Channel Knowledge Acquisition in Relay and Multi-Point to Multi-Point Transmission Systems

Hlaing Minn, *Senior Member, IEEE,* and Daniel Munoz

Abstract—Most of the relaying strategies exploit channel state information (CSI), but there exist only a few limited works on acquisition of the required CSI. In this paper we present general protocols and pilot designs for CSI acquisition in different types of relay systems using multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) transmission. We propose new pilot forwarding mechanisms and a related estimator for obtaining CSIs of individual MIMO channels of cascaded links at the end node(s). We develop new pilot designs with low overhead for channels with different lengths. We present several CSI acquisition protocols for relay MIMO systems with time division duplexing (TDD) as well as frequency division duplexing (FDD). We introduce node multiplexing which reduces overhead and delay for pilot or CSI feedback transmissions from multiple nodes. In addition, we present similar CSI acquisition protocols for multi-point to multi-point systems while keeping overhead low. Simulation results are included to demonstrate the estimation performance of our proposed designs. Discussions on several protocol options are also provided.

Index Terms—CSI acquisition, Relay, Multi-point to multi-point, Pilot designs, Channel estimation

I. INTRODUCTION

Relay systems [1]–[11] are among the promising technologies to address coverage and data rate issues of wireless communication systems. These relaying strategies rely heavily on accurate channel state information (CSI) of individual or cascaded links. For example, [3] considers a multiple source, multiple relay, multiple destination multiple-input multiple-output (MIMO) system. Each relay needs knowledge of forward and reverse channel links. The destinations need information of the cascaded channel between the source, relay, and destination. [6] assumes that perfect CSI of the first and second hop is available at the source. In summary, protocols and pilot designs for CSI acquisition at various nodes are needed in order for these relaying schemes to work. The acquisition of the first hop CSI would be straightforward from previous multi-input multi-output (MIMO) channel estimation designs (see [12]–[15]), but getting the second hop CSI at the source needs further investigation. [16]–[26] have timely addressed channel estimation in relay systems. However, their applications are limited to time division duplexing (TDD) only and for channels with the same length. [25], [26] proposed estimation methods for relay assisted orthogonal frequency

division multiplexing (OFDM) systems with single antenna nodes. [20]–[24] presented estimation methods for single carrier MIMO systems. [16] and [19] proposed relay-inserted pilot tones and related channel estimators. Both of them consider a single antenna OFDM based 3 node relay system with TDD and do not consider pilot overhead cost. New designs are needed to account for more realistic or general system conditions such as MIMO, multiple nodes, unequal channel lengths due to different node locations, and differing duplexing schemes (e.g., frequency division duplexing (FDD)), as well as address the CSI overhead cost, both in terms of tones utilized and delay. We will develop such designs in this paper.

A similar problem of CSI acquisition also arises in multi-point to multi-point (MP-MP) systems where multiple source-destination pairs are directly communicating simultaneously (also see Fig. 1(b)). MP-MP systems do not have intermediate nodes (relays) but to manage potential interferences among multiple simultaneous pairs and to enhance system capacity, MP-MP systems require CSI. The sources can either transmit on separate or the same frequency band to the multiple destinations. If the sources transmit on separate frequency bands, then one node can perform resource allocation, and a strategy to provide this node with the desired CSI is needed. If the sources transmit on the same frequency band, the system utilizes interference alignment (IA) [27]–[31] to decouple the desired signal from the interference. IA needs each node to have knowledge of all CSIs between sources and destinations, but necessary CSI acquisition protocol and pilot design have not been addressed.

This paper addresses CSI acquisition in multi-node relay and MP-MP systems with MIMO OFDM transmission. Our contributions are summarized in the following. 1) We propose new MIMO pilot designs for estimation of channels with unequal lengths and new pilot forwarding mechanisms from multiple antennas and/or nodes together with pilot insertion to estimate individual MIMO channels of all hops. 2) For estimating CSI of indirectly connected MIMO links (e.g., estimating source to relay channels at a destination node), we derive a new improved estimator using the forwarded pilots which incorporates effects of channel estimation errors of the directly connected link (e.g., relay to destination link). 3) During CSI acquisition phase when multiple nodes are involved in transmitting pilots or CSI feedback, we propose a strategy called node multiplexing which reduces overhead and delay. 4) We propose protocols for CSI acquisition in relay systems for both TDD and FDD, and present their associated overhead and delay. 5) New approaches for acquiring CSI in two types of MP-MP systems, characterized by disjoint or non-disjoint resource usage, are outlined for both TDD and FDD and the associated overhead and delay costs are calculated.

H. Minn is and D. Munoz was with the University of Texas at Dallas, Richardson, TX, USA. e-mail: hlaing.minn@utdallas.edu, dmunoz007@gmail.com.

6) We present how to design pilots for minimal overhead with almost optimal performance, equal channel estimation mean-square error (MSE) across links with different lengths, or minimum total channel estimation MSE of all considered channels, as well as how to set pilot signal-to-noise ratios (SNRs) of different hops, and how to select protocol options.

The paper is organized as follows. Section II describes the considered systems. Section III presents several pilot designs, pilot forwarding, and node multiplexing. Section IV derives a new channel estimator based on the forwarded pilots as well as its estimation MSE. Next, we propose several CSI acquisition protocols for relay systems in Section V and for MP-MP systems in Section VI. Section VII discusses protocol choices and simulation results, and Section VIII provides our conclusions.

II. SYSTEM MODEL

We consider relay systems and MP-MP systems as shown in Fig. 1, both using OFDM. As the main application scenario of relays in next generation systems (e.g., 3GPP LTE-Advanced) is to extend coverage range or fill dead spots where the destination nodes cannot receive direct signals from the source, we consider the same scenario of relay systems with no direct link between any source and destination. Such relay systems have been already included in the standards IEEE 802.16j [33] and 3GPP LTE-Advanced Release 10 [34].

Before performing CSI acquisition, any system needs to have already established links and synchronization among involving nodes. In the same manner, we consider that the initial link establishment and synchronization between sources, relays, and destinations have been completed. Then, for CSI acquisition, conventional systems will send control signaling messages (instructing the nodes to acquire CSIs and feedback them) which also include node IDs and related resource assignments for pilots and/or CSI feedback. In contrast, to reduce overhead and delay of CSI acquisition, the proposed approach includes assignment of distinct numbers to the nodes during their initial link establishment procedure. During initial link establishment, corresponding signaling exchanges occur between nodes, and adding number assignment in that process will bear only a small amount of overhead. By such distinct number, each node knows its pilots and its role in CSI acquisition protocol; thus, during CSI acquisition phase, the proposed approach does not require resource assignment for CSI feedback and related node IDs, saving overhead in control signaling (see details in Section VII).

In either type of the systems, we consider multiple nodes of each kind, and every node may have multiple antennas. We use S_i, R_k, and D_l to denote source i, relay k, and destination l, respectively. Similarly, $S_{i,\nu}$, $R_{k,\nu}$, and $D_{l,\nu}$ refer to those nodes with a specific antenna ν, while S, R and D represent all sources, all relays, and all destinations, respectively. The variables i, k, l, m, n, and ν are not specific to certain indexing and the referring index domain is obvious from the main variable. CSI$_{I\text{-}J}$ represents the information of channel impulse responses (CIRs) between I and J, where I and J can be $S_{i,\nu}$, S_i, S, $R_{i,\nu}$, R_i, R, $D_{i,\nu}$, D_i, or D,

depending on the considered link(s). Each node knows the number of sources, $\tilde{N}_S$, and destinations, $\tilde{N}_D$, (and relays, $\tilde{N}_R$, for relay systems) as well as the number of antennas ($N_{S_i} \ \forall \ i$, $N_{D_\nu} \ \forall \ \nu$, and $N_{R_k} \ \forall \ k$, respectively). OFDM system has a total of N subcarriers and the cyclic prefix length is longer than the maximum channel length plus potential asynchronous arrivals and timing synchronization errors to avoid inter-symbol interference. In order to reduce inter-cell interference, we assume an appropriate frequency reuse strategy is in place.

We assume that during the CSI acquisition and data transmission interval, the channels remain static but it can be relaxed for the data transmission interval as some transmission schemes [32] accommodate channel variations within a frame. The channel between $I_{i,\nu}$ and $J_{m,n}$ is represented by a time domain vector $\boldsymbol{h}_{I_{i,\nu}\text{-}J_{m,n}}$ of size $L_{I_{i,\nu}\text{-}J_{m,n}} \times 1$, and its corresponding frequency-domain channel gain vector is denoted by $\boldsymbol{H}_{I_{i,\nu}\text{-}J_{m,n}}$ whose kth element $H_{I_{i,\nu}\text{-}J_{m,n}}[k]$ represents the channel gain on subcarrier k.

For relay systems, there will be $N_{S_{\text{TOT}}}N_{R_{\text{TOT}}}$ first hop links and $N_{R_{\text{TOT}}}N_{D_{\text{TOT}}}$ second hop links to estimate, where $N_{I_{\text{TOT}}} = \sum_{i=0}^{\tilde{N}_I-1} N_{I_i}$ is the total number of antennas of type I (source, relay, or destination). If the desired CSI corresponds to individual cascaded channels, then the number of channels that need to be estimated is $N_{S_{\text{TOT}}}N_{R_{\text{TOT}}}N_{D_{\text{TOT}}}$. The channel length $L_{I_{i,\nu}\text{-}J_{m,n}}$ is in general unknown a priori, however, we can use the maximum number of taps needed to model the link with the largest delay (obtained from measurement in the deployed environment) to model the links. Therefore, we will define $L_1 = \max\{L_{S_{i,\nu}\text{-}R_{m,n}} : \ \forall \ i,\nu,m,n\}$ and $L_2 = \max\{L_{R_{m,n}\text{-}D_{l,k}} : \ \forall \ m,n,l,k\}$. A simple definition of $L = \max\{L_1, L_2\}$ can be used, if desired.

