Precoding-Based Network Alignment For Three Unicast Sessions

Chun Meng, *Student Member, IEEE,* Abhik Kumar Das,
Abinesh Ramakrishnan, *Student Member, IEEE,* Syed Ali Jafar, *Fellow, IEEE,*
Athina Markopoulou, *Senior Member, IEEE,* and Sriram Vishwanath, *Senior Member, IEEE.*

***Abstract*—We consider the problem of network coding across three unicast sessions over a directed acyclic graph, where the sender and the receiver of each unicast session are both connected to the network via a single edge of unit capacity. We consider a network model in which the middle of the network can only perform random linear network coding, and we restrict our approaches to precoding-based linear schemes, where the senders use precoding matrices to encode source symbols. We adapt a precoding-based interference alignment technique, originally developed for the wireless interference channel, to construct a precoding-based linear scheme, which we refer to as *precoding-based network alignment scheme (PBNA).* A primary difference between this setting and the wireless interference channel is that the network topology can introduce dependencies among the elements of the transfer matrix, which we refer to as coupling relations, and can potentially affect the achievable rate of PBNA. We identify all these coupling relations and we interpret them in terms of network topology. We then present polynomial-time algorithms to check the presence of these coupling relations in a particular network. Finally, we show that, depending on the coupling relations present in the network, the optimal symmetric rate achieved by precoding-based linear scheme can take only three possible values, all of which can be achieved by PBNA.**

***Index Terms*—Network Coding, Multiple Unicasts, Interference Alignment.**

I. Introduction

Ever since the development of *network coding* and its success in characterizing the achievable throughput for single multicast scenario [1] [2], there has been hope that the framework can be extended to characterize network capacity in other scenarios, namely inter-session network coding. Of particular practical interest is network coding across *multiple unicast sessions*, as unicast is the dominant type of traffic in today's networks. There have been some successes in this domain, such as the derivation of a sufficient condition for linear network coding to achieve the maximal throughput in networks with multiple unicast sessions [3] [4]. However, finding linear network codes for guaranteeing rates for multiple unicasts is known to be NP-hard [5]. Only sub-optimal and heuristic methods are known today, including methods based on linear optimization [6] [7] and evolutionary approaches [8]. Moreover, scalar or even vector linear network coding [5] [9] alone has been shown to be insufficient for achieving the limits of inter-session network coding [10].

In this paper, we consider the problem of linear network coding across *three* unicast sessions over a network represented by a directed acyclic graph (DAG), where the sender and the receiver of each unicast session are both connected to the network via a single edge of unit capacity. We refer to this communication scenario as a *Single-Input Single-Output* scenario or *SISO* scenario for short (Fig. 1a). This is the smallest, yet highly non-trivial, instance of the problem. Furthermore, we consider a network model, in which the middle of the network only performs random linear network coding, and restrict our approaches to *precoding-based linear schemes*, where the senders use precoding matrices to encode source symbols[1]. Apart from being of interest on its own right, we hope that this can be used as a building block and for better understanding of the general network coding problem across multiple unicasts.

Our approach is motivated by the observation that under the linear network coding framework, a SISO scenario behaves roughly like a wireless interference channel. As shown in Fig. 1, the entire network can be viewed as a channel with a linear transfer function, albeit this function is no longer given by nature, as it is the case in wireless, but is determined by the network topology, routing and coding coefficients. This analogy enables us to apply the technique of precoding-based interference alignment, designed by Cadambe and Jafar [11] for wireless interference channels. We adapt this technique to our problem and refer to it as *precoding-based network*

Manuscript received July 8, 2013; accepted October 7, 2014; date of current version October 20, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jean-Claude Belfiore. This work was partially supported by NSF CAREER Award 1028394, NSF CDI Award 1028394, AFOSR MURI FA9550-09-0643, and NSF grants CCF-1161418 and CCF-1319104.

Preliminary versions of parts of this paper were presented at IEEE International Symposium on Information Theory (ISIT), Austin, TX, U.S.A., June 2010 [34], 48th Allerton Conference on Communication, Control, and Computing, Monticello, IL, U.S.A., Sept. 2010 [12], and IEEE International Symposium on Information Theory (ISIT), Boston, MA, U.S.A., July 2012 [36]. C. Meng, A. Ramakrishnan (email: abinesh.r@uci.edu), A. Markopoulou (email: athina@uci.edu) and S. Jafar (email: syed@uci.edu) were affiliated with the Electrical Engineering and Computer Science Department, and also with the Center for Pervasive Communications and Computing (CPCC), University of California, Irvine, CA 92697 USA, when this work was conducted. C. Meng is currently with Speedy Networks Inc., Pasadena, CA 91101 USA (email: cmeng1@uci.edu). A.K. Das and S. Viswanath (email: sriram@ece.utexas.edu) were with the Electrical and Computer Engineering Dept at UT Austin, when this work was conducted. A.K. Das is currently with Samsung Mobile Solutions Lab., San Diego, CA, USA (email: akdas@mail.utexas.edu).

[1]The precise definition of precoding-based linear scheme is presented in Section III.

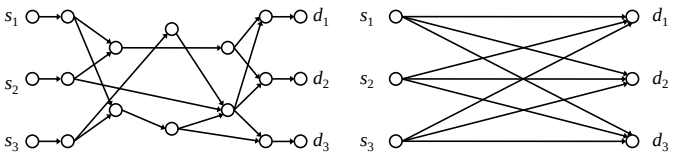


Fig. 1. Analogy between a SISO scenario employing linear network coding and a wireless interference channel, each with three unicast sessions (s_i, d_i), $i = 1, 2, 3$. Both these systems can be treated as linear transform systems and are amenable to interference alignment techniques.

alignment, or *PBNA* for short: precoding occurs only at source nodes, and all the intermediate nodes in the network perform random network coding. One advantage of PBNA is complexity: it significantly simplifies network code design since the nodes in the middle of the network perform random network coding. Another advantage is that PBNA can achieve the optimal symmetrical rate achieved by any precoding-based linear schemes.

An important difference between the SISO scenario and the wireless interference channel is that there may be algebraic dependencies, which we refer to as *coupling relations*, between elements of the transfer matrix, which we refer to as *transfer functions*. These are introduced by the network topology and may affect the achievable rate of PBNA [12]. Such algebraic dependencies are not present in the wireless interference channel, where channel gains are independent from each other such that the precoding-based interference alignment scheme of [11] can achieve 1/2 rate per session almost surely. Therefore, traditional interference alignment techniques, developed for the wireless interference channel, cannot be directly applied to networks with network coding but (i) they need to be properly adapted in the new setting, and (ii) their achievability conditions need to be characterized in terms of the network topology. Towards the second goal, we identify graph-related properties of the transfer functions, which together with a degree-counting technique, enable us to identify the minimal set of coupling relations that might affect the achievable rate of PBNA.

Our main contributions in this paper are the followings:

- *PBNA Design:* We design the first precoding-based interference-alignment scheme for the SISO scenario, in which the senders use precoding matrices to encode source symbols, and the intermediate nodes in the middle of the network perform random linear network coding. The scheme is inspired by the Cadambe and Jafar scheme in [11].
- *Achievability Conditions:* We identify the minimal set of coupling relations between transfer functions, the presence of which will potentially affect the achievable rate of PBNA. We further interpret these coupling relations in terms of network topology, and present polynomial-time algorithms for checking the existence of these coupling relations.
- *Rate Optimality:* We show that for the SISO scenarios where all senders are connected to all receivers via directed paths, depending on the coupling relations present in the network, there are only three possible optimal symmetric rates achieved by any precoding-based linear scheme (namely $1/3$, $2/5$ and $1/2$), all of which are achievable through PBNA.

The rest of the paper is organized as follows. In Section II, we review related work. In Section III, we present the problem setup and formulation. In Section IV, we present our proposed precoding-based interference alignment (PBNA) scheme for the network setting. In Section V, we present an overview of our main results. In Section VI, we discuss in depth the achievability conditions of PBNA. In Section VII, we provide polynomial-time algorithms to check the presence of the coupling relations that may affect the achievable rate of PBNA. In Section VIII, we prove the optimal symmetric rates achieved by any linear precoding-based scheme. Section IX concludes the paper and outlines future directions. In Appendices A-D, we present detailed proofs for the lemmas and the theorems presented in this paper. In Appendix E, we present a comparison between routing and PBNA.

II. Related Work

A. Network Coding

Network coding was first proposed to achieve optimal throughput for single multicast scenario [1] [2] [3], which is a special case of intra-session network coding. The rate region for this setting can be easily calculated by using linear programming techniques [13]. Moreover, the code design for this scenario is fairly simple: Either a polynomial-time algorithm [14] can be used to achieve the optimal throughput in a deterministic manner, or a random network coding scheme [15] can be used to achieve the optimal throughput with high probability.

One case, which is best understood up to now, is network coding across two unicasts. Wang and Shroff provided a graph-theoretical characterization of sufficient and necessary condition for the achievability of symmetrical rate of one for two multicast sessions, of which two unicasts is a special case, over networks with integer edge capacities [16]. They showed that linear network code is sufficient to achieve this symmetrical rate. Wang et al. [17] further pointed out that there are only two possible capacity regions for the network studied in [16]. They also showed that for layered linear deterministic networks, there are exactly five possible capacity regions. Kamath et al. [18] provided a edge-cut outer bound for the capacity region of two unicasts over networks with arbitrary edge capacities.

For network coding across more than two unicasts, there is only limited progress. It is known that there exist networks in which network coding significantly outperforms routing schemes in terms of transmission rate [4]. However, there exist only approximation methods to characterize the rate region for this setting [19]. Moreover, it is known that finding linear network codes for this setting is NP-hard [5]. Therefore, only sub-optimal and heuristic methods exist to construct linear network code for this setting. For example, Ratnakar

et al. [6] considered coding pairs of flows using poison-antidote butterfly structures and packing a network using these butterflies to improve throughput; Traskov et al. [7] further presented a linear programming-based method to find butterfly substructures in the network; Ho et al. [20] developed online and offline back pressure algorithms for finding approximately throughput-optimal network codes within the class of network codes restricted to XOR coding between pairs of flows; Effros et al. [21] described a tiling approach for designing network codes for wireless networks with multiple unicast sessions on a triangular lattice; Kim et al. [8] presented an evolutionary approach to construct linear code. Unfortunately, most of these approaches don't provide any guarantee in terms of performance. Moreover, most of these approaches are concerned about finding network codes by jointly considering code assignment and network topology at the same time. In contrast, our approach is oblivious to network topology in the sense that the design of encoding/decoding schemes is separated from network topology, and is predetermined regardless of network topology. The separation of code design from network topology greatly simplifies the code design of PBNA.

The part of our work that identifies coupling relations is related to some recent work on network coding. Ebrahimi and Fragouli [22] found that the structure of a network polynomial, which is the product of the determinants of all transfer matrices, can be described in terms of certain subgraph structures; Zeng et al. [23] proposed the Edge-Reduction Lemma which makes connections between cut sets and the row and column spans of the transfer matrices.

B. Interference Alignment

The original concept of precoding-based interference alignment was first proposed by Cadambe and Jafar [11] to achieve the optimal degree of freedom (DoF) for K-user wireless interference channel. After that, various approaches to interference alignment have been proposed. For example, Nazer et al. proposed ergodic interference alignment [24]; Bresler, Parekh and Tse proposed lattice alignment [25]; Jafar introduced blind alignment [26] for the scenarios where the actual channel coefficient values are entirely unknown to the transmitters; Maddah-Ali and Tse proposed retrospective interference alignment [27] which exploits only delayed CSIT. Interference alignment has been applied to a wide variety of scenarios, including K-user wireless interference channel [11], compound broadcast channel [28], cellular networks [29], relay networks [30], and wireless networks supported by a wired backbone [31]. Recently, it was shown that interference alignment can be used to achieve exact repair in distributed storage systems [32] [33].

C. Network Alignment

The idea of PBNA was first proposed by Das et al., who also proposed a sufficient condition for PBNA to asymptotically achieve a symmetrical rate of 1/2 per session [34]. However, the sufficient achievability condition proposed in [34] contains an exponential number of constraints, and is very difficult to verify in practice. Later, Ramakrishnan et al. observed that whether PBNA can achieve a symmetrical rate of 1/2 per session depends on network topology [12], and conjectured that the condition proposed in [34] can be reduced to just six constraints. Han et al. [35] proved that this conjecture is true for the special case of three symbol extensions. They also identified some important properties of transfer functions, which are used in this paper. In [36], Meng et al. showed that the conjecture in [12] is false for more than three symbol extensions, and reduced the condition proposed in [34] to just 12 constraints by using two graph-related properties of transfer functions. Later, Meng et al. reduced the 12 constraints to a set of 9 constraints [37] by using a result from [35], and proved that they are also necessary conditions for PBNA to achieve 1/2 rate per session. They also provided an interpretation of all the constraints in terms of graph structure. At the same time and independently, a technical report by Han et al. [38] also provided a similar characterization.

This journal paper combines our previous work in [12], [34], [36], [37], and extends them by finding the optimal symmetrical rates achieved by precoding-based linear schemes, of which PBNA is a special case. Compared to the most closely related work, namely [38], our work addresses a more general setting: (i) it considers the use of any precoding-matrix, not only the one proposed by Cadambe and Jafar [11] and (ii) it applies to all network topologies, which subsume the cases considered in [38]. In addition, we prove that PBNA can achieve all the optimal symmetric rates achieved by precoding-based linear schemes.

III. Problem Formulation

A. Network Model

A network is represented by a directed acyclic graph $\mathcal{G} = (V, E)$, where V is the set of nodes and E the set of edges. We consider the simplest non-trivial communication scenario where there are three unicast sessions in the network. The ith ($i = 1, 2, 3$) unicast session is represented by a tuple $\omega_i = (s_i, d_i, \mathbf{X}_i)$, where s_i and d_i are the sender and the receiver of the ith unicast session, respectively; $\mathbf{X}_i = (X_i^{(1)}, X_i^{(2)}, \cdots, X_i^{(k_i)})^T$ is a vector of independent random variables, each of which represents a packet that s_i sends to d_i. Each sender s_i is connected to the network via a single edge σ_i, called a sender edge, and each receiver node d_i via a single edge τ_i, called a receiver edge. Each edge has unit capacity, *i.e.,* can carry one symbol of $\mathbb{F}_{2^m}$ in a time slot, and represents an error-free and delay-free channel. We group these unicast sessions into a set $\Omega = \{\omega_1, \omega_2, \omega_3\}$. We refer to the tuple $(\mathcal{G}, \Omega)$ as a single-input and single-output communication scenario, or a SISO scenario for short. An example of SISO scenario is shown in Fig. 1a. Clearly, in a SISO scenario, each sender can transmit at most one symbol to its corresponding receiver node in a time slot.

Given an edge $e = (u, v) \in E$, let $u = \mathrm{head}(e)$ and $v = \mathrm{tail}(e)$ denote the head and the tail of e, respectively. Given a node $v \in V$, let $\mathrm{In}(v) = \{e \in E : \mathrm{head}(e) = v\}$ denote the set of incoming edges at v, and $\mathrm{Out}(v) = \{e \in E : \mathrm{tail}(e) = v\}$ the set of outgoing edges at v. Given two distinct

edges $e, e' \in E$, a directed path from e to e' is a subset of edges $P = \{e_1, e_2, \cdots, e_k\}$ such that $e_1 = e$, $e_k = e'$, and $\mathrm{head}(e_i) = \mathrm{tail}(e_{i+1})$ for $i \in \{1, 2, \cdots, k-1\}$. The set of directed paths from e to e' is denoted by $\mathcal{P}_{ee'}$. For $i, j \in \{1, 2, 3\}$, we also use $\mathcal{P}_{ij}$ to represent $\mathcal{P}_{\sigma_i \tau_j}$.

Each node in the network performs scalar linear network coding operations on the incoming symbols [2] [3]. The symbols transmitted in the network are elements of a finite field $\mathbb{F}_{2^m}$. Let $\hat{X}_i$ be the symbol injected at the sender node s_i. Thus, for an edge $e = (u, v) \in E$, the symbol transmitted along e, denoted by Y_e, is a linear combination of the incoming symbols at u:

$$Y_e = \begin{cases} \hat{X}_i & \text{if } e = \sigma_i; \\ \sum_{e' \in \mathrm{In}(u)} x_{e'e} Y_{e'} & \text{otherwise.} \end{cases} \tag{1}$$

where $x_{e'e}$ denotes the coding coefficient that is used to combine the incoming symbol $Y_{e'}$ into Y_e. Following the algebraic framework of [3], we treat the coding coefficients as variables. Let $\mathbf{x}$ denote the vector consisting of all the coding coefficients in the network, *i.e.,* $\mathbf{x} = (x_{e'e} : e', e \in E, \mathrm{head}(e') = \mathrm{tail}(e))$.

Due to the linear operations at each node, the network functions like a linear system such that the received symbol at τ_i is a linear combination of the symbols injected at sender nodes:

$$Y_{\tau_i} = m_{1i}(\mathbf{x})\hat{X}_1 + m_{2i}(\mathbf{x})\hat{X}_2 + m_{3i}(\mathbf{x})\hat{X}_3 \tag{2}$$

In the above formula, $m_{ji}(\mathbf{x})$ $(j = 1, 2, 3)$ is a multivariate polynomial in the ring $\mathbb{F}_2[\mathbf{x}]$, and is defined as follows [3]:

$$m_{ji}(\mathbf{x}) = \sum_{P \in \mathcal{P}_{ji}} t_P(\mathbf{x}) \tag{3}$$

Each $t_P(\mathbf{x})$ denotes a monomial in $m_{ji}(\mathbf{x})$, and is the product of all the coding coefficients along path P, *i.e.,* for a given path $P = \{e_1, e_2, \cdots, e_k\}$,

$$t_P(\mathbf{x}) = \prod_{i=1}^{k-1} x_{e_i e_{i+1}} \tag{4}$$

Thus, $t_P(\mathbf{x})$ represents the signal gain along a path P, and $m_{ji}(\mathbf{x})$ is simply the summation of the signal gains along all possible paths from σ_j to τ_i. We refer to $m_{ji}(\mathbf{x})$ as the *transfer function* from σ_j to τ_i.

We make the following assumptions:

1) The nodes in $V - \{s_i, d_i : 1 \le i \le 3\}$ can only perform random linear network coding, i.e., there is no intelligence in the middle of the network. The variables in $\mathbf{x}$ all take values independently and uniformly at random from $\mathbb{F}_{2^m}$.
2) Except for the senders and the receivers, all other nodes in the network have zero memory, and therefore cannot store any received data.
3) The senders have no incoming edges, and the receivers have no outgoing edges.
4) The random variables in all $\mathbf{X}_i$'s are mutually independent. Each element of $\mathbf{X}_i$ has an entropy of m bits.
5) The transmissions within the network are all synchronized with respect to the symbol timing.

B. Transmission Process

The transmission process in the network continues for $N \in \mathbb{Z}_{>0}$ time slots, where $N \ge \max\{k_1, k_2, k_3\}$. Both N and k_i are parameters of the transmission scheme. We will show how to set these parameters in Section IV. Let $\mathbf{x}^{(t)} = (x_{e'e}^{(t)} : e', e \in E, \mathrm{head}(e') = \mathrm{tail}(e))$ denote the vector of coding coefficients for time slot t, where $x_{e'e}^{(t)}$ represents the coding coefficient used to combine the incoming symbol along e' into the symbol along e for time slot t. For an edge e, let $Y_e^{(t)}$ denote the symbol transmitted along e during time slot t, and $\mathbf{Y}_e = (Y_e^{(1)}, Y_e^{(2)}, \cdots, Y_e^{(N)})^T$ the vector of all the symbols transmitted along e during the N time slots. Define a vector of variables, $\xi = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots, \mathbf{x}^{(N)}, \theta_1, \theta_2, \cdots, \theta_k)$, where $\theta_1, \cdots, \theta_k$ are variables, which take values from $\mathbb{F}_{2^m}$, and are used in the encoding process at the senders.

Each sender s_i first encode $\mathbf{X}_i$ into a vector $\hat{\mathbf{X}}_i$ of N symbols:

$$\hat{\mathbf{X}}_i = \mathbf{V}_i \mathbf{X}_i \tag{5}$$

where $\mathbf{V}_i$ is an $N \times k_i$ matrix, each element of which is a rational function in $\mathbb{F}_{2^m}(\xi)$[2], and is called the *precoding matrix* at s_i. Define the following $N \times N$ diagonal matrix which includes all the transfer functions $m_{ji}(\mathbf{x}^{(t)})$ for the N time slots:

$$\mathbf{M}_{ji} = \begin{pmatrix} m_{ji}(\mathbf{x}^{(1)}) & 0 & \cdots & 0 \\ 0 & m_{ji}(\mathbf{x}^{(2)}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{ji}(\mathbf{x}^{(N)}) \end{pmatrix} \tag{6}$$

Hence, the input-output equation of the network can be formulated in a matrix form as follows:

$$\begin{aligned} \mathbf{Y}_{\tau_i} &= \mathbf{M}_{1i}\hat{\mathbf{X}}_1 + \mathbf{M}_{2i}\hat{\mathbf{X}}_2 + \mathbf{M}_{3i}\hat{\mathbf{X}}_3 \\ &= \mathbf{M}_{1i}\mathbf{V}_1\mathbf{X}_1 + \mathbf{M}_{2i}\mathbf{V}_2\mathbf{X}_2 + \mathbf{M}_{3i}\mathbf{V}_3\mathbf{X}_3 \\ &= \mathbf{M}_i\mathbf{X} \end{aligned} \tag{7}$$

where $\mathbf{M}_i = (\mathbf{M}_{1i}\mathbf{V}_1 \quad \mathbf{M}_{2i}\mathbf{V}_2 \quad \mathbf{M}_{3i}\mathbf{V}_3)$, and $\mathbf{X} = (\mathbf{X}_1^T \quad \mathbf{X}_2^T \quad \mathbf{X}_3^T)^T$. Since the elements of $\mathbf{M}_{ji}$ $(1 \le j \le 3)$ and $\mathbf{V}_j$ are rational functions in $\mathbb{F}_{2^m}(\xi)$, the elements of $\mathbf{M}_i$ are also rational functions in terms of ξ.

C. Precoding-Based Linear Scheme

In this paper, we consider the following transmission scheme, called precoding-based linear scheme:

Definition III.1. Given a SISO scenario $(\mathcal{G}, \Omega)$, a *precoding-based linear scheme* for $(\mathcal{G}, \Omega)$ is a transmission scheme, where each sender s_i $(1 \le i \le 3)$ uses a precoding matrix $\mathbf{V}_i$ to encode source symbols, and the variables in ξ all take values independently and uniformly at random from $\mathbb{F}_{2^m}$. We use a tuple $\lambda = (\xi, \mathbf{V}_i : 1 \le i \le 3)$ to denote a precoding-based linear scheme.

From the above definition, it can be seen that a precoding-based linear scheme is a random linear network coding

[2]Given a field $\mathbb{F}$, $\mathbb{F}(x_1, \cdots, x_k)$ denotes the field consisting of all multivariate rational functions in terms of $(x_1, \cdots, x_k)$ over $\mathbb{F}$.

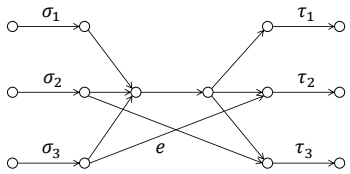


Fig. 2. An illustrative example for precoding-based linear scheme.

scheme. Given a precoding-based linear scheme, let P_{succ} denote the probability that the denominators of the precoding matrices are all evaluated to non-zero values, and all receivers can successfully decode their required source symbols from received symbols.

Definition III.2. Given a precoding-based linear scheme $\lambda = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$, we say that it *achieves* the rate tuple $(\frac{k_1}{N}, \frac{k_2}{N}, \frac{k_3}{N})$, if $\lim_{m\to\infty} P_{succ} = 1$.

Given a precoding-based linear scheme, if the conditions of the above definition is satisfied, by choosing sufficiently large finite field $\mathbb{F}_{2^m}$, a random assignment of values to ξ will enable each receiver to successfully decode its required source symbols with high probability. In this sense, given sufficiently large $\mathbb{F}_{2^m}$, a precoding-based linear scheme works for most random realizations of ξ, but not all realizations.

Before proceeding, we introduce the following Schwartz-Zippel Theorem [39].

Theorem III.1 (Schwartz-Zippel Theorem)**.** Let $Q(x_1, x_2, \cdots, x_n)$ be a non-zero multivariate polynomial of total degree d in the ring $\mathbb{F}[x_1, x_2, \cdots, x_n]$, where $\mathbb{F}$ is a field. Fix a finite set $S \subseteq \mathbb{F}$. Let $r_1, r_2, \cdots, r_n$ be chosen independently and uniformly at random from S. Then,

$$Pr(Q(r_1, r_2, \cdots, r_n) = 0) \leq \frac{d}{|S|}$$

Example III.1. We use an example to illustrate the above concepts. Consider the network in Fig. 2. Note that under the network model considered in the paper, interference is almost unavoidable at the receivers. Consider a receiver d_i. Without loss of generality, assume that the $(1,1)$ element of $\mathbf{V}_j$ $(i \neq j)$ is a non-zero rational function $f_{11}(\xi)$. Thus, the $(1,1)$ element of $\mathbf{M}_{ji}\mathbf{V}_j$ is a non-zero rational function $m_{ji}(\mathbf{x}^{(1)})f_{11}(\xi)$. Due to Theorem III.1, the probability that $m_{ji}(\mathbf{x}^{(1)})f_{11}(\xi)$ is evaluated to zero under a random assignment of values to ξ approaches to zero as $m \to \infty$. Hence, the probability that $\mathbf{M}_{ji}\mathbf{V}_j = \mathbf{0}$ approaches zero as $m \to \infty$. This means that interference is almost unavoidable at d_i.

Next, we present a precoding-based linear scheme that achieves a symmetric rate of $\frac{1}{3}$ per unicast session. Let $N = 3$, and $k_1 = k_2 = k_3 = 1$. Consider the following precoding matrix $\mathbf{V}_1 = (\theta_1^{(1)}\ \theta_1^{(2)}\ \theta_1^{(3)})$. According to Eq. (7), the output vector at d_i is $\mathbf{Y}_{\tau_i} = \mathbf{M}_i\mathbf{X}$, where $\mathbf{M}_i$ is as follows:

$$\mathbf{M}_i = \begin{pmatrix} m_{1i}(\mathbf{x}^{(1)})\theta_1^{(1)} & m_{2i}(\mathbf{x}^{(1)})\theta_2^{(1)} & m_{3i}(\mathbf{x}^{(1)})\theta_3^{(1)} \\ m_{1i}(\mathbf{x}^{(2)})\theta_1^{(2)} & m_{2i}(\mathbf{x}^{(2)})\theta_2^{(2)} & m_{3i}(\mathbf{x}^{(2)})\theta_3^{(2)} \\ m_{1i}(\mathbf{x}^{(3)})\theta_1^{(3)} & m_{2i}(\mathbf{x}^{(3)})\theta_2^{(3)} & m_{3i}(\mathbf{x}^{(3)})\theta_3^{(3)} \end{pmatrix}$$

It can be verified that $\det(\mathbf{M}_i)$ is a non-zero polynomial in $\mathbb{F}_{2^m}(\xi)$[3]. Let d be the total degree of $\det(\mathbf{M}_i)$. Due to Theorem III.1, we have:

$$\begin{aligned} P_{succ} \geq & Pr(\det(\mathbf{M}_i) \neq 0) \\ = & 1 - Pr(\det(\mathbf{M}_i) = 0) \geq 1 - \frac{d}{2^m} \end{aligned}$$

Since $\lim_{m\to\infty}(1 - \frac{d}{2^m}) = 1$, it follows that $\lim_{m\to\infty} P_{succ} = 1$. Hence, the above precoding-based linear scheme achieves a symmetric rate $\frac{1}{3}$ per unicast session. As we will show in Section VI, using precoding-based alignment scheme, which is a special case of precoding-based linear scheme, each unicast session can achieve a symmetric rate $\frac{1}{2}$ per unicast session, which is the optimal symmetric rate achieved by any precoding-based linear schemes. ■

Table I summarizes the notations used in this paper, in which $e', e \in E$ and $1 \leq i, j, k \leq 3$.

