

Signal Alignment: Enabling Physical Layer Network Coding for MIMO Networking

Ruiting Zhou, Zongpeng Li, Chuan Wu, Carey Williamson

Abstract—We apply signal alignment (SA), a wireless communication technique that enables physical layer network coding (PNC) in multi-input multi-output (MIMO) wireless networks. Through calculated precoding, SA contracts the perceived signal space at a node to match its receive capability, and hence facilitates the demodulation of linearly combined data packets. PNC coupled with SA (PNC-SA) has the potential of fully exploiting the precoding space at the senders, and can better utilize the spatial diversity of a MIMO network for higher system degrees-of-freedom (DoF). PNC-SA adopts the idea of ‘demodulating a linear combination’ from PNC. The design of PNC-SA is also inspired by recent advances in IA, though SA aligns signals not interferences. We study the optimal precoding and power allocation problem of PNC-SA, for SNR (signal-to-noise-ratio) maximization at the receiver. The mapping from SNR to BER is then analyzed, revealing that the DoF gain of PNC-SA does not come with a sacrifice in BER. We then design a general PNC-SA algorithm in larger systems, and demonstrate general applications of PNC-SA, and show via network level simulations that it can substantially increase the throughput of unicast and multicast sessions, by opening previously unexplored solution spaces in multi-hop MIMO routing.

Index Terms—Network coding; physical layer network coding; interference alignment; signal alignment; MIMO networks.

I. INTRODUCTION

NEW physical layer techniques and their applications in wireless routing have been active areas of research in the recent past. A salient example is multi-input multi-output (MIMO) communication. A MIMO link employs multiple transmit and receive antennas that operate over the same wireless channel. MIMO transmission brings extra spatial diversity that can be exploited to break through capacity limits inherent in single-input single-output (SISO) channels [1], [2]. Another recent example is physical layer network coding (PNC) [3], which extends the concept of network coding [4] from higher layers to the physical layer. PNC is seminal in that it utilizes the natural additive property of Electro-Magnetic (E-M) waves in space. Viewing collided transmissions simply as superimposed signals, PNC applies tailored demodulation for translating them into linear combinations of transmitted data packets. Such demodulated linear combinations, similar to encoded packets in network coding [4], are then used to facilitate further data routing.

Manuscript received September 23, 2012; revised January 17 and April 7, 2013; accepted April 17, 2013. The associate editor coordinating the review of this paper and approving it for publication was D. Niyato.

R. Zhou, Z. Li, and C. Williamson are with the Department of Computer Science, University of Calgary (e-mail: {rzho, zongpeng, carey}@ucalgary.ca).

C. Wu is with the Department of Computer Science, University of Hong Kong (e-mail: cwu@cs.hku.hk).

Digital Object Identifier 10.1109/TWC.2013.050313.121454

We apply signal alignment (SA) [5], a new technique that enables PNC in wireless networks consisting of MIMO links. A central idea behind SA is to improve network capacity by enabling simultaneous transmissions from multiple MIMO senders. SA performs calculated precoding at the senders, such that the number of dimensions spanned by signals arriving at a receiver is reduced to exactly match its receive diversity, *i.e.*, the dimension of the received signal vector, which is also the number of antennas employed at the receiver. Consequently, the receiver can decode linear combinations of the transmitted packets. This is through classic MIMO detection, such as maximum likelihood detection (ML) or zero forcing (ZF) [1], [6], followed by PNC mapping [3]. In this work, we demonstrate that PNC coupled with SA (PNC-SA) can open new solution spaces for routing in MIMO networks, leading to higher throughput with good bit-error-rate (BER), as compared to previous techniques.

The idea and benefit of PNC-SA can be illustrated in an uplink communication scenario, designed to motivate interference alignment and cancellation (IAC) [2], [7], a technique for improving throughput in MIMO networks. Such a multi-user MIMO (MU-MIMO) architecture represents a trend in cellular communication that seeks further capacity gain over a simple MIMO link. PNC-SA provides a further degrees-of-freedom (DoF) [1], [5], [8] gain over IAC at 33%.

Fig. 1 depicts a MIMO uplink from two clients to two inter-connected APs. Each node is equipped with 2 antennas that operate on the same channel, with flat Rayleigh fading [1], [2]. During propagation, a signal experiences amplitude attenuation and phase shift, which can be modeled using a complex number. \mathbf{H}_{ij} is the 2×2 complex matrix for the channel gains from client i to AP j . An Ethernet link connects the two APs, enabling *limited collaboration*: digital packets can be exchanged, but not analog ones, since otherwise substantial overhead is incurred for no-loss recovery at AP2 using double sampling [2]. The system DoF here becomes the number of data signals or packets that can be simultaneously transmitted from the clients and recovered at the APs, as SNR approaches infinity ([5], Sec. II.A).

A naive solution uses one Tx-Rx antenna pair to avoid any interference at all. Let’s normalize a time unit to be one packet transmission time. For a quick improvement, we can use a 2×2 MIMO link formed by a client-AP pair, to transmit two packets, x_1 and x_2 , simultaneously. Each AP receives two overlapped signals of x_1 and x_2 . ML or ZF detection can be applied to recover x_1 and x_2 , increasing the throughput from 1 to 2 packets (per time unit).

Can we utilize all available antennas to form a 4×4 MIMO

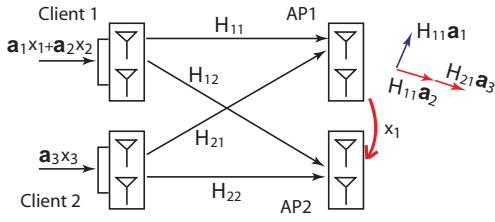


Fig. 1. The 2-client 2-AP MIMO uplink, where the two APs are co-located in the same base station and are interconnected through an Ethernet link. IAC achieves a throughput of 3 packets per time unit. Each \mathbf{a}_i is a 2×1 precoding vector. $\mathbf{H}_{11}\mathbf{a}_1$ is called the *direction* of x_1 at AP1.

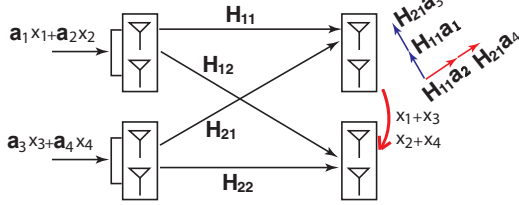


Fig. 2. PNC-SA can achieve a throughput of 4 packets per time unit.

link, to transmit >2 packets? The answer, unfortunately, is ‘no’. Since the four Rx antennas are distributed at two nodes, we do not have all four received analog signals at one location, as required in MIMO decoding.

IAC [2], [7] breaks through this limitation by combining interference alignment (IA) [9] and interference cancellation (IC) techniques. As shown in Fig. 1, IAC first performs precoding over 3 packets x_1 , x_2 and x_3 at the clients, such that x_2 and x_3 arrive along the same direction at AP1. *Direction* here is a signal’s encoding vector when received at AP1. AP1 has two equations of two unknowns, from which it can solve x_1 . Next, AP1 transmits x_1 in digital format to AP2. AP2 subtracts the component of x_1 from its received signals (IC), leaving it with two equations over two unknowns, from which it recovers x_2 and x_3 .

Can we use IAC to transmit 4 packets in one time unit instead of 3? The answer is ‘no’. With IAC, the intended signal has to take its own direction at AP1, while all other ‘interferences’ take another. As a result, the two packets from client 2 have to be aligned to the same direction at AP1. This requires identical precoding vectors for them at the clients, making them impossible to separate at AP2.

Departing from such a requirement of IA and IAC, SA allows multiple *signals* to be aligned to the same direction at a receiver. In fact, there is no *interference* in SA; all data transmissions are treated as *signals*. As shown in Fig. 2, PNC-SA simultaneously transmits 4 packets, x_1, \dots, x_4 . Precoding is performed such that at AP1, x_1 and x_3 are aligned to the same direction, and x_2 and x_4 are aligned to another direction. AP1 has two equations, from which it solves x_1+x_3 , x_2+x_4 to transmit in digital format to AP2. Having accumulated 4 equations, AP2 then solves them to recover the 4 original packets, x_1, \dots, x_4 .

Two ideas work in concert in PNC-SA. One is *demodulating a linear combination*, adapted from PNC. The other is *precoding at the sender for alignment at the receiver*, inspired by IA. PNC-SA helps the exploration of the full precoding space at

the senders, and the full spatial diversity of the system. As we will show, PNC and IAC can indeed be viewed as special cases of PNC-SA. When each node has a single receive diversity, SA degrades into phase synchronization [3], [10], and PNC-SA degrades into PNC. With extra restrictions on precoding and decoding, PNC-SA degrades into IAC.

In wireless transmissions, high DoF is less attractive if it comes with higher BER. The BER of the PNC-SA scheme depends on two factors: (a) the SNR at the receiver, and (b) the function that maps SNR to BER. While (a) depends on precoding (signal pre-rotation and power allocation) at senders, (b) depends on the modulation scheme. We study each factor in detail. We show that SA introduces a new, interesting optimization problem in precoding design, and classic solutions such as singular value decomposition followed by water filling (SVD-WF) does not apply any more. We formulate the optimization as a vector programming problem, and design an efficient solution using orthogonal signal alignment. The SNR-BER performance of PNC-SA is then analyzed, and compared to that of IAC. We observe that the throughput gain of PNC-SA indeed does not come with a cost in error rate.