Regarding MP-MP systems, we consider those which involve simultaneous transmissions from multi-nodes to multi-nodes in the physical layer. For such systems to work, initial synchronization and link establishment need to be performed first. We assume such process has been done and all nodes are properly synchronized and they know what are the other nodes involved in the multiple node transmission. During this initial link establishment stage, a distinct sequence number can be assigned to each node so that there is no conflict in their pilot selection and protocol execution. The resources for different links of an MP-MP system can be disjoint similar to orthogonal frequency division multiple access (OFDMA) or the same using IA. The CSIs of $N_{S_{\text{TOT}}}N_{D_{\text{TOT}}}$ channels may be needed at the node performing resource allocation in the former case, or at each node for the latter case. The maximum length of these channels will be represented by L.

Notations: $\boldsymbol{I}_M$ and $\boldsymbol{1}_K$ denote the $M \times M$ identity matrix and the $K \times 1$ all-one vector while the size indicators will be excluded if the size is obvious from the usage. The superscripts $*$, T and $\dagger$ mean the conjugate, the transpose and the Hermitian transpose. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{Q})$ denotes a complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{Q}$. $\lceil \cdot \rceil$ represents the ceiling operation. $\text{diag}\{\cdot\}$ defines a diagonal (or block-diagonal) matrix with diagonal elements (or blocks) given by the arguments inside the bracket. $\otimes$, $\odot$, and δ_{ij}
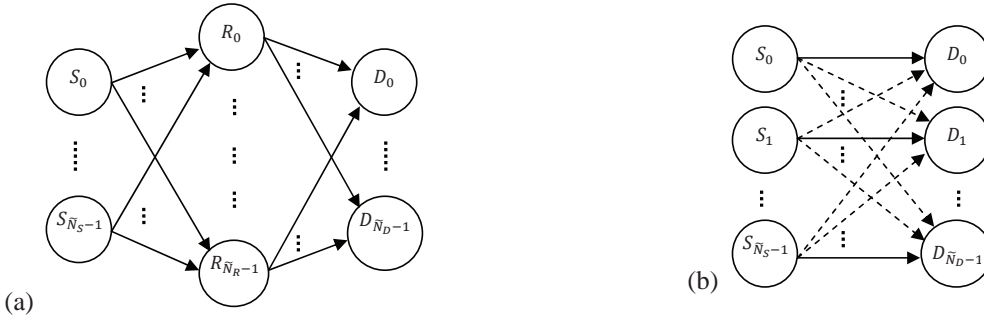
Fig. 1. (a) A general relay system where transmissions from different nodes typically use disjoint resources, (b) An MP-MP system ($\tilde{N}_S = \tilde{N}_D$) where transmissions from different nodes may use the same resource (potential interference links are indicated by dashed lines)

represent the Kronecker product, the Hadamard product, and the Kronecker delta, respectively. $\boldsymbol{F}$ denotes the $N \times N$ unitary discrete Fourier transform (DFT) matrix while $\boldsymbol{F}_{L_i}$ represent the submatrix of $\boldsymbol{F}$ composing of its first L_i columns. The autocovariance matrix of $\boldsymbol{x}$ and cross-covariance matrix of $\boldsymbol{x}$ and $\boldsymbol{y}$ are denoted by $\boldsymbol{Q}_{\boldsymbol{x}}$ and $\boldsymbol{Q}_{\boldsymbol{xy}}$. The cardinality of a set $\mathcal{J}$ is denoted by $|\mathcal{J}|$. vec($\boldsymbol{A}$) represents a column vector obtained by stacking the columns of the matrix $\boldsymbol{A}$. We use $k{:}n$ and $M \times \{k{:}n\}$ to indicate $k, k+1, \cdots, n$ and $\{kM, (k+1)M, \cdots, nM\}$. Tr$\{\boldsymbol{X}\}$ represents the trace of matrix $\boldsymbol{X}$. All tone indexes are assumed to be modulo N.

III. PILOT DESIGNS AND NODE MULTIPLEXING

In this section, first we present pilot designs for direct links. Next, we develop designs of forwarded pilots for cascaded or feedback links. Then we introduce a node multiplexing concept to reduce overhead and delay in CSI acquisition. After that, we propose new pilot designs for channels with different lengths.

A. Pilot Designs for Direct Links

Let us consider the direct links between $\{I_i\}$ and $\{J_m\}$ (say between $\{S_i\}$ and $\{R_m\}$). Define $C_{i,\nu}^{(l)}[k]$ and $B_{i,\nu}^{(l)}[k]$ as the transmitted pilot and data symbol on subcarrier k of OFDM symbol l from $S_{i,\nu}$. Their respective discrete-time signals are denoted by $\{c_{i,\nu}^{(l)}[n]\}$ and $\{b_{i,\nu}^{(l)}[n]\}$. Define circulant matrices $\boldsymbol{\Omega}_{\boldsymbol{c}_{i,\nu}}[l]$ and $\boldsymbol{\Omega}_{\boldsymbol{b}_{i,\nu}}[l]$ of size $N \times L_1$ such that their (m,k)th entries are respectively given by $c_{i,\nu}^{(l)}[m-k \bmod N]$ and $b_{i,\nu}^{(l)}[m-k \bmod N]$ with $m \in \{0{:}N-1\}$ and $k \in \{0{:}L_1-1\}$. Define $\boldsymbol{\Lambda}_{C_{i,\nu}^{(l)}} \triangleq \mathrm{diag}\{C_{i,\nu}^{(l)}[0], C_{i,\nu}^{(l)}[1], \cdots, C_{i,\nu}^{(l)}[N-1]\}$ and $\boldsymbol{\Lambda}_{B_{i,\nu}^{(l)}} \triangleq \mathrm{diag}\{B_{i,\nu}^{(l)}[0], B_{i,\nu}^{(l)}[1], \cdots, B_{i,\nu}^{(l)}[N-1]\}$. Then, we have $\boldsymbol{\Omega}_{\boldsymbol{c}_{i,\nu}}[l] = \sqrt{N}\boldsymbol{F}^H \boldsymbol{\Lambda}_{C_{i,\nu}^{(l)}} \boldsymbol{F}_{L_1}$ and $\boldsymbol{\Omega}_{\boldsymbol{b}_{i,\nu}}[l] = \sqrt{N}\boldsymbol{F}^H \boldsymbol{\Lambda}_{B_{i,\nu}^{(l)}} \boldsymbol{F}_{L_1}$. The received time domain signal vector of V OFDM symbols after cyclic prefix removal at $R_{m,n}$ is

$$\boldsymbol{y}_{R_{m,n}} = \sum_{i=0}^{\tilde{N}_S-1} (\boldsymbol{\Omega}_{\boldsymbol{c}_i} + \boldsymbol{\Omega}_{\boldsymbol{b}_i}) \boldsymbol{h}_{S_i\text{-}R_{m,n}} + \boldsymbol{w}_{R_{m,n}}, \quad (1)$$

where $\boldsymbol{\Omega}_{\boldsymbol{c}_i}$ and $\boldsymbol{\Omega}_{\boldsymbol{b}_i}$ are $NV \times N_{S_i}L_1$ matrices and their (l,ν)th submatrice are respectively given by $\boldsymbol{\Omega}_{\boldsymbol{c}_{i,\nu}}[l]$ and $\boldsymbol{\Omega}_{\boldsymbol{b}_{i,\nu}}[l]$, with $l \in \{0{:}V-1\}$ and $\nu \in \{0{:}N_{S_i}-1\}$,

and $\boldsymbol{h}_{S_i\text{-}R_{m,n}} \triangleq \left[\boldsymbol{h}_{S_{i,0}\text{-}R_{m,n}}^T \boldsymbol{h}_{S_{i,1}\text{-}R_{m,n}}^T \cdots \boldsymbol{h}_{S_{i,N_{S_i}-1}\text{-}R_{m,n}}^T\right]^T$. $\boldsymbol{w}_{R_{m,n}}$ is the noise vector at $R_{m,n}$ with $\mathcal{CN}(\boldsymbol{0}, \sigma_{R_{m,n}}^2 \boldsymbol{I})$. If no pilots (data) are transmitted during lth symbol, $\boldsymbol{\Omega}_{\boldsymbol{c},i,\nu}[l]$ ($\boldsymbol{\Omega}_{\boldsymbol{b},i,\nu}[l]$) will be the zero matrix. The minimum necessary and sufficient number of pilot tones to estimate $\boldsymbol{H}_{S_{i,0}\text{-}R_{m,n}}$ is L_1 since $\boldsymbol{H}_{S_{i,0}\text{-}R_{m,n}}$ is just the Fourier transform of the $L_1 \times 1$ CIR vector $\boldsymbol{h}_{S_{i,0}\text{-}R_{m,n}}$. We build upon the work in [12] as follows. The data and pilot tones are disjoint and the pilot sequences for different transmit antennas are orthogonal and meet the optimal conditions:

$$\boldsymbol{\Omega}_{\boldsymbol{c}_i}^\dagger \boldsymbol{\Omega}_{\boldsymbol{b}_k} = \boldsymbol{0}, \ \forall \ i,k, \ \& \ \boldsymbol{\Omega}_{\boldsymbol{c}_i}^\dagger \boldsymbol{\Omega}_{\boldsymbol{c}_k} = E_{\mathrm{av}}\delta_{ik}\boldsymbol{I} \quad (2)$$

where E_{av} represents the total pilot energy over the V symbols per antenna. The above orthogonality between pilot and data matrices $\boldsymbol{\Omega}_{\boldsymbol{c}_i}$ and $\boldsymbol{\Omega}_{\boldsymbol{b}_k}$ is simply satisfied when the pilot tones are disjoint from the data tones within the same OFDM symbol. Mathematically, for disjoint pilot and data tones, we have $\boldsymbol{\Lambda}_{C_{i,\nu}^{(l)}}^\dagger \boldsymbol{\Lambda}_{B_{k,m}^{(l)}} = \boldsymbol{0}$ and hence $\boldsymbol{\Omega}_{\boldsymbol{c}_{i,\nu}}^\dagger \boldsymbol{\Omega}_{\boldsymbol{b}_{k,m}} = N\boldsymbol{F}_{L_1}^\dagger \boldsymbol{\Lambda}_{C_{i,\nu}^{(l)}}^\dagger \boldsymbol{\Lambda}_{B_{k,m}^{(l)}} \boldsymbol{F}_{L_1} = \boldsymbol{0}$ which yields $\boldsymbol{\Omega}_{\boldsymbol{c}_i}^\dagger \boldsymbol{\Omega}_{\boldsymbol{b}_k} = \boldsymbol{0}$.