IV. Applying Precoding-Based Network Alignment to Networks

In this section, we first present how to utilize precoding-based interference alignment technique to find a precoding-based linear scheme for $(\mathcal{G}, \Omega)$. Then, we present achievability conditions for PBNA. We then introduce the concept of "coupling relations," which are essential in determining the achievability of PBNA.

Throughout this section, we assume that all the senders are connected to all the receivers via directed paths, *i.e.,* $m_{ij}(\mathbf{x})$ is a non-zero polynomial for all $1 \leq i, j \leq 3$. This is the most challenging case, since each receiver may suffer interference from the other two senders. This case also models most practical communication scenarios, in which it is common that all the senders are connected to all the receivers. The other setting, where some sender s_i is disconnected from some receiver d_j $(i \neq j)$, *i.e.,* $m_{ij}(\mathbf{x})$ is a zero polynomial, is easier to deal with, since there is less interference at receivers. We defer the later case to Section VI, where we show that this case can be handled similarly as the first case.

A. Precoding-Based Network Alignment Scheme

In this section, we present how to apply interference alignment to networks to construct a precoding-based linear scheme for $(\mathcal{G}, \Omega)$. The basic idea is that under linear network coding, the network behaves like a wireless interference channel[4], which is shown below:

$$U_i = H_{1i}W_1 + H_{2i}W_2 + H_{3i}W_3 + N_i \quad i = 1, 2, 3 \qquad (8)$$

where W_j, H_{ji}, U_i, and N_i $(j = 1, 2, 3)$ are all complex numbers, representing the transmitted signal at sender j, the channel gain from sender j to receiver i, the received signal at receiver j, and the noise term respectively. As we can see from

[3]It can be seen that each row of $\mathbf{M}_i$ is of the form $(m_{1i}(\mathbf{x})\theta_1\ \ m_{2i}(\mathbf{x})\theta_2\ \ m_{3i}(\mathbf{x})\theta_3)$. Since $m_{1i}(\mathbf{x})\theta_1$, $m_{2i}(\mathbf{x})\theta_2$ and $m_{3i}(\mathbf{x})\theta_3$ are linearly independent, according to Lemma IV.2 (see Subsection IV-B), $\det(\mathbf{M}_i)$ is a non-zero polynomial.

[4]The wireless interference channel that we consider here has only one sub-channel.

TABLE I
SUMMARY OF NOTATIONS

Notations	Meanings
$\omega_i = (s_i, d_i)$	The ith unicast session, where s_i and d_i are the sender and receiver of ω_i respectively.
$(\mathcal{G}, \Omega)$	A SISO scenario, where $\mathcal{G}$ represents the network, and Ω the set of unicast sessions.
σ_i, τ_i	The sender edge and the receiver edge for ω_i.
$\mathbf{X}_i$	A vector that holds all the source symbols transmitted from s_i to d_i.
$\mathbb{F}_{2^m}$	The finite field which forms the support for all the symbols transmitted in the network.
$x_{e'e}$	The coding coefficient used to combine the incoming symbol along e' to the symbol along e.
$\mathbf{x}$	The vector consisting of all the coding coefficients in the network.
$\mathcal{P}_{e'e}$	The set of directed paths from e' to e.
$\mathcal{P}_{ji}$	The set of directed paths from σ_j to τ_i.
$t_P(\mathbf{x})$	The product of coding coefficients along path P. It represents a monomial in a transfer function.
$m_{ji}(\mathbf{x})$	The transfer function from σ_j to τ_i.
$\mathbf{x}^{(t)}$	The vector consisting of all the coding coefficients in the network for time slot t.
ξ	A vector that holds all the coding coefficients in the network for the whole transmission process, and the variables used in the encoding process at all the senders.
$\mathbf{V}_i$	The precoding matrix used to encode the symbols sent by s_i.
$\mathbf{M}_{ji}$	A diagonal matrix, in which the element at coordinate (l, l) is the transfer function $m_{ji}(\mathbf{x}^{(l)})$.
P_{succ}	The probability that the denominators of the elements in the precoding matrices are evaluated to non-zero values, and all receivers can decode their required source symbols.
$\lambda = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$	A precoding-based linear scheme for $(\mathcal{G}, \Omega)$.
$\mathscr{A}_i$, $\mathscr{B}_i$	The alignment condition and the rank condition for ω_i.
$\mathbf{V}_i^*$	The precoding matrix proposed in [11] (see Eq. (12)-(14)).
$\mathbf{P}_i$, $\mathbf{T}$	The diagonal matrices used in the reformulated alignment conditions Eq. (10) and the reformulated rank conditions $\mathscr{B}'_1 \sim \mathscr{B}'_3$.
$\mathbf{I}_n$	The $n \times n$ identity matrix.
$p_i(\mathbf{x})$, $\eta(\mathbf{x})$	The rational functions that form the elements along the diagonals of $\mathbf{P}_i$ and $\mathbf{T}$ respectively
α_{ijk}	The last edge that forms a cut-set between σ_i and $\{\tau_j, \tau_k\}$ in a topological ordering of the edges in the network.
β_{ijk}	The first edge that forms a cut-set between $\{\sigma_j, \alpha_{ijk}\}$ and τ_k in a topological ordering of the edges in the network.
$\mathcal{C}_{e'e}$	The set of edges that forms a cut-set between e' and e.
$\mathcal{C}_{ij}$	The set of edges that forms a cut-set between σ_i and τ_j.
$\gcd(f(x), g(x))$	The greatest common divisor of two polynomials $f(x)$ and $g(x)$.

Eq. (2), in a network equipped with linear network coding, $\hat{X}_j$'s ($j \neq i$) play the roles of interfering signals, and transfer functions the roles of channel gains. This analogy enables us to borrow some techniques, such as precoding-based interference alignment [11], which is originally developed for the wireless interference channel, to the network setting.

A precoding-based network alignment scheme is defined as follows:

Definition IV.1. Given a SISO scenario $(\mathcal{G}, \Omega)$, $n \in \mathbb{Z}_{>0}$, and $s \in \{0, 1\}$, a precoding-based network alignment scheme with $2n+s$ symbol extensions, or a PBNA for short, is a precoding-based linear scheme $\lambda = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$, which satisfies the following conditions:

1) $\mathbf{V}_1$ is a $(2n+s) \times (n+s)$ matrix with rank $n+s$ on $\mathbb{F}_{2^m}(\xi)$, and $\mathbf{V}_2, \mathbf{V}_3$ are both $(2n+s) \times n$ matrices with rank n on $\mathbb{F}_{2^m}(\xi)$.
2) The following equations are satisfied [11]:

$$\mathscr{A}_1 : \text{span}(\mathbf{M}_{21}\mathbf{V}_2) = \text{span}(\mathbf{M}_{31}\mathbf{V}_3)$$
$$\mathscr{A}_2 : \text{span}(\mathbf{M}_{32}\mathbf{V}_3) \subseteq \text{span}(\mathbf{M}_{12}\mathbf{V}_1)$$
$$\mathscr{A}_3 : \text{span}(\mathbf{M}_{23}\mathbf{V}_2) \subseteq \text{span}(\mathbf{M}_{13}\mathbf{V}_1)$$

   where for a matrix $\mathbf{E}$, $\text{span}(\mathbf{E})$ denotes the linear space spanned by the column vectors contained in $\mathbf{E}$.
3) The variables in ξ all take values independently and uniformly at random from $\mathbb{F}_{2^m}$.

Definition IV.2. Given a SISO scenario $(\mathcal{G}, \Omega)$, and a rate tuple $(R_1, R_2, R_3) \in \mathbb{Q}^3_{>0}$, we say that (R_1, R_2, R_3) is asymptotically achievable through PBNA, if there exists a sequence $(\lambda_n)_{n=1}^{\infty}$, where each λ_n is a PBNA for $(\mathcal{G}, \Omega)$, such that each λ_n achieves a rate tuple $\mathbf{r}_n \in \mathbb{Q}^3_{>0}$, and $\lim_{n\to\infty} \mathbf{r}_n = (R_1, R_2, R_3)$.

In the above definition, $\mathscr{A}_i$ ($1 \leq i \leq 3$) is called the alignment condition for ω_i. It guarantees that the undesired symbols or interferences at each receiver are mapped into a single linear space, such that the dimension of received symbols or the number of unknowns is decreased.

B. Achievability Conditions of PBNA

The following lemma provides sufficient conditions for PBNA schemes to achieve the rate tuple $(\frac{n+s}{2n+s}, \frac{n}{2n+s}, \frac{n}{2n+s})$.

Lemma IV.1. Assume that all the senders and all the receivers are connected via directed paths. Consider a PBNA $\lambda = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$. It achieves the rate tuple $(\frac{n+2}{2n+s}, \frac{n}{2n+s}, \frac{n}{2n+s})$, if the following conditions are satisfied [11]:

$$\mathscr{B}_1 : \text{rank}(\mathbf{M}_{11}\mathbf{V}_1 \quad \mathbf{M}_{21}\mathbf{V}_2) = 2n+s$$
$$\mathscr{B}_2 : \text{rank}(\mathbf{M}_{12}\mathbf{V}_1 \quad \mathbf{M}_{22}\mathbf{V}_2) = 2n+s$$
$$\mathscr{B}_3 : \text{rank}(\mathbf{M}_{13}\mathbf{V}_1 \quad \mathbf{M}_{33}\mathbf{V}_3) = 2n+s$$

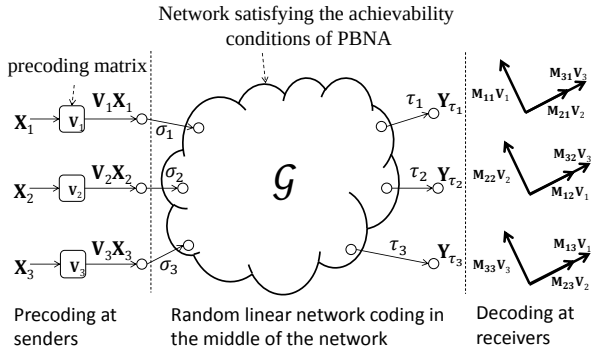


Fig. 3. Applying precoding-based interference alignment to a network which satisfies the rank conditions of PBNA as per Lemma IV.1. At each sender edge σ_i $(i = 1, 2, 3)$, the input vector $\mathbf{X}_i$ is first encoded into $2n + s$ symbols through the precoding matrix $\mathbf{V}_i$; then the encoded symbols are transmitted through the network in $2n + s$ time slots via random linear network coding in the middle of the network; at each receiver edge τ_i, the undesired symbols are aligned into a single linear space, which is linearly indepdent from the linear space spanned by the desired signals, such that the receiver can decode all the desired symbols.

Proof. Suppose $\mathscr{B}_1 \sim \mathscr{B}_3$ are satisfied. Define the following matrices:

$$\mathbf{D}_1 = (\mathbf{M}_{11}\mathbf{V}_1 \quad \mathbf{M}_{21}\mathbf{V}_2)^{-1}$$
$$\mathbf{D}_2 = (\mathbf{M}_{12}\mathbf{V}_1 \quad \mathbf{M}_{22}\mathbf{V}_2)^{-1}$$
$$\mathbf{D}_3 = (\mathbf{M}_{13}\mathbf{V}_1 \quad \mathbf{M}_{33}\mathbf{V}_3)^{-1}$$

Let $f_i(\xi)$ denote the product of the denominators of all the elements in $\mathbf{V}_i$, and $g_i(\xi)$ the product of the denominators of all the elements in $\mathbf{D}_i$. Thus, $f_i(\xi), g_i(\xi)$ are both non-zero polynomials in $\mathbb{F}_{2^m}[\xi]$. Define $q(\xi) = \prod_{i=1}^{3} f_i(\xi)g_i(\xi)$. Let d denote the total degree of $q(\xi)$. Suppose ξ_0 is an assignment of values to ξ such that $q(\xi_0) \neq 0$. Hence, the denominators of the elements in $\mathbf{V}_i$'s and $\mathbf{D}_i$'s are evaluated to non-zeros. Moreover, $\mathbf{X}_i$ is a sub-vector of $\mathbf{D}_i|_{\xi_0}\mathbf{Y}_{\tau_i}$, where $\mathbf{D}_i|_{\xi_0}$ is a matrix acquired through evaluating each element of $\mathbf{D}_i$ under the assignment $\xi = \xi_0$. Thus, all the receivers can decode their required source symbols. Hence, the probability P_{succ} that all the receivers can decoded their required source symbols satisfies the following inequalities:

$$P_{succ} \geq Pr(q(\xi) \neq 0) = 1 - Pr(q(\xi) = 0) \geq 1 - \frac{d}{2^m}$$

where the last inequality follows from Theorem III.1. Since $\lim_{m\to\infty}(1 - \frac{d}{2^m}) = 1$, we have $\lim_{m\to\infty} P_{succ} = 1$. Hence, λ achieves $(\frac{n+s}{2n+s}, \frac{n}{2n+s}, \frac{n}{2n+s})$. ■

In Lemma IV.1, $\mathscr{B}_i$ $(1 \leq i \leq 3)$ is called the rank condition for ω_i. $\mathscr{B}_i$ guarantees that d_i can decode its required source symbols with high probability when the the size of $\mathbb{F}_{2^m}$ is sufficiently large. In Fig. 3, we use a figure to illustrate how to apply PBNA to a network which satisfies the rank conditions.

We can further simplify the alignment conditions as follows. First, we reformulate $\mathscr{A}_1 \sim \mathscr{A}_3$ as follows:

$$\mathscr{A}_1' : \mathbf{M}_{21}\mathbf{V}_2 = \mathbf{M}_{31}\mathbf{V}_3\mathbf{A}$$
$$\mathscr{A}_2' : \mathbf{M}_{32}\mathbf{V}_3 = \mathbf{M}_{12}\mathbf{V}_1\mathbf{B}$$
$$\mathscr{A}_3' : \mathbf{M}_{23}\mathbf{V}_2 = \mathbf{M}_{13}\mathbf{V}_1\mathbf{C}$$

where $\mathbf{A}$ is an $n \times n$ invertible matrix, and $\mathbf{B}$, $\mathbf{C}$ are both $(n+s) \times n$ matrices with rank n. A direct consequence of $\mathscr{A}_2'$ and $\mathscr{A}_3'$ is that the precoding matrices are not independent from each other: Both $\mathbf{V}_2$ and $\mathbf{V}_3$ are determined by $\mathbf{V}_1$ through the following equations:

$$\mathbf{V}_2 = \mathbf{M}_{13}\mathbf{M}_{23}^{-1}\mathbf{V}_1\mathbf{C} \quad \mathbf{V}_3 = \mathbf{M}_{12}\mathbf{M}_{32}^{-1}\mathbf{V}_1\mathbf{B} \tag{9}$$

Substituting the above equations into $\mathscr{A}_1'$, the three alignment conditions can be further consolidated into a single equation:

$$\mathbf{T}\mathbf{V}_1\mathbf{C} = \mathbf{V}_1\mathbf{B}\mathbf{A} \tag{10}$$

where $\mathbf{T} = \mathbf{M}_{13}\mathbf{M}_{21}\mathbf{M}_{32}\mathbf{M}_{12}^{-1}\mathbf{M}_{23}^{-1}\mathbf{M}_{31}^{-1}$. Eq. (10) suggests that alignment conditions introduce constraint on $\mathbf{V}_1$. Thus, in general, we cannot choose $\mathbf{V}_1$ freely.

Finally, using Eq. (9) and Eq. (10), the rank conditions are transformed into the following equivalent equations:

$$\mathscr{B}_1' : \quad \text{rank}(\mathbf{V}_1 \quad \mathbf{P}_1\mathbf{V}_1\mathbf{C}) = 2n + s$$
$$\mathscr{B}_2' : \quad \text{rank}(\mathbf{V}_1 \quad \mathbf{P}_2\mathbf{V}_1\mathbf{C}) = 2n + s$$
$$\mathscr{B}_3' : \quad \text{rank}(\mathbf{V}_1 \quad \mathbf{P}_3\mathbf{V}_1\mathbf{C}\mathbf{A}^{-1}) = 2n + s$$

where $\mathbf{P}_1 = \mathbf{M}_{13}\mathbf{M}_{21}\mathbf{M}_{11}^{-1}\mathbf{M}_{23}^{-1}$, $\mathbf{P}_2 = \mathbf{M}_{13}\mathbf{M}_{22}\mathbf{M}_{12}^{-1}\mathbf{M}_{23}^{-1}$, and $\mathbf{P}_3 = \mathbf{M}_{21}\mathbf{M}_{33}\mathbf{M}_{23}^{-1}\mathbf{M}_{31}^{-1}$. Recalling each $\mathbf{M}_{kl}$ $(1 \leq k, l \leq 3)$ is a diagonal matrix (see Eq. (6)) with the elements along the diagonal being of the form $m_{kl}(\mathbf{x})$, $\mathbf{P}_i$ and $\mathbf{T}$ are both diagonal matrices. Define the following functions:

$$p_1(\mathbf{x}) = \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}{m_{11}(\mathbf{x})m_{23}(\mathbf{x})} \quad p_2(\mathbf{x}) = \frac{m_{13}(\mathbf{x})m_{22}(\mathbf{x})}{m_{12}(\mathbf{x})m_{23}(\mathbf{x})}$$
$$p_3(\mathbf{x}) = \frac{m_{21}(\mathbf{x})m_{33}(\mathbf{x})}{m_{23}(\mathbf{x})m_{31}(\mathbf{x})} \quad \eta(\mathbf{x}) = \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})m_{32}(\mathbf{x})}{m_{12}(\mathbf{x})m_{23}(\mathbf{x})m_{31}(\mathbf{x})} \tag{11}$$

It can been seen that $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$ form the elements along the diagonals of $\mathbf{P}_i$ and $\mathbf{T}$ respectively.

Next, we reformulate the rank conditions in terms of $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$. To this end, we need to know the internal structure of $\mathbf{V}_1$. We distinguish the following two cases:

Case I: $\eta(\mathbf{x})$ is non-constant, and thus $\mathbf{T}$ is not an identity matrix. For this case, Eq. (10) becomes non-trivial, and we cannot choose $\mathbf{V}_1$ freely. We use the following precoding matrices proposed by Cadambe and Jafar [11]:

$$\mathbf{V}_1^* = (\mathbf{w} \quad \mathbf{T}\mathbf{w} \quad \cdots \quad \mathbf{T}^n\mathbf{w}) \tag{12}$$
$$\mathbf{V}_2^* = \mathbf{M}_{13}\mathbf{M}_{23}^{-1}(\mathbf{w} \quad \mathbf{T}\mathbf{w} \quad \cdots \quad \mathbf{T}^{n-1}\mathbf{w}) \tag{13}$$
$$\mathbf{V}_3^* = \mathbf{M}_{12}\mathbf{M}_{32}^{-1}(\mathbf{T}\mathbf{w} \quad \mathbf{T}^2\mathbf{w} \quad \cdots \quad \mathbf{T}^n\mathbf{w}) \tag{14}$$

where $\mathbf{w}$ is a column vector of $2n + 1$ ones. The above precoding matrices correspond to the configuration where $s = 1$, $\mathbf{A} = \mathbf{I}_n$, $\mathbf{C}$ consists of the left n columns of $\mathbf{I}_{n+1}$, and $\mathbf{B}$ the right n columns of $\mathbf{I}_{n+1}$. It is straightforward to verify that the above precoding matrices satisfy the alignment conditions.

We consider the following matrix,

$$\mathbf{H} = \begin{pmatrix} f_1(\mathbf{y}_1) & f_2(\mathbf{y}_1) & \cdots & f_r(\mathbf{y}_1) \\ f_1(\mathbf{y}_2) & f_2(\mathbf{y}_2) & \cdots & f_r(\mathbf{y}_2) \\ \cdots & \cdots & \cdots & \cdots \\ f_1(\mathbf{y}_r) & f_2(\mathbf{y}_r) & \cdots & f_r(\mathbf{y}_r) \end{pmatrix}$$

where $f_i(\mathbf{y})$ $(i = 1, 2, \cdots, r)$ is a rational function in terms of a vector of variables $\mathbf{y} = (y_1, \cdots, y_k)$ in $\mathbb{F}_{2^m}(\mathbf{y})$, and the jth row of $\mathbf{H}$ is simply a repetition of the vector $(f_1(\mathbf{y}), \cdots, f_r(\mathbf{y}))$, with $\mathbf{y}$ being replaced by a vector of variables $\mathbf{y}_j = (y_{j1}, \cdots, y_{jk})$. Due to the particular structure of $\mathbf{H}$, the problem of checking whether $\mathbf{H}$ is full rank can be simplified to checking whether $f_1(\mathbf{y}), \cdots, f_r(\mathbf{y})$ are linearly independent, as stated in the following lemma. Here, $f_1(\mathbf{y}), \cdots, f_r(\mathbf{y})$ are said to be linearly independent, if for any scalars $a_1, \cdots, a_r \in \mathbb{F}_q$, which are not all zeros, $a_1 f_1(\mathbf{y}) + \cdots + a_r f_r(\mathbf{y}) \neq 0$.

Lemma IV.2. $\det(\mathbf{H}) \neq 0$ if and only if $f_1(\mathbf{y}), \cdots, f_r(\mathbf{y})$ are linearly independent.

Proof. See Theorem 1 of [35]. ∎

An important observation is that using the precoding matrices defined in Eq. (12)-(14), all of the matrices involved in $\mathscr{B}'_1, \mathscr{B}'_2, \mathscr{B}'_3$ have the same form as $\mathbf{H}$. Specifically, each row of the matrix in $\mathscr{B}'_i$ is of the form:

$$\begin{matrix} (1 & \eta(\mathbf{x}) & \cdots & \eta^n(\mathbf{x}) \\ & p_i(\mathbf{x}) & \cdots & p_i(\mathbf{x})\eta^{n-1}(\mathbf{x})) \end{matrix} \tag{15}$$

where for $1 \leq j \leq n+1$, the jth element is $\eta^{j-1}(\mathbf{x})$, and for $n+2 \leq j \leq 2n+1$, the jth element is $p_i(\mathbf{x})\eta^{j-n-2}(\mathbf{x})$. Hence, using Lemma IV.2, we can quickly derive:

Lemma IV.3. Assume that all the senders are connected to all the receivers via directed paths, and $\eta(\mathbf{x})$ is non-constant. Consider a PBNA $\lambda_n = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$, where $\mathbf{V}_i$ is defined in Eq. (12)-(14). λ_n achieves the rate tuple $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$, if for each $1 \leq i \leq 3$, the following condition is satisfied:[5]

$$p_i(\mathbf{x}) \notin \mathcal{S}_n = \left\{ \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} : f(z), g(z) \in \mathbb{F}_q[z], f(z)g(z) \neq 0, \gcd(f(z), g(z)) = 1, d_f \leq n, d_g \leq n-1 \right\} \tag{16}$$

Proof. If Eq. (16) is satisfied, the rational functions in Eq. (15) are linearly independent. Therefore, due to Lemma IV.2, condition $\mathscr{B}'_i$ is satisfied. Hence, due to Lemma IV.1, $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$ is achieved by λ_n. ∎

Note that each rational function $\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}_n$ represents a constraint on $p_i(\mathbf{x})$, i.e., $p_i(\mathbf{x}) \neq \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$, the violation of which invalidates the use of the PBNA for achieving the rate tuple $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$ through the precoding matrices defined in Eq. (12)-(14). Also note that Eq. (16) only guarantees that PBNA achieves a symmetrical rate close to one half. In order for each unicast session to asymptotically achieve a transmission rate of one half, we simply combine the conditions of Lemma IV.3 for all possible values of n, and get the following result:

Theorem IV.1. Assume that all the senders are connected to all the receivers via directed paths, and $\eta(\mathbf{x})$ is non-constant. The symmetrical rate $\frac{1}{2}$ is asymptotically achievable through PBNA, if for each $1 \leq i \leq 3$,

$$p_i(\mathbf{x}) \notin \mathcal{S}' = \left\{ \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} : f(z), g(z) \in \mathbb{F}_q[z], f(z)g(z) \neq 0, \gcd(f(z), g(z)) = 1 \right\} \tag{17}$$

Proof. Consider the PBNA scheme λ_n defined in Lemma IV.3. If Eq. (17) is satisfied, Eq. (16) is satisfied, and thus λ_n achieves the rate tuple $(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1})$. Since $\lim_{n\to\infty}(\frac{n+1}{2n+1}, \frac{n}{2n+1}, \frac{n}{2n+1}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. This implies that the symmetrical rate $\frac{1}{2}$ is asymptotically achievable through PBNA. ∎

Case II: $\eta(\mathbf{x})$ is constant, and thus $\mathbf{T}$ is an identity matrix. For this case, Eq. (10) becomes trivial. In fact, we set $n = 1$, $s = 0$, and $\mathbf{BA} = \mathbf{C}$, and hence Eq. (10) can be satisfied by any *arbitrary* $\mathbf{V}_1$. Specifically, we use the following precoding matrices:

$$\mathbf{V}_1 = (\theta_1 \quad \theta_2)^T \tag{18}$$

$$\mathbf{V}_2 = \mathbf{M}_{13}\mathbf{M}_{23}^{-1}(\theta_1 \quad \theta_2)^T \tag{19}$$

$$\mathbf{V}_3 = \mathbf{M}_{12}\mathbf{M}_{32}^{-1}(\theta_1 \quad \theta_2)^T \tag{20}$$

where θ_1, θ_2 are variables. The above precoding matrices correspond to the configuration where $\mathbf{A} = \mathbf{B} = \mathbf{C} = \mathbf{I}_2$. Using the above precoding matrices, $\mathscr{A}_1 \sim \mathscr{A}_3$ all become equalities, *i.e.,* the interfering signals are perfectly aligned into a single linear space. Meanwhile, using these precoding matrices, each row of the matrix in $\mathscr{B}'_i$ is of the following form:

$$(\theta \quad p_i(\mathbf{x})\theta) \tag{21}$$

Hence, using Lemma IV.2, we can quickly derive:

Theorem IV.2. Assume that all the senders are connected to all the receivers via directed paths, and $\eta(\mathbf{x})$ is constant. Consider the PBNA scheme $\lambda = (\xi, \mathbf{V}_i : 1 \leq i \leq 3)$, where the precoding matrices are defined in Eq. (18)-(20). Then λ achieves the symmetrical rate $\frac{1}{2}$, if for each $1 \leq i \leq 3$, $p_i(\mathbf{x})$ is non-constant.