For a larger system with $N > 2$ client AP pairs where each node has $M > 2$ antennas, we design a heuristic PNC-SA algorithm that searches for a feasible precoding and signal alignment solution towards a target DoF X . The application of PNC-SA is not limited to scenarios of limited receiver collaboration. We study general applications of PNC-SA in multi-hop MIMO networks, for routing tasks including information exchange, unicast, and multicast/broadcast. We show that PNC-SA opens previously unexplored solution spaces for MIMO routing, and can augment the capacity region of a MIMO network. Via packet-level simulations, throughput gains up to a factor of 2 are observed, especially at high SNR. In both unicast and multicast routing, PNC-SA can lead to a natural fusion of PNC and digital network coding (DNC). We finally demonstrate that SA can even be applied independent of PNC, in supporting simple and efficient broadcast algorithm design in MIMO networks.

In the rest of the paper, we review previous research in Sec. II, outline the system model in Sec. III, present a detailed PNC-SA solution in Sec. IV, analyze its BER performance in Sec. V, and consider more general PNC-SA schemes in Sec. VI and Sec. VII. Sec. VIII concludes the paper.

II. PREVIOUS RESEARCH

Cadambe and Jafar [9] studied interference alignment for the k -pair interference channel. They demonstrated that such a system allows a DoF of $k/2$. Intuitively, if a single node pair can communicate at rate C , then the k pairs can simultaneously communicate at a rate of $C/2$ each. This discovery of *everyone gets half of the pie* has since spurred considerable interest in the wireless communication community. The underlying technique, aligning unwanted signals and contracting their dimensions perceived at a receiver, has spawned further applications [2], [7], [11]. In comparison with IA, SA does not necessarily differentiate between *wanted signals* and *unwanted interferences*. In SA, there is usually no single signal of focus, which requires demodulation in uncoded form. Gollakota *et*

al. [2] combined IA with IC in their IAC scheme, tailored for the scenario of multi-user MIMO transmission with limited receiver collaboration (Fig. 1). Li *et al.* [7] studied the application of IAC in more general, multi-hop wireless networks. The problem of appropriately applying IAC across a network is formulated and solved through a convex programming approach. Unlike PNC-SA, IA and IAC demodulate original packets but not their linear combinations. The IA phase in IAC can be viewed as a special case of PNC-SA precoding, and the IC phase is a special case of *decoding via remodulation* in PNC-SA (Sec. IV).

Zhang *et al.* [3] initiated the study of physical layer network coding, where entangled E-M signals are viewed as new, linearly combined packets. Focusing on the basic two-way relay channel, they showed how a PNC-demodulation algorithm can be implemented at the relay, to extract the digital version of two colliding data packets. PNC is new both in utilizing collided transmissions as useful encoded signals, and in demodulating a linear combination of transmitted packets. Zhang and Liew further studied PNC in the two-way relay channel with two antennas at the relay [12]. Compute-and-forward (C&F) [13] is a parallel work to PNC that also proposes to compute a function of the collided packets to further transmit in digital form. MIMO compute-and-forward [14] studies the theoretical achievable rates of a many-to-one transmissions, with multiple antennas at each node. Assuming all senders employ the same lattice code for modulation, the authors demonstrate that the idea of demodulating a linear combination can improve the achievable rates. They also point out the importance of optimal precoding at the senders, but leave such non-convex optimization as an open problem. In this paper, the optimal precoding problem of PNC-SA is formulated and solved in Sec. IV.

The technique of signal space alignment (SSA) is proposed by Lee *et al.* [5] in the context of the MIMO-Y channel, where three users multicast to each other with the help of a relay in the middle. The DoF of multi-link two way channels under SSA is studied by Lee *et al.* [15]. A more general model with $K \geq 3$ users is analyzed by Lee *et al.*, with amplify&forward and SSA combined for showing feasibility conditions in DoF [16], [17], and by Lee and Chun [18], which shows that the *DoF* is at least half the number of users K when the relay has $K - 1$ antennas. Liu and Yang apply SA in multi carrier CDMA systems, and propose a spectral-efficient SA signaling scheme for MC-CDMA two-way relay systems [19]. Park *et al.* [20] study power allocation and SA in the MIMO two-way relay channel setting, and propose a channel diagonalization scheme using generalized singular value decomposition.

The relation between the system DoF and the alignment constraints in IA has also been studied for the k -pair MIMO interference channel [8], [21], [22]. A series of recent work [22]–[24] design heuristic algorithms, often iterative in nature, for computationally efficient interference alignment solutions. Bresler *et al.* [8] prove that in a k -pair MIMO interference channel where every node has N antennas, the degrees of freedom of the system is tightly upper-bounded by $2N/(k+1)$.

III. MODEL AND NOTATION

We consider a multi-hop wireless network where each node is equipped with one or more antennas. Flat Rayleigh channel fading [1], [2], [7] is assumed, in which a signal experiences amplitude attenuation and phase shift through a channel. In each one-hop transmission, the sender transmits an N_t -dimensional signal vector \mathbf{x} , using the same carrier frequency. The receiver records an N_r -dimensional complex signal vector $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$. Here \mathbf{H} is the channel matrix of dimension $N_r \times N_t$, and each entry $h_{i,j}$ is the channel gain from Tx antenna i to Rx antenna j . All entries in \mathbf{H} , \mathbf{x} and \mathbf{y} are complex numbers. The length and direction of the vector representation of the complex number represent the amplitude (or amplitude attenuation) and phase (or phase shift) of the signal, respectively. An additive white Gaussian noise (AWGN) \mathbf{n} with zero mean and variance σ_n^2 is assumed.

We assume that full channel state information (CSI) is available, *i.e.*, each node knows the channel matrices of all adjacent (MIMO) links. A rich-scattering environment is assumed, such that channel matrices are of full rank.

The *trace* of a matrix \mathbf{A} is $Tr(\mathbf{A}) = \sum_i A_{ii}$. \mathbf{A}^* denotes the *conjugate transpose* of a matrix \mathbf{A} , obtained by transposing \mathbf{A} first, and then negating the imaginary component of each entry. The *Frobenius norm* of a matrix \mathbf{A} is $\|\mathbf{A}\|_F = (\sum_i \sum_j |A_{ij}|^2)^{\frac{1}{2}} = (Tr(\mathbf{A}^*\mathbf{A}))^{\frac{1}{2}}$. The *Euclidean norm* of a vector \mathbf{v} is $\|\mathbf{v}\| = (\sum_i |v_i|^2)^{\frac{1}{2}}$. A matrix \mathbf{A} is a *unitary matrix* if it satisfies $\mathbf{A}^* = \mathbf{A}^{-1}$. A unitary matrix \mathbf{A} preserves the Frobenius norm, *i.e.*, $\|\mathbf{A}\mathbf{B}\|_F = \|\mathbf{B}\|_F$.

Throughout this paper, matrices are denoted with boldface capital letters, vectors with boldface lowercase letters, and scalars with non-boldface letters.

IV. A DETAILED PNC-SA SCHEME DESIGN

A detailed PNC-SA solution that can work in the MIMO uplink in Fig. 2 includes two components: a precoding scheme at the clients, and a decoding scheme at the APs. We next present a detailed design of the two schemes.

A. PNC-SA Precoding at Clients

Let \mathbf{v}_1 and \mathbf{v}_2 be two 2×1 vectors that denote the target directions at AP1 for signal alignment and $\mathbf{v}_1 \neq \mathbf{v}_2$. We have the following *alignment constraint*:

$$\mathbf{H}_{11}\mathbf{a}_1 = \mathbf{H}_{21}\mathbf{a}_3 = \mathbf{v}_1, \quad \mathbf{H}_{11}\mathbf{a}_2 = \mathbf{H}_{21}\mathbf{a}_4 = \mathbf{v}_2$$

Another type of constraint in PNC-SA comes from the power budget available at each client, E_T . Let $\mathbf{A}_1 = (\mathbf{a}_1, \mathbf{a}_2)$ and $\mathbf{A}_2 = (\mathbf{a}_3, \mathbf{a}_4)$ be the 2×2 precoding matrices at clients 1 and 2, respectively. The *nodal power constraint* requires:

$$\|\mathbf{A}_1\|_F^2 = Tr(\mathbf{A}_1^*\mathbf{A}_1) \leq E_T,$$

$$\|\mathbf{A}_2\|_F^2 = Tr(\mathbf{A}_2^*\mathbf{A}_2) \leq E_T.$$

Optimal PNC-SA Precoding: Formulation

Given the two types of constraints, the client-side precoding aims to maximize the SNR of x_1+x_3 and x_2+x_4 , for demodulation at AP1, leading to the following optimal PNC-SA precoding problem:

Maximize $f(\mathbf{V}) = |\mathbf{v}_1^\dagger \cdot \mathbf{v}_2|$ (1)
 Subject to:

$$\begin{cases} \mathbf{H}_{11}\mathbf{A}_1 = \mathbf{V} = \mathbf{H}_{21}\mathbf{A}_2 & (2) \\ \|\mathbf{A}_1\|_F^2 \leq E_T & (3) \\ \|\mathbf{A}_2\|_F^2 \leq E_T & (4) \end{cases}$$

Here \mathbf{v}_1^\dagger is an orthogonal vector to \mathbf{v}_1 with equal length: if $\mathbf{v}_1 = (c_1, c_2)^T$, then $\mathbf{v}_1^\dagger = (c_2^*, -c_1^*)^T$, and $\mathbf{v}_1 \cdot \mathbf{v}_1^\dagger = 0$. The inner product $f(\mathbf{V}) = |\mathbf{v}_1^\dagger \cdot \mathbf{v}_2|$ targets two goals. The first is maximizing $|\mathbf{v}_1|$ and $|\mathbf{v}_2|$, for large received signal strength at AP1. The second is to make \mathbf{v}_1 and \mathbf{v}_2 as orthogonal as possible. The two goals together help maximize the SNR of detecting x_1+x_3 and x_2+x_4 .