First, let us consider pilots to be transmitted from a single node (say S_i) with N_{S_i} antennas. Define $\bar{L}_1 \triangleq 2^{\lceil \log_2(L_1) \rceil}$. The pilot designs from [12] use N_{S_i} sets of cyclically equi-spaced equi-energy $\bar{L}_1$ pilot tones. Define $\mathcal{T} \triangleq N/\bar{L}_1 \times \{0{:}\bar{L}_1 - 1\}$ and $\mathcal{T}_n \triangleq \mathcal{T} + \tau_n$, $n = 0{:}N_{S_i} - 1$, with $\tau_n \in \{0{:}(N/\bar{L}_1) - 1\}$ and $\tau_n \neq \tau_k$ if $n \neq k$. Let $\mathcal{J}_{I_{i,\nu}}^{(l)}$ denote the set of non-zero pilot tone indexes at OFDM symbol l for $I_{i,\nu}$, and its mth element is represented by $\mathcal{J}_{I_{i,\nu}}^{(l)}(m)$. In the following, we introduce a variable $n_{i,\nu}$ which is written as a function of i and ν for a later use but for Section III-A and III-B, $n_{i,\nu} = \nu$. Then, several pilot designs can be expressed as follows:

Frequency division multiplexing (FDM) design: It meets optimality conditions in (2) through frequency disjointness and the pilots are given by $\mathcal{J}_{I_{i,\nu}}^{(0)} = \mathcal{T}_{n_{i,\nu}}$, $C_{i,\nu}^{(l)}[k] = a_{i,\nu,m}$ for $k = \mathcal{J}_{I_{i,\nu}}^{(0)}(m)$ with $l = 0$ and $C_{i,\nu}^{(l)}[k] = 0$ otherwise. Note that $\{a_{i,\nu,m}\}$ in the above and below as well as $\{a_{l,m}\}$ in the following are arbitrary symbols with amplitude $\sqrt{E_{\mathrm{av}}/\bar{L}_1}$ (i.e., to yield equal amplitude pilots and the same total pilot energy per antenna).

Time division multiplexing (TDM) design: It satisfies (2) through time disjointness and the pilots are defined by $\mathcal{J}_{I_{i,\nu}}^{(l)} = \mathcal{T}_{\bar{i}}$, $\bar{i} \in \{0 : N_{S_i} - 1\}$, $l = 0 : N_{S_i} - 1$, $C_{i,\nu}^{(l)}[k] = a_{i,\nu,m}$ for $n_{i,\nu} = l$ with $k = \mathcal{J}_{I_{i,\nu}}^{(l)}(m)$, and $C_{i,\nu}^{(l)}[k] = 0$ otherwise.
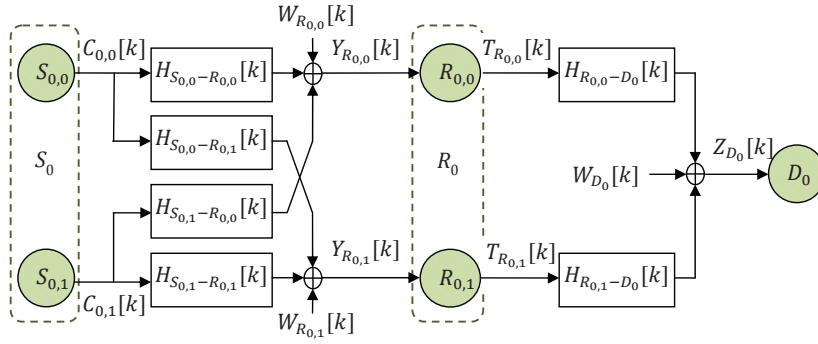
Fig. 2. Signal model of a relay system with one source and one relay, each with two antennas

Code division multiplexing across time (CDM-T) design: It applies coding across time amongst antennas to satisfy (2) and the pilots are defined by $\mathcal{J}_{I_{i,\nu}}^{(l)} = \mathcal{T}_{\bar{i}}$, $\bar{i} \in \{0 : N_{S_i} - 1\}$, $l = 0 : N_{S_i} - 1$, $C_{i,\nu}^{(l)}[k] = (a_{i,\nu,m}/\sqrt{N_{S_i}})e^{j2\pi n_{i,\nu}l/N_{S_i}}$ or $(a_{l,m}/\sqrt{N_{S_i}})\,e^{j2\pi n_{i,\nu}l/N_{S_i}}$ for $k = \mathcal{J}_{I_{i,\nu}}^{(l)}(m)$.

Code division multiplexing across frequency (CDM-F) design: It applies coding across frequency amongst antennas to satisfy (2) and the pilots are given by $\mathcal{J}_{I_{i,\nu}}^{(0)} = \cup_{l=0}^{N_{S_i}-1}\mathcal{T}_l$ and $C_{i,\nu}^{(0)}[k] = (a_{l,m}/\sqrt{N_{S_i}})e^{j2\pi n_{i,\nu}l/N_{S_i}}$ for $k = \mathcal{T}_l(m)$.

The proofs of the above designs satisfying the optimal conditions in (2) are referred to [12]. Note that the reason for presenting several designs is to provide solutions for various scenarios or system constraints. For example, if more pilot tones ($> N$) are needed to support more antennas or users, the pilot designs can be extended across time using TDM or CDM-T. If stringent peak power constraint per tone per antenna is imposed (due to RF device and spectrum emission requirement), CDM-F and CDM-T would be better, but if device power saving via sleep cycle is preferred, TDM would be better. If reducing CSI acquisition delay is desired, FDM or CDM-F could be used. If there are stringent constraints on total power per OFDM symbol as well as per tone peak power per antenna (due to spectrum emission requirement and large path loss), CDM-T would suit the situation.

B. Reduced Overhead Pilot Designs

All of the above designs use a total of $N_{S_i}\bar{L}_1$ pilot tones. If L_1 is not a power of 2 (i.e., $L_1 < \bar{L}_1$) and minimum pilot overhead is desired, we propose to change the above designs by replacing each pilot set of cyclically equi-spaced equi-energy $\bar{L}_1$ pilot tones with that of approximately cyclically equi-spaced equi-energy L_1 pilot tones while maintaining disjointness among the pilot sets. For *example*, consider three channels with $L_1 = 5$ taps each and $N = 64$; the minimum overhead pilot design can use 15 pilot tones with $\mathcal{T}_0 = \{0, 13, 26, 39, 52\}$ and $\mathcal{T}_i = \mathcal{T}_0 + i$ for $i = 1, 2$, while the existing designs use 24 pilot tones (i.e., $\mathcal{T}_0 = 8 \times \{0:7\}, \mathcal{T}_i = \mathcal{T}_0 + i$). Apart from $\{\mathcal{T}_n\}$, the pilot designs remain the same as those presented above, and hence they are not repeated here. A slight degradation in estimation performance may incur for the minimum overhead pilot design but it is insignificant as will be shown in Section VII.

We can also develop a new CDM-F design which maintains estimation optimality while requiring less overhead than the existing designs. It uses cyclically equi-spaced equi-energy $L_1' = 2^{\lceil \log_2(N_{S_i} L_1) \rceil}$ pilot tones as

$$\mathcal{J}_{S_i,\nu} = N/L_1' \times \{0 : (L_1' - 1)\} + \tau, \tag{3}$$

$$C_{i,\nu}^{(0)}[k] = \frac{e^{j2\pi k n_{i,\nu} L_1/N}}{\sqrt{N_{S_i}}} D_i[k], \ \nu = 0 : N_{S_i} - 1 \tag{4}$$

$$D_i[k] = \begin{cases} a_{i,m}, & k = \mathcal{J}_{S_i,\nu}(m) \\ 0, & \text{else,} \end{cases} \tag{5}$$

where $\tau \in \{0 : (N/L_1') - 1\}$ and $\{a_{i,m}\}$ are arbitrary symbols with amplitude $\sqrt{E_{\text{av}}/L_1'}$. With regard to overhead, for the parameters in the above example, this design requires 16 pilot tones (with $\mathcal{J}_{S_i,\nu} = 4 \times \{0:15\}, \forall \nu$), thus 8 tones less than the existing designs. The proof for the optimality of the above design is provided in Appendix A.