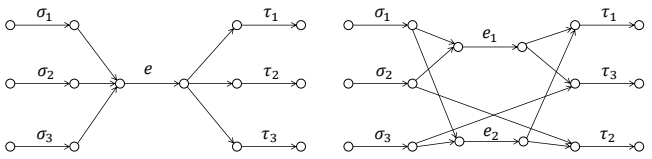
Proof. If $p_i(\mathbf{x})$ is not constant, the functions in Eq. (21) are linearly independent, and therefore $\mathscr{B}'_i$ is satisfied due to Lemma IV.2. Thus, $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is achieved by λ according to Lemma IV.1. ∎

As shown in the above theorem, if $\eta(\mathbf{x})$ is constant, each unicast session can achieve one half rate in exactly two time slots by using PBNA.

C. Coupling Relations and Achievability of PBNA

In the previous section, we reformulated the achievability conditions of PBNA in terms of the functions $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$. One critical question is: What is the connection between the reformulated conditions and network topology? We start by illustrating that through examples of networks whose structure violates these conditions. Let's first consider the network

[5]Notation: For two polynomials $f(x)$ and $g(x)$, let $\gcd(f(x), g(x))$ denote their greatest common divisor, and d_f the degree of $f(x)$.



(a) $p_1(\mathbf{x}) = p_2(\mathbf{x}) = p_3(\mathbf{x}) = \eta(\mathbf{x}) = 1$)

(b) $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$

Fig. 4. Examples of realizable coupling relations: The left network realizes the coupling relations $p_i(\mathbf{x}) = \eta(\mathbf{x}) = 1$ such that the conditions of Theorem IV.2 are violated; in the right network, $\eta(\mathbf{x}) \neq 1$, but $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, which violates the conditions of Theorem IV.1.

shown in Fig. 4a. Due to the bottleneck e, it can be easily verified that $p_1(\mathbf{x}) = p_2(\mathbf{x}) = p_3(\mathbf{x}) = \eta(\mathbf{x}) = 1$, and thus the conditions of Theorem IV.2 are violated. Moreover, consider the network shown in Fig. 4b. It can be easily verified that for this network, $\eta(\mathbf{x}) \neq 1$, and $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$. Thus the conditions of Theorem IV.1 are violated. Moreover, by exchanging $\sigma_1 \leftrightarrow \sigma_2$ and $\tau_1 \leftrightarrow \tau_2$, we obtain another example, where $p_2(\mathbf{x}) = 1 + \eta(\mathbf{x})$, and thus the conditions of Theorem IV.1 are again violated. While the key feature of the first example can be easily identified, it is not obvious what are the defining features of the second example. Nevertheless, both examples demonstrate an important difference between networks and wireless interference channel: In networks, due to the internal structure of transfer functions, network topology might introduce dependence between different transfer functions, e.g., $p_1(\mathbf{x}) = 1$ or $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$; in contrast, in wireless channel, channel gains are algebraically independent almost surely.

The above dependence relations can be seen as special cases of coupling relations, as defined below.

Definition IV.3. A *coupling relation* is an equation in the following form:

$$f(m_{i_1j_1}(\mathbf{x}), m_{i_2j_2}(\mathbf{x}), \cdots, m_{i_kj_k}(\mathbf{x})) = 0 \tag{22}$$

where $f(z_1, z_2, \cdots, z_k)$ is a polynomial in $\mathbb{F}_{2^m}[z_1, \cdots, z_k]$, $1 \leq i_l, j_l \leq 3$ for $1 \leq l \leq k$. If there exists a network $\mathcal{G}$ such that the transfer functions $m_{i_1j_1}(\mathbf{x}), m_{i_2j_2}(\mathbf{x}), \cdots, m_{i_kj_k}(\mathbf{x})$ satisfy the above equation, we say that the coupling relation Eq. (22) is *realizable*, or $\mathcal{G}$ *realizes* the coupling relation Eq. (22).

As shown in Theorem IV.1, each rational function $\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'$ represents a coupling relation $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$.

The existence of coupling relations greatly complicates the achievability problem of PBNA. As shown previously, most of the coupling relations, such as $p_1(\mathbf{x}) = 1$ and $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, are harmful to PBNA, because their presence violates the conditions of Theorems IV.1 and IV.2. The only exception is $\eta(\mathbf{x}) = 1$, which does help simplify the construction of precoding matrices, and thus is beneficial to PBNA. Indeed, as shown in Theorem IV.2, this coupling relation allows interferences to be perfectly aligned at each receiver, and each unicast session can achieve one half rate in exactly two time slots. Unfortunately, as we will see in Section VII, this coupling relation requires that the network possesses particular structures, which are absent in most networks. For this reason, we will mainly focus on the case $\eta(\mathbf{x}) \neq 1$, which is applicable for most networks.

One interesting observation is that not all coupling relations are realizable. For example, consider the coupling relation $p_1(\mathbf{x}) = \eta^3(\mathbf{x})$, where both $p_1(\mathbf{x})$ and $\eta(\mathbf{x})$ are non-constants. Let $p_1(\mathbf{x}) = \frac{u(\mathbf{x})}{v(\mathbf{x})}$, $\eta(\mathbf{x}) = \frac{s(\mathbf{x})}{t(\mathbf{x})}$ denote the *unique forms* [6] of $p_1(\mathbf{x})$ and $\eta(\mathbf{x})$ respectively. Consider a coding variable $x_{ee'}$ that appears in both $\frac{u(\mathbf{x})}{v(\mathbf{x})}$ and $\frac{s(\mathbf{x})}{t(\mathbf{x})}$. Because the maximum degree of each coding variable in a transfer function is at most one, according to Eq. (11), the maximum of the degrees of $x_{ee'}$ in $u(\mathbf{x})$ and $v(\mathbf{x})$ is at most two. However, it can be easily seen that the maximum of the degrees of $x_{ee'}$ in $s^3(\mathbf{x})$ and $t^3(\mathbf{x})$ is at least three. Therefore, it is impossible that $p_1(\mathbf{x}) = \eta^3(\mathbf{x})$. This example suggests that there exists significant redundancy in the conditions of Theorem IV.1. More formally, it raises the following important question:

Q1: Which coupling relations $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'$ are realizable?

The answer to this question allows us to reduce the set $\mathcal{S}'$ defined in Theorem IV.1 to its minimal size. For $i = 1, 2, 3$, we define the following set, which represents the minimal set of coupling relations we need to consider:

$$\mathcal{S}'_i = \left\{ \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}' : p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \text{ is realizable} \right\} \tag{23}$$

Then the next important question is:

Q2: Given $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'_i$, what are the defining features of the networks for which this coupling relation holds?

As we will see in the rest of this paper, the answers to Q1 and Q2 both lie in a deeper understanding of the properties of transfer functions. Intuitively, because each transfer function is defined on a graph, it usually possesses special properties. The graph-related properties not only allow us to reduce $\mathcal{S}'$ to the minimal set $\mathcal{S}'_i$, but also enable us to identify the defining features of the networks which realize the coupling relations represented by $\mathcal{S}'_i$.

In the derivation of Theorem IV.1, we only consider the precoding matrices defined in Eq. (12)-(14). However, the choices of precoding matrices are not limited to these matrices. In fact, as we will see in Section VI, given different $\mathbf{A}, \mathbf{B}$, and $\mathbf{C}$, we can derive different precoding matrix $\mathbf{V}_1$ such that Eq. (10) is satisfied. This raises the following interesting question:

Q3: Assume some coupling relation $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'_i$ is present in the network. Is it still possible to utilize PBNA via other precoding matrices instead of those defined in Eq. (12)-(14)?

As we will see in Section VI, the answer to this question is negative. The basic idea is that each precoding matrix $\mathbf{V}_1$ that satisfies Eq. (10) can be transformed into the precoding matrix in Eq. (12) through a transform equation $\mathbf{V}_1^* = \mathbf{G}^{-1}\mathbf{V}_1\mathbf{F}^{-1}$, where $\mathbf{G}$ is a diagonal matrix and $\mathbf{F}$ a full-rank matrix (See Lemma VI.3). Using this transform equation, we can prove that if the precoding matrices cannot be used due to the

[6] For a non-zero rational function $h(\mathbf{y}) \in \mathbb{F}_q(\mathbf{y})$, its unique form is defined as $h(\mathbf{y}) = \frac{f(\mathbf{y})}{g(\mathbf{y})}$, where $f(\mathbf{y}), g(\mathbf{y}) \in \mathbb{F}_q[\mathbf{y}]$ and $\gcd(f(\mathbf{y}), g(\mathbf{y})) = 1$.

presence of a coupling relation, then any precoding matrices cannot be used.

V. Overview of Main Results

In this section, we state our main results. Proofs are deferred to Sections VI and VIII, and Appendices.

A. Sufficient and Necessary Conditions for PBNA to Achieve Symmetrical Rate $\frac{1}{2}$

Since the construction of $\mathbf{V}_1$ depends on whether $\eta(\mathbf{x})$ is constant, we distinguish two cases.

1) $\eta(\mathbf{x})$ Is Not Constant :

Theorem V.1 (The Main Theorem)**.** Assume that all the senders are connected to all the receivers via directed paths, and $\eta(\mathbf{x})$ is not constant. The three unicast sessions can asymptotically achieve the rate tuple $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ through PBNA if and only if the following conditions are satisfied:

$$m_{11}(\mathbf{x}) \neq \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}{m_{23}(\mathbf{x})}, \frac{m_{12}(\mathbf{x})m_{31}(\mathbf{x})}{m_{32}(\mathbf{x})}, \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}{m_{23}(\mathbf{x})} + \frac{m_{12}(\mathbf{x})m_{31}(\mathbf{x})}{m_{32}(\mathbf{x})} \tag{24}$$

$$m_{22}(\mathbf{x}) \neq \frac{m_{12}(\mathbf{x})m_{23}(\mathbf{x})}{m_{13}(\mathbf{x})}, \frac{m_{32}(\mathbf{x})m_{21}(\mathbf{x})}{m_{31}(\mathbf{x})}, \frac{m_{12}(\mathbf{x})m_{23}(\mathbf{x})}{m_{13}(\mathbf{x})} + \frac{m_{32}(\mathbf{x})m_{21}(\mathbf{x})}{m_{31}(\mathbf{x})} \tag{25}$$

$$m_{33}(\mathbf{x}) \neq \frac{m_{23}(\mathbf{x})m_{31}(\mathbf{x})}{m_{21}(\mathbf{x})}, \frac{m_{13}(\mathbf{x})m_{32}(\mathbf{x})}{m_{12}(\mathbf{x})}, \frac{m_{23}(\mathbf{x})m_{31}(\mathbf{x})}{m_{21}(\mathbf{x})} + \frac{m_{13}(\mathbf{x})m_{32}(\mathbf{x})}{m_{12}(\mathbf{x})} \tag{26}$$

Proof. See Appendix B. ∎

Eq. (24)-(26) can be reformulated into the following equivalent conditions:

$$p_1(\mathbf{x}) \notin \mathcal{S}_1' = \left\{1, \eta(\mathbf{x}), \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}\right\} \tag{27}$$

$$p_2(\mathbf{x}) \notin \mathcal{S}_2' = \{1, \eta(\mathbf{x}), 1+\eta(\mathbf{x})\} \tag{28}$$

$$p_3(\mathbf{x}) \notin \mathcal{S}_3' = \{1, \eta(\mathbf{x}), 1+\eta(\mathbf{x})\} \tag{29}$$

Note that in Theorem V.1, we reduce the conditions of Theorem IV.1 to its minimal size, such that each $\mathcal{S}_i'$ as defined in Eq. (27)-(29) represents the minimal set of coupling relations that are realizable. Moreover, as we will see later, each of these coupling relations has a unique interpretation in terms of the network topology. The interpretations further provide polynomial-time algorithms to check the existence of these coupling relations.

The conditions of the Theorem V.1 can be understood from the perspective of the interference channel. As shown in Section IV-A, under linear network coding, the network behaves as a 3-user wireless interference channel, where the channel coefficients $m_{ij}(\mathbf{x})$ are all non-zeros. Let $\mathbf{H}$ denote the matrix with the (i,j)-element being $m_{ij}(\mathbf{x})$. It is easy to see that the first two inequalities in Eq. (24)-(26) can be rewritten as $M_{kl}(\mathbf{H}) \neq 0$ for some $k \neq l$, where $M_{kl}(\mathbf{H})$ denotes the (k,l)-Minor of $\mathbf{H}$. For example, $m_{11}(\mathbf{x}) \neq \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}{m_{23}(\mathbf{x})}$ is equivalent to $M_{32}(\mathbf{H}) \neq 0$, and $m_{11}(\mathbf{x}) \neq \frac{m_{12}(\mathbf{x})m_{31}(\mathbf{x})}{m_{32}(\mathbf{x})}$ is equivalent to $M_{23}(\mathbf{H}) \neq 0$. Suppose that there exists $M_{kl}(\mathbf{H}) = 0$ for some $k \neq l$. For such a channel, it is known that the sum-rate achieved by the three unicast sessions cannot be more than 1 in the information theoretical sense (see Lemma 1 of [40]), i.e., no precoding-based linear scheme can achieve a rate beyond 1/3 per user. Therefore, given that all senders are connected to all receivers, the condition $M_{kl}(\mathbf{H}) \neq 0$ is information theoretically necessary for achievable rate 1/2 per session. Hence, the first two inequalities of Eq. (24)-(26) are simply the information theoretic necessary conditions, so they must hold for any precoding-based linear schemes.

2) $\eta(\mathbf{x})$ Is Constant : In this case, we can choose $\mathbf{V}_1$ freely by setting $\mathbf{BA} = \mathbf{C}$. As stated in the following theorem, each unicast session can achieve one half rate in exactly two time slots.

Theorem V.2. Assume that all the senders are connected to all the receivers via directed paths, and $\eta(\mathbf{x})$ is constant. The three unicast sessions can achieve the rate tuple $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ in exactly two time slots through PBNA if and only if the following conditions are satisfied:

$$m_{11}(\mathbf{x}) \neq \frac{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}{m_{23}(\mathbf{x})} \tag{30}$$

$$m_{22}(\mathbf{x}) \neq \frac{m_{12}(\mathbf{x})m_{23}(\mathbf{x})}{m_{13}(\mathbf{x})} \tag{31}$$

$$m_{33}(\mathbf{x}) \neq \frac{m_{23}(\mathbf{x})m_{31}(\mathbf{x})}{m_{21}(\mathbf{x})} \tag{32}$$

Proof. See Section VI-B. ∎

Eq. (30)-(32) can be reformulated into the following equivalent conditions:

$$p_i(\mathbf{x}) \neq 1 \quad \forall 1 \leq i \leq 3$$

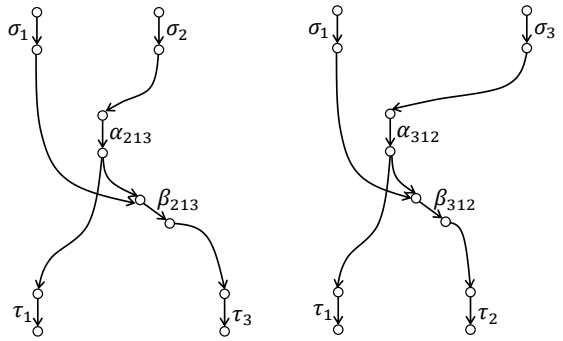
B. Topological Interpretations of the Feasibility Conditions

As we have seen, the following coupling relations are important for the achievability of PBNA: 1) $\eta(\mathbf{x}) = 1$; 2) $p_i(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) = \eta(\mathbf{x})$ where $i = 1, 2, 3$; 3) $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, $p_i(\mathbf{x}) = 1+\eta(\mathbf{x})$, where $i = 2, 3$. As we will see, the networks that realize these coupling relations have special topological properties. We defer all the proofs to Appendix C.

We assume that all the edges in E are arranged in a topological ordering such that if $\mathrm{head}(e) = \mathrm{tail}(e')$, e must precede e' in this ordering.

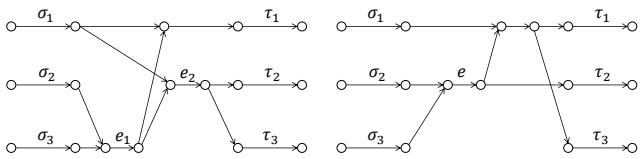
Definition V.1. Given two subsets of edges S and D, we define an edge e as a *bottleneck* between S and D if the removal of e will disconnect every directed path from S to D.

Given $1 \leq i, j, k \leq 3$, let α_{ijk} denote the last bottleneck between σ_i and $\{\tau_j, \tau_k\}$ in this topological ordering, and β_{ijk} the first bottleneck between $\{\sigma_j, \alpha_{ijk}\}$ and τ_k.



(a) α_{213} and β_{213} (b) α_{312} and β_{312}

Fig. 5. A graphical illustration of the four edges, α_{213}, β_{213}, α_{312}, and β_{312}, which are important in defining the networks that realize $\eta(\mathbf{x}) = 1$.



(a) $\eta(\mathbf{x}) = 1$ (b) $p_2(\mathbf{x}) = \eta(\mathbf{x})$

Fig. 6. Additional examples of coupling relations

As shown below, the four edges, α_{213}, β_{213}, α_{312}, and β_{312}, are important in defining the networks that realize $\eta(\mathbf{x}) = 1$. A graphical illustration of the four edges is shown in Fig. 5.

Theorem V.3. $\eta(\mathbf{x}) = 1$ if and only if $\alpha_{213} = \alpha_{312}$ and $\beta_{213} = \beta_{312}$.

In [35], the authors independently discovered a similar result. Consider the example shown in Fig. 4a. It is easy to see that in this example, $\alpha_{213} = \alpha_{312} = \beta_{213} = \beta_{312} = e$, and thus $\eta(\mathbf{x}) = 1$. In Fig. 6a, we show another example, where $\alpha_{213} = \alpha_{312} = e_1$, $\beta_{213} = \beta_{312} = e_2$, and thus $\eta(\mathbf{x}) = 1$.

Given two subsets of edges, S and D, a cut-set C between S and D is a subset of edges, the removal of which will disconnect every directed path from S to D. The capacity of cut-set C is defined as the summation of the capacities of the edges contained in C. The minimum cut between S and D is the minimum capacity of all cut-sets between S and D.

Theorem V.4. The following statements hold:

1) $p_1(\mathbf{x}) = 1$ if and only if the minimum cut between $\{\sigma_1, \sigma_2\}$ and $\{\tau_1, \tau_3\}$ equals one; $p_1(\mathbf{x}) = \eta(\mathbf{x})$ if and only if the minimum cut between $\{\sigma_1, \sigma_3\}$ and $\{\tau_1, \tau_2\}$ equals one.
2) $p_2(\mathbf{x}) = 1$ if and only if the minimum cut between $\{\sigma_1, \sigma_2\}$ and $\{\tau_2, \tau_3\}$ equals one; $p_2(\mathbf{x}) = \eta(\mathbf{x})$ if and only if the minimum cut between $\{\sigma_2, \sigma_3\}$ and $\{\tau_1, \tau_2\}$ equals one.
3) $p_3(\mathbf{x}) = 1$ if and only if the minimum cut between $\{\sigma_2, \sigma_3\}$ and $\{\tau_1, \tau_3\}$ equals one; $p_3(\mathbf{x}) = \eta(\mathbf{x})$ if and only if the minimum cut between $\{\sigma_1, \sigma_3\}$ and $\{\tau_2, \tau_3\}$ equals one.

For instance, in Fig. 4a, the cut-set with minimum capacity between $\{\sigma_2, \sigma_3\}$ and $\{\tau_1, \tau_2\}$ contains only one edge e, and thus $p_2(\mathbf{x}) = \eta(\mathbf{x})$.

Given two edges e_1 and e_2, we say that they are parallel with each other if there is no directed paths from e_1 to e_2, or from e_2 to e_1. As shown below, two edges are important in defining the networks that realizes the third coupling relation in Eq. (27)-(29), *e.g.,* α_{213} and α_{312} are used to define the networks that realize $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, and so on.

Theorem V.5. The following statements hold:

1) $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ if and only if the following conditions are satisfied: a) α_{312} is a bottleneck between σ_1 and τ_2; b) α_{213} is a bottleneck between σ_1 and τ_3; c) α_{312} is parallel with α_{213}; d) $\{\alpha_{312}, \alpha_{213}\}$ forms a cut-set between σ_1 from τ_1.
2) $p_2(\mathbf{x}) = 1+\eta(\mathbf{x})$ if and only if the following conditions are satisfied: a) α_{123} is a bottleneck between σ_2 and τ_3; b) α_{321} is a bottleneck between σ_2 and τ_1; c) α_{123} is parallel with α_{321}; d) $\{\alpha_{123}, \alpha_{321}\}$ forms a cut-set between σ_2 from τ_2.
3) $p_3(\mathbf{x}) = 1+\eta(\mathbf{x})$ if and only if the following conditions are satisfied: a) α_{231} is a bottleneck between σ_3 and τ_1; b) α_{132} is a bottleneck between σ_3 and τ_2; c) α_{231} is parallel with α_{132}; d) $\{\alpha_{231}, \alpha_{132}\}$ forms a cut-set between σ_3 from τ_3.

Consider the network as shown in Fig. 4b. It is easy to see that $e_2 = \alpha_{312}$ and $e_1 = \alpha_{213}$, and all the conditions in 1) of Theorem V.5 are satisfied. Therefore, this network realizes the coupling relation $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$. Note that these three coupling relations are mutually exclusive when $\eta(\mathbf{x})$ is not constant. If any two of these coupling relation were to occur in the same network, then it would induce a graph structure that forces $\eta(\mathbf{x})$ to be a constant [35].

C. Optimal Symmetric Rates Achieved by Precoding-Based Linear Schemes

For SISO scenarios where all senders are connected to all receivers, there are only three possible rates achievable through any precoding-based network coding schemes.

Definition V.2. We classify the networks based on the coupling relations present in the network as follows:

- $Type\ I$: Networks in which at least one of the coupling relations, $p_i(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) = \eta(\mathbf{x})$ $(1 \leq i \leq 3)$, is present.
- $Type\ II$: Networks in which $p_i(\mathbf{x}) \notin \{1, \eta(\mathbf{x})\}$ for $1 \leq i \leq 3$, but one of the three mutually exclusive coupling conditions, $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, $p_2(\mathbf{x}) = 1 + \eta(\mathbf{x})$, and $p_3(\mathbf{x}) = 1 + \eta(\mathbf{x})$, is present.
- $Type\ III$: Networks in which none of the above coupling relations is present.

Theorem V.6. Assume that all the senders are connected to all the receivers via directed paths. The following statements hold:

1) The optimal symmetric rate achieved by precoding-based linear schemes for $Type\ I$ networks is $1/3$ per unicast session.

2) The optimal symmetric rate achieved by precoding-based linear schemes for $Type\ II$ networks is $2/5$ per unicast session.
3) The optimal symmetric rate achieved by precoding-based linear schemes for $Type\ III$ networks is $1/2$ per unicast session.

Moreover, all of the above optimal symmetric rate is achievable through PBNA schemes.

Proof. See Section VIII. ■

VI. Sufficient and Necessary Conditions for PBNA to Achieve Symmetric Rate $\frac{1}{2}$

In this section, we explain the main ideas behind the proofs of Theorem V.1 and V.2. Consistent with Section V, we distinguish two cases based on whether $\eta(\mathbf{x})$ is constant.

A. $\eta(\mathbf{x})$ Is Not Constant

In this subsection, we first present a simple method to quickly identify a class of networks, for which PBNA can asymptotically achieve symmetric rate $\frac{1}{2}$. Then, we sketch the outline of the proof for the sufficiency of Theorem V.1. Next, we explain the main idea behind the proof for the necessity of Theorem V.1.

1) A Simple Method Based on Theorem IV.1: As shown in Theorem IV.1, the set $\mathcal{S}'$ contains an exponential number of rational functions, and thus it is very difficult to check the conditions of Theorem IV.1 in practice. Interestingly, the theorem directly yields a simple method to quickly identify a class of networks for which PBNA is feasible. The major idea of the method is to exploit the asymmetry between $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$ in terms of effective variables. Here, given a rational function $f(\mathbf{y})$, we define a variable as an effective variable of $f(\mathbf{y})$ if it appears in the unique form of $f(\mathbf{y})$. Let $\mathcal{V}(f(\mathbf{y}))$ denote the set of effective variables of $f(\mathbf{y})$. Intuitively, this asymmetry allows us more freedom to control the values of $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$ such that they can change independently, which makes the network behave more like a wireless channel. The formal description of the method is presented below:

Corollary VI.1. Assume all $m_{ij}(\mathbf{x})$'s $(i,j=1,2,3)$ are non-zeros, and $\eta(\mathbf{x})$ is not constant. Each unicast session can asymptotically achieve one half rate through PBNA if for $i=1,2,3$, $p_i(\mathbf{x})\neq 1$ and $\mathcal{V}(\eta(\mathbf{x}))\neq\mathcal{V}(p_i(\mathbf{x}))$.

Proof. If the above conditions are satisfied, we must have $p_i(\mathbf{x})\neq\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}\in\mathcal{S}'$. Thus, the theorem holds. ■

Consider the networks shown in Fig. 7a and Fig. 2, which we replicate in Fig. 7b for easy review. As shown in these examples, due to edge e, $\eta(\mathbf{x})$ contains effective variables $x_{\sigma_3 e}, x_{e\tau_2}$, which are absent in the unique form of $p_i(\mathbf{x})$ $(i=1,2,3)$. Thus, by Corollary VI.1, each unicast session can asymptotically achieve one half rate through PBNA. However, Corollary VI.1 doesn't subsume all possible networks for which PBNA can achieve one half rate. For instance, in Fig. 7c, we show a counter example, where $\mathcal{V}(\eta(\mathbf{x}))=\mathcal{V}(p_1(\mathbf{x}))$, and thus Corollary VI.1 is not applicable. Nevertheless, it is easy to verify the network satisfies the conditions of Theorem V.1, and thus PBNA can still achieve one half rate.

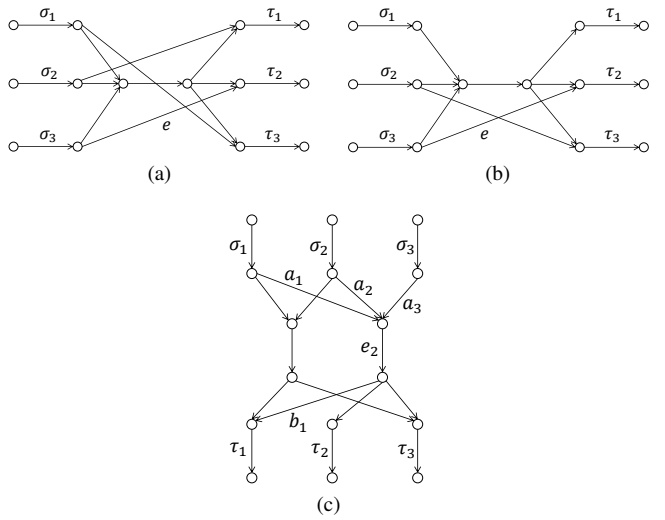


Fig. 7. Illustration of type III networks. (i) It can be seen that for all the three examples, PBNA can achieve one half rate. (ii) The three examples can be verified by using different methods: for (a) and (b), due to edge e, $\eta(\mathbf{x})$ contains coding variables $x_{\sigma_3 e}, x_{e\tau_2}$, which are absent in the unique forms of $p_1(\mathbf{x}), p_2(\mathbf{x})$ and $p_3(\mathbf{x})$, and thus Corollary VI.1 applies to both cases; Corollary VI.1 doesn't apply to (c), but PBNA can still achieve a symmetric rate $\frac{1}{2}$ for this network according to Theorem V.1. (iii) For both (a) and (b), routing can only achieve a symmetrical rate $\frac{1}{3}$; for (c), PBNA and routing can both achieve a symmetrical rate $\frac{1}{2}$.