PNC-SA Precoding: Solution

Solving the vector programming problem in (1) is in general computationally expensive [14], especially when the number of antennas is large. In particular, the classic water filling approach [1] does not directly apply, due to the extra alignment constraints in (2). We design an efficient approximate solution instead, which leads to a closed-form representation of the precoding scheme, and becomes optimal with two reasonable restrictions on the precoding solution space: (a) \mathbf{v}_1 and \mathbf{v}_2 are orthogonal. Having orthogonal signals for x_1+x_3 and x_2+x_4 is in general beneficial to their detection; (b) $\|\mathbf{v}_1\| = \|\mathbf{v}_2\|$, which is also reasonable since information contained in x_1+x_3 and in x_2+x_4 are equally important in general.

Given (a) and (b) above, \mathbf{V} can be scaled to a unitary matrix \mathbf{V}_0 with total power of 2. We compute how much power is required at each client, for its transmitted signals to fade into a unitary \mathbf{V}_0 at AP1. The power required at client 1 is:

$$\|\mathbf{A}_1\|_F^2 = \|\mathbf{H}_{11}^{-1}\mathbf{V}_0\|_F^2$$

Since \mathbf{V}_0 is unitary, it preserves the Frobenius norm of \mathbf{H}_{11}^{-1} , hence $\|\mathbf{A}_1\|_F^2 = \|\mathbf{H}_{11}^{-1}\|_F^2$. This significantly simplifies the precoding design, by decoupling joint precoding at both clients to independent precoding at each of them. Similarly, the power required at AP2 is $\|\mathbf{A}_2\|_F^2 = \|\mathbf{H}_{21}^{-1}\|_F^2$. Let

$$\xi = \max(\|\mathbf{H}_{11}^{-1}\|_F^2, \|\mathbf{H}_{21}^{-1}\|_F^2),$$

we can set the precoding matrices by first picking an arbitrary unitary matrix \mathbf{V}_0 , and set:

$$\mathbf{A}_1 = \sqrt{\frac{E_T}{\xi}}\mathbf{H}_{11}^{-1}\mathbf{V}_0, \mathbf{A}_2 = \sqrt{\frac{E_T}{\xi}}\mathbf{H}_{21}^{-1}\mathbf{V}_0.$$

The solution above satisfies both the alignment constraint in (2), and the power constraints in (3)-(4) (at least one of them is tight), and maximizes the objective function in (1) under the two simplifying assumptions in (a) and (b).

B. PNC-SA Demodulation at AP1

The digital packets x_1+x_3 and x_2+x_4 are demodulated at AP1 in two steps. Assuming BPSK modulation (+1 for 1, -1 for 0) at the clients, AP1 first detects ternary values in $\{-2, 0, +2\}$, then maps them to binary values in $\{0, 1\}$ through PNC mapping. We next discuss two detection schemes, ZF and ML, followed by PNC mapping. ZF and ML are representative detection methods in the literature: the former has low computational complexity, and the latter has optimal BER performance among all detection schemes.

ZF Detection. Conceptually, AP1 can view x_1+x_3 and x_2+x_4 as two variables, and solve them through the two received signals at its antennas. ZF detection does so by projecting the combined signals to a direction orthogonal to x_2+x_4 (or x_1+x_3), for detecting x_1+x_3 (or x_2+x_4). ZF is particularly well-suited for PNC-SA, if we have restricted \mathbf{v}_1 and \mathbf{v}_2 to be orthogonal, as described in Sec. IV-A. The ZF projection matrix is a scaled conjugate transpose of \mathbf{V}_0 selected in Sec. IV-A, $\sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^*$:

$$\begin{aligned} \tilde{\mathbf{y}} &= \sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^*\mathbf{y} \\ &= \sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^*(\mathbf{H}_{11}\mathbf{A}_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbf{H}_{21}\mathbf{A}_2 \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + \mathbf{n}) \\ &= \sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^* \left(\sqrt{\frac{E_T}{\xi}}\mathbf{V}_0 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \sqrt{\frac{E_T}{\xi}}\mathbf{V}_0 \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + \mathbf{n} \right) \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + \sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^*\mathbf{n} \\ &= \begin{pmatrix} x_1+x_3 \\ x_2+x_4 \end{pmatrix} + \tilde{\mathbf{n}} \end{aligned}$$

Since the projection is linear, the projected noise $\tilde{\mathbf{n}} = \sqrt{\frac{\xi}{E_T}}\mathbf{V}_0^*\mathbf{n}$ is still AWGN.

ML Detection. Alternatively, we can apply the *a posteriori* method of ML detection. ML infers which source vector is most likely to have been transmitted, based on receiver side information. ML has a higher computational complexity than ZF, but provides optimal BER performance.

A salient difference between a standard ML scheme and ML for PNC-SA is that the former ‘guesses’ what’s transmitted at each Tx antenna, while the latter ‘guesses’ the most probable linear combinations of the transmitted data. Equivalently, ML for PNC-SA views the multi-user MIMO channel from both clients to AP1 as a *virtual* 2×2 MIMO channel, with channel matrix $\sqrt{\frac{E_T}{\xi}}\mathbf{V}_0$ and ternary modulation, and detects the desired linear combination as:

$$\begin{pmatrix} x_1+x_3 \\ x_2+x_4 \end{pmatrix} = \arg \min_{\mathbf{x} \in \{-2, 0, 2\}^2} \|\mathbf{y} - \sqrt{\frac{E_T}{\xi}}\mathbf{V}_0\mathbf{x}\|$$

PNC Mapping. While BPSK demodulation simply maps from $\{-1, 1\}$ to $\{0, 1\}$, PNC demodulation maps from $\{+2, 0, -2\}$ to $\{0, 1\}$ [3]. The basic rule is: +2 and -2 map to 0, and 0 maps to 1. The intuition is that when +2 (-2) is seen, x_1 and x_3 (or x_2 and x_4) must have both been +1 (-1), and x_1+x_3 (or x_2+x_4) should be 0. Otherwise, x_1+x_3 (or x_2+x_4) should be 1. In the case of ZF detection, one may merge the ternary detection and ternary-to-binary mapping into a single step. Based on a maximum posterior probability criterion, Zhang and Liew [3] derived the following optimal decision rule for such direct mapping: map values between $-1 - \alpha$ and $1 + \alpha$ to 1, and other values to 0, for $\alpha = \frac{\sigma_n^2}{2} \ln(1 + \sqrt{1 - e^{-4/\sigma_n^2}})$.

C. PNC-SA Decoding at AP2

After receiving x_1+x_3 and x_2+x_4 from AP1, AP2 has accumulated four packets, two digital ones from AP1, two analog ones from its own antennas:

$$\begin{cases} \begin{pmatrix} x_1 + x_3 \\ x_2 + x_4 \end{pmatrix} \\ y' \triangleq \mathbf{H}_{12}\mathbf{A}_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbf{H}_{22}\mathbf{A}_2 \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} + \mathbf{n} \end{cases}$$

We describe two methods below for AP2 to solve the above four equations: adapted ML decoding, and decoding via remodulation. The former provides better BER performance, while the latter scales better with the source symbol space and the number of antennas.

Adapted ML Decoding. AP2 traverses all possible combinations of (x_1, x_2, x_3, x_4) . Before applying the normal minimum distance criterion in ML (selecting the source vector whose faded version has the minimum geometric distance from the received signals), it first filters out the enumerated vectors that are not in agreement with the known values for x_1+x_3 and x_2+x_4 . Consequently, adapted ML reduces the computational complexity of ML by a factor of 2^{N_r} , or a factor of 4 for the uplink in Fig. 2.

Decoding via Remodulation. Alternatively, AP2 may first re-construct the analog version of x_1+x_3 and x_2+x_4 after modulation. Next, AP2 can apply low-complexity MIMO decoding methods (*e.g.*, ZF or MMSE-SIC [1]) to decode x_1, \dots, x_4 as at a 4×4 MIMO receiver. The IC technique, as in IAC, is essentially decoding via remodulation in its simplest form, where only subtracting the remodulation of an uncoded packet is performed.

D. Discussions

PNC-SA provides full flexibility in precoding. Unlike IA or IAC, it places no restrictions on the precoding matrix, except that each sender can only encode locally available data. PNC-SA also opens new solution spaces for fully exploring the spatial diversity of a MIMO network, augmenting its achievable capacity region. This will be further demonstrated in Sections VII-A, VII-B and VII-C. PNC alone can be viewed as a special case of PNC-SA, where each node has a receive diversity of 1, and SA degrades into signal synchronization. IAC can be viewed as a special case of PNC-SA, which further restricts the way SA is performed, precludes the application of PNC demodulation, and applies decoding via remodulation in its simplest form only.