C. Designs for Forwarded Pilots

As an example of forwarded pilot scenario, a relay node R_m receives pilots from source S_i on tones $\mathcal{J}_{S_i} = \{\mathcal{T}_\nu : \nu = 0 : N_{S_i} - 1\}$ and forwards them to a destination D_l. A signal model is illustrated in Fig. 2. The received pilot on tone $k \in \mathcal{J}_{S_i}$ at $R_{m,n}$ is given by

$$Y_{R_{m,n}}[k] = \sum_{\nu=0}^{N_{S_i}-1} H_{S_{i,\nu}\text{-}R_{m,n}}[k]C_{S_{i,\nu}}[k] + W_{R_{m,n}}[k], \tag{6}$$

where $\{W_{R_{m,n}}[k]\}$ are independent and identically distributed with $\mathcal{CN}(0, \sigma_{R_m}^2)$. In order for D_l to estimate individual channels $\{\boldsymbol{h}_{S_{i,\nu}\text{-}R_{m,n}}\}$, R_m constructs $\{T_{R_{m,n}}[k]\}$ based on $\{Y_{R_{m,n}}[k]\}$ and forwards them to D_l. First, let us consider R_m with a single antenna. Then in general the received source pilots on $\mathcal{J}_{S_i}$ can be forwarded on any $|\mathcal{J}_{S_i}|$ tones, thus there are $N!/(|\mathcal{J}_{S_i}|!(N - |\mathcal{J}_{S_i}|)!)$ ways. We explore the choice of tones for pilot forwarding by minimizing the mean-square error (MSE) in estimating $\boldsymbol{h}_{S_i\text{-}R_m}$ at the destination (See Appendix B for details). An exact closed-form analytical solution is intractable but the analysis in Appendix B and numerical simulation reults in Section VII-C show that maintaining the relative positions as in individual $\{\mathcal{T}_\nu\}$ (in simplest form, using $\tau + \mathcal{J}_{S_i}$) gives the best result. Using this design for a single antenna at R_m, we propose several designs for multiple antennas at R_m in the following, where the remaining task is
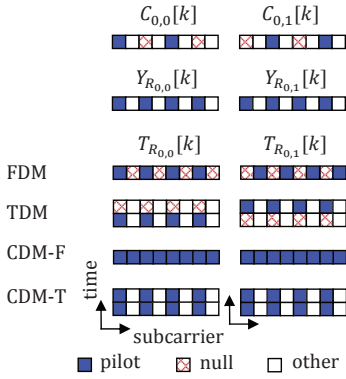
Fig. 3. Design examples of forwarded MIMO pilots for one source and one relay with two antennas each (OFDM with 8 subcarriers in two-tap channels. α is a fixed amplification factor at the relay.)

maintaining orthogonality among forwarded pilots of different relay antennas. The rationale for variety of designs is the same as that mentioned in Section III-A.

For better illustration, we present some design examples in Fig. 3 which can facilitate easy understanding of the general design descriptions to be given below where α is a fixed amplification factor at the relay.

FDM design: $R_{m,n}$ forwards its received source pilots on a set of shifted tones $\mathcal{J}_{R_{m,n}}$ as

$$T_{R_{m,n}}^{(0)}[k + \tau_{m,n}] = \begin{cases} \alpha Y_{R_{m,n}}[k], & k \in \mathcal{J}_{S_i} \\ 0, & \text{else} \end{cases} \quad (7)$$

$$\mathcal{J}_{R_{m,n}} \triangleq \mathcal{J}_{S_i} + \tau_{m,n}, \quad \bigcap_{n=0}^{N_{R_m}-1} \mathcal{J}_{R_{m,n}} = \emptyset. \quad (8)$$

TDM design: Different antennas of R_m transmit their received source pilots in different symbols but on the same tones $\mathcal{J}_{R_{m,n}} = \mathcal{J}_{S_i} + \tau$ (any τ) as

$$T_{R_{m,n}}^{(p)}[k + \tau] = \begin{cases} \alpha Y_{R_{m,n}}[k], & k \in \mathcal{J}_{S_i}, \ n = p \\ 0, & \text{else.} \end{cases} \quad (9)$$

CDM-T design: $R_{m,n}$ transmits its received source pilots on $\mathcal{J}_{R_{m,n}} = \mathcal{J}_{S_i} + \tau$ (any τ) over N_{R_m} symbols as

$$T_{R_{m,n}}^{(p)}[k + \tau] = \begin{cases} \dfrac{\alpha e^{j2\pi pn/N_{R_m}}}{\sqrt{N_{R_m}}} Y_{R_{m,n}}[k], & k \in \mathcal{J}_{S_i} \\ 0, & \text{else.} \end{cases} \quad (10)$$

CDM-F design: $R_{m,n}$ applies a length-N_{R_m} antenna-specific orthonormal spreading to each received source pilot and transmits them on the same N_{R_m} disjoint sets of tones as

$$T_{R_{m,n}}^{(0)}[k + \tau_\nu] = \begin{cases} \dfrac{\alpha e^{\frac{j2\pi\nu n}{N_{R_m}}}}{\sqrt{N_{R_m}}} Y_{R_{m,n}}[k], & k \in \mathcal{J}_{S_i}, \\ 0, & \text{else} \end{cases} \quad (11)$$

$$\{\tau_\nu\} = \{\tau : \tau + N_{R_m} - 1\}, \quad \bigcap_{\nu=0}^{N_{R_m}-1} \{\mathcal{J}_{S_i} + \tau_\nu\} = \emptyset. \quad (12)$$

Note that for TDM or CDM-T source pilots, the above pilot forwarding schemes can be applied to each OFDM symbol of received source pilots. Alternatively, if all $N_{S_i}L_1$ received source pilots can fit within one OFDM symbol, they can be treated as if they were received in one OFDM symbol when applying the above pilot forwarding schemes.

The forwarded pilot sets in FDM, TDM and CDM-T forwarding schemes strictly maintain orthogonality and they can be decoupled at the receiver without knowing the forwarded channel gains. Such is not the case for CDM-F forwarding scheme since the channel gains on tones $\{k + \tau_0, \cdots, k + \tau_{N_{R_m}-1}\}$ are not exactly the same (although highly correlated) and their knowledge is required in decoupling the pilot sets. Consequently, the random estimation errors of the forwarded channels will affect the orthogonality of the CDM-F pilot sets.

In the above designs, each antenna forwards its received source pilots and hence D_l also needs knowledge of $\{\boldsymbol{h}_{R_{m,n}\text{-}D_l}\}$ to estimate $\{\boldsymbol{h}_{S_{i,\nu}\text{-}R_{m,n}}\}$. If such knowledge is not available yet, R_m can insert its own pilots (designed according to one of the direct link pilot designs and their locations are disjoint from the forwarded pilots over frequency or time) and transmit to D_l. Those received source pilots at R_m can also be collectively forwarded by a single predefined antenna $R_{m,\nu}$ if D_l does not require $\{\boldsymbol{h}_{R_{m,n}\text{-}D_l} : \forall n \neq \nu\}$. Note that different from [16], our approach does not need to reserve (waste) null tones at the source for the relay-inserted pilots, thus saving pilot overhead.

D. Node Multiplexing

CSI acquisitions in relay and MP-MP systems involve transmissions of pilots or quantized CSI bits from multiple nodes. Instead of transmitting from one node after another in a sequential manner, we introduce a node multiplexing concept which allows multiple nodes to transmit simultaneously if certain conditions are met. The motivation is to save overhead and delay in the CSI acquisition phase by exploiting already-accomplished synchronization and node sequence assignment done during the initial link establishment. The nodes involved in the node multiplexing need not be of the same type, e.g., transmissions from the sources and destinations to the relays can be node-multiplexed together.

The conditions for the node multiplexing are that pilots from different nodes should satisfy the pilot conditions described in previous sections while any modulation symbols carrying quantized CSI bits should be transmitted on disjoint tones in a similar manner without requiring equal power and cyclically equal spacing condition of the pilots. This is to satisfy (2) except that $\boldsymbol{\Omega_c}$ and $\boldsymbol{\Omega_b}$ contain all nodes involved in the
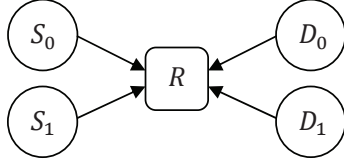
Transmission Sequence:

(a) Conventional/Sequential transmission

    1) $S_0 \to R : \{C_{S_0}[k]\}$,   2) $S_1 \to R : \{C_{S_1}[k]\}$

    3) $D_0 \to R : \{C_{D_0}[k]\}$,   4) $D_1 \to R : \{C_{D_1}[k]\}$

(b) Node multiplexing option 1: TDM between $\{C_S[k]\}$ & $\{C_D[k]\}$

    1) $S_i \to R : \{C_{S_i}[k]\}$; FDM or CDM-F between $\{C_{S_0}[k]\}$ & $\{C_{S_1}[k]\}$

    2) $D_i \to R : \{C_{D_i}[k]\}$; FDM or CDM-F between $\{C_{D_0}[k]\}$ & $\{C_{D_1}[k]\}$

(c) Node multiplexing option 2: CDM-T between $\{C_S[k]\}$ & $\{C_D[k]\}$ with $\mathcal{J}_S = \mathcal{J}_D$

    1) $S_i \to R : \{C_{S_i}[k]\}$; $D_i \to R : \{C_{D_i}[k]\}$; FDM or CDM-F between $\{C_{S_0}[k]\}$ & $\{C_{S_1}[k]\}$;

    2) $S_i \to R : \{C_{S_i}[k]\}$; $D_i \to R : \{-C_{D_i}[k]\}$; FDM or CDM-F between $\{C_{D_0}[k]\}$ & $\{C_{D_1}[k]\}$

(d) Node multiplexing option 3: FTDM between $\{C_S[k]\}$ & $\{C_D[k]\}$

    1) $S_{i,0} \to R : \{C_{S_{i,0}}[k]\}$; $D_{i,0} \to R : \{C_{D_{i,0}}[k]\}$; (FDM)

    2) $S_{i,1} \to R : \{C_{S_{i,1}}[k]\}$; $D_{i,1} \to R : \{C_{D_{i,1}}[k]\}$; (FDM)

Fig. 4. Illustration of node multiplexing where two sources and two destinations transmit pilots to the relays. Each node has two antennas, each channel has 4 taps (hence 4 pilot tones per channel), and OFDM has $N = 16$ subcarriers

node multiplexing. The pilot designs are given by those in Section III-A and III-B except N_{S_i} is replaced with $N_{S_{\mathrm{TOT}}}$ and $n_{i,\nu}$ is defined as $\nu + \sum_{m=0}^{i-1} N_{S_m}$. $\boldsymbol{c}$ and $\boldsymbol{b}$ are disjoint over the frequency and time domain and each node does not need to know pilots or data of other nodes. The pilots and protocols are pre-designed for a few number of system condition sets (number of nodes, number of antennas at each node, and nominal channel lengths involved in the CSI acquisition), and the global information about the system condition set (a few bits) is broadcast to all nodes while individual pilot assignment and role in the protocol are conveyed to each node via a sequence number in the initial link establishment. In the same manner, node multiplexing can be applied for the forwarded pilots of several nodes.