2) Sufficiency of Theorem V.1: As shown in Section IV, not all coupling relations $p_i(\mathbf{x})=\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}\in\mathcal{S}'$ are realizable due to the special properties of transfer functions. Indeed, since the transfer functions are defined on graphs, they exhibit special properties due to the graph structure. As we will see, these properties are essential in identifying the minimal sub-set of realizable coupling relations. In fact, we only need two such properties, namely Linearization Property and Square-Term Property.

The proof consists of three steps. First, we use Linearization Property and a simple degree-counting technique to reduce $\mathcal{S}'$ to the following set $\mathcal{S}_1''$: We consider the general form of $p_i(\mathbf{x})$ as below

$$h(\mathbf{x})=\frac{m_{ab}(\mathbf{x})m_{pq}(\mathbf{x})}{m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})} \tag{33}$$

Note that $\mathcal{S}_1''$ only includes a finite number of rational functions. where $a,b,p,q=1,2,3$ and $a\neq p, b\neq q$. Moreover, by the definition of transfer function, the numerator and denominator of $h(\mathbf{x})$ can be expanded respectively as follows:

$$m_{ab}(\mathbf{x})m_{pq}(\mathbf{x})=\sum\nolimits_{(P_1,P_2)\in\mathcal{P}_{ab}\times\mathcal{P}_{pq}} t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x})$$
$$m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})=\sum\nolimits_{(P_3,P_4)\in\mathcal{P}_{aq}\times\mathcal{P}_{pb}} t_{P_3}(\mathbf{x})t_{P_4}(\mathbf{x})$$

Hence, each path pair in $\mathcal{P}_{ab}\times\mathcal{P}_{pq}$ contributes a term in $m_{ab}(\mathbf{x})m_{pq}(\mathbf{x})$, and each path pair in $\mathcal{P}_{aq}\times\mathcal{P}_{pb}$ contributes a term in $m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})$.

The first property, the Linearization Property, is stated in the following lemma. According to this property, if $p_i(\mathbf{x})\neq 1$, it can be transformed into its simplest non-trivial form, i.e., a linear function or the inverse of a linear function, through a partial assignment of values to $\mathbf{x}$.

Lemma VI.1 (The Linearization Property)**.** Assume $h(\mathbf{x})$ is not constant. Let $h(\mathbf{x}) = \frac{u(\mathbf{x})}{v(\mathbf{x})}$ such that $\gcd(u(\mathbf{x}), v(\mathbf{x})) = 1$. Then, we can assign values to $\mathbf{x}$ other than a variable $x_{ee'}$ such that $u(\mathbf{x})$ and $v(\mathbf{x})$ are transformed into either $u(x_{ee'}) = c_1 x_{ee'} + c_0$, $v(x_{ee'}) = c_2$ or $u(x_{ee'}) = c_2, v(x_{ee'}) = c_1 x_{ee'} + c_0$, where c_0, c_1, c_2 are constants in $\mathbb{F}_{2^m}$, and $c_1 c_2 \neq 0$.

Proof. See Appendix A. ∎

The second property, namely the Square-Term Property, is presented in the following lemma. According to this property, the coefficient of $x_{ee'}^2$ in the numerator of $h(\mathbf{x})$ equals its counter-part in the denominator of $h(\mathbf{x})$. Thus, if $x_{ee'}^2$ appears in the numerator of $h(\mathbf{x})$ under some assignment to $\mathbf{x}$, it must also appear in the denominator of $h(\mathbf{x})$, and vice versa.

Lemma VI.2 (The Square-Term Property)**.** Given a coding variable $x_{ee'}$, let $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ be the coefficients of $x_{ee'}^2$ in $m_{ab}(\mathbf{x})m_{pq}(\mathbf{x})$ and $m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})$ respectively. Then $f_1(\mathbf{x}) = f_2(\mathbf{x})$.

Proof. See Appendix A ∎

Now, we sketch the outline for the proof of the sufficiency of Theorem V.1. The proof consists of three steps:

First, we use the Linearization Property and a simple degree-counting technique to reduce $\mathcal{S}'$ to the following set $\mathcal{S}_1''$:

$$\mathcal{S}_1'' = \left\{ \frac{a_0 + a_1\eta(\mathbf{x})}{b_0 + b_1\eta(\mathbf{x})} \in \mathcal{S}' : a_0, a_1, b_0, b_1 \in \mathbb{F}_q \right\} \tag{34}$$

Next, we iterate through all possible configurations of a_0, a_1, b_0, b_1, and utilize the Linearization Property and the Square-Term Property to further reduce $\mathcal{S}_1''$ to just four rational functions:

$$\mathcal{S}_2'' = \left\{ 1, \eta(\mathbf{x}), 1 + \eta(\mathbf{x}), \frac{\eta(\mathbf{x})}{1 + \eta(\mathbf{x})} \right\} \tag{35}$$

Finally, we use a recent result from [35] to rule out the fourth redundant rational function in $\mathcal{S}_2''$, resulting in the minimal set $\mathcal{S}_i'$ defined in Theorem V.1. The detailed proof is deferred to Appendix B.

3) Necessity of the Conditions of Theorem V.1: We first show how to get a precoding matrix $\mathbf{V}_1$ that satisfies Eq. (12). The construction of $\mathbf{V}_1$ involves solving a system of linear equations defined on $\mathbb{F}_{2^m}(\xi)(z)$:

$$\mathbf{r}(z)(z\mathbf{C} - \mathbf{BA}) = 0 \tag{36}$$

In the above equation, $\mathbf{r}(z) = (r_1(z), \cdots, r_{n+s}(z))$, where $r_i(z) \in \mathbb{F}_{2^m}(\xi)(z)$ for $1 \leq i \leq n+s$. Assume $\mathbf{r}_0(z)$ is a non-zero solution to Eq. (36). Substitute z with $\eta(\mathbf{x})$, and we have $\eta(\mathbf{x})\mathbf{r}_0(\eta(\mathbf{x}))\mathbf{C} = \mathbf{r}_0(\eta(\mathbf{x}))\mathbf{BA}$. Finally, construct the following precoding matrix

$$\mathbf{V}_1^T = (\mathbf{r}_0^T(\eta(\mathbf{x}^{(1)})) \quad \mathbf{r}_0^T(\eta(\mathbf{x}^{(2)})) \quad \cdots \quad \mathbf{r}_0^T(\eta(\mathbf{x}^{(2n+s)}))) \tag{37}$$

Apparently, $\mathbf{V}_1$ satisfies Eq. (10). Hence, each non-zero solution to Eq. (36) corresponds to a row of $\mathbf{V}_1$ satisfying Eq. (10). Conversely, it is straightforward to see that each row of $\mathbf{V}_1$ satisfying Eq. (10) corresponds to a solution to Eq. (36).

As we will prove in Appendix B, $\text{rank}(z\mathbf{C} - \mathbf{BA}) = n$. If $s = 0$, $z\mathbf{C} - \mathbf{BA}$ becomes an invertible square matrix, and Eq. (36) only has zero solution. Thus, in order for Eq. (12) to have a non-zero solution, s must equal 1.

As an example, consider the case where $s = 1$, $n = 2$, and $2^m = 4$. Let α be the primitive element of $\mathbb{F}_4$ such that $\alpha^3 = 1$ and $\alpha^2 + \alpha + 1 = 0$. Moreover, let $\mathbf{A} = \mathbf{I}_2$ and

$$\mathbf{C} = \begin{pmatrix} 1 & \alpha \\ \alpha & 1 \\ \alpha^2 & 1 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} \alpha^2 & \alpha \\ 1 & 1 \\ 1 & \alpha \end{pmatrix}$$

It's easy to verify that $\mathbf{r}(z) = (\alpha^2 z^2 + \alpha, z + \alpha, z^2 + \alpha z + \alpha^2)$ satisfies Eq. (36). Thus, we substitute z with $\eta(\mathbf{x}^j)$ and construct $\mathbf{V}_1^T = (\mathbf{r}^T(\eta(\mathbf{x}^1)) \ \mathbf{r}^T(\eta(\mathbf{x}^2)) \ \cdots \ \mathbf{r}^T(\eta(\mathbf{x}^5)))$. Apparently, Eq. (10) is satisfied. From this example, we can see that given different $\mathbf{A}, \mathbf{B}, \mathbf{C}$, we can construct different precoding matrix $\mathbf{V}_1$, and thus the choices of precoding matrices are not limited to those defined in Eq. (12)-(14). An interesting observation is that the above precoding matrix $\mathbf{V}_1$ is closely related to Eq. (12) through a transform equation: $\mathbf{V}_1 = \mathbf{V}_1^*\mathbf{F}$, where

$$\mathbf{F} = \begin{pmatrix} \alpha & \alpha & \alpha^2 \\ 0 & 1 & \alpha \\ \alpha^2 & 0 & 1 \end{pmatrix}$$

Actually, this observation can be generalized to the following Lemma.

Lemma VI.3. Assume $s = 1$. Any $\mathbf{V}_1$ satisfying Eq. (10) is related to $\mathbf{V}_1^*$ through the following transform equation

$$\mathbf{V}_1 = \mathbf{G}\mathbf{V}_1^*\mathbf{F} \tag{38}$$

where $\mathbf{V}_1^*$ is defined in Eq. (12), $\mathbf{F}$ is an $(n+1) \times (n+1)$ matrix, and $\mathbf{G}$ is a $(2n+1) \times (2n+1)$ diagonal matrix, with the (i,i) element being $f_i(\eta(\mathbf{x}^i))$, where $f_i(z)$ is an arbitrary non-zero rational function in $\mathbb{F}_{2^m}(\xi)(z)$. Moreover, the $(n+1)$th row of $\mathbf{FC}$ and the 1st row of $\mathbf{FBA}$ are both zero vectors.

Proof. See Appendix B. ∎

Using Lemma VI.3, we can prove that if a coupling relation $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'$ is present in the network, any PBNA cannot achieve one half rate per unicast session. This implies that the conditions of Theorem V.1 are also necessary for PBNA to achieve one half rate per unicast session. We defer the detailed proof to Appendix B.

B. $\eta(\mathbf{x})$ Is Constant

Proof of Theorem V.2. In the proof of Theorem IV.2, we've proved the sufficiency of Theorem V.2. If $p_i(\mathbf{x}) = 1$, $\mathbf{P}_i$ becomes an identity matrix. We will show that it is impossible for PBNA to achieve one half rate for each unicast session. We only prove the case for $i = 1$. The other cases $i = 2, 3$ can be proved similarly, and are omitted. The matrix in the reformulated rank condition $\mathcal{B}_1'$ becomes $(\mathbf{V}_1 \quad \mathbf{V}_1\mathbf{C})$. Since $\text{rank}(\mathbf{V}_1\mathbf{C}) = n$, there are n columns in $\mathbf{V}_1$ that are linearly dependent of the columns in $\mathbf{V}_1\mathbf{C}$. Thus, it is impossible for PBNA to achieve one half rate for ω_1. ∎

In Fig. 8, we show an example of this case. Note that the network in Fig. 8 has rich connectivity such that each

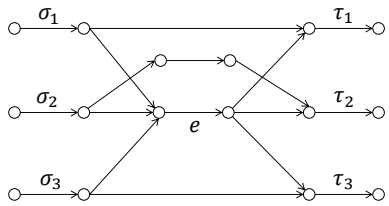


Fig. 8. An example where $\eta(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) \neq 1$ for $i \in \{1, 2, 3\}$, and thus each unicast session can achieve one half rate in exactly two time slots due to Theorem V.2. For this example, routing achieves symmetric rate of one

sender is connected to its corresponding receiver via a disjoint directed path. Thus, there is no coding opportunity that can be exploited, and routing is sufficient to achieve rate 1 per unicast session, which is the maximum symmetric rate achieved by any network coding schemes. Hence, this class of networks is of less significance than the class of networks considered in Theorem V.1.

C. Some s_i Is Disconnected from Some d_j ($i \neq j$)

In this case, since the number of interfering signals is reduced, at least one alignment condition can be removed, and thus the restriction on $\mathbf{V}_1$ imposed by Eq. (10) vanishes. Therefore, we can choose $\mathbf{V}_1$ freely, and the feasibility conditions of PBNA can be greatly simplified. For example, assume $m_{21}(\mathbf{x}) = 0$ and all other transfer functions are non-zeros. Hence, the alignment condition for the first unicast session vanishes. Using a scheme similar to above, we set $\mathbf{V}_1 = (\theta_1 \quad \theta_2)^T$, $\mathbf{V}_2 = \mathbf{M}_{13}\mathbf{M}_{23}^{-1}(\theta_1 \quad \theta_2)^T$ and $\mathbf{V}_3 = \mathbf{M}_{12}\mathbf{M}_{32}^{-1}(\theta_1 \quad \theta_2)^T$, and thus the interferences at τ_2 and τ_3 are all perfectly aligned. It is easy to see that $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ is feasible through PBNA if and only if $p_i(\mathbf{x})$ is not constant for every $i = 1, 2, 3$. Using similar arguments, we can discuss other cases.

VII. Checking the Feasibility Conditions of PBNA

In this section, we propose a polynomial-time algorithm to check the feasibility conditions of PBNA. We use $\mathcal{C}_{e_1e_2}$ to denote the set of bottlenecks between two edges e_1 and e_2, and use $\mathcal{C}_{ij}$ to represent $\mathcal{C}_{\sigma_i\tau_j}$. Using this notation, it can be easily seen that α_{ijk} is the last edge of the topological ordering of the edges in $\mathcal{C}_{ij} \cap \mathcal{C}_{ik}$, and β_{ijk} is the first edge of the topological ordering of the edges in $\mathcal{C}_{jk} \cap \mathcal{C}_{\alpha_{ijk},\tau_k}$.

We assume $\mathcal{G}$ is stored as an adjacency list, i.e., for each node $v \in V'$, we associate it with the set of its incoming edges and the set of its outgoing edges. Moreover, we assume all the edges in $\mathcal{G}$ have been arranged in topological order.

The checking process consists of the following steps: 1) Check if $\eta(\mathbf{x}) = 1$; 2) if $\eta(\mathbf{x}) = 1$, check the conditions of Theorem V.2; 3) otherwise, check the conditions of Theorem V.1. In the following discussion, we present the building blocks involved in these steps.

1) Calculating $\mathcal{C}_{ee'}$: We use Algorithm 1 to calculate the set of bottlenecks $\mathcal{C}_{ee'}$ which separates e from e'. The algorithm consists of two steps: 1) Lines 1-3 are used to calculate the set of edges traversed by the paths in $P_{ee'}$, denoted by

Algorithm 1: Calculate $\mathcal{C}_{ee'}$

```
1  Use BFS (Breadth First Search) algorithm to calculate the set
   of edges reachable from e, denoted by E_1;
2  Use reverse BFS algorithm to calculate the set of edges which
   is connected to e', denoted by E_2;
3  E_ee' ← E_1 ∩ E_2;
4  C_ee' ← {e}, C ← {e};
5  for each e_1 ∈ E_ee' in the topological order do
6  |   C ← C − {e_1};
7  |   for each e_2 such that head(e_1) = tail(e_2) do
8  |   |   if e_2 ∈ E_ee' then C ← C ∪ {e_2} ;
9  |   end
10 |   if C contains one edge then C_ee' ← C_ee' ∪ C ;
11 end
```

$E_{ee'}$. Note that in the reverse BFS algorithm, we start from e' and move upwards by following the incoming edges associated with each node. 2) Lines 4-11 are used to calculate $\mathcal{C}_{ee'}$. In this step, we iterate through each edge $e \in E_{ee'}$ in the topological order. In each iteration, we calculate $\mathcal{C}$, which forms a cut separating e from e'. If $\mathcal{C}$ contains only one edge, we then incorporate $\mathcal{C}$ into $\mathcal{C}_{ee'}$. The running time of the algorithm is $O(h|E|)$, where h is the maximum in-degree of nodes in G'.

2) Checking if $\eta(\mathbf{x}) = 1$: Using algorithm 1 and Theorem V.3, we can easily check whether this coupling relation holds. First, we calculate $\mathcal{C}_{31} \cap \mathcal{C}_{32}$, $\mathcal{C}_{21} \cap \mathcal{C}_{23}$, from which we get the two edges α_{312} and α_{213}. Then, we calculate $\mathcal{C}_{12} \cap \mathcal{C}_{\alpha_{312},\tau_2}$, $\mathcal{C}_{13} \cap \mathcal{C}_{\alpha_{213},\tau_3}$, from which we get β_{312} and β_{213}. Finally, we use Theorem V.3 to check if $\eta(\mathbf{x}) = 1$ by checking whether $\alpha_{312} = \beta_{312}$ and $\alpha_{213} = \beta_{213}$.

3) Checking if $p_i(\mathbf{x}) = 1$ or $p_i(\mathbf{x}) = \eta(\mathbf{x})$: Due to Theorem V.4, we use Ford-Fulkerson Algorithm to check these coupling relations. For example, in order to check whether $p_1(\mathbf{x}) = 1$, we add a super sender node s', which is connected to s_1 and s_2 via two directed edges of capacity one, and a super receiver node d', to which d_1 and d_3 are connected via two directed edges of capacity one. We then use Ford-Fulkerson Algorithm to calculate the maximum flow from s' to d', which is identical to the minimum capacity of cut-sets between $\{s_1, s_2\}$ and $\{d_1, d_2\}$, denoted by $C_{12,13}$. Thus, by checking whether $C_{12,13} = 1$, we can identify whether $p_1(\mathbf{x}) = 1$. Similarly, we can check other coupling relations.

4) Checking if $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ or $p_2(\mathbf{x}), p_3(\mathbf{x}) = 1+\eta(\mathbf{x})$: We use Algorithm 2 to check if $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$. The other two coupling relations can be checked similarly. Note that Line 4 consists of two steps: First, we start from α_{312} and use BFS to check if α_{213} is reachable from α_{312}; then we start from α_{213} and use BFS to check if α_{312} is reachable from α_{213}. The running time of the algorithm is $O(h|E|)$.

VIII. Optimal Linear Precoding-Based Rates

In this section, we prove that for SISO scenarios where all the senders are connected to all the receivers via directed paths, there are only three possible symmetric rates achieved by any precoding-based linear schemes. We'll also show that PBNA can achieve the optimal symmetric rate achieved by precoding-based linear schemes. In order to show this, we first prove that for the networks that violate one of the following

Algorithm 2: Check if $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$

$\alpha_{312} \leftarrow$ the last edge of $\mathcal{C}_{31} \cap \mathcal{C}_{32}$;
$\alpha_{213} \leftarrow$ the last edge of $\mathcal{C}_{21} \cap \mathcal{C}_{23}$;
if $\alpha_{312} \notin \mathcal{C}_{12}$ **or** $\alpha_{213} \notin \mathcal{C}_{13}$ **then return false** ;
Use BFS algorithm to check whether α_{312} is connected with α_{213} by a directed path;
if α_{312} is connected with α_{213} **then return false**;
Let G_1 denote the subgraph of G' induced by $E' - \{\alpha_{312}, \alpha_{213}\}$;
Use BFS algorithm to check whether τ_1 is connected to σ_1 in G_1;
if τ_1 is connected to σ_1 in G_1 **then return false** ;
else return true ;

three conditions: $p_1(\mathbf{x}) \neq \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, $p_2(\mathbf{x}) \neq 1 + \eta(\mathbf{x})$, and $p_3(\mathbf{x}) \neq 1 + \eta(\mathbf{x})$, it is not possible to achieve a symmetric rate of more than 2/5 per user, through any precoding-based scheme (the proof follows from [41]). We also show that this outer bound of $2/5$ is achievable through our PBNA scheme and thus it is tight.

Consider any precoding-based linear scheme over N channel uses. Let $\tilde{v}_1, \tilde{v}_2, \tilde{v}_3$ be vectors from the spaces span($\mathbf{V}_1$), span($\mathbf{V}_2$), and span($\mathbf{V}_3$), respectively. Consider a $Type\ II$ network, without loss of generality , we assume that the network realizes $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ (see Fig. 4b). This relation can be equivalently represented in matrix form as

$$\mathbf{M}_{11} = \mathbf{M}_{31}{\mathbf{M}_{32}}^{-1}\mathbf{M}_{12} + \mathbf{M}_{21}{\mathbf{M}_{23}}^{-1}\mathbf{M}_{13} \tag{39}$$

Lemma VIII.1. If $\tilde{v}_1$ aligns with $\tilde{v}_3$ at d_2 and with $\tilde{v}_2$ at d_3, then $\tilde{v}_1$ must align in the space spanned by $\tilde{v}_2$ and $\tilde{v}_3$ at d_1.

Proof. Since $\tilde{v}_1$ aligns with $\tilde{v}_3$ at d_2 and with $\tilde{v}_2$ at d_3, it follows that,

$$d_2: \quad \mathbf{M}_{12}\tilde{v}_1 = a\ \mathbf{M}_{32}\tilde{v}_3 \tag{40}$$
$$d_3: \quad \mathbf{M}_{13}\tilde{v}_1 = b\ \mathbf{M}_{23}\tilde{v}_3 \tag{41}$$

where a, b are scalars. At d_1, we see the vector $\mathbf{M}_{11}\tilde{v}_1$. Using (39), (40) and (41) we get,

$$\begin{aligned}\mathbf{M}_{11}\tilde{v}_1 &= \mathbf{M}_{31}{\mathbf{M}_{32}}^{-1}\mathbf{M}_{12}\tilde{v}_1 + \mathbf{M}_{21}{\mathbf{M}_{23}}^{-1}\mathbf{M}_{13}\tilde{v}_1 \\ &= a\ \mathbf{M}_{31}\tilde{v}_2 + b\ \mathbf{M}_{21}\tilde{v}_2\end{aligned}$$

This shows that the desired vector at d_1 aligns with the space spanned by the interference. ■

Theorem VIII.1. For a $Type\ II$ network the symmetric rate achievable per user through any precoding-based scheme cannot be more than $2/5$.

Proof. Suppose every sender sends d symbols over n dimensions, through any linear precoding scheme. Consider ω_1 , lets use l_{12} and l_{13} to represent the number of dimensions of signal space of d_1 that align with ω_2 at d_3 and ω_3 at d_2 respectively, and V_{12} and V_{13} to represent their corresponding spaces. From Lemma VIII.1, we know that V_{12} and V_{13} must have no intersection, otherwise the intersection part will contain vectors that will align with interference at d_1. Therefore, we must have $l_{12} + l_{13} \leq d$. Now consider ω_2, we already know that there is a l_{13} dimensional space where interference from ω_1 and ω_3 are aligned. So the number of interference dimension is given as $(d+d-l_{13}) = 2d-l_{13}$. The number of desired dimensions at d_2 is d, and this d dimensional desired signal space should remain resolvable from the interference space, so we we have $3d - l_{13} \leq n$. Similarly, consider User 3 to obtain another inequality : $3d - l_{12} \leq n$. Combining these inequalities we get $6d - (l_{13} + l_{12}) \leq n$. But we know $l_{12} + l_{13} \leq d$, so $6d - d \leq 2n \Rightarrow d/n \leq 2/5$, which implies it is not possible to achieve a symmetric rate more than $2/5$ per user. ■

Corollary VIII.1. For $Type\ II$ networks, it is possible to achieve a rate of $2/5$ per user through through a finite time-slot precoding based network alignment scheme, *i.e.*, the outer bound is tight.

Proof. Without loss of generality, assume the $Type\ II$ networks has a coupling relation $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$. This scheme can be easily modified to fit the other coupling relations too. Suppose we use a $2n+1 = 5$ symbol extension, then according to the PBNA scheme in Section IV we have precoding vectors $\mathbf{V}_1 = (\mathbf{w}\ \mathbf{Tw}\ \mathbf{T}^2\mathbf{w})$, $\mathbf{V}_2 = (\mathbf{w}\ \mathbf{Tw})$ and $\mathbf{V}_3 = (\mathbf{Tw}\ \mathbf{T}^2\mathbf{w})$. The given coupling relation only affects User 1, so the rates at Receiver 2 and 3 will remain unaffected. The matrix equivalent of the coupling relation is given in (39), which can be rewritten as,

$$\mathbf{M}_{11} = \mathbf{M}_{31}{\mathbf{M}_{32}}^{-1}\mathbf{M}_{12} + \mathbf{M}_{31}{\mathbf{M}_{32}}^{-1}\mathbf{M}_{12}\mathbf{T} \tag{42}$$

At Receiver 1, the desired signal space is given $\mathbf{M}_{11}\mathbf{V}_1$ and the interference space is given by $\mathbf{M}_{31}\mathbf{V}_3$ (Note: The interference from transmitter 2 and 3 are aligned, *i.e.*, $\mathbf{M}_{21}\mathbf{V}_2$ = $\mathbf{M}_{31}\mathbf{V}_3$). Substituting the alignment equation from Receiver 2 for $\mathbf{V}_3$ we get,

$$\mathbf{M}_{31}\mathbf{V}_3 = \mathbf{M}_{31}{\mathbf{M}_{32}}^{-1}\mathbf{M}_{12}(\mathbf{Tw}\ \mathbf{T}^2\mathbf{w}) \tag{43}$$

From (42) and (43), it can be seen that the second column of the desired signal space ($\mathbf{M}_{11}\mathbf{V}_1$) can be written as a linear combination of the two columns of the interference space. The other two columns of the desired space are linearly independent of the column of interference space. User 1 could use these two dimension to send its signal without interference. In other words, each user would be able to achieve a rate of $2/5$ ■

Proof of Theorem V.6. $Type\ I$ networks fail to satisfy certain conditions which are information theoretically necessary to achieve any rate more than $1/3$ user per session, this was explained in a remark under Theorem V.1 in Section V. The outer bound for $Type\ II$ networks was derived in Theorem VIII.1 and the achievability was shown in Corollary VIII.1. $Type\ III$ networks were the main focus of this paper, previous sections discussed in detail about schemes and their feasibility for achieving $1/2$ rate per user in detail and it is a well known fact that it is not possible to achieve more than $1/2$ per user for SISO scenarios in fully connected networks [11]. ■

IX. Conclusion and Future Directions

In this paper, we consider the problem of network coding for the SISO scenarios with three unicast sessions. We consider a

network model, in which the middle of the network performs random linear network coding. We apply precoding-based interference alignment [11] to this network setting. We show that network topology may introduce algebraic dependence ("coupling relations") between different transfer functions, which can potentially affect the rate achieved by PBNA. Using two graph-related properties and a recent result from [35], we identify the minimal set of coupling relations that are realizable in networks. Moreover, we show that each of these coupling relations has a unique interpretation in terms of network topology. Based on these interpretations, we present a polynomial-time algorithm to check the existence of these coupling relations.

This work is limited to three unicast sessions in the SISO scenario (i.e., with min-cut one per session) and following a precoding-based approach (all precoding is performed at the end nodes, while intermediate nodes perform random network coding). This is the simplest, yet highly non-trivial instance of the general problem of network coding across multiple unicasts. Apart from being of interest on its own right, we hope that it can be used as a building block and provide insight into the general problem.