The technique of PNC-SA is independent of the modulation scheme. We have referred to BPSK for ease of exposition. Similar to PNC [3], PNC-SA can be applied with more sophisticated modulation schemes such as QPSK or QAM (quadrature amplitude modulation) 16.

The precoding optimization in Sec. IV-A in general under-utilizes the available power at one of the clients, for exact signal matching between x_1 (x_2) and x_3 (x_4). It is possible to relax exact matching, and fully utilize all available power. An adapted PNC detection scheme will be required, with 4 instead of 3 possible values for combined signal strength. The current precoding optimization focuses on SNR at AP1 only. As a more comprehensive solution, one may formulate a global optimization that further considers the signal reception

at AP2. We leave such a formulation and its solution as future work.

Our optimization in (1)-(4) considers a Tx-side precoding scheme with fixed ZF decoding and the Rx-side. Such a scheme in general delivers sub-optimal performance, especially when the channel matrix is ill-conditioned. An interesting future direction, as pointed by one of the anonymous referees, is to consider a joint precoder-decoder design, for maximizing the SNR of the network coding based scheme.

V. BER ANALYSIS AND COMPARISON

We next analyze the BER performance of PNC-SA, and compare it with the BER of IAC. We first review the BER analysis of a general ML decoder, which will be helpful in analyzing the BER of PNC-SA and IAC.

A. BER of ML Detection

For a $N_r \times N_t$ MIMO channel. ML Detection searches for a source vector that was most likely to have been transmitted, based on information available at the receiver side:

$$\tilde{\mathbf{x}}_{\text{ml}} = \arg \max_{\tilde{\mathbf{x}}_i} p(\mathbf{y}|\mathbf{H}, \tilde{\mathbf{x}}_i) = \arg \min_{\tilde{\mathbf{x}}_i} \|\mathbf{y} - \mathbf{H}\tilde{\mathbf{x}}_i\|^2$$

where the search space of the $N_t \times 1$ source vector $\tilde{\mathbf{x}}_i$ has a size of M^{N_t} , M being the modulation alphabet cardinality. For flat Rayleigh fading with AWGN, the pairwise error probability (PEP), *i.e.*, the probability that MLD mistakenly outputs $\tilde{\mathbf{x}}_k$ when a different source vector $\tilde{\mathbf{x}}_i$ is transmitted, is ([25], Ch 4.2.2)

$$Pr(\tilde{\mathbf{x}}_i \rightarrow \tilde{\mathbf{x}}_k) = Q\left(\sqrt{\frac{\|\mathbf{H}(\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_k)\|^2}{2\sigma_n^2}}\right) \quad (5)$$

Function Q computes the area under the tail of a Gaussian PDF. Using Boole's inequality, one can derive the average MIMO vector error probability ([25], Ch 4.2.1):

$$Pr_s \leq \frac{1}{M^{N_t}} \sum_{\tilde{\mathbf{x}}_i} \sum_{\tilde{\mathbf{x}}_k \neq \tilde{\mathbf{x}}_i} Pr(\tilde{\mathbf{x}}_i \rightarrow \tilde{\mathbf{x}}_k), \quad (6)$$

and, an approximation on BER can be found with

$$Pr_b \approx Pr_s / (N_t \log_2 M). \quad (7)$$

B. BER Analysis of PNC-SA

The analysis of the BER performance of PNC-SA involves two phases. In phase one, we study the BER at AP1, for decoding x_1+x_3 and x_2+x_4 . In phase two, we study the BER at AP2, using adapted ML for decoding x_1, \dots, x_4 .

BER at AP1. As discussed in Sec. IV-B, AP1 can demodulate x_1+x_3 and x_2+x_4 by applying ML detection over a virtual 2×2 MIMO channel. Let $\mathbf{c} = (c_t, c_b)^T$, where $c_t = x_1 + x_3$ and $c_b = x_2 + x_4$ are in the $\{-2, 0, 2\}$ domain, before PNC mapping. Let \mathbf{c}_i and \mathbf{c}_k be two possible 2×1 transmit vectors, with $i, k \in \{1, \dots, 9\}$. Assume \mathbf{c}_i is transmitted; from (5), the probability that AP1 incorrectly outputs \mathbf{c}_k is:

$$Pr(\mathbf{c}_i \rightarrow \mathbf{c}_k) = Q\left(\sqrt{\frac{E_T/\xi \|\mathbf{V}_0\|^2 \lambda_{ik}}{2\sigma_n^2}}\right) = Q\left(\sqrt{\frac{\lambda_{ik} \rho_{S1}}{2}}\right),$$

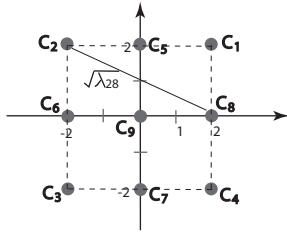


Fig. 3. Constellation diagram for PNC-SA, at AP1.

where $\lambda_{ik} = (\mathbf{c}_i - \mathbf{c}_k)^T(\mathbf{c}_i - \mathbf{c}_k)$, and ρ_{S1} is SNR at AP1.

Let's define constellation points $\mathbf{c}_1, \dots, \mathbf{c}_9$ as shown in Fig. 3. Assuming 0 and 1 are equally likely to appear in the source packets, the ternary values in $\{-2, 0, 2\}$ appear in \mathbf{c} with probabilities of 25%, 50%, and 25%, respectively. As a result, $P(\mathbf{c}_1) = P(\mathbf{c}_2) = P(\mathbf{c}_3) = P(\mathbf{c}_4) = 1/12$; $P(\mathbf{c}_5) = P(\mathbf{c}_6) = P(\mathbf{c}_7) = P(\mathbf{c}_8) = 1/8$; $P(\mathbf{c}_9) = 1/6$. AP1 wishes to demodulate the digital bits $\mathbf{d} = (d_t, d_b)^T$, where $d_t = x_1 + x_3$ and $d_b = x_2 + x_4$. Thus, $Pr(\mathbf{c}_i \rightarrow \mathbf{c}_k) = 0$ when both \mathbf{c}_i and \mathbf{c}_k are in $(\pm 2, \pm 2)^T$. In other words, judging -2 to be $+2$ or vice versa does not lead to an error in \mathbf{d} . The average vector error probability for \mathbf{d} is

$$Pr_s(\mathbf{d}) = 4P(\mathbf{c}_1) \sum_{i=5}^9 Pr(\mathbf{c}_1 \rightarrow \mathbf{c}_i) + 4P(\mathbf{c}_5) \sum_{i \neq 5} Pr(\mathbf{c}_5 \rightarrow \mathbf{c}_i) + P(\mathbf{c}_9) \sum_{i=1}^8 Pr(\mathbf{c}_9 \rightarrow \mathbf{c}_i)$$

BER at AP2. Consider applying adapted ML to decode x_1, \dots, x_4 at AP2. We first study the case that $x_1 + x_3$ and $x_2 + x_4$ from AP1 are correct. We only need to search over source vectors that agree with the given $x_1 + x_3$ and $x_2 + x_4$ values. Under BPSK modulation, there are 4 such vectors, with dimension 4×1 . Let $\tilde{\mathbf{x}}_i$ and $\tilde{\mathbf{x}}_k$ ($i, k \in \{1, \dots, 4\}$) be two distinct vectors among the four. Assume $\tilde{\mathbf{x}}_i$ is transmitted. By (5), the probability that AP2 outputs $\tilde{\mathbf{x}}_k$ erroneously equals:

$$Pr(\tilde{\mathbf{x}}_i \rightarrow \tilde{\mathbf{x}}_k | \mathbf{d}_c) = Q\left(\sqrt{\frac{\lambda'_{ik} \rho_{S2}}{2}}\right).$$

Here $\lambda'_{ik} = (\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_k)^T(\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_k)$, and ρ_{S2} is the SNR at AP2. Let \mathbf{d}_c and \mathbf{d}_w denote the events that AP2 receives the correct and wrong data in \mathbf{d} from AP1, respectively. The average vector error probability is:

$$Pr_s(\tilde{\mathbf{x}} | \mathbf{d}_c) = \frac{1}{4} \sum_{i=1}^4 \sum_{k=1, k \neq i}^4 Q\left(\sqrt{\frac{\lambda'_{ik} \rho_{S2}}{2}}\right).$$

Further including the case that $x_1 + x_3$ and $x_2 + x_4$ transmitted from AP1 contain errors, we have $Pr_s(\tilde{\mathbf{x}}) = Pr_s(\tilde{\mathbf{x}} | \mathbf{d}_c) Pr_s(\mathbf{d}_c) + Pr_s(\tilde{\mathbf{x}} | \mathbf{d}_w) Pr_s(\mathbf{d}_w)$. When information from AP1 is wrong, AP2 outputs a wrong vector with probability 1, *i.e.*, $Pr_s(\tilde{\mathbf{x}} | \mathbf{d}_w) = 1$. Therefore, the vector error rate of the overall PNC-SA scheme is:

$$Pr_s(\tilde{\mathbf{x}}) = Pr_s(\tilde{\mathbf{x}} | \mathbf{d}_c)(1 - Pr_s(\mathbf{d})) + Pr_s(\mathbf{d}). \quad (9)$$