In brief, the direct or forwarded pilot designs under node multiplexing can take one of the two approaches: i) considering all nodes collectively as a single virtual node (the designs presented in the previous sections still hold but with a larger number of antennas) and ii) designing pilots for individual nodes and then applying FDM, TDM, frequency and time division multiplexing (FTDM)[1], or CDM-T across pilots of different nodes. For reducing delay, the design should use as small number of OFDM symbols as possible for pilot transmissions.

As an example, we will illustrate the FTDM node multiplexing. Suppose K nodes, $\{I_i\}$, are involved, each OFDM symbol has M sets of cyclically equi-spaced equi-energy L_1 pilots, and node i needs M_{I_i} such pilot sets. Then the FTDM design uses $\lceil \sum_{i=0}^{K-1} M_{I_i}/M \rceil$ OFDM symbols and all of the pilot sets of K nodes are disjoint over the frequency or time. For example, consider $K = 3$ nodes with $M_{I_0} = 3$, $M_{I_1} = 3$, $M_{I_2} = 2$, and $M = 4$. With FTDM, I_0 and I_1 will use the first 3 pilot sets from the first and second symbol, respectively, while I_2 will exploit the 4-th pilot set from the first and second symbol. The overhead and delay is only two OFDM symbols instead of three incurred in typical sequential transmission.

Fig. 4 illustrates that several options of node multiplexing

are possible. There, all three multiplexing options finish pilot transmissions in two OFDM symbols while the conventional sequential transmission will take 4 OFDM symbols.

E. Pilot Designs for MIMO Channels with Different Lengths

MIMO channels of a single node to single node link are commonly assumed to have the same channel length. However, for channels corresponding to different node pairs, e.g., $\boldsymbol{h}_{S_i\text{-}R_m}$ versus $\boldsymbol{h}_{D_l\text{-}R_m}$, or even $\boldsymbol{h}_{S_i\text{-}R_m}$ versus $\boldsymbol{h}_{S_\nu\text{-}R_m}$ for $i \neq \nu$, their channel lengths can be different. This characteristic is inherent in the proposed node multiplexing framework and it creates a new pilot design problem for MIMO channels with different channel lengths. In fact, even for a single node to single node link, if distributed antennas (e.g., remote radio heads in LTE) are used, the MIMO channels may have different lengths. To discuss further, suppose there are K channels with channel lengths $\{L_i\}$, and we like to design their pilots $\{C_i[k]\}$ which minimize the sum MSE of these channel estimates. For the case with $L_{\text{total}} \triangleq \sum_{i=0}^{K-1} L_i > N$, we can divide these channels into TDM or CDM-T groups such that each group has a sum of channel lengths being less than or equal to N. Thus, without loss of generality, in the following we just consider $L_{\text{total}} \leq N$.

We use the same pilot optimization criterion as in [12] which minimizes the total MSE of least-square or maximum likelihood channel estimation as

$$\arg\min_{\{C_i[k]\}} \text{Tr}[\sigma_{J_\nu}^2 (\boldsymbol{\Omega}_{\boldsymbol{c}}^\dagger \boldsymbol{\Omega}_{\boldsymbol{c}})^{-1}], \quad \text{s.t.} \ \sum_{i=0}^{K-1} E_i = E_{\text{total}} \quad (13)$$

where $\boldsymbol{\Omega}_{\boldsymbol{c}} = [\boldsymbol{\Omega}_{\boldsymbol{c}_0}, \boldsymbol{\Omega}_{\boldsymbol{c}_1}, \cdots, \boldsymbol{\Omega}_{\boldsymbol{c}_{K-1}}]$, $\boldsymbol{\Omega}_{\boldsymbol{c}_i}$ is of size $N \times L_i$, $E_i \triangleq \sum_{k=0}^{N-1} |C_i[k]|^2$, and $\sigma_{J_\nu}^2$ is the noise variance at the estimating node J_ν. A difference from [12] is that the constraint is no longer equivalent to setting $\text{Tr}[\boldsymbol{\Omega}_{\boldsymbol{c}}^\dagger \boldsymbol{\Omega}_{\boldsymbol{c}}]$ to a constant due to different $\{L_i\}$. An optimal design is given by (see Appendix C for the proof)

$$\boldsymbol{\Omega}_{\boldsymbol{c}_i}^\dagger \boldsymbol{\Omega}_{\boldsymbol{c}_i} = E_i \boldsymbol{I}_{L_i}, \quad \boldsymbol{\Omega}_{\boldsymbol{c}_i}^\dagger \boldsymbol{\Omega}_{\boldsymbol{c}_k} = \boldsymbol{0}_{L_i \times L_k}, \forall i \neq k, \quad (14)$$

$$E_i = \frac{\sqrt{L_i}}{\sum_{l=0}^{K-1} \sqrt{L_l}} E_{\text{total}}. \quad (15)$$

---

[1]It can also be applied without node multiplexing but was skipped in the previous sections for conciseness.

Compared with the condition $\mathbf{\Omega}_c^\dagger \mathbf{\Omega}_c = E_{\mathrm{av}} \mathbf{I}$ of [12], the differences are the unequal energy allocation in (15) and different sizes $\{\mathbf{I}_{L_i}\}$ in (14). Our previous design concepts can be applied with appropriate modification as will be described in the following. The MSE for the channel i is $\sigma_{J_\nu}^2 \mathrm{Tr}[(\mathbf{\Omega}_{c_i}^\dagger \mathbf{\Omega}_{c_i})^{-1}] = \sigma_{J_\nu}^2 L_i / E_i$.

We can apply an existing pilot design from [12] using $\bar{L} \triangleq 2^{\lceil \log_2(L_{\max}) \rceil}$ pilot tones per channel where $L_{\max} \triangleq \max_i\{L_i\}$ but its estimation performance will not be optimal. An easy fix for estimation optimality is to adjust $\{E_i\}$ according to (15). This solution maintains estimation optimality but requires a total of $K\bar{L}$ pilot tones which are larger than the minimum required amount of L_{total} tones. If minimum pilot overhead is desired, the FDM with approximate equi-spacing from Section III-B can be applied where L_i pilot tones are used for channel i and $\{E_i\}$ are set as in (15); its estimation performance may slightly deviate from the optimality.

Alternatively, similar to the new CDM-F design in Section III-B, we can develop a new design which maintains optimal estimation performance while reducing the pilot overhead of the existing design. This CDM-F design uses a total of $\tilde{L}_{\mathrm{total}} = 2^{\lceil \log_2 L_{\mathrm{total}} \rceil}$ pilot tones with cyclically equal spacing of $\tilde{M} \triangleq N/\tilde{L}_{\mathrm{total}}$ tones, i.e., at tone index set $\mathcal{J} = \tilde{M} \times \{0{:}\tilde{L}_{\mathrm{total}} - 1\}$ or any of its shifted version, and the pilots are given by

$$C_0[k] = \begin{cases} a_l, & k = \mathcal{J}(l), \ |a_l|^2 = E_0/\tilde{L}_{\mathrm{total}}, \\ 0, & \text{else} \end{cases} \quad (16)$$

$$C_i[k] = \sqrt{\frac{E_i}{E_0}} C_0[k] e^{\frac{j2\pi\tau_i k}{N}}, \ \tau_i = \tau_{i-1} + L_i + \Delta_i \quad (17)$$

$$\text{or } C_i[k] = \sqrt{\frac{E_i}{E_0}} C_0[k] e^{\frac{-j2\pi\tau_i k}{N}}, \ \tau_i = \tau_{i-1} + L_{i-1} + \Delta_i \quad (18)$$

where $\{\Delta_i\}$ are any non-negative integers satisfying $0 \leq \sum_{i=1}^{K-1} \Delta_i \leq \tilde{L}_{\mathrm{total}} - L_{\mathrm{total}}$, $\tau_0 = 0$, and $\{E_i\}$ are given by (15). The optimality of this design is proved in Appendix C.