There are still many problems that remain to be solved regarding applying interference alignment techniques to the network setting. For example, one important problem is the complexity of PBNA, which arises in two aspects, i.e., precoding matrix and field size, and is inherent in the framework of PBNA. One direction for future work is to apply other alignment techniques (with lower complexity) to the network setting. The extensions to other network scenarios beyond SISO with more than three unicast sessions are highly non-trivial. Finally, the current paper applies precoding at the sources only, while intermediate nodes performed simply random network coding; an open direction for future work is alignment by network code design in the middle of the network as well.

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[2] S.-Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Transactions on Information Theory*, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[3] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[4] Z. Li and B. Li, "Network coding: The case of multiple unicast sessions," in *the Proceedings of the 42nd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, U.S.A., Sept. 2004, pp. 11–19.

[5] A. R. Lehman and E. Lehman, "Complexity classification of network information flow problems," in *the Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*, Philadelphia, PA, U.S.A., 2004, pp. 142–150.

[6] N. Ratnakar, R. Koetter, and T. Ho, "Linear flow equations for network coding in the multiple unicast case," in *the Proceedings of DIMACS Working Group on Network Coding*, Piscataway, NJ, U.S.A., Jan. 2005.

[7] D. Traskov, N. Ratnakar, D. S. Lun, R. Koetter, and M. Médard, "Network coding for multiple unicasts: An approach based on linear optimization," in *the Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Seattle, WA, U.S.A., July 2006, pp. 1758–1762.

[8] M. Kim, M. Médard, U.-M. O'Reilly, and D. Traskov, "An evolutionary approach to inter-session network coding," in *the Proceedings of the 28th IEEE Conference on Computer Communications (INFOCOM)*, Rio de Janeiro, Brazil, Apr. 2009, pp. 450–458.

[9] M. Médard, M. Effros, D. Karger, and T. Ho, "On coding for non-multicast networks," in *the Proceedings of the 41st Annual Allerton Conference on Communication Control and Computing (Allerton)*, vol. 41, no. 1, Monticello, IL, U.S.A., 2003, pp. 21–29.

[10] D. Traskov, N. Ratnakar, D. S. Lun, R. Koetter, and M. Mdard, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745–2759, Aug. 2005.

[11] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[12] A. Ramakrishnan, A. Das, H. Maleki, A. Markopoulou, S. Jafar, and S. Vishwanath, "Network coding for three unicast sessions: Interference alignment approaches," in *the Proceedings of the 48th Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, U.S.A., Sept. 2010, pp. 1054–1061.

[13] Z. Li, B. Li, and L. C. Lau, "On achieving maximum multicast throughput in undirected networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2467–2485, June 2006.

[14] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. M. G. M. Tolhuizen, "Polynomial time algorithms for multicast network code construction," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1973–1982, June 2005.

[15] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol. 52, no. 10, pp. 4413–4430, Oct. 2006.

[16] C.-C. Wang and N. B. Shroff, "Pairwise intersession network coding on directed networks," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3879–3900, Aug. 2010.

[17] I.-H. Wang, S. U. Kamath, and D. N. Tse, "Two unicast information flows over linear deterministic networks," in *the Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Saint-Petersburg, Russia, July 2011, pp. 2462–2466.

[18] S. U. Kamath, D. N. Tse, and V. Anantharam, "Generalized network sharing outer bound and the two-unicast problem," in *the Proceedings of International Symposium on Network Coding (NetCod)*, Boston, MA, U.S.A., June 2011, pp. 1–6.

[19] N. J. A. Harvey, R. Kleinberg, and A. R. Lehman, "On the capacity of information networks," *IEEE/ACM Transactions on Networking (TON) - Special issue on networking and information theory*, vol. 52, no. 6, pp. 2345–2364, June 2006.

[20] T. Ho, Y. Chang, and K. J. Han, "On constructive network coding for multiple unicasts," in *the Proceeding of the 44th Allerton Conference on Communication, Control and Computing (Allerton)*, Monticello, IL, U.S.A., Sept. 2006, pp. 779–788.

[21] M. Effros, T. Ho, and S. Kim, "A tiling approach to network code design for wireless networks," in *the Proceedings of IEEE Information Theory Workshop (ITW)*, Punta del Este, Uruguay, March 2006, pp. 62–66.

[22] J. B. Ebrahimi and C. Fragouli, "Properties of network polynomials," in *the Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Cambridge, MA, U.S.A., July 2012, pp. 1306–1310.

[23] W.Zeng, C. Viveck., and M. Médard, "An edge reduction lemma and application to linear network coding for two-unicast networks," in *the Proceedings of 50th Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, U.S.A., Oct. 2012, pp. 509–516.

[24] B. Nazer, S. Jafar, M. Gastpar, and S. Vishwanath, "Ergodic interference alignment," in *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6355-6371, Oct. 2012.

[25] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4566–4592, Sept. 2010.

[26] S. Jafar, "Blind Interference Alignment," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 216-227, June 2012.

[27] M. Maddah-Ali and D. Tse, "On the degrees of freedom of MISO broadcast channels with delayed feedback," Tech. Rep., 2010. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-122.html

[28] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel," in *the Proceedings of Annual Information Theory and Applications Workshop (ITA)*, San Diego, CA, U.S.A., Jan. 2007.

[29] C. Suh and D. Tse, "Interference alignment for cellular networks," in *the Proceedings of the 46th Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, U.S.A., Sept. 2008, pp. 1037–1044.

[30] N. Lee and J. Lim, "A novel signaling for communication on MIMO Y channel: Signal space alignment for network coding," in *the Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Seoul, South Korea, June 2009, pp. 2892–2896.

[31] S. Gollakota, S. Perli, and D. Katabi, "Interference alignment and cancellation," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 4, pp. 159–170, Oct. 2009.

[32] C. Suh and K. Ramchandran, "Exact-repair MDS code construction using interference alignment," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1425–1442, March 2011.

[33] V. R. Cadambe, S. A. Jafar, H. Maleki, K. Ramchandran, and C. Suh, "Asymptotic interference alignment for optimal repair of MDS codes in distributed storage," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2974–2987, May 2013.

[34] A. Das, S. Vishwanath, S. Jafar, and A. Markopoulou, "Network coding for multiple unicasts: An interference alignment approach," in *the Proceedings of IEEE International Symposium on Information Theory (ISIT)*, Austin, TX, USA, June 2010, pp. 1878–1882.

[35] J. Han, C. C. Wang, and N. B. Shroff, "Analysis of precoding-based intersession network coding and the corresponding 3-unicast interference alignment scheme," Purdue University, Tech. Rep., 2011. [Online]. Available: http://web.ics.purdue.edu/~han83/

[36] C. Meng, A. Ramakrishnan, A. Markopoulou, and S. Jafar, "On the feasibility of precoding-based network alignment for three unicast sessions," in *the Proceedings of IEEE International Symposium on Information Theory Proceedings (ISIT)*, Boston, MA, U.S.A., July 2012, pp. 1907–1911.

[37] C. Meng, A. K. Das, A. Ramakrishnan, S. A. Jafar, A. Markopoulou, and S. Vishwanath, "Precoding-based network alignment for three unicast sessions," Tech. Rep., May 2013. [Online]. Available: http://arxiv.org/abs/1305.0868

[38] J. Han, C.-C. Wang, and N. B. Shroff, "Graph-theoretic characterization of the feasibility of the precoding-based 3-unicast interference alignment scheme," Tech. Rep., May 2013. [Online]. Available: http://arxiv.org/abs/1305.0503

[39] R. Motwani and P. Raghavan, *Randomized Algorithms*. Cambridge Univ. Press, 1995.

[40] V. R. Cadambe and S. A. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 3983–3990, Sept. 2009.

[41] V. Cadambe, S. Jafar, and C. Wang, "Interference alignment with asymmetric complex signaling settling the Høst-Madsen–Nosratinia conjecture," *IEEE Transactions on Information Theory*, vol. 56, no. 9, pp. 4552–4565, Sept. 2010.

Chun Meng received his B.S. degree from the Department of Computer Science at Peking University in 1998, his Master's degree from the Department of Electrical Engineering at Beijing Univ. of Posts and Telecommunications, in 2009, and his PhD degree from the Electrical Engineering and Computer Sience Department with a major in Computer Network and Distributed Computing at UC Irvine in Sept. 2014. He worked as a research intern at the Institute of Network Coding of the Chinese University of Hong Kong. He is now a research scientist at Speedy Packets Inc. in Pasadena, CA. His research interests include network coding, wireless networks, and protocol design.

Abhik Kumar Das received the B.Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT) Kanpur in 2008, and the M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Texas at Austin in 2010 and 2013 respectively. He is currently a Senior Systems Engineer at Samsung Mobile Solutions Lab in San Diego, CA. His research interests include network information theory, wireless communication and graphical models.

Abinesh Ramakrishnan received his B.E. degree in Electronics and Communications Engineering from the Madras Institute of Technology, Anna University, India in 2008 and his Master's degree from EECS department at the University of California, Irvine. He is currently a Ph.D. candidate in the EECS department at UCI. He has worked as an intern at Huawei US R&D Center. His research interests include network coding, information theory, interference alignment, index coding, distributed storage and content delivery.

Syed Ali Jafar received his B. Tech. from IIT Delhi, India, in 1997, M.S. from Caltech, USA, in 1999, and Ph.D. from Stanford, USA, in 2003, all in Electrical Engineering. His industry experience includes positions at Lucent Bell Labs , Qualcomm Inc. and Hughes Software Systems. He is a Professor in the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory, wireless communications and network coding.

Dr. Jafar received the NSF CAREER award in 2006, the ONR Young Investigator Award in 2008, the Information Theory Society paper award in 2009, the School of Engineering Maseeh Outstanding Research Award in 2010, an IEEE GLOBECOM CTS Best Paper Award in 2012 and an IEEE Communication Society Best Tutorial Paper Award in 2013. Dr. Jafar received the UC Irvine EECS Professor of the Year award five times, in 2006, 2009, 2011, 2012, and 2014, from the Engineering Students Council and the Teaching Excellence Award in 2012 from the School of Engineering. He was a University of Canterbury Erskine Fellow in 2010 and is an IEEE Communications Society Distinguished Lecturer for 2013-2014. Dr. Jafar was recognized as a Thomson Reuters Highly Cited Researcher and included by Sciencewatch among The World's Most Influential Scientific Minds in 2014. He served as Associate Editor for IEEE Transactions on Communications 2004-2009, for IEEE Communications Letters 2008-2009 and for IEEE Transactions on Information Theory 2009-2012. Dr. Jafar was elevated to IEEE Fellow, Class of 2014, for contributions to analyzing the capacity of wireless communication networks.

Athina Markopoulou (S'98-M'02-SM'13) received the Diploma degree in Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 1996, and the M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 1998 and 2003, respectively.

She joined the EECS Department, at the University of California, Irvine in 2006, where she is currently an Associate Professor. She also holds a joint appointment with the Computer Science Department, she is affiliated with CalIT2 and CPCC, as well as with the Networked Systems Program whose she is currently the Director. Outside UC Irvine, she has held positions as a postdoctoral research fellow at Sprintlabs (2003) and at Stanford University (2004-2005), as a member of the technical staff at Arista Networks (2005), and as a visiting faculty at the IT University of Copenhagen (2012-2013). Her research interests are in the area of computer networks and include mobile devices, network coding, network measurement, online social networks, and network security.

She received the NSF CAREER award in 2008, and the Henry Samueli School of Engineering Mid-career Research Award in 2014. She is a senior member of the IEEE and she currently serves as an Associate Editor for IEEE/ACM Transactions on Networking.

Sriram Vishwanath received B. Tech. degree in Electrical Engineering from the Indian Institute of Technology (IIT), Madras, India in 1998, the M.S. degree in Electrical Engineering from California Institute of Technology (Caltech), Pasadena USA in 1999, and the Ph.D. degree in Electrical Engineering from Stanford University, Stanford, CA USA in 2003. He is currently a Professor in the Department of Electrical & Computer Engineering at The University of Texas, Austin, USA. His research interests include network information theory, wireless systems and mobile systems. He has served on the board of directors and/or co-founded several startups, including M87 Inc., Lynx Laboratories Inc., Azul Mobile Inc., Agilux Systems Inc., Accordion Health Inc., and Sunfish Inc.; and was the 2014 University of Texas Entrepreneur of the Year.

Sriram received the NSF CAREER award in 2005 and the ARO Young Investigator Award in 2008. He is the co-recipient of the 2005 IEEE Joint Information Theory Society and Communications Society best paper award. He has served as the general chair of IEEE Wireless Network Coding Conference (WiNC) in 2010, the general co-chair of the IEEE Information Theory School in 2011, the local arrangements chair of ISIT 2010 and the guest editor-in-chief of Transactions on Information Theory special issue on interference networks.

APPENDIX A
PROOFS OF GRAPH-RELATED PROPERTIES

A. Linearization Property and Square-Term Property

The following lemma plays an important role in the proof of Linearization Property and the interpretation of the coupled relations, $p_i(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) = \eta(\mathbf{x})$. The basic idea of this lemma is that we can multicast two symbols from two senders to two receivers via network coding if and only if the minimum cut separating the senders from the receivers is greater than one.

Lemma A.1. $m_{ab}(\mathbf{x})m_{pq}(\mathbf{x}) \neq m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})$ if and only if there is disjoint path pair $(P_1, P_2) \in \mathcal{P}_{ab} \times \mathcal{P}_{pq}$ or $(P_3, P_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb}$.

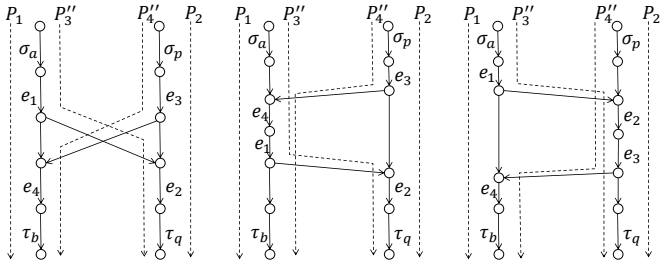
Proof. We add a super sender s and connect it to s'_a and s'_p via two edges of unit capacity, and a super receiver d, to which we connect d'_b and d'_q via two edges of unit capacity. Thus, the transfer matrix at d is

$$\mathbf{M} = \begin{pmatrix} m_{ab}(\mathbf{x}) & m_{aq}(\mathbf{x}) \\ m_{pb}(\mathbf{x}) & m_{pq}(\mathbf{x}) \end{pmatrix}$$

It is easy to see $\det(\mathbf{M}) = m_{ab}(\mathbf{x})m_{pq}(\mathbf{x}) - m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})$. Hence, we can multicast two symbols from s to d, i.e., $\det(\mathbf{M}) \neq 0$, if and only if the minimum cut separating s from d is at least two, or equivalently there is a disjoint path pair $(P_1, P_2) \in \mathcal{P}_{ab} \times \mathcal{P}_{pq}$ or $(P_3, P_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb}$. ■

The proof of Lemma VI.1 involves finding a subgraph H such that some coding variable appears exclusively in the denominator or numerator of $h(\mathbf{x}_H)$, i.e., $h(\mathbf{x}_H)$ restricted to H. In fact, due to Lemma A.1, such subgraph H always exists, if $h(\mathbf{x})$ is not constant.

Proof of Lemma VI.1. In this proof, given a path P, let $P[e : e']$ denote the path segment of P between two edges e and e', including e, e'. We arrange the edges of G' in topological order, and for $e \in E'$, let $o(e)$ denote e's position in this ordering. Moreover, denote $h_1(\mathbf{x}) = m_{ab}(\mathbf{x})m_{pq}(\mathbf{x})$, $h_2(\mathbf{x}) = m_{aq}(\mathbf{x})m_{pb}(\mathbf{x})$ and $d(\mathbf{x}) = \gcd(h_1(\mathbf{x}), h_2(\mathbf{x}))$. Let $s_1(\mathbf{x}) = \frac{h_1(\mathbf{x})}{d(\mathbf{x})}$ and $s_2(\mathbf{x}) = \frac{h_2(\mathbf{x})}{d(\mathbf{x})}$. Hence $\gcd(s_1(\mathbf{x}), s_2(\mathbf{x})) = 1$. It follows $u(\mathbf{x}) = cs_1(\mathbf{x}), v(\mathbf{x}) = cs_2(\mathbf{x})$, where c is a non-zero constant in $\mathbb{F}_{2^m}$. By Lemma A.1, there exists disjoint path pair $(P_1, P_2) \in \mathcal{P}_{ab} \times \mathcal{P}_{pq}$ or $(P_3, P_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb}$. Now we consider the first case.



(a) $o(e_2) > o(e_3)$ and $o(e_1) < o(e_4)$ (b) $o(e_2) > o(e_3)$ and $o(e_1) > o(e_4)$ (c) $o(e_2) < o(e_3)$

Fig. 9. The construction of H (in the proof of the Linearization Property) enabled by Lemma A.1 (P_1 is disjoint with P_2)

We arbitrarily select another path pair $(P'_3, P'_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb}$. Since P_1, P'_3 both originate at σ_a, and P_2, P'_3 both terminate at τ_q, there exist $e_1 \in P_1 \cap P'_3$ and $e_2 \in P_2 \cap P'_3$ such that the path segment along P'_3 between e_1 and e_2 is disjoint with $P_1 \cup P_2$. Similarly, there exist $e_3 \in P_2 \cap P'_4$ and $e_4 \in P_1 \cap P'_4$ such that the path segment between e_3 and e_4 along P'_4 is disjoint with $P_1 \cup P_2$. Construct the following two paths: $P''_3 = P_1[\sigma_a : e_1] \cup P'_3[e_1 : e_2] \cup P_2[e_2 : \tau_q]$ and $P''_4 = P_2[\sigma_p : e_3] \cup P'_4[e_3 : e_4] \cup P_1[e_4 : \tau_b]$ (see Fig. 9). Let H denote the subgraph of G' induced by $P_1 \cup P_2 \cup P''_3 \cup P''_4$.

We then prove that the theorem holds for H. If $o(e_2) > o(e_3)$ (Fig. 9a and 9b), the variables along $P_2[e_3 : e_2]$ are absent in $h_2(\mathbf{x}_H)$. We then arbitrarily select a variable $x_{ee'}$ from $P_2[e_3 : e_2]$, and write $h_1(\mathbf{x}_H)$ as $f(\mathbf{x}'_H)x_{ee'} + g(\mathbf{x}'_H)$, where $\mathbf{x}'_H$ includes all the variables in $\mathbf{x}_H$ other than $x_{ee'}$. Meanwhile, $h_2(\mathbf{x}_H)$ can be written as $h_2(\mathbf{x}'_H)$. Clearly, $x_{ee'}$ will not show up in $d(\mathbf{x}_H)$ and thus it can also be written as $d(\mathbf{x}'_H)$. We then find values for $\mathbf{x}'_H$, denoted by $\mathbf{r}$, such that $f(\mathbf{r})h_2(\mathbf{r})d(\mathbf{r}) \neq 0$. Finally, denote $c_0 = cg(\mathbf{r})d^{-1}(\mathbf{r})$, $c_1 = cf(\mathbf{r})d^{-1}(\mathbf{r})$ and $c_2 = ch_2(\mathbf{r})d^{-1}(\mathbf{r})$ and the theorem holds. On the other hand, if $o(e_2) < o(e_3)$ (see Fig. 9c), the variables along $P_1[e_1 : e_4]$ are absent in $h_2(\mathbf{x}_H)$. We then select a variable $x_{ee'}$ from $P_1[e_1 : e_4]$. Similar to above, it's easy to see that $u(\mathbf{x})$ and $v(\mathbf{x})$ can be transformed into $c_1x_{ee'} + c_0$ and c_2 respectively.

For the case where $(P_3, P_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb}$ is a disjoint path pair, we can show that $u(\mathbf{x})$ and $v(\mathbf{x})$ can be transformed into c_2 and $c_1x_{ee'} + c_0$ respectively. ■

The basic idea of Lemma VI.2 is to construct a one-to-one mapping between the square terms in the numerator of $h(\mathbf{x})$ and those in the denominator of $h(\mathbf{x})$.

Proof of Lemma VI.2. First, we define two sets $\mathcal{Q}_1 = \{(P_1, P_2) \in \mathcal{P}_{ab} \times \mathcal{P}_{pq} : x^2_{ee'} \mid t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x})\}$ and $\mathcal{Q}_2 = \{(P_3, P_4) \in \mathcal{P}_{aq} \times \mathcal{P}_{pb} : x^2_{ee'} \mid t_{P_3}(\mathbf{x})t_{P_4}(\mathbf{x})\}$. Consider a path pair $(P_1, P_2) \in \mathcal{Q}_1$. Since the degree of $x_{ee'}$ in $t_{P_1}(\mathbf{x})$ and $t_{P_2}(\mathbf{x})$ is at most one, we must have $x_{ee'} \mid t_{P_1}(\mathbf{x})$ and $x_{ee'} \mid t_{P_2}(\mathbf{x})$. Thus $e, e' \in P_1 \cap P_2$. Let P^1_1, P^2_1 be the parts of P_1 before e and after e' respectively.

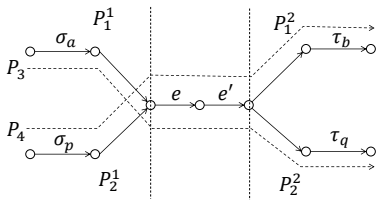


Fig. 10. Illustration of Square-Term Property. A term with $x_{ee'}^2$ introduced by (P_1, P_2) in the numerator of $h(\mathbf{x})$ equals another term introduced by (P_3, P_4) in the denominator of $h(\mathbf{x})$.

Similarly, define P_2^1 and P_2^2. Then construct two new paths: $P_3 = P_1^1 \cup \{e, e'\} \cup P_2^2$ and $P_4 = P_2^1 \cup \{e, e'\} \cup P_1^2$ (see Fig. 10). Clearly, $t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x}) = t_{P_3}(\mathbf{x})t_{P_4}(\mathbf{x})$, and thus $(P_3, P_4) \in \mathcal{Q}_2$. The above method establishes a one-to-one mapping $\phi : \mathcal{Q}_1 \to \mathcal{Q}_2$, such that for $\phi((P_1, P_2)) = (P_3, P_4)$, $t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x}) = t_{P_3}(\mathbf{x})t_{P_4}(\mathbf{x})$. Hence, $f_1(\mathbf{x}) = \frac{1}{x_{ee'}^2}\sum_{(P_1,P_2)\in\mathcal{Q}_1} t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x}) = \frac{1}{x_{ee'}^2}\sum_{(P_3,P_4)\in\mathcal{Q}_2} t_{P_3}(\mathbf{x})t_{P_4}(\mathbf{x}) = f_2(\mathbf{x})$. ■

B. Other Graph-Related Properties

In this section, we present other graph-related properties, which reveal more microscopic structures of transfer functions, and are to be used in the proofs of Theorems V.3 and V.5. Before proceeding, we first extend the concept of transfer function to any two edges $e, e' \in E'$, i.e., $m_{ee'}(\mathbf{x}) = \sum_{P\in\mathcal{P}_{ee'}} t_P(\mathbf{x})$, where $\mathcal{P}_{ee'}$ is the set of paths from e to e'.

The following lemma states that any transfer function $m_{ee'}(\mathbf{x})$ is fully determined by the two edges e, e'.

Lemma A.2. Consider two transfer functions $m_{e_1e_2}(\mathbf{x})$ and $m_{e_3e_4}(\mathbf{x})$. Then $m_{e_1e_2}(\mathbf{x}) = m_{e_3e_4}(\mathbf{x})$ if and only if $e_1 = e_3$ and $e_2 = e_4$.

Proof. Apparently, the "if" part holds trivially. Now assume $e_1 \neq e_3$ or $e_2 \neq e_4$. Then, there must be some edge which appears exclusively in $\mathcal{P}_{e_1e_2}$ or $\mathcal{P}_{e_3e_4}$, implying $m_{e_1e_2}(\mathbf{x}) \neq m_{e_3e_4}(\mathbf{x})$. Thus, the lemma holds. ■

The following result was first proved by Han et al. [35]. It states that each transfer function $m_{ee'}(\mathbf{x})$ can be uniquely factorized into a product of irreducible polynomials according to the bottlenecks between e and e'.

Lemma A.3. We arrange the bottlenecks in $\mathcal{C}_{ee'}$ in topological order: $e_1, e_2, \cdots, e_k$, such that $e = e_1$, $e' = e_k$. Then, $m_{ee'}(\mathbf{x})$ can be factorized as $m_{ee'}(\mathbf{x}) = \prod_{i=1}^{k-1} m_{e_ie_{i+1}}(\mathbf{x})$, where $m_{e_ie_{i+1}}(\mathbf{x})$ is an irreducible polynomial.

In addition, as shown below, any transfer function $m_{ee'}(\mathbf{x})$ can be partitioned into a summation of products of transfer functions according to a cut between e and e'.

Lemma A.4. Assume $\mathcal{U} = \{e_1, e_2, \cdots, e_k\}$ is a cut which separates e from e'. If $e_i || e_j$ for $e_i \neq e_j \in \mathcal{U}$, we have $m_{ee'}(\mathbf{x}) = \sum_{i=1}^{k} m_{ee_i}(\mathbf{x})m_{e_ie'}(\mathbf{x})$. Otherwise, the above equality doesn't hold.

Proof. For $e_i \in \mathcal{U}$, let $\mathcal{P}_{ee'}^i$ denote the set of paths in $\mathcal{P}_{ee'}$ which pass through e_i. Because $e_i || e_j$ for $e_i \neq e_j \in \mathcal{U}$, $\mathcal{P}_{ee'}^i$ is disjoint with $\mathcal{P}_{ee'}^j$. Hence, $m_{ee'}(\mathbf{x}) = \sum_{i=1}^{k}\sum_{P\in\mathcal{P}_{ee'}^i} t_P(\mathbf{x})$. Note that $m_{ee_i}(\mathbf{x})m_{e_ie'}(\mathbf{x}) = \sum_{(P_1,P_2)\in\mathcal{P}_{ee_i}\times\mathcal{P}_{e_ie'}} t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x})$. Moreover, each monomial $t_P(\mathbf{x})$ in $m_{ee'}(\mathbf{x})$ corresponds to a monomial $t_{P_1}(\mathbf{x})t_{P_2}(\mathbf{x})$ in $m_{ee_i}(\mathbf{x})m_{e_ie'}(\mathbf{x})$. Hence, $m_{ee_i}(\mathbf{x})m_{e_ie'}(\mathbf{x}) = \sum_{P\in\mathcal{P}_{ee'}^i} t_P(\mathbf{x})$, and the lemma holds. On the other hand, if some e_i is upstream of e_j, $P_{ee'}^i \cap P_{ee'}^j \neq \emptyset$, and thus $m_{ee'}(\mathbf{x}) \neq \sum_{i=1}^{k}\sum_{P\in\mathcal{P}_{ee'}^i} t_P(\mathbf{x})$, indicating that the lemma doesn't hold. ■

APPENDIX B
PROOFS OF FEASIBILITY CONDITIONS OF PBNA

A. Reducing $\mathcal{S}'$ to $\mathcal{S}_i'$

In order to utilize the degree-counting technique, we use the following lemma. Basically, it allows us to reformulate each $\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'$ to its unique form $\frac{\alpha(\mathbf{x})}{\beta(\mathbf{x})}$, such that we can compare the degrees of a coding variable in $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ with its degrees in the numerator and denominator of $p_i(\mathbf{x})$ respectively.