The probability of more than two bit errors happening in the same vector can be ignored. In adapted ML decoding, if there is a decoding error in $(x_1, x_3)^T$, it must have been the

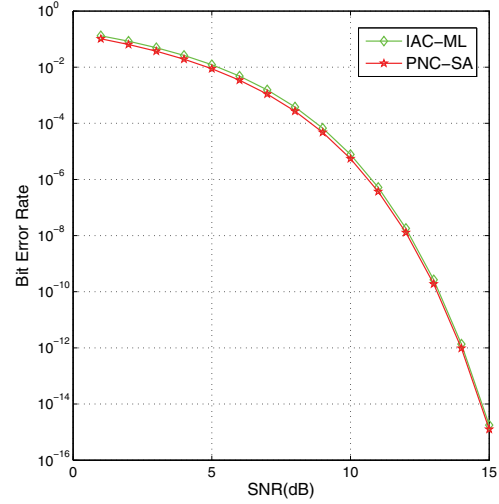


Fig. 4. BER performance comparison: PNC-SA vs IAC.

case that $x_1 + x_3$ was not received in its correct form from AP1, and one of x_1 and x_3 is decoded correctly and the other incorrectly. Similar for $(x_2, x_4)^T$. Consequently, when an error occurs in the vector (x_1, x_2, x_3, x_4) , half of its bits are still correct. Thus the average bit error probability is half of the vector error probability:

$$Pr_b(\tilde{\mathbf{x}}) = Pr_s(\tilde{\mathbf{x}})/2. \quad (10)$$

C. Comparison of BER Performance

The analysis of BER performance for IAC is also carried out in two steps (at AP1 and AP2), similar to the case of PNC-SA in Sec. V-B. Below we omit the intermediate steps and provide the result only:

$$Pr_b(\mathbf{x}) = \frac{1}{8} \sum_{i=1}^4 \sum_{k \neq i}^4 Q\left(\sqrt{\frac{\lambda_{ik}^{I2} \rho_{I2}}{2}}\right) (1 - Pr_s(\mathbf{x}_1)) + Pr_s(\mathbf{x}_1)$$

where $\lambda_{ik}^{I1} = (\mathbf{e}_i - \mathbf{e}_k)^T(\mathbf{e}_i - \mathbf{e}_k)$, ρ_{I1} is SNR at AP1, and

$$Pr_s(x_1) = 4P(\mathbf{e}_1) \sum_{i=4}^6 Pr(\mathbf{e}_1 \rightarrow \mathbf{e}_i) + 2P(\mathbf{e}_2) \sum_{i=4}^6 Pr(\mathbf{e}_2 \rightarrow \mathbf{e}_i),$$

$$Pr(\mathbf{e}_i \rightarrow \mathbf{e}_k) = Q\left(\sqrt{\frac{\lambda_{ik}^{I1} \rho_{I1}}{2}}\right).$$

Fig. 4 shows the comparison of the BER performance of PNC-SA and IAC, under varying SNR levels. The BER of PNC-SA is always slightly better than that of IAC, under the same SNR at the receiver's antennas.

VI. GENERAL PNC-SA SCHEMES

A. General Degrees-of-Freedom of PNC-SA

The analysis in Sec. IV focuses on a 2-client 2-AP scenario, where each node is equipped with two antennas. The DoF of

TABLE I
AN $N \times N \times M$ PNC-SA ALGORITHM

(1) initialize the set of encoded packets accumulated from SA:
 $\Delta \leftarrow \phi$

(2) for each AP_i :
 $\Theta = \{x_1, \dots, x_X\}$
for each antenna k at AP_i , $1 \leq k \leq M$:
 Choose $\Lambda \subseteq \Theta$, such that:
 $y \triangleq \sum_{x \in \Lambda} x \notin SPAN(\Delta)$
 $\Theta \leftarrow \Theta - \Lambda$
 $\Delta \leftarrow \Delta \cup \{y\}$

if no such Λ exists :
if $i = M - 1$: terminate and declare failure
else:
 Allocate Θ to the rest of antennas in AP_i evenly
 proceed to next AP

if $|\Delta| \geq X - M$: go to (3)
 \rightarrow feasible signal alignment scheme obtained

(3) Compute precoding vectors $\mathbf{a}_1, \dots, \mathbf{a}_{NM}$, for desired SA
 computed in Step (2)
 \rightarrow feasible precoding scheme obtained

(4) AP_N collects all $X - M$ digital packets, and combine them
 with its M analog signals for decoding all X original signals,
 using Adapted ML or Decoding via Remodulation.
 \rightarrow source signals decoded

PNC-SA apparently depends on the number of client-AP pairs, as well as the number of antennas each node has. We now study such a dependence, and provide a constructive proof for an inner-bound on the DoF of general PNC-SA.

Consider a general $N \times N \times M$ uplink communication scenario with N clients on the Tx side and N APs on the Rx side, each equipped with M antennas. Each client has up to M packets for precoding and transmission. The APs again co-locate within the same base station, and are inter-connected with Ethernet cables feasible for transmitting digital packets.

Table I shows an algorithm for achieving a DoF of X , i.e., simultaneously transmitting X source signals from the clients to the APs. The algorithm either succeed with a signal alignment scheme and a precoding scheme discovered, or fails to achieve DoF X and declares failure. The algorithm is heuristic in nature and is not always optimal, in the sense that it does not guarantee finding a feasible solution whenever DoF is above X . We leave the question of computing the exact DoF of a general $N \times N \times M$ system, which involves both wireless and wireline channels, as future work.

As shown in Table I, the general PNC-SA algorithm first initializes the set Δ , which will accumulate digital packets to be decoded from APs where signal alignment successfully happens, to an empty set ϕ , in Step (1). Step (2) contains a double loop and constitutes the core of the algorithm that searches for a feasible signal alignment scheme. A valid signal alignment solution must satisfy two constraints: (i) the set of encoded packets decoded from signal alignment, Δ , has cardinality $X - M$, so that together with analog signals at AP_N , sufficient information is available for decoding all X source

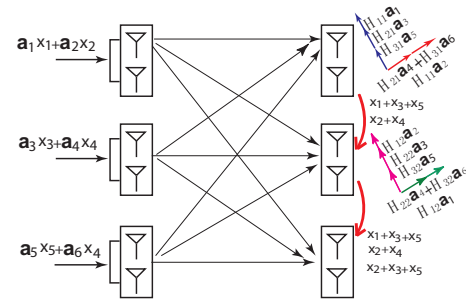


Fig. 5. The DoF of PNC-SA is 5 in a $3 \times 3 \times 2$ system, by Theorem 6.1. A detailed signal alignment scheme can be obtained by solving the alignment equations.

signals. (ii) Different encoded signals to be demodulated at the same AP do not involve a common source signal. For example, if AP_i can not demodulate both $x_1 + x_3$ and $x_2 + x_3$, since that will impose an infeasible requirement on the precoding vectors.

In the outer **for** loop that iterates over every AP, we first initialize the set of source signals not aligned to a direction yet, Θ , to the full set. Then the inner **for** loop attempts to take variables from Θ for constructing encoded signals that is linearly independent to the set Δ . Here $SPAN(\Delta)$ is the linear subspace spanned by vectors in Δ . Upon success, we update Δ and Θ accordingly. Upon failure, if the current AP is the last AP possible for signal alignment, the algorithm terminates and declares failure; otherwise, the algorithm allocates the remaining source signals to the other directions at the current AP in an arbitrary manner, e.g., as even as possible. Once the cardinality of Δ reaches $X - M$, the double **for** loop terminates, and the algorithm jumps to Step (3), precoding vector computation. Since the signal alignment computed in Step (2) satisfies conditions (i) and (ii), Step (3) is guaranteed to succeed.

As an example input to our general PNC-SA algorithm, Fig. 5 shows a 3-client 3-AP system with 2 antennas per node, where the DoF is 5. A sample precoding and signal alignment solution for simultaneously transmitting 5 source signals is illustrated.

B. PNC-SA with QPSK Modulation

So far we introduced PNC-SA decoding and its BER performance by assuming BPSK modulation. The technique of PNC-SA is, however, independent of the modulation scheme. We referred to BPSK simply for ease of exposition. Similar to PNC, PNC-SA can be applied with more sophisticated modulation schemes such as QPSK or 16QAM. In this section, we will discuss in detail how PNC-SA works with the QPSK modulation scheme.

Quadrature Phase-Shift Keying (QPSK) is a digital modulation scheme that conveys data by changing the phase of a reference carrier wave. QPSK modulates by changing the phase of the in-phase (I) carrier from 0° to 180° and the quadrature-phase (Q) carrier between 90° and 270° . As shown in the constellation diagram in Fig. 6, QPSK uses four points around a circle to represent digital data. With four phases, QPSK can encode two bits per symbol.

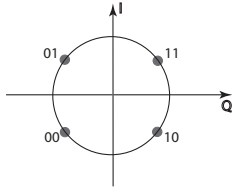


Fig. 6. Constellation diagram for QPSK.