With the above optimal design, the MSE_i for the channel i is $\sigma_{J_\nu}^2 \sqrt{L_i} \sum_{k=0}^{K-1} \sqrt{L_k}/E_{\mathrm{total}}$. For the existing design (i.e., with constant E_i), MSE_i is $\sigma_{J_\nu}^2 L_i K/E_{\mathrm{total}}$. Both designs yield different MSE values for channels with different lengths. If we want the same MSE_i for different channels, we need to set $E_i = \frac{L_i}{\sum_{k=0}^{K-1} L_k} E_{\mathrm{total}}$ which yields $\mathrm{MSE}_i = \sigma_{J_\nu}^2 \sum_{k=0}^{K-1} L_k/E_{\mathrm{total}}$.

IV. CHANNEL ESTIMATION

We will start with obtaining channel estimates for directly connected links, and then using them we will develop a channel estimator for indirectly connected links (e.g., estimating the first hop channel at the ending node of the second hop). Suppose R_m estimates $\mathbf{h}_{S_i\text{-}R_{m,n}}$ from $\mathbf{y}_{R_{m,n}}$ received at its antenna m (see (1)). Then the least-square or maximum likelihood estimates of $\{\mathbf{h}_{S_{i,\nu}\text{-}R_{m,n}} : \forall i\}$ can be obtained separately or jointly as

$$\hat{\mathbf{h}}_{S_{i,\nu}\text{-}R_{m,n}}^{\mathrm{LS}} = (\mathbf{\Omega}_{c_{i,\nu}}^\dagger \mathbf{\Omega}_{c_{i,\nu}})^{-1} \mathbf{\Omega}_{c_{i,\nu}}^\dagger \mathbf{y}_{R_{m,n}} \quad (19)$$

$$\hat{\mathbf{h}}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}} = (\mathbf{\Omega}_{c_i}^\dagger \mathbf{\Omega}_{c_i})^{-1} \mathbf{\Omega}_{c_i}^\dagger \mathbf{y}_{R_{m,n}}. \quad (20)$$

The error vectors $\Delta\mathbf{h}_{S_{i,\nu}\text{-}R_{m,n}}^{\mathrm{LS}} \triangleq \hat{\mathbf{h}}_{S_{i,\nu}\text{-}R_{m,n}}^{\mathrm{LS}} - \mathbf{h}_{S_{i,\nu}\text{-}R_{m,n}}$ and $\Delta\mathbf{h}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}} \triangleq \hat{\mathbf{h}}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}} - \mathbf{h}_{S_i\text{-}R_{m,n}}$ have $\mathbf{Q}_{\Delta\mathbf{h}_{S_{i,\nu}\text{-}R_{m,n}}^{\mathrm{LS}}} = \sigma_{R_m}^2 (\mathbf{\Omega}_{c_{i,\nu}}^\dagger \mathbf{\Omega}_{c_{i,\nu}})^{-1}$ and $\mathbf{Q}_{\Delta\mathbf{h}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}}} = \sigma_{R_m}^2 (\mathbf{\Omega}_{c_i}^\dagger \mathbf{\Omega}_{c_i})^{-1}$. If $\mathbf{Q}_{\mathbf{h}_{S_i\text{-}R_{m,n}}}$ and $\sigma_{R_m}^2$ are known, an MMSE estimator [35] can be used as

$$\hat{\mathbf{h}}_{S_i\text{-}R_{m,n}}^{\mathrm{MMSE}} = \mathbf{G}_{S_i\text{-}R_{m,n}} \hat{\mathbf{h}}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}} \quad (21)$$

where $\mathbf{G}_{S_i\text{-}R_{m,n}} \triangleq \mathbf{Q}_{\mathbf{h}_{S_i\text{-}R_{m,n}}} (\mathbf{Q}_{\mathbf{h}_{S_i\text{-}R_{m,n}}} + \mathbf{Q}_{\Delta\hat{\mathbf{h}}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}}})^{-1}$. Its error covariance matrix is $\mathbf{Q}_{\Delta\mathbf{h}_{S_i\text{-}R_{m,n}}^{\mathrm{MMSE}}} = (\mathbf{Q}_{\mathbf{h}_{S_i\text{-}R_{m,n}}}^{-1} + \mathbf{Q}_{\Delta\mathbf{h}_{S_i\text{-}R_{m,n}}^{\mathrm{LS}}}^{-1})^{-1}$. Similarly, D_l can obtain $\hat{\mathbf{h}}_{R_m\text{-}D_{l,k}}^{\mathrm{LS}}$ or $\hat{\mathbf{h}}_{R_m\text{-}D_{l,k}}^{\mathrm{MMSE}}$ using pilots transmitted from R_m.

Next, for indirectly connected links, suppose D_l estimates $\mathbf{h}_{S_i\text{-}R_m}$ using $\hat{\mathbf{h}}_{R_m\text{-}D_{l,k}}$ and the relay-forwarded source pilots. Suppose S_i transmits $\tilde{L}_{S_i} = N_{S_i} L_1$ pilots on $\mathcal{J}_{S_i}$ (corresponding time-domain matrix is $\mathbf{\Omega}_{c_i}$) and R_m uses $N_{R_m} \tilde{L}_{S_i}$ tones (counted over frequency and time) to forward source pilots to D_l. Then, for FDM or CDM-F source pilots, $\mathcal{J}_{S_i}$ consists of N_{S_i} disjoint index sets and $\mathbf{\Omega}_{c_i}$ is an $N \times N_{S_i} L_1$ matrix. But for TDM or CDM-T source pilots, $\mathcal{J}_{S_i}$ is composed of N_{S_i} duplicates of an index set (over N_{S_i} symbols) and $\mathbf{\Omega}_{c_i}$ is an $NN_{S_i} \times N_{S_i} L_1$ matrix. Define $\mathbf{\Phi}_{\mathcal{J}}$ for an index set $\mathcal{J}$ to be a $|\mathcal{J}| \times N$ matrix whose kth row is the $\mathcal{J}(k)$-th row of $\mathbf{I}_N$. Thus, left-multiplying with $\mathbf{\Phi}_{\mathcal{J}}$ to a vector or matrix returns a sub vector or matrix whose elements or rows are defined by $\mathcal{J}$. Define $\bar{\mathbf{\Omega}}_{c_i} \triangleq \mathbf{I}_{N_{R_m}} \otimes (\mathbf{\Phi}_{\mathcal{J}_{S_i}} \mathbf{F} \mathbf{\Omega}_{c_i})$ for FDM or CDM-F source pilots and $\bar{\mathbf{\Omega}}_{c_i} \triangleq \mathbf{I}_{N_{R_m}} \otimes (\mathbf{\Phi}_{\mathcal{J}_{S_i}} (\mathbf{1}_{N_{S_i}} \otimes \mathbf{F}) \mathbf{\Omega}_{c_i})$ for TDM or CDM-T source pilots. Then, the total received source pilots in frequency domain at R_m are given by

$$\mathbf{Y} = \bar{\mathbf{\Omega}}_{c_i} \mathbf{h}_{S_i\text{-}R_m} + \mathbf{W}_{R_m} \quad (22)$$

where $\mathbf{W}_{R_m}$ is $\mathcal{CN}(\mathbf{0}, \mathbf{Q}_{\mathbf{W}_{R_m}})$. The channel gains on the forwarded pilot tones can be defined by $\tilde{\mathbf{H}}_{R_{m,n}\text{-}D_{l,k}} = \mathbf{\Phi}_{\mathcal{J}_n} \mathbf{H}_{R_{m,n}\text{-}D_{l,k}}$ where $\mathbf{\Phi}_{\mathcal{J}_n}$ is equal to $\mathbf{\Phi}_{\mathcal{J}_{S_i}}$ for TDM and CDM-T, $\mathbf{\Phi}_{\mathcal{J}_{S_i}+\tau_n}$ for FDM, and $[\mathbf{\Phi}_{\mathcal{J}_{S_i}}^T, \mathbf{\Phi}_{\mathcal{J}_{S_i}+\tau_1}^T, \cdots, \mathbf{\Phi}_{\mathcal{J}_{S_i}+\tau_{N_{R_p}-1}}^T]^T$ for CDM-F. Define $\mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,n}\text{-}D_{l,k}}} \triangleq \mathrm{diag}\{\tilde{\mathbf{H}}_{R_{m,n}\text{-}D_{l,k}}\}$. Then, the frequency-domain forwarded source pilots received at $D_{l,k}$ can be collected as a $N_{R_m} \tilde{L}_{S_i} \times 1$ vector as

$$\mathbf{Z}_{D_{l,k}} = \bar{\mathbf{\Lambda}}_{\mathbf{H}_{R_m\text{-}D_{l,k}}} \mathbf{P} \left(\bar{\mathbf{\Omega}}_{c_i} \mathbf{h}_{S_i\text{-}R_m} + \mathbf{W}_{R_m} \right) + \mathbf{W}_{D_{l,k}} \quad (23)$$

where the noise $\mathbf{W}_{D_{l,k}}$ is $\mathcal{CN}(\mathbf{0}, \mathbf{Q}_{\mathbf{W}_{D_l}})$, and $\mathbf{P}$ and $\bar{\mathbf{\Lambda}}_{\mathbf{H}_{R_m\text{-}D_{l,k}}}$ are defined as follows.