Lemma B.1. Let $\mathbb{F}$ be a field. z is a variable and $\mathbf{y} = (y_1, y_2, \cdots, y_k)$ is a vector of variables. Consider four non-zero polynomials $f(z), g(z) \in \mathbb{F}[z]$ and $s(\mathbf{y}), t(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$, such that $\gcd(f(z), g(z)) = 1$ and $\gcd(s(\mathbf{y}), t(\mathbf{y})) = 1$. Denote $d = \max\{d_f, d_g\}$. Define two polynomials in $\mathbb{F}[\mathbf{y}]$: $\alpha(\mathbf{y}) = f\big(\frac{s(\mathbf{y})}{t(\mathbf{y})}\big)t^d(\mathbf{y})$ and $\beta(\mathbf{y}) = g\big(\frac{s(\mathbf{y})}{t(\mathbf{y})}\big)t^d(\mathbf{y})$. Then $\gcd(\alpha(\mathbf{y}), \beta(\mathbf{y})) = 1$.

Proof. See Appendix D. ■

We use the following three steps to reduce $\mathcal{S}'$ to $\mathcal{S}_i'$.

Step 1: $\mathcal{S}' \Rightarrow \mathcal{S}_1'' = \{\frac{a_0+a_1\eta(\mathbf{x})}{b_0+b_1\eta(\mathbf{x})} : a_0, a_1, b_0, b_1 \in \mathbb{F}_{2^m}\}$. Assume $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}'$. We will prove that $d = \max\{d_f, d_g\} = 1$. Let $p_i(\mathbf{x}) = \frac{u(\mathbf{x})}{v(\mathbf{x})}$, $\eta(\mathbf{x}) = \frac{s(\mathbf{x})}{t(\mathbf{x})}$ denote the unique forms of $p_i(\mathbf{x})$ and $\eta(\mathbf{x})$ respectively. Without loss of generality, let $f(z) = \sum_{j=0}^{k} a_j z^j$, $g(z) = \sum_{j=0}^{l} b_j z^j$ where $a_k b_l \neq 0$. We first consider the case where $l \leq k$ and thus $d = k$. Define the following two polynomials:

$$\alpha(\mathbf{x}) = f(\eta(\mathbf{x}))t^k(\mathbf{x}) = \sum\nolimits_{j=0}^{k} a_j t^{k-j}(\mathbf{x})s^j(\mathbf{x})$$
$$\beta(\mathbf{x}) = g(\eta(\mathbf{x}))t^k(\mathbf{x}) = \sum\nolimits_{j=0}^{l} b_j t^{k-j}(\mathbf{x})s^j(\mathbf{x})$$

Due to Lemma B.1, we have $\alpha(\mathbf{x}) = cu(\mathbf{x}), \beta(\mathbf{x}) = cv(\mathbf{x})$, where c in a non-zero constant in $\mathbb{F}_q$. Moreover, according to Lemma VI.1, we assign values to $\mathbf{x}$ other than a coding variable $x_{ee'}$ such that $u(\mathbf{x})$ and $v(\mathbf{x})$ are transformed into:

$$u(x_{ee'}) = c_1x_{ee'} + c_0 \quad v(x_{ee'}) = c_2$$
$$\text{or } u(x_{ee'}) = c_2 \quad v(x_{ee'}) = c_1x_{ee'} + c_0$$

where $c_0, c_1, c_2 \in \mathbb{F}_q$ and $c_1c_2 \neq 0$. We only consider the first case. The proof for the other case is similar. In this case, $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are transformed into $\alpha(x_{ee'}) = cc_1x_{ee'} + cc_0$ and $\beta(x_{ee'}) = cc_2$ respectively.

By contradiction, assume $d \geq 2$. We first consider the case where $l \leq k$ and thus $d = k$. In this case, we have

$$\alpha(x_{ee'}) = \sum_{j=0}^{k} a_j t^{k-j}(x_{ee'}) s^j(x_{ee'}) = cc_1 x_{ee'} + cc_0$$
$$\beta(x_{ee'}) = \sum_{j=0}^{l} b_j t^{k-j}(x_{ee'}) s^j(x_{ee'}) = cc_2$$

Assume $s(x_{ee'}) = \sum_{j=0}^{r} s_j x_{ee'}^j$ and $t(x_{ee'}) = \sum_{j=0}^{r'} t_j x_{ee'}^j$, where $s_r t_{r'} \neq 0$. Thus $\max\{r, r'\} \geq 1$. Note that the degree of $x_{ee'}$ in $t^{k-j}(x_{ee'})s^j(x_{ee'})$ is $kr' + j(r - r')$. We consider the following two cases:

Case I: $r \neq r'$. If $r > r'$, $d_\alpha = kr \geq 2$, contradicting that $d_\alpha = 1$. Now assume $r < r'$. Let l_1 and l_2 be the minimum exponents of z in $f(z)$ and $g(z)$ respectively. It follows that $d_\alpha = kr' - l_1(r' - r) = 1$ and $d_\beta = kr' - l_2(r' - r) = 0$. Clearly, $l_2 > 0$ due to $d_\beta = 0$. If $r > 0$, $kr' - l_2(r' - r) > kr' - l_2 r' \geq 0$, contradicting $d_\beta = 0$. Hence, $r = 0$, and $l_2 = k$ due to $d_\beta = 0$. Meanwhile, $d_\alpha = (k - l_1)r' = 1$, which implies that $l_1 = k - 1$ and $r' = 1$. Thus, z^{k-1} is a common divisor of $f(z)$ and $g(z)$, contradicting $\gcd(f(z), g(z)) = 1$.

Case II: $r = r'$. Since $d_\alpha = 1$ and $d_\beta = 0$, all the terms in $\alpha(x_{ee'})$ and $\beta(x_{ee'})$ containing $x_{ee'}^{kr}$ must be cancelled out, implying that

$$\sum_{j=0}^{k} a_j t_r^{k-j} s_r^j = t_r^k \sum_{j=0}^{k} a_j \left(\frac{s_r}{t_r}\right)^j = t_r^k f\left(\frac{s_r}{t_r}\right) = 0$$
$$\sum_{j=0}^{l} b_j t_r^{k-j} s_r^j = t_r^k \sum_{j=0}^{l} b_j \left(\frac{s_r}{t_r}\right)^j = t_r^k g\left(\frac{s_r}{t_r}\right) = 0$$

Hence $z - \frac{s_r}{t_r}$ is a common divisor of $f(z)$ and $g(z)$, contradicting $\gcd(f(z), g(z)) = 1$.

Therefore, we have proved $d = 1$ when $l \leq k$. Using similar technique, we can prove that $d = 1$ when $l \geq k$. This implies that $\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$ can only be of the form $\frac{a_0 + a_1\eta(\mathbf{x})}{b_0 + b_1\eta(\mathbf{x})}$. Hence, we have reduced $\mathcal{S}'$ to $\mathcal{S}_1''$.

Step 2: $\mathcal{S}_1'' \Rightarrow \mathcal{S}_2'' = \{1, \eta(\mathbf{x}), 1 + \eta(\mathbf{x}), \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}\}$. We consider the coupling relation $p_1(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$. The coupling relations $p_2(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$ and $p_3(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$ can be dealt with similarly. Define $q_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{p_1(\mathbf{x})} = \frac{m_{11}(\mathbf{x})m_{23}(\mathbf{x})}{m_{13}(\mathbf{x})m_{21}(\mathbf{x})}$. Assume the characteristic of $\mathbb{F}_q$ is p. Given an integer m, let m_p denote the remainder of m divided by p. Since $\mathcal{S}_1''$ only consists of a finite number of rational functions, we iterate all possible configurations of a_0, a_1, b_0, b_1 as follows:

Case I: $\frac{f(z)}{g(z)} = \frac{a_0 + a_1 z}{b_0 + b_1 z}$, where $a_1 a_0 b_1 b_0 \neq 0$, and $a_0 b_1 \neq a_1 b_0$. For this case, we have $p_1(x_{ee'}) = \frac{a_0 + a_1 p_1(x_{ee'}) q_1(x_{ee'})}{b_0 + b_1 p_1(x_{ee'}) q_1(x_{ee'})}$. It immediately follows

$$q_1(x_{ee'}) = \frac{a_0 c_2^2 - b_0 c_0 c_2 - b_0 c_1 c_2 x_{ee'}}{b_1 c_1^2 x_{ee'}^2 + (2_p b_1 c_0 c_1 - a_1 c_1 c_2) x_{ee'} + b_1 c_0^2 - a_1 c_0 c_2}$$

Let $u_1(x_{ee'}), v_1(x_{ee'})$ denote the numerator and denominator of the above equation respectively. Assume $u_1(x_{ee'}) \mid v_1(x_{ee'})$ and thus $x_{ee'} = \frac{a_0 c_2 - b_0 c_0}{b_0 c_1}$ is a solution to $v_1(x_{ee'}) = 0$. However, $v_1(\frac{a_0 c_2 - b_0 c_0}{b_0 c_1}) = \frac{a_0 c_2^2}{b_0^2}(a_0 b_1 - a_1 b_0) \neq 0$. Hence, $u_1(x_{ee'}) \nmid v_1(x_{ee'})$. Thus, by the definition of $q_1(\mathbf{x})$ and Lemma VI.2, $x_{ee'}^2$ must appear in $u_1(x_{ee'})$, which contradicts the formulation of $u_1(x_{ee'})$.

Case II: $\frac{f(z)}{g(z)} = \frac{a_0 + a_1 z}{b_1 z}$, where $a_0 a_1 b_0 \neq 0$. Similar to Case I, we can derive

$$q_1(x_{ee'}) = \frac{a_0 c_2^2}{b_1 c_1^2 x_{ee'}^2 + (2_p b_1 c_0 c_1 - a_1 c_1 c_2) x_{ee'} + b_1 c_0^2 - a_1 c_0 c_2}$$

which contradicts Lemma VI.2.

Case III: $\frac{f(z)}{g(z)} = \frac{a_1 z}{b_0 + b_1 z}$, where $a_1 b_0 b_1 \neq 0$. Thus $\frac{1}{p_1(\mathbf{x})} = \frac{b_0}{a_1}\frac{1}{\eta(\mathbf{x})} + \frac{b_1}{a_1}$. Since the coefficient of each monomial in the denominators and numerators of $p_1(\mathbf{x})$ and $\eta(\mathbf{x})$ equals one, it follows $\frac{a_0}{b_1} = \frac{b_1}{a_1} = 1$. This indicates that $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{\eta(\mathbf{x})+1}$.

Case IV: $\frac{f(z)}{g(z)} = \frac{a_0}{b_0 + b_1 z}$, where $a_0 b_0 b_1 \neq 0$. It follows that

$$q_1(x_{ee'}) = \frac{a_0 c_2^2 - b_0 c_0 c_2 - b_0 c_1 c_2 x_{ee'}}{b_1 c_0^2 + 2_p b_1 c_0 c_1 x_{ee'} + b_1 c_1^2 x_{ee'}^2}$$

Similar to Case I, this also contradicts Lemma VI.2.

Case V: $\frac{f(z)}{g(z)} = \frac{a_0}{z}$, where $a_0 \neq 0$. Hence, $q_1(x_{ee'}) = \frac{a_0 c_2^2}{c_1^2 x_{ee'}^2 + 2_p c_0 c_1 x_{ee'} + c_0^2}$, contradicting Lemma VI.2.

Case VI: $\frac{f(z)}{g(z)} = a_0 + a_1 z$, where $a_0 a_1 \neq 0$. Thus, it follows $p_1(\mathbf{x}) = a_0 + a_1\eta(\mathbf{x})$. Similar to Case III, $a_1 = a_0 = 1$, implying that $p_1(\mathbf{x}) = 1 + \eta(\mathbf{x})$.

Case VII: $\frac{f(z)}{g(z)} = a_1 z$, where $a_1 \neq 0$. Similar to Case III, $a_1 = 1$ and hence $p_1(\mathbf{x}) = \eta(\mathbf{x})$.

Therefore, we have proved that $\frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))}$ can only take the form of the four rational functions in $\mathcal{S}_2''$. Thus, we have reduced $\mathcal{S}_1''$ to $\mathcal{S}_2''$.

Step 3: $\mathcal{S}_2'' \Rightarrow \mathcal{S}_i'$. We note that in Proposition 3 of [35], it was proved that $p_1(\mathbf{x}) \neq 1 + \eta(\mathbf{x})$, $p_2(\mathbf{x}) \neq \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ and $p_3(\mathbf{x}) \neq \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$. Combined with the above results, we have reduced $\mathcal{S}_2''$ to $\mathcal{S}_i'$.

In summary, according to Theorem IV.1, if the conditions of Theorem V.1 are satisfied, the three unicast sessions can asymptotically achieve the rate tuple $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ through PBNA.

B. Necessity of the Conditions in Theorem V.1

As shown previously, each row of $\mathbf{V}_1$ satisfying the alignment conditions corresponds to a non-zero solution to Eq. (36).

Lemma B.2. $\text{rank}(z\mathbf{C} - \mathbf{BA}) = n$.

Proof. Denote $\mathbf{D} = \mathbf{BA}$. Let $\mathbf{c}_i$ and $\mathbf{d}_i$ denote the ith column of $\mathbf{C}$ and $\mathbf{D}$ respectively. Hence, $\mathbf{c}_1, \cdots, \mathbf{c}_n$ are linearly independent and so are $\mathbf{d}_1, \cdots, \mathbf{d}_n$. Assume there exist $f_1(z), \cdots, f_n(z) \in \mathbb{F}_{2^m}(\xi)(z)$ such that $\sum_{i=1}^{n} f_i(z)(z\mathbf{c}_i - \mathbf{d}_i) = 0$. Without loss of generality, assume $f_i(z) = \frac{g_i(z)}{h(z)}$ for $i \in \{1, 2, \cdots, n\}$, where $g_i(z), h(z) \in \mathbb{F}_{2^m}(\xi)[z]$. Thus, $\sum_{i=1}^{n} g_i(z)(z\mathbf{c}_i - \mathbf{d}_i) = 0$. Let $k = \max_{i \in \{1,2,\cdots,n\}}\{d_{g_i}\}$ and assume $g_i(z) = \sum_{l=0}^{k} a_{l,i}(\xi) z^l$, where $a_{l,i}(\xi) \in \mathbb{F}_{2^m}(\xi)$. Then, it follows

$$\sum_{i=1}^{n} g_i(z)(z\mathbf{c}_i - \mathbf{d}_i) = \sum_{l=0}^{k}\sum_{i=1}^{n}(a_{l,i}(\xi) z^{l+1}\mathbf{c}_i - a_{l,i}(\xi) z^l \mathbf{d}_i)$$
$$= z^{k+1}\sum_{i=1}^{n} a_{k,i}(\xi)\mathbf{c}_i + \sum_{l=0}^{k-1} z^{l+1}\sum_{i=1}^{n}(a_{l,i}(\xi)\mathbf{c}_i - a_{l+1,i}(\xi)\mathbf{d}_i)$$
$$- \sum_{i=1}^{n} a_{0,i}(\xi)\mathbf{d}_i = \mathbf{0}$$

Therefore, the following equations must hold:

$$\sum_{i=1}^{n} a_{k,i}(\xi)\mathbf{c}_i = 0 \quad \sum_{i=1}^{n} a_{0,i}(\xi)\mathbf{d}_i = 0$$
$$\sum_{i=1}^{n}(a_{l,i}(\xi)\mathbf{c}_i - a_{l+1,i}(\xi)\mathbf{d}_i) = 0 \quad \forall l \in \{0,\cdots,k-1\}$$

Thus $a_{l,i}(\xi) = 0$ for any $i \in \{1,\cdots,n\}, l \in \{0,\cdots,k\}$, implying $f_i(z) = 0$. Hence, $\text{rank}(z\mathbf{C}-\mathbf{D}) = n$. ■

The following lemma reveals that any non-zero solution to Eq. (36) is linearly dependent on the particular vector $(1, z, z^2, \cdots, z^n)$, which forms each row of the precoding matrix $\mathbf{V}_1^*$.

Corollary B.1. Eq. (36) has a non-zero solution if and only if $s = 1$. Moreover, when $s = 1$, Eq. (36) has a non-zero solution in the form of $\mathbf{r}(z) = (1, z, z^2, \cdots, z^n)\mathbf{F}$, where $\mathbf{F}$ is an $(n+1)\times(n+1)$ matrix over $\mathbb{F}_{2^m}(\xi)$. Moreover, any solution to Eq. (36) is linearly dependent on $(1, z, \cdots, z^n)\mathbf{F}$.

Proof. We first prove the "only if" part. If $s = 0$, $z\mathbf{C}-\mathbf{BA}$ is an invertible square matrix. Thus, Eq. (36) has only zero solution. Hence, if Eq. (36) has only non-zero solution, it must be that $s = 1$.

We then prove the "if" part. Assume $s = 1$. We will construct a non-zero solution to Eq. (36) as follows. There must be an $n\times n$ invertible submatrix in $z\mathbf{C}-\mathbf{D}$. Without loss of generality, assume this submatrix consists of the top n rows of $z\mathbf{C}-\mathbf{D}$ and denote this submatrix by $\mathbf{E}_{n+1}$. Let $\mathbf{b}$ denote the $(n+1)$th row of $z\mathbf{C}-\mathbf{D}$. In order to get a non-zero solution to equation (36), we first fix $r_{n+1}(z) = -1$. Therefore, equation (36) is transformed into $(r_1(z),\cdots,r_n(z))\mathbf{E}_{n+1} = \mathbf{b}$. Let $\mathbf{E}_i$ denote the submatrix acquired by replacing the ith row of $\mathbf{E}_{n+1}$ with $\mathbf{b}$. Hence, we get a non-zero solution to (36), $\mathbf{r}(z) = (\frac{\det\mathbf{E_1}}{\det\mathbf{E}_{n+1}},\cdots,\frac{\det\mathbf{E}_n}{\det\mathbf{E}_{n+1}},-1)$. Moreover, $\bar{\mathbf{r}}(z) = (\det\mathbf{E}_1,\cdots,\det\mathbf{E}_n,-\det\mathbf{E}_{n+1})$ is also a solution. Note that the degree of z in each $\det\mathbf{E}_i$ is at most n. Thus, $\bar{\mathbf{r}}(z)$ can be formulated as $(1, z, \cdots, z^n)\mathbf{F}$, where $\mathbf{F}$ is an $(n+1)\times(n+1)$ matrix. Since $\text{rank}(z\mathbf{C}-\mathbf{D}) = n$, all the solutions to equation (36) form a one-dimensional linear space. Thus, all solutions must be linearly dependent on $\bar{\mathbf{r}}(z)$. ■

Based on Corollary B.1, we can easily derive that each $\mathbf{V}_1$ satisfying Eq. (10) is related to $\mathbf{V}_1^*$ through a transform equation, as defined in Lemma VI.3.

Proof of Lemma VI.3. Let $\mathbf{r}_i$ be the ith row of $\mathbf{V}_1$, which satisfies Eq. (10). According to Corollary B.1, $\mathbf{r}_i$ must have the form $f_i(\eta(\mathbf{x}^i))(1,\eta(\mathbf{x}^i),\cdots,\eta^n(\mathbf{x}^i))\mathbf{F}$, where $f_i(z)$ is a non-zero rational function in $\mathbb{F}_{2^m}(\xi)(z)$. Hence, $\mathbf{V}_1$ can be written as $\mathbf{G}\mathbf{V}_1^*\mathbf{F}$. Moreover, Eq. (36) can be rewritten as follows:

$$(z, z^2, \cdots, z^{n+1})\mathbf{FC} = (1, z, \cdots, z^n)\mathbf{FBA}$$

The right side of the above equation contains no z^{n+1}, and thus the $(n+1)$th row of $\mathbf{FC}$ must be zero. Similarly, there is no constant term on the left side of the above equation, implying that the 1st row of $\mathbf{FBA}$ is zero. ■

In the followings, we will prove the necessity of the conditions in Theorem V.1. Assume a coupling relation $p_i(\mathbf{x}) = \frac{f(\eta(\mathbf{x}))}{g(\eta(\mathbf{x}))} \in \mathcal{S}_i'$ is present in the network. Without loss of generality, assume $f(z) = \sum_{k=0}^{p} a_k z^k$ and $g(z) = \sum_{k=0}^{q} b_k z^k$, where $a_p \neq 0$ and $b_q \neq 0$. We'll prove that it is impossible for ω_i to asymptotically achieve one half rate by using any PBNA. We only consider the case $i = 1$. The other cases $i = 2, 3$ can be proved similarly, and are omitted.

Consider a PBNA $\lambda = (\xi, \mathbf{V}_i : 1 \le i \le 3)$ with $2n+s$ symbol extensions, where $n > \max\{p, q\} + 1$. According to Corollary B.1, s must equal 1, and thus $\mathbf{V}_1$ is a $(2n+1)\times(n+1)$ matrix. By Lemma VI.3, $\mathbf{V}_1 = \mathbf{G}\mathbf{V}_1^*\mathbf{F}$, where $\mathbf{F}$ is an $(n+1)\times(n+1)$ invertible matrix. The jth row of $\mathbf{V}_1$ is $\mathbf{r}_j = f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^n(\mathbf{x}^{(j)}))\mathbf{F}$. Since the $(n+1)$th row of $\mathbf{FC}$ is zero, we have

$$\mathbf{r}_j\mathbf{C} = f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^{n-1}(\mathbf{x}^{(j)}))\mathbf{H} \qquad (44)$$

where $\mathbf{H}$ consists of the top n rows of $\mathbf{FC}$ and $\text{rank}(\mathbf{H}) = n$. For $0 \le l \le n-p-1$, define the following vector:

$$\mathbf{a}_l = (\overbrace{0\ \cdots\ 0}^{l}\ a_0\ \cdots\ a_p\ \overbrace{0\ \cdots\ 0}^{n-p-l})^T$$
$$\mathbf{b}_l = (\overbrace{0\ \cdots\ 0}^{l}\ b_0\ \cdots\ b_q\ \overbrace{0\ \cdots\ 0}^{n-p-l-1})^T$$

It follows that

$$f(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)}) = (1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^n(\mathbf{x}^{(j)}))\mathbf{a}_l \qquad (45)$$
$$g(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)}) = (1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^{n-1}(\mathbf{x}^{(j)}))\mathbf{b}_l \qquad (46)$$

Define $\mathbf{a}_l' = \mathbf{F}^{-1}\mathbf{a}_l$ and $\mathbf{b}_l' = \mathbf{H}^{-1}\mathbf{b}_l$. We can derive:

$$\begin{aligned}
\mathbf{r}_j\mathbf{a}_l' &= f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^n(\mathbf{x}^{(j)}))\mathbf{F}\mathbf{a}_l' \\
&= f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^n(\mathbf{x}^{(j)}))\mathbf{a}_l \\
&\overset{(a)}{=} f_j(\eta(\mathbf{x}^{(j)}))f(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)}) \\
&\overset{(b)}{=} f_j(\eta(\mathbf{x}^{(j)}))p_1(\mathbf{x}^{(j)})g(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)}) \\
&\overset{(c)}{=} p_i(\mathbf{x}^{(j)})f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^{n-1}(\mathbf{x}^{(j)}))\mathbf{b}_l \\
&= p_i(\mathbf{x}^{(j)})f_j(\eta(\mathbf{x}^{(j)}))(1\ \eta(\mathbf{x}^{(j)})\ \cdots\ \eta^{n-1}(\mathbf{x}^{(j)}))\mathbf{H}\mathbf{b}_l' \\
&\overset{(d)}{=} p_i(\mathbf{x}^{(j)})\mathbf{r}_j\mathbf{C}\mathbf{b}_l'
\end{aligned}$$

where (a) follows from Eq. (45); (b) follows because $p_1(\mathbf{x}^{(j)}) = \frac{f(\eta(\mathbf{x}^{(j)}))}{g(\eta(\mathbf{x}^{(j)}))} = \frac{f(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)})}{g(\eta(\mathbf{x}^{(j)}))\eta^l(\mathbf{x}^{(j)})}$; (c) is due to Eq. (46); (d) follows from Eq. (44). Let $\mathbf{H}_1 = (\mathbf{V}_1\ \ \mathbf{P}_1\mathbf{V}_1\mathbf{C})$ denote the matrix in the reformulated rank condition $\mathscr{B}_1'$. Since $\mathbf{a}_0,\cdots,\mathbf{a}_{n-p-1}$ are linearly independent, the above equation means that there are at most $n+1-(n-p) = p+1$ columns in $\mathbf{V}_1$ that are linearly independent of the columns in $\mathbf{P}_1\mathbf{V}_1\mathbf{C}$. Therefore, d_1 can decode at most $p+1$ source symbols. This means that it is impossible for ω_1 to achieve one half rate by using any PBNA. ■

APPENDIX C
PROOFS OF INTERPRETATIONS OF COUPLING RELATIONS

A. $\eta(\mathbf{x}) = 1$

First, note that $\eta(\mathbf{x})$ can be rewritten as a ratio of two rational functions $\eta(\mathbf{x}) = \frac{f_{213}(\mathbf{x})}{f_{312}(\mathbf{x})}$, where $f_{ijk}(\mathbf{x}) \triangleq \frac{m_{ij}(\mathbf{x})m_{jk}(\mathbf{x})}{m_{ik}(\mathbf{x})}$.

Hence, in order to interpret $\eta(\mathbf{x}) = 1$, we first study the properties of $f_{ijk}(\mathbf{x})$.

The following lemma is to be used to derive the general structure of $f_{ijk}(\mathbf{x})$. Basically, it provides an easy method to calculate the greatest common divisor of two transfer functions with one common starting edge or ending edge.

Lemma C.1. The following statements hold:

1) For $e_1, e_2, e_3 \in E'$ such that e_2, e_3 are both downstream of e_1. Let e be the last edge of the topological ordering of the edges in $\mathcal{C}_{e_1e_2} \cap \mathcal{C}_{e_1e_3}$. Then $m_{e_1e}(\mathbf{x}) = \gcd(m_{e_1e_2}(\mathbf{x}), m_{e_1e_3}(\mathbf{x}))$.
2) For $e_1, e_2, e_3 \in E'$ such that e_1, e_2 are both upstream of e_3. Let e be the first edge of the topological ordering of the edges in $\mathcal{C}_{e_1e_3} \cap \mathcal{C}_{e_2e_3}$. Then $m_{ee_3}(\mathbf{x}) = \gcd(m_{e_1e_3}(\mathbf{x}), m_{e_2e_3}(\mathbf{x}))$.