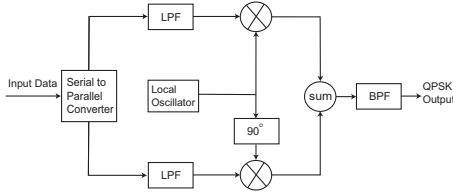


Fig. 7. Block diagram of a QPSK transmitter

Fig. 7 shows a block diagram of a typical QPSK transmitter. The input binary data stream is split into the in-phase and quadrature-phase components by a serial to parallel converter. Then the two bit streams are fed to two orthogonal modulators after passing through the low pass filter (LPF). In the last step, the two modulated bit streams are summed and fed to the band pass filter (BPF) for producing the QPSK output.

When PNC-SA works with QPSK modulation at the client side in Fig. 2, each transmitted signal includes two substreams: the in-phase stream and the quadrature-phase stream. However, the client actually transmits the sum of the in-phase and quadrature-phase waves, which is a composite wave with the same frequency. Furthermore, when we align x_1 and x_3 to the same direction at AP1, it is the composite signal, rather than the in-phase or quadrature-phase signal, that is being aligned. Recall that *direction* here refers to a signal’s encoding vector when received at the Rx node. When we restrict the alignment directions at AP1, \mathbf{v}_1 and \mathbf{v}_2 , to be orthogonal, the directions of the composite signals become orthogonal.

Now consider PNC-SA demodulation at AP1. Because the in-phase and quadrature-phase components of a combined QPSK signal propagate through the same fading channel, they arrive with the same amplitude attenuation and phase shift, and hence the I and Q components are still orthogonal to each other. Therefore, if two composite signals are aligned to the same direction, their I and Q components are also aligned to the same direction. With ZF detection, we can first separate the two combined QPSK signals by projection, then apply QPSK demodulation and PNC-mapping to obtain the I and Q substreams that together form the digital version of $(x_1 + x_3)$ and $(x_2 + x_4)$. Alternatively, we can also apply ML detection to “guess” the most probable linear combinations of the transmitted data.

Similar to the case of QPSK, PNC-SA can be adapted to work with more complex schemes such as 16QAM. The higher the data rate that a modulation scheme can provide, the worse its BER performance is. There is always a tradeoff between the BER performance and the raw data rate.

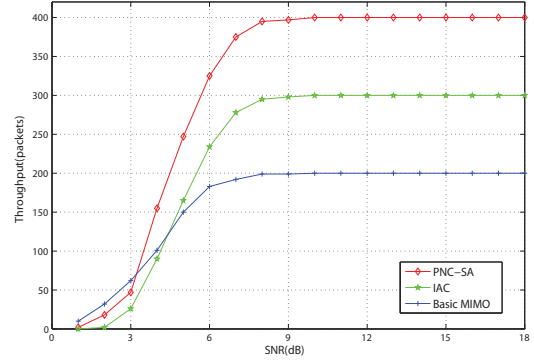


Fig. 8. Packet-level throughput for multi-AP uplink communication, PNC-SA vs. IAC vs. MIMO alone.

VII. GENERAL APPLICATIONS OF PNC-SA AND PACKET-LEVEL THROUGHPUT

Applications of PNC-SA in wireless routing can be broad, and are not restricted to cases where receivers have limited collaboration (Fig. 2). In this section, we first present Matlab simulation results on packet level comparisons between PNC-SA and alternative solutions for the uplink scenario in Fig. 2. We then extend the discussions to more general applications of PNC-SA, for information exchange, unicast and multi-cast/broadcast.

Fig. 8 shows the comparison of packet-level throughput achieved by PNC-SA, IAC and MIMO, respectively. We assume a synchronized environment where nodes transmit packets in a total of 100 rounds. During each round, PNC-SA, IAC and MIMO transmit 4, 3, and 2 raw packets of 50% bits each, respectively. At the receiver side, an error detection code helps identify bit errors. A packet received with 1 or more bits in error is discarded and not counted towards total throughput. BER is computed from SNR as discussed in Sec. V. The noise level is equal at all nodes.

Fig. 8 shows that at high SNR (> 9), the ratio of throughput achieved by the three schemes converges to 4 : 3 : 2, with PNC-SA performing the best. As SNR decreases, the gap between PNC-SA and IAC slightly increases, due to the slightly better SNR-BER performance of PNC-SA, as shown in Fig. 4. It is interesting to note that at the very low SNR regime, basic MIMO actually performs the best, because basic MIMO strikes a better balance between system DoF and error rate at the very low SNR regime, leading to a better BER performance.

A. PNC-SA for Info Exchange

Fig. 9 shows the two-way relay channel in a wireless network, where Alice and Bob wish to exchange data packets with the help of a relay [12], [26]. Each node is equipped with 3 antennas. Transmitting simultaneously, Alice and Bob can align their six signals to three common directions at the relay. The relay then demodulates $x_1 + x_4$, $x_2 + x_5$ and $x_3 + x_6$, and broadcasts them to both Alice and Bob. Alice and Bob each subtract their known packets from the three combined signals received, and apply normal demodulation to recover the other three packets.

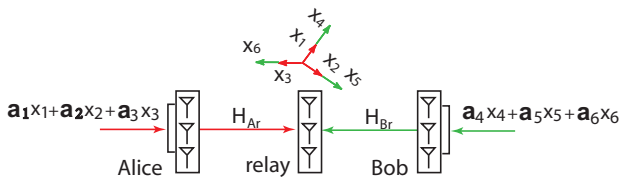


Fig. 9. PNC-SA with three antennas per node. Here and in the rest of the paper, we label an aligned direction with the corresponding signal instead of its vector direction, for simplicity. For example, the direction of $\mathbf{H}_{11}\mathbf{a}_1$ is simply labelled as x_1 .

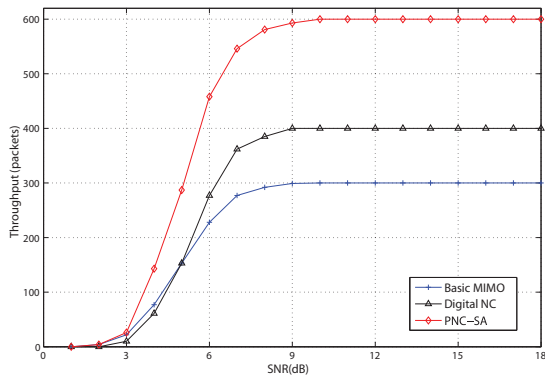


Fig. 10. Packet-level throughput for information exchange, PNC-SA vs. DNC vs. MIMO alone.

With PNC-SA, 6 packets can be exchanged in 2 time slots. Without PNC-SA, it takes 3 time slots with digital network coding, and 4 time slots with no coding at all [26]. Without SA, PNC alone does not fully exploit the full DoF of such a MIMO network. For example, Zhang and Liew [12] studied the utilization of multiple antennas at the relay, by combining its received signals for generating a single encoded packet, for better BER.

We can see that the application of PNC-SA is not limited to scenarios with limited receiver collaboration; nor is it limited to 2 antennas per node. Examples shown in this paper can all be generalized to work with 3 or more antennas per node.

Fig. 10 shows the packet-level throughput comparison between PNC-SA, DNC and basic MIMO. Here the system is run for 200 time slots, with normalized length for a SISO channel capacity to be 50 bits. We can observe that at high SNR, the throughput ratio converges to 6 : 4 : 3, with PNC-SA leading the alternatives. At low SNR, DNC performs the worst. The main reason is that DNC needs to succeed in all transmissions in 3 time slots for successful packet reception and decoding, while PNC-SA and MIMO only need 2 time slots each.

B. PNC-SA for Unicast Routing

PNC-SA for Cross Unicasts

Fig. 11 depicts two unicast sessions, from S_1 to T_1 and from S_2 to T_2 , whose routes intersect at a relay. Each sender cannot directly reach its intended receiver, and needs to resort to the help of the relay node in the middle.

With PNC-SA, the two senders can transmit simultaneously, aligning the signals for reception at the relay: x_1 is aligned

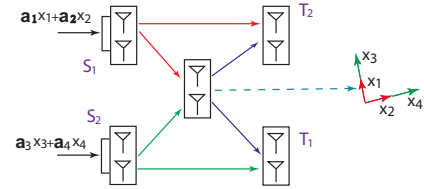


Fig. 11. PNC-SA with PNC performed at the relay node in the middle.

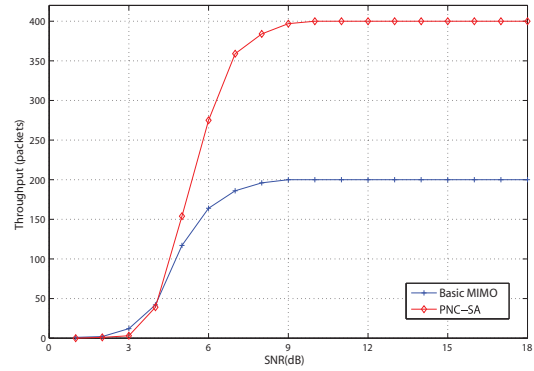


Fig. 12. Packet-level throughput for cross unicasts, PNC-SA vs. MIMO alone.

with x_3 , and x_2 with x_4 . The relay decodes and broadcasts x_1+x_3 and x_2+x_4 . Only 3 transmissions in 2 time slots are required. T_1 can first decode x_3 and x_4 overhead from S_1 , and then combine them with x_1+x_3 and x_2+x_4 to recover x_1 and x_2 . T_2 recovers x_3 and x_4 similarly.