$\bar{\mathbf{\Lambda}}_{\mathbf{H}_{R_m\text{-}D_{l,k}}}$ is given by $\mathrm{diag}\{\mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,0}\text{-}D_{l,k}}}, \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,1}\text{-}D_{l,k}}}, \cdots, \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,N_{R_m}-1}\text{-}D_{l,k}}}\}$ for TDM and FDM, $[\mathbf{I}_{N_{R_m}} \otimes \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,0}\text{-}D_{l,k}}}, \mathbf{I}_{N_{R_m}} \otimes \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,1}\text{-}D_{l,k}}}, \cdots, \mathbf{I}_{N_{R_m}} \otimes \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,N_{R_m}-1}\text{-}D_{l,k}}}]$ for CDM-T, and $[\mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,0}\text{-}D_{l,k}}}, \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,1}\text{-}D_{l,k}}}, \cdots, \mathbf{\Lambda}_{\tilde{\mathbf{H}}_{R_{m,N_{R_m}-1}\text{-}D_{l,k}}}]$ for CDM-F. Let $\mathbf{v}_i$ represent the ith column of the unitary DFT matrix of size $N_{R_m} \times N_{R_m}$. Then, $\mathbf{P}$ is equal to $\mathrm{diag}\{\alpha_0, \alpha_1, \cdots, \alpha_{N_{R_m}-1}\} \otimes \mathbf{I}_{\tilde{L}_{S_i}}$ for TDM and FDM, and $\mathrm{diag}\{\alpha_0 \mathbf{v}_0, \alpha_1 \mathbf{v}_1, \cdots, \alpha_{N_{R_m}-1} \mathbf{v}_{N_{R_m}-1}\} \otimes \mathbf{I}_{\tilde{L}_{S_i}}$ for

CDM-T and CDM-F. $\{\alpha_n\}$ represent fixed constants to adjust SNRs and they could be different if R_m represents a virtual node consisting of several nodes with different SNRs.

With perfect knowledge of $\tilde{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}$, the MMSE estimate of $\boldsymbol{h}_{S_i\text{-}R_m}$ is obtained as

$$\hat{\boldsymbol{h}}_{S_i\text{-}R_m}^{\text{MMSE}} = \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}} \boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}}^{-1} \boldsymbol{Z}_{D_{l,k}} \qquad (24)$$

where

$$\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}} = \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}}^{\dagger} \boldsymbol{P}^{\dagger} \bar{\boldsymbol{\Lambda}}_{\boldsymbol{H}_{R_m\text{-}D_{l,k}}}^{\dagger}, \qquad (25)$$

$$\boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}} = \bar{\boldsymbol{\Lambda}}_{\boldsymbol{H}_{R_m\text{-}D_{l,k}}} \boldsymbol{P} \boldsymbol{Q}_{\boldsymbol{Y}} \boldsymbol{P}^{\dagger} \bar{\boldsymbol{\Lambda}}_{\boldsymbol{H}_{R_m\text{-}D_{l,k}}}^{\dagger} + \boldsymbol{Q}_{\boldsymbol{W}_{D_l}}, \quad (26)$$

with $\boldsymbol{Q}_{\boldsymbol{Y}} = \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}} \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}}^{\dagger} + \boldsymbol{Q}_{\boldsymbol{W}_{R_m}}$. In practice, $D_{l,k}$ has $\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}$ instead of $\boldsymbol{H}_{R_m\text{-}D_{l,k}}$, and existing approaches (e.g., [16]) uses $[\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}}]_{\text{ref}}$ and $[\boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}}]_{\text{ref}}$ which replace $\bar{\boldsymbol{\Lambda}}_{\boldsymbol{H}_{R_m\text{-}D_{l,k}}}$ in (25) and (26), respectively, with $\bar{\boldsymbol{\Lambda}}_{\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}}$.

We propose an alternative approach, where the effect of the error in $\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}$ denoted by $\boldsymbol{\Delta H}_{R_m\text{-}D_{l,k}}$ is incorporated by replacing $\bar{\boldsymbol{\Lambda}}_{\boldsymbol{H}_{R_m\text{-}D_{l,k}}}$ in (23) with $\bar{\boldsymbol{\Lambda}}_{\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}} - \bar{\boldsymbol{\Lambda}}_{\boldsymbol{\Delta H}_{R_m\text{-}D_{l,k}}}$, and obtain the estimator conditioned on $\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}$ which takes the same form as in (24) but with

$$\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}} = [\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}}]_{\text{ref}} \qquad (27)$$

$$\boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}} = [\boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}}]_{\text{ref}} + \boldsymbol{\Delta Q}_{\boldsymbol{Z}_{D_{l,k}}} \qquad (28)$$

$$\boldsymbol{\Delta Q}_{\boldsymbol{Z}_{D_{l,k}}} \triangleq E[\bar{\boldsymbol{\Lambda}}_{\boldsymbol{\Delta H}_{R_m\text{-}D_{l,k}}} \boldsymbol{P} \boldsymbol{Q}_{\boldsymbol{Y}} \boldsymbol{P}^{\dagger} \bar{\boldsymbol{\Lambda}}_{\boldsymbol{\Delta H}_{R_m\text{-}D_{l,k}}}^{\dagger}]$$

$$= \begin{cases} \boldsymbol{Q}_{\boldsymbol{\Delta \bar{H}}_{R_m\text{-}D_{l,k}}} \odot (\boldsymbol{P} \boldsymbol{Q}_{\boldsymbol{Y}} \boldsymbol{P}^{\dagger}), & \text{TDM, FDM} \\ \boldsymbol{U}^{\dagger} [\boldsymbol{Q}_{\boldsymbol{\Delta \bar{H}}_{R_m\text{-}D_{l,k}}} \odot (\boldsymbol{P} \boldsymbol{Q}_{\boldsymbol{Y}} \boldsymbol{P}^{\dagger})] \boldsymbol{U}, & \text{CDM-T, CDM-F} \end{cases}$$
$$(29)$$

where $\boldsymbol{U} \triangleq \boldsymbol{1}_{N_{R_m}} \otimes \boldsymbol{I}_{N_{R_m} \tilde{L}_{S_i}}$, and $\boldsymbol{\Delta \bar{H}}_{R_m\text{-}D_{l,k}}$ is defined as $[\boldsymbol{\Delta \tilde{H}}_{R_m,0\text{-}D_{l,k}}^T, \boldsymbol{\Delta \tilde{H}}_{R_m,1\text{-}D_{l,k}}^T, \cdots, \boldsymbol{\Delta \tilde{H}}_{R_m,N_{R_m}-1\text{-}D_{l,k}}^T]^T$ for TDM, FDM, and CDM-F, and $[(\boldsymbol{1}_{N_{R_m}} \otimes \boldsymbol{\Delta \tilde{H}}_{R_m,0\text{-}D_{l,k}})^T, \cdots, (\boldsymbol{1}_{N_{R_m}} \otimes \boldsymbol{\Delta \tilde{H}}_{R_m,N_{R_m}-1\text{-}D_{l,k}})^T]^T$ for CDM-T. Here, $\boldsymbol{\Delta \tilde{H}}_{R_m,n\text{-}D_{l,k}}$ represents error on $\tilde{\boldsymbol{H}}_{R_m,n\text{-}D_{l,k}}$. The above estimator incorporates the effect of channel estimation errors of the directly connected link in estimating the channels of indirectly connected link. This yields performance advantage over the existing approaches as will be shown later. The MSE of the proposed estimator in (24) is straightly obtained as [35]

$$\text{MSE}_{\boldsymbol{h}_{S_i\text{-}R_m}} = \text{Tr}\left\{\boldsymbol{Q}_{\Delta \boldsymbol{h}_{S_i\text{-}R_m}^{\text{MMSE}}}\right\} = \text{Tr}\left\{E[\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}} \right.$$
$$\left. -\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}} \boldsymbol{Q}_{\boldsymbol{Z}_{D_{l,k}}}^{-1} \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_m}\boldsymbol{z}_{D_{l,k}}}^H]\right\} \quad (30)$$

where the expectation is over $\hat{\boldsymbol{H}}_{R_m\text{-}D_{l,k}}$.

V. CSI ACQUISITION PROTOCOL FOR RELAY SYSTEMS

Regarding CSI acquisition between multiple nodes that can communicate directly, we present two solutions: pilot transmission (Section V-A) and explicit CSI feedback (Section V-B). For CSI acquisition of indirectly connected links (e.g. obtaining $\text{CSI}_{R_k\text{-}D_i}$ at S_m), we develop protocols under two scenarios: acquisition at the source or destination (Section V-C) and at the relays (Section V-D). In the protocols, pilots used are designed according to Section III and node multiplexing, if applicable, is according to Section III-D.

A. CSI Acquisition via Pilot Transmission

To acquire $\text{CSI}_{S\text{-}R_k}$ at each R_k, sources transmit pilots and each R_k estimates the CSI (e.g., using (20) or (21)). By changing the roles of sources and relays, together with the corresponding pilot design, each S_i can obtain estimates of $\text{CSI}_{R\text{-}S_i}$. The same pilot design, channel estimation, and transmission protocol can be applied for the relay to destination hop, with a possible change in channel length.

For TDD with reciprocal channels, transmission of pilots in both directions as described above will provide required CSIs to both sides. For FDD, the above scheme alone will not provide all necessary CSIs; the approach from Section V-C1 or V-C2 can be used in this case.

B. CSI Acquisition via Explicit CSI Feedback

Another approach for directly connected links is the use of explicit CSI feedback, where a node quantizes, encodes, and transmits back the CSI as data. Consider the MIMO link between source S_i and relays $\{R_k\}$. For TDD, there are two options, both starting with transmission of pilots. In the first option, each S_i transmits $N_{S_i}L_1$ pilots, $\forall i$, to $\{R_k\}$ and each R_k estimates $\text{CSI}_{S\text{-}R_k}$ and feeds it back to $\{S_i\}$. In the second option, each R_k transmits $N_{R_k}L_1$ pilots to $\{S_i\}$, each S_i then estimates $\text{CSI}_{R\text{-}S_i}$ and feeds it back to all $\{R_k\}$. The choice between these two options can be determined by the least pilot overhead usage. Suppose the first option is adopted (for the second option, sources and relays switch roles). Then each R_k estimates $\text{CSI}_{S\text{-}R_k}$ based on the pilot sets sent by S, quantizes and encodes it and transmits back to S. This explicit CSI feedback requires $2L_1 N_{\text{bits}}/\beta$ bits where N_{bits} is the effective number of bits used for a real valued channel tap gain, β is the coding rate ($\beta = 1$ if no error correcting code is applied), and the factor of 2 is due to the complex component. We define $\bar{N}_{\text{bits}} = 2N_{\text{bits}}/\beta$, so $L_1 \bar{N}_{\text{bits}}$ bits per channel are needed for CSI feedback.

