahaei, "An Adaptive Optimal Mapping Selection Algorithm for PNC using Variable QAM Modulation, "*IEEE Wireless Communications Letters*, accepted.

[25] Y. Chu, T. Peng, D. Grace and A. G. Burr, "Implementation of Uplink Network Coded Modulation for Two-Hop Networks, "*IEEE Access*, submitted. Available on: https://arxiv.org/abs/1808.07354

[26] William Brown, "*Matrices over Commutative Rings (Chapman & Hall Pure and Applied Mathematics)* ,"Taylor & Francis, 1992.