Without any coding, it takes 4 transmissions in 4 time slots to send 2 packets in each session: each sender transmits once (using both antennas), and the relay transmits twice. With DNC, it takes 3 transmissions in 3 time slots — the relay can transmit just once, broadcasting two encoded packets.

The PNC-SA precoding optimization discussed in Sec. IV-A still applies here. SA enables PNC in this MIMO network, and PNC further enables demodulate-and-forward at the relay, which provides an alternative to amplify-and-forward for cooperative communication [27]. In general multi-session unicast routing, such a cross-unicast topology can be applied as a gadget, embedded into larger unicast sessions [26].

Fig. 12 shows packet-level throughput comparison between PNC-SA and a basic MIMO solution. Again the network is run for 200 time slots, with the same node transmission capacity and packet lengths as previously assumed. At high SNR, the throughput gap between PNC-SA and MIMO is a factor of 2, confirming the analysis above. As SNR decreases, however, MIMO catches up with PNC-SA and eventually outperforms, due to its better SNR-BER performance. This suggests that a good design of error-correction code in combination with PNC-SA is important at the low SNR regime.

The Zig-Zag Unicast Flow: PNC Meets DNC

Existing literature on the application of network coding in wireless routing often focuses on identifying local gadgets, such as the two-way relay channel and the cross-unicast topology [3], [26]. These gadgets usually involve multiple unicast sessions with reverse or crossing routes. It is often

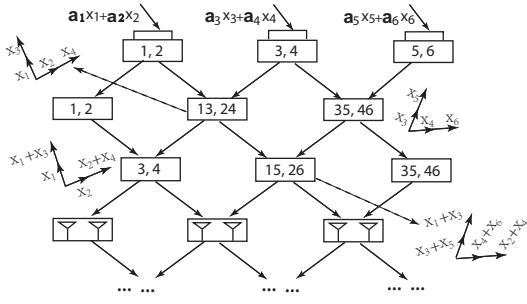


Fig. 13. The zig-zag unicast flow using PNC-SA. Here 35, 46 in a node represents x_3+x_5 and x_4+x_6 . The first row transmits 6 packets simultaneously. The signals are aligned at the second row for demodulating (x_5, x_6) , (x_3+x_5, x_4+x_6) and (x_1+x_3, x_2+x_4) . In the odd (even) rows, the left-most (right-most) node receive from one sender in the previous row only, without PNC.

believed that network coding provides little benefit to a single unicast session, when links are lossless [28], [29]. We present an application of PNC-SA, where PNC and DNC work in concert to enable a new, efficient wireless unicast routing algorithm.

Consider a large wireless sensor network with two antennas per sensor, where information is to be routed from the top of the network to the bottom [30]. What multi-hop unicast routing scheme can we use, to achieve a high throughput? Fig. 13 illustrates a PNC-SA based solution: a *zig-zag unicast flow*.

The zigzag solution routes k parallel data streams side by side, employing k nodes for transmission at each row ($k=3$ in Fig. 13). The resulting unicast flow exhibits a zigzag topology. The following theorem shows that the packets at each row can be used to recover the $2k$ original packets.

Theorem 7.1. *At each row in the zigzag unicast flow, the $2k$ data packets are linearly independent, and can be used to recover the original packets x_1, \dots, x_{2k} .*

Proof: We prove the theorem using a row-by-row induction. As the basis, the $2k$ packets at the first row are the original ones, and are independent. Assume the packets at row i , y_1, \dots, y_{2k} , are independent. Number the nodes in each row from left to right. Without loss of generality, assume the left-most node (node 1) in row $i+1$ receives packets without PNC coding. Packets at node 1 in row $i+1$ are y_1 and y_2 . Packets at node 2 in row $i+1$ are y_1+y_3 and y_2+y_4 and can be used to further recover y_3 and y_4 . Similarly, each node $j \in [2 \dots k]$ in row $i+1$ possesses packets that can be used to further recover y_{2j-1} and y_{2j} . In conclusion, packets at row $i+1$ can be used to recover all packets in row i . Since the latter are linearly independent, so are the former. \square

The table below lists the packets received by nodes at each row, for $k=3$. The intra-row linear independence can be verified. It is also interesting to observe that after every 7 rows, the 6 data packets in routing return to uncoded form.

Compared to a basic single-chain unicast solution, the zigzag flow represents a k -fold throughput gain. Unlike traditional multi-path wireless routing, the k parallel data streams in the zigzag flow do not need to be spatially far apart to avoid interference, and is in that sense more practical to deploy. The rationale behind the zigzag structure guarantees that a node at

row	node 1	node 2	node 3
0	x_1, x_2	x_3, x_4	x_5, x_6
1	x_1, x_2	x_1+x_3, x_2+x_4	x_3+x_5, x_4+x_6
2	x_3, x_4	x_1+x_5, x_2+x_6	x_3+x_5, x_4+x_6
3	x_3, x_4	$x_1+x_3+x_5,$ $x_2+x_4+x_6$	x_1+x_3, x_2+x_4
4	x_1+x_5, x_2+x_6	x_5, x_6	x_1+x_3, x_2+x_4
5	x_1+x_5, x_2+x_6	x_1, x_2	$x_1+x_3+x_5,$ $x_2+x_4+x_6$
6	x_5, x_6	x_3+x_5, x_4+x_6	$x_1+x_3+x_5,$ $x_2+x_4+x_6$
7	x_5, x_6	x_3, x_4	x_1, x_2

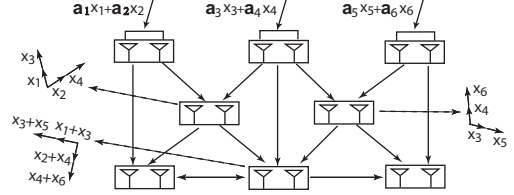


Fig. 14. Multicast from top layer to bottom layer. PNC-SA doubles throughput.

the border obtains data without PNC, which can be used to bootstrap the decoding process along that row.

C. PNC-SA for Multicast/Broadcast Routing

Network coding is naturally well-suited for multicast and broadcast routing in wireless networks. The local broadcast nature of omnidirectional antennas is well suited for simultaneously transmitting an encoded packet to multiple receivers. PNC-SA extends such benefit of DNC to information dissemination in MIMO networks.

Multi-Sender Multicast

Fig. 14 depicts a multi-sender multicast in an 8-node MIMO network. The 3 top nodes are senders, and the 3 bottom nodes are receivers. Each sender wishes to multicast to all receivers. As another natural fusion of PNC and DNC, the application of PNC-SA here doubles the achievable multicast throughput.

With PNC-SA, 6 packets can be multicast to all receivers in 4 time slots. (i) The three senders align their six signals at the two relays in the middle, such that they can successfully demodulate $\{x_1+x_3, x_2+x_4\}$ and $\{x_3+x_5, x_4+x_6\}$, respectively. At the same time, the three receivers obtain $\{x_1, x_2\}$, $\{x_3, x_4\}$ and $\{x_5, x_6\}$, respectively. (ii) The two relays transmit x_1+x_3, x_2+x_4 , respectively, simultaneously. Their signals are aligned so the middle receiver can demodulate $x_1+x_3+x_3+x_5 = x_1+x_5$ and $x_2+x_4+x_4+x_6 = x_2+x_6$. From left to right, the three receivers accumulate $\{x_1, x_2, x_3, x_4\}$, $\{x_3, x_4, x_1+x_5, x_2+x_6\}$ and $\{x_3, x_4, x_5, x_6\}$, respectively. (iii) The middle receiver broadcasts x_1+x_5 and x_2+x_6 , the other two receivers can now recover all 6 packets via DNC decoding. (iv) The left receiver transmits x_1 and x_2 to the middle receiver, who can now decode all 6 original packets too.

Using a straightforward multicast scheme without network coding, we need 7 time slots instead. x_1 and x_2 require 3 broadcasts to reach all receivers, the same for x_5 and x_6 . x_3

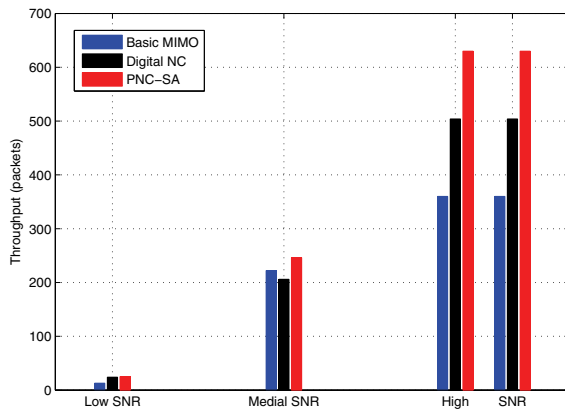


Fig. 15. Packet-level throughput for multicast, PNC-SA vs. DNC vs. MIMO alone.

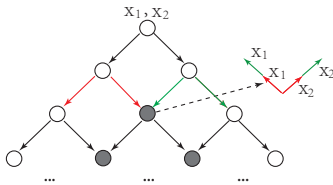


Fig. 16. Cascading signal alignment for multi-hop broadcast. Note that the two x_1 's reinforce instead of cancel out each other, since we apply normal BPSK instead of PNC demodulation. Signals are aligned at dark nodes.

and x_4 require two broadcasts. Among these 8 broadcast transmissions, only two can be scheduled concurrently, resulting in a total of 7 time slots. With DNC, the number of time slots required is between that of PNC-SA and a no coding solution, at 5.