Proof. First, consider the first statement. By Lemma A.3, the following equations hold: $m_{e_1e_2}(\mathbf{x}) = m_{e_1e}(\mathbf{x})m_{ee_2}(\mathbf{x})$ and $m_{e_1e_3}(\mathbf{x}) = m_{e_1e}(\mathbf{x})m_{ee_3}(\mathbf{x})$. Thus $m_{e_1e}(\mathbf{x}) \mid \gcd(m_{e_1e_2}(\mathbf{x}), m_{e_1e_3}(\mathbf{x}))$. Assume $\gcd(m_{ee_2}(\mathbf{x}), m_{ee_3}(\mathbf{x})) \neq 1$. By Lemma A.3, there exists bottlenecks e_4, e_5 such that $m_{e_4e_5}(\mathbf{x}) \mid \gcd(m_{ee_2}(\mathbf{x}), m_{ee_3}(\mathbf{x}))$. Clearly, $e_5 \in \mathcal{C}_{e_1e_2} \cap \mathcal{C}_{e_1e_3}$ and e_5 is downstream of e, which contradicts that e is the last edge of the topological ordering of $\mathcal{C}_{e_1e_2} \cap \mathcal{C}_{e_1e_3}$. Hence, we have proved that $\gcd(m_{ee_2}(\mathbf{x}), m_{ee_3}(\mathbf{x})) = 1$, which in turn implies that $m_{e_1e}(\mathbf{x}) = \gcd(m_{e_1e_2}(\mathbf{x}), m_{e_1e_3}(\mathbf{x}))$. Similarly, we can prove the other statement. ■

Using the above lemma, $f_{ijk}(\mathbf{x})$ can be reformulated as a fraction of two coprime polynomials, as shown below.

Corollary C.1. $f_{ijk}(\mathbf{x})$ can be formulated as

$$f_{ijk}(\mathbf{x}) = \frac{m_{\sigma_j,\beta_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x})}{m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})} \tag{47}$$

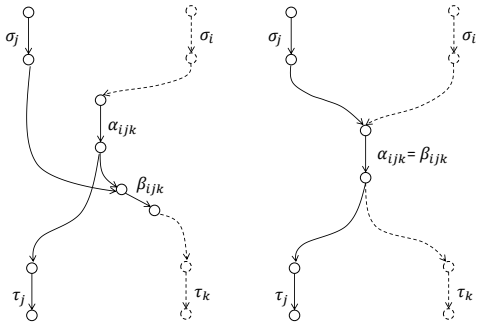
where $\gcd(m_{\sigma_j,\beta_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x}), m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})) = 1$.

Proof. $f_{ijk}(\mathbf{x})$ can be calculated as

$$\begin{aligned} f_{ijk}(\mathbf{x}) &= \frac{m_{\sigma_i,\alpha_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x})m_{jk}(\mathbf{x})}{m_{\sigma_i,\alpha_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_k}(\mathbf{x})} \\ &= \frac{m_{\alpha_{ijk},\tau_j}(\mathbf{x})m_{jk}(\mathbf{x})}{m_{\alpha_{ijk},\tau_k}(\mathbf{x})} \\ &= \frac{m_{\alpha_{ijk},\tau_j}(\mathbf{x})m_{\sigma_j,\beta_{ijk}}(\mathbf{x})m_{\beta_{ijk},\tau_k}(\mathbf{x})}{m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})m_{\beta_{ijk},\tau_k}(\mathbf{x})} \\ &= \frac{m_{\sigma_j,\beta_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x})}{m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})} \end{aligned}$$

By Lemma , $\gcd(m_{\alpha_{ijk},\tau_k}(\mathbf{x}), m_{\alpha_{ijk},\tau_j}(\mathbf{x})) = 1$ and thus $\gcd(m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x}), m_{\alpha_{ijk},\tau_j}(\mathbf{x})) = 1$. Meanwhile, $\gcd(m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x}), m_{\sigma_j,\beta_{ijk}}(\mathbf{x})) = 1$. Hence, we must have $\gcd(m_{\sigma_j,\beta_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x}), m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})) = 1$. ■

According to Corollary C.1, the structure of $f_{ijk}(\mathbf{x})$ must fall into one of the two types, as shown in Fig. 11. In Fig. 11a, $\alpha_{ijk} \neq \beta_{ijk}$ and $f_{ijk}(\mathbf{x})$ is a rational function, the denominator of which is a non-constant polynomial $m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x})$. On the other hand, when $\alpha_{ijk} \in \mathcal{C}_{jk}$ and thus $\alpha_{ijk} = \beta_{ijk}$, as shown in Fig. 11b, $f_{ijk}(\mathbf{x})$ becomes a polynomial $m_{\sigma_j,\alpha_{ijk}}(\mathbf{x})m_{\alpha_{ijk},\tau_j}(\mathbf{x})$.



(a) $\alpha_{ijk} \neq \beta_{ijk}$ (b) $\alpha_{ijk} = \beta_{ijk}$

Fig. 11. The structure of $f_{ijk}(\mathbf{x})$ can be classified into two types: 1) $\alpha_{ijk} \neq \beta_{ijk}$ such that $f_{ijk}(\mathbf{x})$ is a rational function with non-constant denominator; 2) $\alpha_{ijk} = \beta_{ijk}$ such that $f_{ijk}(\mathbf{x})$ is a polynomial.

Moreover, using Corollary C.1, we can easily check whether two $f_{ijk}(\mathbf{x})$'s are equivalent, as shown in the next corollary. It is easy to see that Theorem V.3 is just a special case of this corollary.

Corollary C.2. Assume $i, j, k, i', k' \in \{1, 2, 3\}$ such that $i \neq j, j \neq k$ and $i' \neq j, j \neq k'$. $f_{ijk}(\mathbf{x}) = f_{i'jk'}(\mathbf{x})$ if and only if $\alpha_{ijk} = \alpha_{i'jk'}$ and $\beta_{ijk} = \beta_{i'jk'}$.

Proof. By Corollary C.1, if $\alpha_{ijk} = \alpha_{i'jk'}$ and $\beta_{ijk} = \beta_{i'jk'}$, we must have $f_{ijk}(\mathbf{x}) = f_{i'jk'}(\mathbf{x})$. Conversely, if $f_{ijk}(\mathbf{x}) = f_{i'jk'}(\mathbf{x})$, $m_{\alpha_{ijk},\beta_{ijk}}(\mathbf{x}) = m_{\alpha_{i'jk'},\beta_{i'jk'}}(\mathbf{x})$. Thus $\alpha_{ijk} = \alpha_{i'jk'}$ and $\beta_{ijk} = \beta_{i'jk'}$ by Lemma A.2. ■

B. $p_i(\mathbf{x}) = 1$ *and* $p_i(\mathbf{x}) = \eta(\mathbf{x})$

Using Lemma A.1, we can easily prove Theorem V.4, as shown below.

Proof of Theorem V.4. Apparently, by Lemma A.1 and the definition of $p_1(\mathbf{x})$, $p_1(\mathbf{x}) = 1$ if and only if the minimum cut separating σ_1, σ_2 from τ_1 and τ_3 is one, i.e., $C_{12,13} = 1$. In order to interpret $p_1(\mathbf{x}) = \eta(\mathbf{x})$, we consider $q_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{p_1(\mathbf{x})} = \frac{m_{11}(\mathbf{x})m_{32}(\mathbf{x})}{m_{12}(\mathbf{x})m_{31}(\mathbf{x})}$. Hence $p_1(\mathbf{x}) = \eta(\mathbf{x})$ is equivalent to $q_1(\mathbf{x}) = 1$. Similarly, using Lemma A.1, it is easy to see that $p_1(\mathbf{x}) = \eta(\mathbf{x})$ if and only if the minimum cut separating σ_1, σ_3 from τ_1, τ_2 is one, i.e., $C_{13,12} = 1$. ■

C. $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ *and* $p_2(\mathbf{x}), p_3(\mathbf{x}) = 1 + \eta(\mathbf{x})$

Note that the three coupling relations can be respectively reformulated in terms of $f_{ijk}(\mathbf{x})$ as follows:

$$\begin{aligned} m_{11}(\mathbf{x}) &= f_{312}(\mathbf{x}) + f_{213}(\mathbf{x}) \\ m_{22}(\mathbf{x}) &= f_{123}(\mathbf{x}) + f_{321}(\mathbf{x}) \\ m_{33}(\mathbf{x}) &= f_{231}(\mathbf{x}) + f_{132}(\mathbf{x}) \end{aligned}$$

Thus, as shown below, the three coupling relations can also be interpreted by using the properties of $f_{ijk}(\mathbf{x})$.

Proof of Theorem V.5. We only prove statement 1). The other statements can be proved similarly. First, we prove the "if" part. Due to $\alpha_{312} \in \mathcal{C}_{12}$ and

$\alpha_{213} \in \mathcal{C}_{13}$, $f_{312}(\mathbf{x}) = m_{\sigma_1,\alpha_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})$ and $f_{213}(\mathbf{x}) = m_{\sigma_1,\alpha_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})$. Hence, $f_{312}(\mathbf{x})+f_{213}(\mathbf{x}) = m_{\sigma_1,\alpha_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x}) + m_{\sigma_1,\alpha_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})$. On the other hand, because $\alpha_{312}||\alpha_{213}$ and $\{\alpha_{312},\alpha_{213}\}$ forms a cut which separates σ_1 from τ_1, $m_{11}(\mathbf{x}) = m_{\sigma_1,\alpha_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})+m_{\sigma_1,\alpha_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})$ by Lemma A.4. Therefore, $m_{11}(\mathbf{x}) = f_{312}(\mathbf{x}) + f_{213}(\mathbf{x})$.

Next we prove the "only if" part. Assume $m_{11}(\mathbf{x}) = f_{312}(\mathbf{x}) + f_{213}(\mathbf{x})$. If $\alpha_{312} \notin \mathcal{C}_{12}$ but $\alpha_{213} \in \mathcal{C}_{13}$, $f_{312}(\mathbf{x})$ is a rational function whose denominator is a non-constant polynomial, while $f_{213}(\mathbf{x})$ is a polynomial. Hence $f_{312}(\mathbf{x})+f_{213}(\mathbf{x})$ must be a rational function with non-constant denominator, and thus $m_{11}(\mathbf{x}) \neq f_{312}(\mathbf{x})+f_{213}(\mathbf{x})$. Similarly, if $\alpha_{312} \in \mathcal{C}_{12}$ but $\alpha_{213} \notin \mathcal{C}_{13}$, we can also prove that $m_{11}(\mathbf{x}) \neq f_{312}(\mathbf{x}) + f_{213}(\mathbf{x})$.

Now assume $\alpha_{312} \notin \mathcal{C}_{12}$ and $\alpha_{213} \notin \mathcal{C}_{13}$. It follows that $f_{312}(\mathbf{x}) = \frac{m_{\sigma_1,\beta_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})}{m_{\alpha_{312},\beta_{312}}(\mathbf{x})}$ and $f_{213}(\mathbf{x}) = \frac{m_{\sigma_1,\beta_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})}{m_{\alpha_{213},\beta_{213}}(\mathbf{x})}$. Because $\eta(\mathbf{x}) \neq 1$, we have $f_{312}(\mathbf{x}) \neq f_{213}(\mathbf{x})$, which indicates that $\alpha_{312} \neq \alpha_{213}$ or $\beta_{312} \neq \beta_{213}$ by Corollary C.2, and $m_{\alpha_{312},\beta_{312}}(\mathbf{x}) \neq m_{\alpha_{213},\beta_{213}}(\mathbf{x})$. Therefore, by Lemma A.3, one of the following cases must hold: 1) There exists an irreducible polynomial $m_{ee'}(\mathbf{x})$ such that $m_{ee'}(\mathbf{x}) \mid m_{\alpha_{312},\beta_{312}}(\mathbf{x})$ but $m_{ee'}(\mathbf{x}) \nmid m_{\alpha_{213},\beta_{213}}(\mathbf{x})$; 2) there exists an irreducible polynomial $m_{ee'}(\mathbf{x})$ such that $m_{ee'}(\mathbf{x}) \nmid m_{\alpha_{312},\beta_{312}}(\mathbf{x})$ but $m_{ee'}(\mathbf{x}) \mid m_{\alpha_{213},\beta_{213}}(\mathbf{x})$.

Consider case 1). Define the following polynomials: $f(\mathbf{x}) = \text{lcm}(m_{\alpha_{312},\beta_{312}}(\mathbf{x}), m_{\alpha_{213},\beta_{213}}(\mathbf{x}))$ [7] and $f_1(\mathbf{x}) = f(\mathbf{x})/m_{\alpha_{312},\beta_{312}}(\mathbf{x})$ and $f_2(\mathbf{x}) = f(\mathbf{x})/m_{\alpha_{213},\beta_{213}}(\mathbf{x})$. Hence, we have $m_{ee'}(\mathbf{x}) \nmid f_1(\mathbf{x})$, $m_{ee'}(\mathbf{x}) \mid f_2(\mathbf{x})$, and $f_{312}(\mathbf{x}) + f_{213}(\mathbf{x}) = [m_{\sigma_1,\beta_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})f_1(\mathbf{x}) + m_{\sigma_1,\beta_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})f_2(\mathbf{x})]/f(\mathbf{x})$. Moreover, due to $\gcd(m_{\alpha_{312},\beta_{312}}(\mathbf{x}), m_{\sigma_1,\beta_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})) = 1$, it follows that $m_{ee'}(\mathbf{x}) \nmid m_{\sigma_1,\beta_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})$. This implies that $m_{ee'}(\mathbf{x}) \nmid m_{\sigma_1,\beta_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x})f_1(\mathbf{x}) + m_{\sigma_1,\beta_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})f_2(\mathbf{x})$. However, $m_{ee'}(\mathbf{x}) \mid f(\mathbf{x})$. This indicates that $f_{312}(\mathbf{x})+f_{213}(\mathbf{x})$ is a rational function with non-constant denominator. Thus $m_{11}(\mathbf{x}) \neq f_{312}(\mathbf{x})+f_{213}(\mathbf{x})$. Similarly, for case 2), we can also prove that $m_{11}(\mathbf{x}) \neq f_{312}(\mathbf{x}) + f_{213}(\mathbf{x})$.

Thus, we have proved that $\alpha_{312} \in \mathcal{C}_{12}$ and $\alpha_{213} \in \mathcal{C}_{13}$. It immediately follows that $m_{11}(\mathbf{x}) = m_{\sigma_1,\alpha_{312}}(\mathbf{x})m_{\alpha_{312},\tau_1}(\mathbf{x}) + m_{\sigma_1,\alpha_{213}}(\mathbf{x})m_{\alpha_{213},\tau_1}(\mathbf{x})$. Hence each path P in $\mathcal{P}_{\sigma_1\tau_1}$ either pass through α_{312} or α_{213}, implying that $\{\alpha_{312},\alpha_{213}\}$ forms a cut separating σ_1 from τ_1. Moreover, according to Lemma A.4, $\alpha_{312}||\alpha_{213}$. ■

APPENDIX D
PROOFS OF LEMMAS ON MULTIVARIATE POLYNOMIALS

In this section, we present the proof of Lemma B.1. We first prove that Lemma B.1 holds for the case where $s(\mathbf{x})$ and $t(\mathbf{x})$ are both univariate polynomials. In order to extend this result to multivariate polynomials, we employ a simple idea that each multivariate polynomial can be viewed as an equivalent univariate polynomial on a field of rational functions. Specifically, we prove that the problem of checking if two multivariate polynomials are co-prime is equivalent to checking if their equivalent univariate polynomials are co-prime. Finally, based on this result, we prove that Lemma B.1 also holds for the multivariate case.

[7] We use $\text{lcm}(f(\mathbf{x}), g(\mathbf{x}))$ to denote the least common multiple of two polynomials $f(\mathbf{x})$ and $g(\mathbf{x})$.

A. The Univariate Case

In the following lemma, we show that Lemma B.1 holds for the univariate case.

Lemma D.1. Let $\mathbb{F}$ be a field, and z, y are two variables. Consider four non-zero polynomials $f(z), g(z) \in \mathbb{F}[z]$ and $s(y), t(y) \in \mathbb{F}[y]$, such that $\gcd(f(z), g(z)) = 1$ and $\gcd(s(y), t(y)) = 1$. Denote $d = \max\{d_f, d_g\}$. Define two polynomials $\alpha(y) = f(\frac{s(y)}{t(y)})t^d(y)$ and $\beta(y) = g(\frac{s(y)}{t(y)})t^d(y)$. Then $\gcd(\alpha(y), \beta(y)) = 1$.

Proof. Assume $w(x) = \gcd(\alpha(x), \beta(x))$ is non-trivial. Thus we can find an extension field $\bar{\mathbb{F}}$ of $\mathbb{F}$ such that there exists $x_0 \in \bar{\mathbb{F}}$ which satisfies $w(x_0) = 0$ and hence $\alpha(x_0) = \beta(x_0) = 0$. In the rest of this proof, we restrict our discussion in $\bar{\mathbb{F}}$. Note that $\gcd(f(z), g(z)) = 1$ and $\gcd(s(x), t(x)) = 1$ also hold for $\bar{\mathbb{F}}$. Assume $t(x_0) = 0$ and thus $x - x_0 \mid t(x)$. Since $\gcd(s(x), t(x)) = 1$, it follows that $x - x_0 \nmid s(x)$ and thus $s(x_0) \neq 0$. Hence, either $\alpha(x_0) \neq 0$ or $\beta(x_0) \neq 0$, contradicting that $\alpha(x_0), \beta(x_0)$ are both zeros. Hence, we have proved that $t(x_0) \neq 0$. Then we have $f\big(\frac{s(x_0)}{t(x_0)}\big) = \frac{\alpha(x_0)}{t^d(x_0)} = 0$ and $g\big(\frac{s(x_0)}{t(x_0)}\big) = \frac{\beta(x_0)}{t^d(x_0)} = 0$, which implies that $z - \frac{s(x_0)}{t(x_0)}$ is a common divisor of $f(z)$ and $g(z)$, contradicting $\gcd(f(z), g(z)) = 1$. Thus, we have proved that $\gcd(\alpha(y), \beta(y)) = 1$. ■

B. Viewing Multivariate as Univariate

In order to extend Lemma D.1 to the multivariate case, we first show that each multivariate polynomial can be viewed as an equivalent univariate polynomial on a field of rational functions. Let $\mathbf{y} = (y_1, y_2, \cdots, y_k)$ be a vector of variables. For any $i \in \{1, 2, \cdots, k\}$, define $\mathbf{y}_i = (y_1, \cdots, y_{i-1}, y_{i+1}, \cdots, y_k)$, i.e., the vector consisting of all variables in $\mathbf{y}$ other than y_i. Note that any polynomial $f(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ can be formulated as $f(\mathbf{y}) = f_0(\mathbf{y}_i) + f_1(\mathbf{y}_i)y_i + \cdots + f_p(\mathbf{y}_i)y_i^p$, where each $f_j(\mathbf{y}_i)$ is a polynomial in $\mathbb{F}[\mathbf{y}_i]$. Because $\mathbb{F}[\mathbf{y}_i]$ is a subset of $\mathbb{F}(\mathbf{y}_i)$, $f(\mathbf{y})$ can also be viewed as a univariate polynomial in $\mathbb{F}(\mathbf{y}_i)[y_i]$. We use $f(y_i)$ to denote $f(\mathbf{y})$'s equivalent counterpart in $\mathbb{F}(\mathbf{y}_i)[y_i]$. To differentiate these two concepts, we reserve the notations, such as "$|$", "gcd" and "lcm" for field $\mathbb{F}$, and append "1" as a subscript to these notations to suggest they are specific to field $\mathbb{F}(\mathbf{y}_i)$. For example, for $f(\mathbf{y}), g(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ and $u(y_i), v(y_i) \in \mathbb{F}(\mathbf{y}_i)[y_i]$, $g(\mathbf{y}) \mid f(\mathbf{y})$ means that there exists $h(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ such that $f(\mathbf{y}) = h(\mathbf{y})g(\mathbf{y})$, and $u(y_i) \mid_1 v(y_i)$ means that there exists $w(y_i) \in \mathbb{F}[\mathbf{y}_i](y_i)$ such that $v(y_i) = w(y_i)u(y_i)$.

Lemma D.2. Assume $g(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$ and $f(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ is of the form $f(\mathbf{y}) = \sum_{j=0}^{p} f_j(\mathbf{y}_i)y_i^j$, where $f_j(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$. Then $g(\mathbf{y}_i) \mid f(\mathbf{y})$ if and only if $g(\mathbf{y}_i) \mid f_j(\mathbf{y}_i)$ for each $j \in \{0, 1, \cdots, p\}$.

Proof. Apparently, if $g(\mathbf{y}_i) \mid f_j(\mathbf{y}_j)$ for any $j \in \{0, 1, \cdots, p\}$, $g(\mathbf{y}_i) \mid f(\mathbf{y})$. Now assume $g(\mathbf{y}_i) \mid f(\mathbf{y})$. Thus there exists $h(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ such that $f(\mathbf{y}) = g(\mathbf{y}_i)h(\mathbf{y})$. Let $h(\mathbf{y}) = \sum_{j=0}^{p} h_j(\mathbf{y}_i)y_i^j$. Hence, it follows that $f_j(\mathbf{y}_i) = h_j(\mathbf{y}_i)g(\mathbf{y}_i)$ and thus $g(\mathbf{y}_i) \mid f_j(\mathbf{y}_i)$. ∎

The following result follows immediately from Lemma D.2.

Corollary D.1. Let $g(\mathbf{y}_i)$ and $f(\mathbf{y})$ be defined as Lemma D.2. Then $\gcd(g(\mathbf{y}_i), f(\mathbf{y})) = \gcd(g(\mathbf{y}_i), f_0(\mathbf{y}_i), \cdots, f_p(\mathbf{y}_i))$.

Proof. Note that any divisor of $g(\mathbf{y}_i)$ must be a polynomial in $\mathbb{F}[\mathbf{y}_i]$. Let $d(\mathbf{y}_i) = \gcd(g(\mathbf{y}_i), f(\mathbf{y}))$ and $d'(\mathbf{y}_i) = \gcd(g(\mathbf{y}_i), f_0(\mathbf{y}_i), \cdots, f_p(\mathbf{y}_i))$. By Lemma D.2, $d(\mathbf{y}_i) \mid f_j(\mathbf{y}_i)$ for any $j \in \{0, 1, \cdots, p\}$, implying that $d(\mathbf{y}_i) \mid d'(\mathbf{y}_i)$. On the other hand, $d'(\mathbf{y}_i) \mid f(\mathbf{y})$, and thus $d'(\mathbf{y}_i) \mid d(\mathbf{y}_i)$. Hence, $d(\mathbf{y}_i) = d'(\mathbf{y}_i)$. ∎

Corollary D.2. For $t \in \{1, 2, \cdots, s\}$, let $f_t(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ be defined as $f_t(\mathbf{y}) = \sum_{j=0}^{p_t} f_{tj}(\mathbf{y}_i)y_i^j$, where $f_{tj}(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$. Let $g(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$. It follows

$$\begin{aligned}&\gcd(g(\mathbf{y}_i), f_1(\mathbf{y}), \cdots, f_t(\mathbf{y}))\\ =&\gcd(g(\mathbf{y}_i), f_{10}(\mathbf{y}_i), \cdots, f_{1p_1}(\mathbf{y}_i), \cdots,\\ &\qquad f_{s0}(\mathbf{y}_i), \cdots, f_{sp_s}(\mathbf{y}_i))\end{aligned}$$

Proof. We have the following equations

$$\begin{aligned}&\gcd(g(\mathbf{y}_i), f_1(\mathbf{y}), \cdots, f_t(\mathbf{y}))\\ =&\gcd(g(\mathbf{y}_i), f_1(\mathbf{y}), \cdots, g(\mathbf{y}_i), f_t(\mathbf{y}))\\ =&\gcd(\gcd(g(\mathbf{y}_i), f_1(\mathbf{y})), \cdots, \gcd(g(\mathbf{y}_i), f_s(\mathbf{y})))\\ =&\gcd(g(\mathbf{y}_i), f_{10}(\mathbf{y}_i), \cdots, f_{1p_1}(\mathbf{y}_i), \cdots,\\ &\qquad g(\mathbf{y}_i), f_{s0}(\mathbf{y}_i), \cdots, f_{sp_s}(\mathbf{y}_i))\\ =&\gcd(g(\mathbf{y}_i), f_{10}(\mathbf{y}_i), \cdots, f_{1p_1}(\mathbf{y}_i), \cdots,\\ &\qquad f_{s0}(\mathbf{y}_i), \cdots, f_{sp_s}(\mathbf{y}_i))\end{aligned}$$

∎

Lemma D.3. For $t \in \{1, 2, \cdots, s\}$, let $a_t(\mathbf{y}), b_t(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$ such that $b_t(\mathbf{y}) \neq 0$ and $\gcd(a_t(\mathbf{y}), b_t(\mathbf{y})) = 1$. For $t \in \{1, 2, \cdots, s\}$, let $v_t(\mathbf{y}) = \mathrm{lcm}(b_1(\mathbf{y}), \cdots, b_t(\mathbf{y}))$. Then we have

$$\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_s(\mathbf{y})\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}, v_s(\mathbf{y})\right) = 1$$

Proof. We use induction on s to prove this lemma. Apparently, the lemma holds for $s = 1$ due to $\gcd(a_1(\mathbf{y}), b_1(\mathbf{y})) = 1$. Assume it holds for $s - 1$. Thus it follows

$$\begin{aligned}&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_s(\mathbf{y})\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}, v_s(\mathbf{y})\right)\\ =&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_s(\mathbf{y})\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}, b_s(\mathbf{y})\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ =&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, \gcd(a_s(\mathbf{y}), b_s(\mathbf{y}))\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ \overset{(a)}{=}&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_{s-1}(\mathbf{y})\frac{v_s(\mathbf{y})}{b_{s-1}(\mathbf{y})}, \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ \overset{(b)}{=}&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_{s-1}(\mathbf{y})\frac{v_s(\mathbf{y})}{b_{s-1}(\mathbf{y})},\right.\\ &\qquad\left.\gcd\left(v_{s-1}(\mathbf{y}), \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\right)\\ =&\gcd\left(a_1(\mathbf{y})\frac{v_s(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_{s-1}(\mathbf{y})\frac{v_s(\mathbf{y})}{b_{s-1}(\mathbf{y})}, v_{s-1}(\mathbf{y}), \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ =&\gcd\left(\frac{v_s(\mathbf{y})}{v_{s-1}(\mathbf{y})}\gcd\left(a_1(\mathbf{y})\frac{v_{s-1}(\mathbf{y})}{b_1(\mathbf{y})}, \cdots, a_{s-1}(\mathbf{y})\frac{v_{s-1}(\mathbf{y})}{b_{s-1}(\mathbf{y})}\right),\right.\\ &\qquad\left.v_{s-1}(\mathbf{y}), \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ \overset{(c)}{=}&\gcd\left(\frac{v_s(\mathbf{y})}{v_{s-1}(\mathbf{y})}, v_{s-1}(\mathbf{y}), \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})}\right)\\ \overset{(d)}{=}&\gcd\left(\frac{b_s(\mathbf{y})}{\gcd(v_{s-1}(\mathbf{y}), b_s(\mathbf{y}))}, v_{s-1}(\mathbf{y}), \frac{v_{s-1}(\mathbf{y})}{\gcd(v_{s-1}(\mathbf{y}), b_s(\mathbf{y}))}\right)\\ =&\gcd(1, v_{s-1}(\mathbf{y})) = 1\end{aligned}$$

In the above equations, (a) is due to $\gcd(a_s(\mathbf{y}), b_s(\mathbf{y})) = 1$; (b) follows from the fact that $\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})} \mid v_{s-1}(\mathbf{y})$ and thus $\frac{v_s(\mathbf{y})}{b_s(\mathbf{y})} = \gcd(v_{s-1}(\mathbf{y}), \frac{v_s(\mathbf{y})}{b_s(\mathbf{y})})$; (c) follows from the inductive assumption; (d) is due to the equality: $v_s(\mathbf{y}) = \mathrm{lcm}(v_{s-1}(\mathbf{y}), b_s(\mathbf{y})) = \frac{v_{s-1}(\mathbf{y})b_s(\mathbf{y})}{\gcd(v_{s-1}(\mathbf{y}), b_s(\mathbf{y}))}$. ∎

In general, each polynomial $h(y_i) \in \mathbb{F}(\mathbf{y}_i)[y_i]$ is of the form $h(y_i) = \frac{a_0(\mathbf{y}_i)}{b_0(\mathbf{y}_i)} + \frac{a_1(\mathbf{y}_i)}{b_1(\mathbf{y}_i)}y_i + \cdots + \frac{a_p(\mathbf{y}_i)}{b_p(\mathbf{y}_i)}y_i^p$, where for each $j \in \{0, 1, \cdots, p\}$, $a_j(\mathbf{y}_i), b_j(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$, $b_j(\mathbf{y}_i) \neq 0$, $\gcd(a_j(\mathbf{y}_i), b_j(\mathbf{y}_i)) = 1$, and $a_p(\mathbf{y}_i) \neq 0$. Note that for each y_i^j which is absent in $h(y_i)$, we let $a_j(\mathbf{y}_i) = 0$ and $b_j(\mathbf{y}_i) = 1$. Moreover, define the following polynomial $\mu_h(\mathbf{y}_i) = \mathrm{lcm}(b_0(\mathbf{y}_i), b_1(\mathbf{y}_i), \cdots, b_p(\mathbf{y}_i))$.