Fig. 15 shows packet-level throughput achieved by PNC-SA, DNC and MIMO. The network is simulated for 140 time slots, with identical node transmission capacity and packet length as previously assumed. At high SNR, PNC-SA again demonstrates a marked throughput gain. DNC slightly leads MIMO at high SNR, but becomes inferior when SNR decreases due to its relatively worse SNR-BER performance.

Cascading SA for Multi-hop Broadcast

In this final application, we show that SA can be applied independently, without coupling with PNC. When signals of distinct packets are aligned to the same direction, PNC demodulation is required; when signals of the same packet are aligned, normal demodulation suffices.

In Fig. 16, the sender at the top wishes to broadcast to the entire network, with m rows. Each node has 2 antennas. The source data is divided into 2 packets, x_1 and x_2 . The goal is to finish broadcast routing in as few time slots as possible.

The SA solution is rather simple: have each row of nodes transmit concurrently, and disseminate the data item in $m - 1$ rounds. Signals are aligned for reception at inner nodes in black. The two signals for x_1 (x_2) augment each other, yielding a power gain. For the k nodes at row k , SA is applied in a cascading fashion: we can first decide the precoding vector for the left-most node. Consequently, all other precoding vectors at the same row are determined. Each node aligns

its signal according to its neighbor on the left. The number of time slots, $m - 1$, is the minimum possible, since under any routing scheme, data can propagate only one row per time slot.

A non-SA solution schedules individual transmissions to avoid interference. It not only takes at least $m - 1$ time slots, but also requires a complex scheduling algorithm, in contrast to the simple row-by-row structure of SA. For the same BER, SA does not consume significantly more energy, even by having all nodes except the bottom row transmit. This can be verified by checking the following facts (assume each node transmits with power P in the non-SA solution). (i) In the optimal non-SA solution, each transmission, with power P , covers ≤ 2 nodes. (ii) With SA, each transmission covers > 1 nodes on average. (iii) With SA, for the same BER, only border nodes in white need to transmit at power P . Inner nodes in black can transmit at roughly $P/2$ due to the MISO power gain. (iv) Border nodes only represent a $O(1/m)$ fraction of the network.

VIII. CONCLUSION

We showed that PNC-SA, SA coupled with PNC, can open new design spaces for routing in MIMO wireless networks, and can hence augment the network capacity region. The design of PNC-SA has been inspired by recent advances in PNC and IA research, yet PNC-SA can better exploit the spatial diversity and precoding opportunities of a MIMO network, leading to a higher system DoF. We studied the new problem of optimal precoding introduced by PNC-SA, formulated it into a vector programming problem, and designed a solution for maximizing SNR at the receiver. The SNR-BER performance of PNC-SA was then analyzed. General applications of both PNC-SA and SA alone were demonstrated, in various multi-hop MIMO routing scenarios, including information exchange, unicast and multicast/broadcast. Throughput gain of up to a factor of 2 was observed, compared to simple solutions without coding.

REFERENCES

- [1] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [2] S. Gollakota, S. D. Perli, and D. Katabi, "Interference alignment and cancellation," in *Proc. 2009 ACM SIGCOMM*, 2009.
- [3] S. Zhang, S. C. Liew, and P. P. Lam, "Physical-layer network coding," in *Proc. 2006 ACM MobiCom*, 2006.
- [4] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.
- [5] N. Lee, J. Lim, and J. Chun, "Degrees of freedom of the MIMO Y channel: signal space alignment for network coding," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3332–3342, July 2010.
- [6] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2, pp. 1740–1751, Mar. 1994.
- [7] L. E. Li, R. Alimi, D. Shen, H. Viswanathan, and Y. R. Yang, "A general algorithm for interference alignment and cancellation in wireless networks," in *Proc. 2010 IEEE INFOCOM*.
- [8] G. Bresler, D. Cartwright, and D. Tse, "Settling the feasibility of interference alignment for the MIMO interference channel: the symmetric square case," University of California, Berkeley, Tech. Rep., Apr. 2011, arXiv:1104.0888.
- [9] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

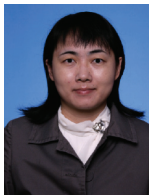
- [10] S. Zhang, S. C. Liew, and P. P. Lam, "On the synchronization of physical-layer network coding," in *Proc. 2006 IEEE Information Theory Workshop*.
- [11] A. Özgür and D. Tse, "Achieving linear scaling with interference alignment," in *Proc. 2009 IEEE International Symposium on Information Theory*.
- [12] S. Zhang and S. C. Liew, "Physical-layer network coding with multiple antennas," in CoRR abs/0910.2603, 2009.
- [13] 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.
- [14] J. Zhan, B. Nazer, M. Gastpar, and U. Erez, "MIMO compute-and-forward," in *Proc. 2009 IEEE International Symposium on Information Theory*.
- [15] K. Lee, S.-H. Park, J.-S. Kim, and I. Lee, "Degrees of freedom on MIMO multi-link two-way relay channels," in *Proc. 2010 IEEE Globecom*.
- [16] K. Lee, N. Lee, and I. Lee, "Feasibility conditions of signal space alignment for network coding on K-user MIMO Y channels," in *Proc. 2011 IEEE ICC*.
- [17] —, "Achievable degrees of freedom on K-user Y channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1210–1219, Mar. 2012.
- [18] N. Lee and J. Chun, "Signal space alignment for an encryption message and successive network code decoding on the MIMO K-way relay channel," in *Proc. 2011 IEEE ICC*.
- [19] T. Liu and C. Yang, "Signal alignment for multicarrier code division multiple user two-way relay systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 11, pp. 3700–3710, Nov. 2011.
- [20] H. Park, H. J. Yang, J. Chun, and R. Adve, "A closed-form power allocation and signal alignment for a diagonalized MIMO two-way relay channel with linear receivers," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5948–5962, Nov. 2012.
- [21] C. Yetis, T. Gou, S. Jafar, and A. Kayram, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, pp. 4771–4782, Sept. 2010.
- [22] S. Peters and R. Heath, "Interference alignment via alternating minimization," in *Proc. 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*.
- [23] M. Razaviyayn, M. Sanjabi, and Z. Q. Luo, "Linear transceiver design for interference alignment: complexity and computation," in *Proc. 2010 IEEE International Workshop on Signal Processing Advances in Wireless Communications*.
- [24] I. Santamaria, O. Gonzalez, R. Heath, and S. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. 2010 IEEE Globecom*.
- [25] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge University Press, 2003.
- [26] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: practical wireless network coding," *IEEE/ACM Trans. Networking*, vol. 16, no. 3, pp. 497–510, June 2008.
- [27] A. Scaglione, D. L. Goeckel, and J. N. Laneman, "Cooperative communications in mobile ad hoc networks: rethinking the link abstraction," *Distributed Antenna Systems*. Auerbach Publications, 2007, pp. 87–111.
- [28] R. Gummadi, L. Massoulie, and R. Sreenivas, "The role of feedback in the choice between routing and coding for wireless unicast," in *2010 IEEE Symposium on Network Coding*.
- [29] Z. Li, B. Li, and L. C. Lau, "A constant bound on throughput improvement of multicast network coding in undirected networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 997–1015, Mar. 2009.
- [30] M. Gastpar and M. Vetterli, "On the capacity of wireless networks: the relay case," in *Proc. 2002 IEEE INFOCOM*.



Ruiting Zhou received a B.E. degree in telecommunication engineering from Nanjing University of Post and Telecommunication, China, in 2007, a M.S. degree in telecommunications from Hong Kong University of Science and Technology, Hong Kong, in 2008 and a M.S. degree in computer science from University of Calgary, Canada, in 2012. She was with Shinetown Telecommunication Ltd (Hong Kong) during 2008–2010. Her research interests are in wireless networking and communications. Ruiting is a student member of IEEE.



Zongpeng Li received a B.E. in Computer Science and Technology from Tsinghua University (Beijing) in 1999, a M.S. in Computer Science from University of Toronto in 2001, and a Ph.D. in Electrical and Computer Engineering from University of Toronto in 2005. He has been with the Department of Computer Science in the University of Calgary since 2005. In 2011–2012, Zongpeng was a visitor at the Institute of Network Coding, Chinese University of Hong Kong. His research interests are in computer networks and network coding.



Chuan Wu received her B.E. and M.E. degrees in 2000 and 2002 from Department of Computer Science and Technology, Tsinghua University, China, and her Ph.D. degree in 2008 from the Department of Electrical and Computer Engineering, University of Toronto, Canada. She is currently an Assistant Professor in the Department of Computer Science, the University of Hong Kong, China. Her research interests include cloud computing, online/mobile social network, and wireless networks. She is a member of IEEE and ACM.



Carey Williamson is a Professor in the Department of Computer Science at the University of Calgary. He has a B.Sc.(Honours) in Computer Science from the University of Saskatchewan, and a Ph.D. in Computer Science from Stanford University. His research interests include Internet protocols, wireless networks, network traffic measurement, network simulation, and Web performance.