Corollary D.3. For $j \in \{1, 2, \cdots, s\}$, let $f_j(y_i) \in \mathbb{F}(\mathbf{y}_i)[y_i]$. Define $v(\mathbf{y}_i) = \mathrm{lcm}(\mu_{f_1}(\mathbf{y}_i), \cdots, \mu_{f_s}(\mathbf{y}_i))$ and $\bar{f}_j(\mathbf{y}) = v(\mathbf{y}_i)f_j(y_i)$. Thus $\gcd(v(\mathbf{y}_i), \bar{f}_1(\mathbf{y}), \cdots, \bar{f}_s(\mathbf{y})) = 1$

Proof. Assume $f_j(y_i)$ has the following form:

$$f_j(y_i) = \frac{a_{j0}(\mathbf{y}_i)}{b_{j0}(\mathbf{y}_i)} + \frac{a_{j1}(\mathbf{y}_i)}{b_{j1}(\mathbf{y}_i)}y_i + \cdots + \frac{a_{jp_j}(\mathbf{y}_i)}{b_{jp_j}(\mathbf{y}_i)}y_i^{p_j}$$

where for any $j \in \{1, 2, \cdots, s\}$ and $t \in \{0, 1, \cdots, p_j\}$, $a_{jt}(\mathbf{y}_i), b_{jt}(\mathbf{y}_i) \in \mathbb{F}[\mathbf{y}_i]$, $b_{jt}(\mathbf{y}_i) \neq 0$ and $\gcd(a_{jt}(\mathbf{y}_i), b_{jt}(\mathbf{y}_i)) = 1$. Apparently, $v(\mathbf{y}_i)$ is the least common multiple of all $b_{jt}(\mathbf{y}_i)$'s. Define $u_{jt}(\mathbf{y}_i) = \frac{v(\mathbf{y}_i)}{b_{jt}(\mathbf{y}_i)} \in \mathbb{F}[\mathbf{y}_i]$. Hence, we have $\bar{f}_j(\mathbf{y}) = \sum_{t=0}^{p_j} a_{jt}(\mathbf{y}_i)u_{jt}(\mathbf{y}_i)y_i^t$. Then it follows

$$\begin{aligned}&\gcd(v(\mathbf{y}_i), \bar{f}_1(\mathbf{y})), \cdots, \bar{f}_s(\mathbf{y}))\\ \overset{(a)}{=}&\gcd(v(\mathbf{y}_i), a_{10}(\mathbf{y}_i)u_{10}(\mathbf{y}_i), \cdots, a_{1p_1}(\mathbf{y}_i)u_{1p_1}(\mathbf{y}_i), \cdots,\\ &\qquad a_{s0}(\mathbf{y}_i)u_{s0}(\mathbf{y}_i), \cdots, a_{sp_s}(\mathbf{y}_i)u_{sp_s}(\mathbf{y}_i))\\ \overset{(b)}{=}&1\end{aligned}$$

where (a) is due to Corollary D.2 and (b) follows from Lemma D.3. ∎

Generally, the definitions of division in $\mathbb{F}[\mathbf{y}]$ and $\mathbb{F}(\mathbf{y}_i)[y_i]$ are different. However, the following theorem reveals the two definitions are closely related.

Theorem D.1. Consider two polynomials $f(\mathbf{y}), g(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$, where $g(\mathbf{y}) \neq 0$. Then $g(\mathbf{y}) \mid f(\mathbf{y})$ if and only if $g(y_i) \mid_1 f(y_i)$ for every $i \in \{1, 2, \cdots, k\}$.

Proof. The division equation between $f(y_i)$ and $g(y_i)$ is as follows

$$f(y_i) = h_i(y_i)g(y_i) + r_i(y_i) \tag{48}$$

where $h_i(y_i), r_i(y_i) \in \mathbb{F}(\mathbf{y}_i)[y_i]$, and either $r_i(y_i) = 0$ or $d_{r_i} < d_g$. Due to the uniqueness of Equation (48), $f(\mathbf{y}) \mid g(\mathbf{y})$ immediately implies that for any $i \in \{1, 2, \cdots, k\}$, $r_i(y_i) = 0$ and thus $g(y_i) \mid_1 f(y_i)$.

Conversely, assume for every $i \in \{1, \cdots, k\}$, $g(y_i) \mid_1 f(y_i)$ and hence $r_i(y_i) = 0$. Denote $\bar{h}_i(\mathbf{y}) = \mu_{h_i}(\mathbf{y}_i)h_i(y_i)$. Clearly, $\bar{h}_i(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$. Then, the following equation holds

$$\mu_{h_i}(\mathbf{y}_i)f(\mathbf{y}) = \bar{h}_i(\mathbf{y})g(\mathbf{y})$$

By Corollary D.3, $\gcd(\mu_{h_i}(\mathbf{y}_i), \bar{h}_i(\mathbf{y})) = 1$. Thus, $\mu_{h_i}(\mathbf{y}_i) \mid g(\mathbf{y})$. Define $\bar{g}(\mathbf{y}) = \frac{g(\mathbf{y})}{\mu_{h_i}(\mathbf{y}_i)}$. By Lemma D.2, $\bar{g}(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$. Define $u(\mathbf{y}) = \frac{g(\mathbf{y})}{\gcd(f(\mathbf{y}), g(\mathbf{y}))} \in \mathbb{F}[\mathbf{y}]$. It follows that

$$\begin{aligned} u(\mathbf{y}) &= \frac{g(\mathbf{y})}{\gcd(f(\mathbf{y}), g(\mathbf{y}))} \\ &= \frac{\mu_{h_i}(\mathbf{y}_i)\bar{g}(\mathbf{y})}{\gcd(\bar{h}_i(\mathbf{y})\bar{g}(\mathbf{y}), \mu_{h_i}(\mathbf{y}_i)\bar{g}(\mathbf{y}))} \\ &= \frac{\mu_{h_i}(\mathbf{y}_i)\bar{g}(\mathbf{y})}{\bar{g}(\mathbf{y})\gcd(\bar{h}_i(\mathbf{y}), \mu_{h_i}(\mathbf{y}_i))} \\ &= \frac{\mu_{h_i}(\mathbf{y}_i)\bar{g}(\mathbf{y})}{\bar{g}(\mathbf{y})} \\ &= \mu_{h_i}(\mathbf{y}_i) \end{aligned}$$

Note that variable y_i is absent in $u(\mathbf{y})$. Because y_i can be any arbitrary variable in $\mathbf{y}$, it immediately follows that all the variables in $\mathbf{y}$ must be absent in $u(\mathbf{y})$, implying that $u(\mathbf{y})$ is a constant in $\mathbb{F}$. Hence $g(\mathbf{y}) \mid f(\mathbf{y})$. ∎

Moreover, in the next theorem, we will prove that checking if two multivariate polynomials are co-prime is equivalent to checking if their equivalent univariate polynomials are co-prime.

Theorem D.2. Let $f(\mathbf{y}), g(\mathbf{y})$ be two non-zero polynomials in $\mathbb{F}[\mathbf{y}]$. Then $\gcd(f(\mathbf{y}), g(\mathbf{y})) = 1$ if and only if $\gcd_1(f(y_i), g(y_i)) = 1$ for any $i \in \{1, 2, \cdots, k\}$.

Proof. First, assume for any $i \in \{1, 2, \cdots, k\}$, $\gcd_1(f(y_i), g(y_i)) = 1$. We use contradiction to prove that $\gcd(f(\mathbf{y}), g(\mathbf{y})) = 1$. Assume $u(\mathbf{y}) = \gcd(f(\mathbf{y}), g(\mathbf{y}))$ is not constant. Let y_i be a variable which is present in $u(\mathbf{y})$. By Theorem D.1, $u(y_i) \mid_1 f(y_i)$ and $u(y_i) \mid_1 g(y_i)$, which contradicts that $\gcd_1(f(y_i), g(y_i)) = 1$.

Then, assume $\gcd(f(\mathbf{y}), g(\mathbf{y})) = 1$. We also use contradiction to prove that for any $i \in \{1, 2, \cdots, k\}$, $\gcd_1(f(y_i), g(y_i)) = 1$. Assume there exists $i \in \{1, \cdots, k\}$ such that $v(y_i) = \gcd_1(f(y_i), g(y_i))$ is non-trivial. Define $w(\mathbf{y}) = \mu_v(\mathbf{y}_i)v(y_i) \in \mathbb{F}[\mathbf{y}]$. Clearly, $w(y_i) \mid_1 f(y_i)$ and $w(y_i) \mid_1 g(y_i)$. Thus, there exists $p(y_i), q(y_i) \in \mathbb{F}(\mathbf{y}_i)[y_i]$ such that

$$f(y_i) = w(y_i)p(y_i) \qquad g(y_i) = w(y_i)q(y_i)$$

Let $s(\mathbf{y}_i) = \text{lcm}(\mu_p(\mathbf{y}_i), \mu_q(\mathbf{y}_i))$. Define $\bar{p}(\mathbf{y}) = s(\mathbf{y}_i)p(y_i)$ and $\bar{q}(\mathbf{y}) = s(\mathbf{y}_i)q(y_i)$. Apparently, $\bar{p}(\mathbf{y}), \bar{q}(\mathbf{y}) \in \mathbb{F}[\mathbf{y}]$. It follows that

$$s(\mathbf{y}_i)f(\mathbf{y}) = w(\mathbf{y})\bar{p}(\mathbf{y}) \qquad s(\mathbf{y}_i)g(\mathbf{y}) = w(\mathbf{y})\bar{q}(\mathbf{y})$$

Then the following equation holds

$$s(\mathbf{y}_i)\gcd(f(\mathbf{y}), g(\mathbf{y})) = w(\mathbf{y})\gcd(\bar{p}(\mathbf{y}), \bar{q}(\mathbf{y}))$$

Due to Corollary D.3, $\gcd(s(\mathbf{y}_i), \gcd(\bar{p}(\mathbf{y}), \bar{q}(\mathbf{y}))) = \gcd(s(\mathbf{y}_i), \bar{p}(\mathbf{y}), \bar{q}(\mathbf{y})) = 1$. Hence $s(\mathbf{y}_i) \mid w(\mathbf{y})$. Let $\bar{w}(\mathbf{y}) = \frac{w(\mathbf{y})}{s(\mathbf{y}_i)}$. According to Lemma D.2, $\bar{w}(\mathbf{y})$ is a non-trivial polynomial in $\mathbb{F}[\mathbf{y}]$. Thus, $\bar{w}(\mathbf{y}) \mid \gcd(f(\mathbf{y}), g(\mathbf{y}))$, contradicting $\gcd(f(\mathbf{y}), g(\mathbf{y})) = 1$. ∎

C. The Multivariate Case

Now, we are in the place of extending Lemma D.1 to the multivariate case.

Proof of Lemma B.1. Note that if we substitute $\mathbb{F}$ with $\mathbb{F}(\mathbf{y}_i)$ and gcd with $\gcd_1$ in Lemma D.1, the lemma also holds. Apparently, $f(z), g(z) \in \mathbb{F}(\mathbf{y}_i)[z]$. We will prove that $\gcd_1(f(z), g(z)) = 1$. By contradiction, assume $r(z) = \gcd_1(f(z), g(z)) \in \mathbb{F}(\mathbf{y}_i)[z]$ is non-trivial. Let $\bar{f}(z) = \frac{f(z)}{r(z)}$ and $\bar{g}(z) = \frac{g(z)}{r(z)}$. Clearly, $\bar{f}(z)$ and $\bar{g}(z)$ are both non-zero polynomials in $\mathbb{F}(\mathbf{y}_i)[z]$. Then we can find an assignment to $\mathbf{y}_i$, denoted by $\mathbf{y}_i^*$, such that the coefficients of the maximum powers of z in $r(z), \bar{f}(z)$ and $\bar{g}(z)$ are all non-zeros. Let $\bar{r}(z)$ denote the univariate polynomial acquired by assigning $\mathbf{y}_i = \mathbf{y}_i^*$ to $r(z)$. Clearly, $\bar{r}(z)$ is a common divisor of $f(z)$ and $g(z)$ in $\mathbb{F}[z]$, contradicting $\gcd(f(z), g(z)) = 1$. Moreover, due to $\gcd(s(\mathbf{y}), t(\mathbf{y})) = 1$ and Theorem D.2, $\gcd_1(s(y_i), t(y_i)) = 1$. Thus, by Lemma D.1, $\gcd_1(\alpha(y_i), \beta(y_i)) = 1$. Since i can be any integer in $\{1, 2, \cdots, k\}$, it follows that $\gcd(\alpha(\mathbf{y}), \beta(\mathbf{y})) = 1$ by Theorem D.2. ∎

APPENDIX E
PBNA VS. ROUTING

In Section VIII, we characterized the optimal rates for different network topologies and under the network model considered in this paper (precoding and RLNC). In this appendix, we provide a comparison of the rate achieved by PBNA to that achieved by routing. [8] Depending on the network structure one scheme can perform better than the other. In Fig. 12, we provide a taxonomy of the networks based on their structure and we provide the rates achievable by routing and PBNA. In particular, we classify networks based on the coupling relations in Section V-C, repeated here for convenience.

- *Type I*: Networks in which at least one of the coupling relations, $p_i(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) = \eta(\mathbf{x})$ $(1 \leq i \leq 3)$, is present. This network structure makes it information theoretically impossible for any precoding-based linear schemes to achieve a rate of more that $\frac{1}{3}$ per session, under the considered network model (precoding at the edge and RLNC in the middle).

[8] We would like to point out that the two schemes are not directly comparable under the model we consider. Routing involves intelligence inside the network, whereas in our problem setup, the internal nodes have no intelligence and can only perform random linear network coding. Therefore, routing by definition is not included in the problem we study in this paper.

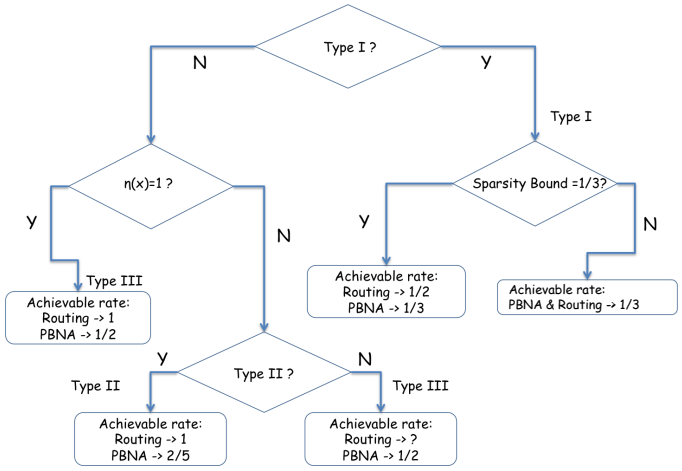


Fig. 12. A comparison between PBNA and routing in terms of achievable symmetric rate for various types of networks.

- *Type II*: Networks in which $p_i(\mathbf{x}) \notin \{1, \eta(\mathbf{x})\}$ for $1 \leq i \leq 3$, but one of the three mutually exclusive coupling conditions, $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$, $p_2(\mathbf{x}) = 1 + \eta(\mathbf{x})$, and $p_3(\mathbf{x}) = 1 + \eta(\mathbf{x})$, is present. The structure of these networks makes it impossible for any precoding-based linear schemes to achieve rate above $\frac{2}{5}$ per user under the considered network setting.
- *Type III*: Networks in which none of the above coupling relations is present and PBNA achieves rate $\frac{1}{2}$.

Type III networks, the ones with $\eta(\mathbf{x}) \neq 1$, are the main focus of this paper. For this type of networks, the performance of routing varies for different networks. In contrast, PBNA always achieves a guaranteed rate $\frac{1}{2}$ per session.

The following points can be noted from Fig. 12:

- *Type I* networks can be further classified into two cases based on the sparsity bound. When the sparsity bound equals $\frac{1}{3}$, both PBNA and routing can only achieve a symmetric rate of $\frac{1}{3}$ per user. An example of such network is shown in Fig. 13a. When the sparsity bound is greater than $\frac{1}{3}$, routing can achieve a symmetric rate of $\frac{1}{2}$ per user. However, PBNA can only achieve a symmetric rate of $\frac{1}{3}$ per user, which is the optimal symmetric rate achieved by any precoding-based linear schemes. Fig. 13b illustrates such an example.
- Type II networks, due to the presence of the coupling relations, $p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{1+\eta(\mathbf{x})}$ or $p_2(\mathbf{x}) = 1 + \eta(\mathbf{x})$ or $p_3(\mathbf{x}) = 1 + \eta(\mathbf{x})$, will have a network structure where each source has a disjoint path to its corresponding receiver, making it possible to achieve a rate of 1 per user with routing. In contrast, PBNA can only achieve a symmetric rate of $\frac{2}{5}$ for these networks. An example of such a network is shown in Fig. 13c.
- For Type III networks, PBNA can achieve a symmetric rate of $\frac{1}{2}$. Consider the special case of $\eta(\mathbf{x}) = 1$, it can be shown (see Subsection E-A) that in these networks, there are disjoint paths from each source to its corresponding receiver, and routing can always achieve a symmetric rate of 1 per user here. An example is shown in Fig. 13f.

For the less constrained case of $\eta(\mathbf{x}) \neq 1$, however, the performance of routing depends on additional properties. We can see that there are networks in which routing can only achieve a symmetric rate of $\frac{1}{3}$ (see Fig. 13d); and there are also networks where routing can achieve a symmetric rate of one due to the rich connectivity in the network (see Fig. 13e).

A. Characterizing the Routing Rate for Type III Networks with $\eta(\mathbf{x}) = 1$

In this subsection, we prove that for Type III networks with $\eta(\mathbf{x}) = 1$, routing can always achieve a symmetric rate of one.

We will first define the following polynomials:

$$L(\mathbf{x}) = m_{13}(\mathbf{x})m_{32}(\mathbf{x})m_{21}(\mathbf{x}) \quad R(\mathbf{x}) = m_{12}(\mathbf{x})m_{23}(\mathbf{x})m_{31}(\mathbf{x})$$

Thus, $\eta(\mathbf{x}) = \frac{L(\mathbf{x})}{R(\mathbf{x})}$. Given two distinct edges/nodes e_1, e_2, if there exists a directed path from e_1 to e_2, we say e_1 is upstream of e_2 (or e_1 is downstream of e_2), and denote this relation by $e_1 \prec e_2$. Similarly, $e_1 \not\prec v_2$ implies that there is no directed path from e_1 to e_2.

Given two subsets of nodes $S, D \subseteq V$, let $EC(S; D)$ denote the minimum capacity of all the edge cuts separating S from D. Define the following subsets of edges:

$$\bar{S}_i \triangleq \{e \in E - \{\sigma_i\} \ : \ e \in C_{ij} \cap C_{ik}, j \neq k, \\ j, k \in \{1, 2, 3\} - \{i\}\}$$

$$\bar{D}_i \triangleq \{e \in E - \{\tau_i\} \ : \ e \in C_{jj} \cap C_{kj}, j \neq k, \\ j, k \in \{1, 2, 3\} - \{i\}\}$$

The following proposition was stated in [35], which gives a graph theoretic interpretation of the condition $\eta(\mathbf{x}) = 1$.

Proposition E.1. $L(\mathbf{x}) \equiv R(\mathbf{x})$ if and only if there exists two distinct integers $i, j \in \{1, 2, 3\}$ such that $\bar{S}_i \cap \bar{S}_j \neq \emptyset$ and $\bar{D}_i \cap \bar{D}_j \neq \emptyset$.

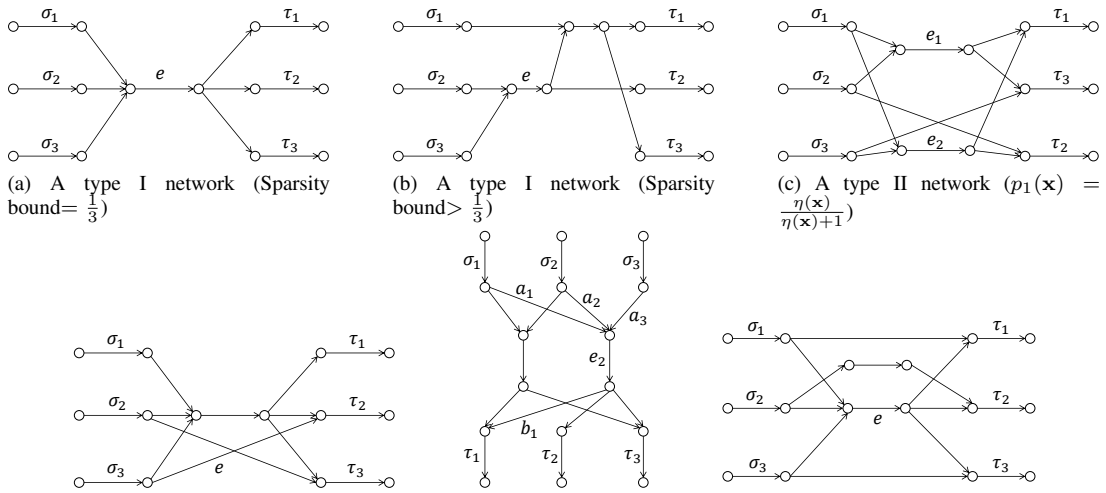
Lemma E.1. Let i, j be two distinct integers in $\{1, 2, 3\}$, and $e_2 \in \bar{D}_i \cap \bar{D}_j$. If $\bar{S}_i \cap \bar{S}_j \neq \emptyset$, then there exists $e_1 \in \bar{S}_i \cap \bar{S}_j$ such that $e_1 \prec e_2$ or $e_1 = e_2$.

Proof. Same as *lemma* 5 in [35] . ■

Lemma E.2. For a given $i, j, k \in \{1, 2, 3\}$ and $i \neq j \neq k$, if $\bar{S}_i \cap \bar{S}_j \neq \emptyset$; $\bar{D}_i \cap \bar{D}_j \neq \emptyset$ and $EC(\{s_i, s_j\}; \{d_i, d_k\}) > 1$, then there exists a path P'_{ii} from s_i to d_i such that for each $e' \in P'_{ii}$, $s_j \not\prec e'$, $s_k \not\prec e'$, and $e' \not\prec d_j$, $e' \not\prec d_k$.

Proof. Without loss of generality, suppose $i = 1$, $j = 2$ and $k = 3$. We can choose two edges $e_1 \in \bar{S}_1 \cap \bar{S}_2$ and $e_2 \in \bar{D}_1 \cap \bar{D}_2$ such that $e_1 \prec e_2$ or $e_1 = e_2$ (from *lemma* 1). Now consider the edge e_1, by definition cutting this edge would cut the flows $s_1 \to d_2$, $s_1 \to d_3$, $s_2 \to d_1$ and $s_2 \to d_3$. Since we also have $EC(\{s_i, s_j\}; \{d_i, d_k\}) > 1$, we can see that there should exist a path P'_{11} such that $e_1 \notin P'_{11}$. Consider any edge $e' \in P_{11}$,

- If this edge e' has d_2 (or d_3) as a *downstream* node, then there will exist a path P_{12} (or P_{13}) such that $e_1 \notin P_{12}$ (or $e_1 \notin P_{13}$), which contradicts the definition of edge e_1 (or e_2). Thus $e' \not\prec d_2$, $e' \not\prec d_3$.



(a) A type I network (Sparsity bound= $\frac{1}{3}$)

(b) A type I network (Sparsity bound> $\frac{1}{3}$)

(c) A type II network ($p_1(\mathbf{x}) = \frac{\eta(\mathbf{x})}{\eta(\mathbf{x})+1}$)

(d) A type III network ($\eta(\mathbf{x}) \neq 1$)

(e) A type III network ($\eta(\mathbf{x}) \neq 1$)

(f) A type III network ($\eta(\mathbf{x}) = 1$)

Fig. 13. Example networks. (a) shows a Type I network, for which PBNA and routing both achieve symmetric rate $\frac{1}{3}$. (b) shows another Type I network, for which routing can achieve symmetric rate $\frac{1}{2}$, and PBNA can only achieve symmetric rate $\frac{1}{3}$. (c) shows a Type II network, for which routing can achieve symmetric rate one, and PBNA can only achieve symmetric rate $\frac{2}{5}$. (d) shows a Type III network, for which routing can only achieve symmetric rate $\frac{1}{3}$, and PBNA can achieve symmetric rate $\frac{1}{2}$. In (e), we show another Type III network, for which routing can achieve symmetric rate one, and PBNA can only achieve symmetric rate $\frac{1}{2}$. (f) shows a Type III network, for which routing can always achieve symmetric rate one, and PBNA can only achieve symmetric rate $\frac{1}{2}$.

- Similarly, if edge e' is *downstream* of s_2, it would result in a path P_{21} such that $e_1 \notin P_{21}$, which again will contradict the definition of e_1. Thus $s_2 \not\prec e'$.
- If edge e' is *downstream* of s_3, it would result in a path P_{31}, where $e_1 \notin P_{31}$. But by definition of e_2, $e_2 \in P_{31}$, this in turn would result in paths P'_{12} and P'_{13} that does not go through edge e_1. Thus $s_3 \not\prec e'$.

■

Theorem E.1. Assume that all the senders are connected to all the receivers via directed paths. If $\eta(\mathbf{x}) = 1$ and $p_i(\mathbf{x}) \neq 1$ for $1 \leq i \leq 3$, then routing can achieve the rate tuple $(1, 1, 1)$.

Proof. Without loss of generality, suppose $i = 1, j = 2, k = 3$ and $\bar{S}_1 \cap \bar{S}_2 \neq \emptyset$; $\bar{D}_1 \cap \bar{D}_2 \neq \emptyset$. From Lemma E.2, we can see that there exist two disjoint paths, $P_1 \in P_{11}$ and $P_2 \in P_{22}$. Therefore, ω_1 and ω_2 can transmit one unit flow through P_1 and P_2 respectively. Meanwhile, ω_3 can route one unit flow through the rest of the network. This implies that routing can achieve the rate tuple $(1, 1, 1)$. ■