When sending CSI feedback, the transmitting node has three options: i) Choose the best antenna for transmission based on SNR calculated by the estimated CSI, ii) Designate a predetermined antenna for CSI feedback, and iii) Utilize multiple antennas using space-time/space-frequency block coding (STBC/SFBC), space division multiplexing (SDM), or another multiple antenna transmission scheme.

For any option, using a predetermined subcarrier set $\mathcal{J}_{R_k,\text{FEED}}$ known by S and R_k, a pilot set (whose length equals the number of transmit antennas used in the feedback times L_1) is transmitted by R_k along with the encoded $\text{CSI}_{S\text{-}R_k}$. If each S_i needs to only know $\text{CSI}_{S_i\text{-}R_k}$, not $\text{CSI}_{S\text{-}R_k}$, and the channels are reciprocal, then the CSI feedback can exclude $\text{CSI}_{S\text{-}R_{k,\hat{m}}}$, where $\hat{m}$ is the index(es) of the transmitting antenna(s). The CSI feedback transmissions from all $\{R_k\}$ can apply node multiplexing. Using the relay transmitted pilot set, the receiving node S_i can estimate $\text{CSI}_{R_{k,\hat{m}}\text{-}S_i}$ using (20) or (21) (which gives $\text{CSI}_{S_i\text{-}R_{k,\hat{m}}}$ for reciprocal channels) and decode the required CSI feedback.

For FDD, both sides of the link need to send pilots. The transmissions from both sides can be simultaneous. Each node estimates CSI of its receiving MIMO link, quantizes,

encodes it and then feeds back. Any of the three options for CSI feedback transmission in TDD can be used here. The node multiplexing can be designed in the same way as in TDD. Another difference from TDD is that when sending CSI feedback, no pilot insertion is necessary for data detection as the receiving nodes already know the receiving channels if a predefined set of transmit antennas is used in the CSI feedback transmission.

The same protocols provided in this subsection can be applied for CSI acquisition of the relay to destination link.

C. Acquiring $\mathrm{CSI}_{R_k\text{-}D}$ at the Source and $\mathrm{CSI}_{S\text{-}R_k}$ at the Destination

This is a scenario of CSI acquisition of indirectly connected links. The problem is the same for either at the source or destination. The options described below can be applied at either node. We will discuss how the destination obtains $\mathrm{CSI}_{S\text{-}R_k}$, as some relaying strategies (e.g. [7], [9]) need this information.

1) Explicit Feedforward of CSI: Suppose R_k has obtained $\mathrm{CSI}_{S\text{-}R_k}$ using pilots or a protocol from Section V-B, then it can provide this information to $\{D_i\}$ after quantizing and encoding it into $N_{R_k} N_{S_{\mathrm{TOT}}} L_1 \bar{N}_{\mathrm{bits}}$ bits. If the destination D_i already knows $\mathrm{CSI}_{R_k\text{-}D_i}$, no pilots are necessary during the CSI feedforward, otherwise R_k needs to insert its own pilots whose length equals the total number of transmit antennas used in CSI feedforward times L_2. The CSI feedforward can adopt one of the three transmission options of Section V-B together with the node multiplexing mechanism for transmissions from different relay nodes.

2) CSI Estimation using Forwarded Pilots: S transmits pilots to $\{R_k\}$ and then each R_k forwards to D its received source pilots. If D_i does not have $\mathrm{CSI}_{R_k\text{-}D_i}$ yet, R_k also needs to insert its own pilots when forwarding the source pilots. Node multiplexing is applied between $\{R_k\}$. Next, D_i estimates $\mathrm{CSI}_{R_k\text{-}D_i}$ (e.g., using (20) or (21)) based on the inserted pilots of R_k and then $\mathrm{CSI}_{S\text{-}R_k}$ (e.g., using (24) with (27) and (28)).

D. Acquiring $\mathrm{CSI}_{S\text{-}R_l}$, $\mathrm{CSI}_{R_l\text{-}S}$, $\mathrm{CSI}_{D\text{-}R_l}$, and $\mathrm{CSI}_{R_l\text{-}D}$ $\forall l \neq k$, at Relay R_k, $\forall k$

This is another scenario of CSI acquisition of indirectly connected links where CSI sharing is required between relays to optimize the overall performance (e.g. [11]). Due to several CSI requirement scenarios with multiple options as will be discussed below, the main protocols for TDD will be presented in sequence in Table I for clarity. The table shows each action needed at each step along with the result from that action. Each step represents a separate transmission, unless otherwise noted. For FDD, we will present protocols for obtaining both forward and reverse link CSIs. If a considered relay system requires only a subset of the above CSIs, the required protocols can be easily extracted from our protocols.

1) Relay CSI sharing together with acquiring $\mathrm{CSI}_{S\text{-}R}$, $\mathrm{CSI}_{R\text{-}S}$, $\mathrm{CSI}_{D\text{-}R}$ and $\mathrm{CSI}_{R\text{-}D}$ at S and D: In Table I we present two options, 1A-TDD and 1B-TDD, to use when TDD is utilized. If FDD is used, additional steps are needed and

the relays should be able to transmit and receive on both frequency bands. There are two scenarios for FDD, based on whether the carrier frequencies of the relays' transmissions to the sources and to the destinations are the same or not, and they are denoted by FDD-1 and FDD-2, respectively. Considering FDD-1, we first apply either 1A-TDD or 1B-TDD (Options 1A-FDD and 1B-FDD, respectively) from Table I except the final step. Then, all S and D get $\mathrm{CSI}_{S\text{-}R}$ and $\mathrm{CSI}_{D\text{-}R}$; and in addition, the relay-inserted pilots provide $\mathrm{CSI}_{R\text{-}S_i}$ to S_i and $\mathrm{CSI}_{R\text{-}D_m}$ to D_m. After that, as step 4, each S_i broadcasts $\mathrm{CSI}_{S_i\text{-}R}$ and $\mathrm{CSI}_{R\text{-}S_i}$ and each D_m broadcasts $\mathrm{CSI}_{D_m\text{-}R}$ and $\mathrm{CSI}_{R\text{-}D_m}$, where pilot insertion is not needed. At this stage, each R_k obtains $\mathrm{CSI}_{S\text{-}R}$, $\mathrm{CSI}_{R\text{-}S}$, $\mathrm{CSI}_{D\text{-}R}$ and $\mathrm{CSI}_{R\text{-}D}$ while each S_i and D_m have $\mathrm{CSI}_{S\text{-}R}$ and $\mathrm{CSI}_{D\text{-}R}$. If S_i and D_m also need $\mathrm{CSI}_{R\text{-}S}$ and $\mathrm{CSI}_{R\text{-}D}$, as step 5, either each R_k broadcasts $\mathrm{CSI}_{R_k\text{-}S}$ and $\mathrm{CSI}_{R_k\text{-}D}$ or a predefined R_ν broadcasts $\mathrm{CSI}_{R\text{-}S}$ and $\mathrm{CSI}_{R\text{-}D}$; no pilot insertion is needed here. Node multiplexing can be appropriately applied in all of the above transmissions if it saves overhead. For FDD-2, the procedures remain the same as in FDD-1 except with the following changes. Any transmission from R_k now needs to be on two transmission bands (one to S and the other to D). Relay pilot insertion, if needed, will be L_1 tones per antenna to S and L_2 tones per antenna to D. If $\{R_k\}$ can transmit/receive on both bands simultaneously, any communications of $\{R_k\}$ to the two sides S and D in the above protocols can be simultaneous.

2) Relay CSI sharing without acquiring $\mathrm{CSI}_{S\text{-}R}$, $\mathrm{CSI}_{R\text{-}S}$, $\mathrm{CSI}_{D\text{-}R}$ and $\mathrm{CSI}_{R\text{-}D}$ at S and D: For TDD, instead of the above protocols, alternative options can be implemented to reduce CSI delay and pilot overhead. The steps required for TDD for these options are presented as 2A-TDD and 2B-TDD in Table I. For FDD-1, we can apply 2A-TDD or 2B-TDD first (dubbed as 2A-FDD or 2B-FDD, respectively). Then as step 3, each R_k broadcasts $\mathrm{CSI}_{S\text{-}R_k}$ and $\mathrm{CSI}_{D\text{-}R_k}$ where pilot insertion is not needed. As step 4, each S_i broadcasts $\mathrm{CSI}_{S_i\text{-}R}$ and each D_m broadcasts $\mathrm{CSI}_{D_m\text{-}R}$ without pilot insertion. For FDD-2, the same procedures as in FDD-1 can be used with the following changes. At step 1, $\{R_k\}$ transmit $N_{R_{\mathrm{TOT}}} L_1$ pilot tones to S and $N_{R_{\mathrm{TOT}}} L_2$ pilot tones to D. At step 3, each R_k broadcasts $\mathrm{CSI}_{S\text{-}R_k}$ to S and $\mathrm{CSI}_{D\text{-}R_k}$ to D. Note that as an alternative to the above 2A-FDD or 2B-FDD, we can also apply 1A-FDD or 1B-FDD up to step 4, which are identified as 2A′-FDD and 2B′-FDD, respectively.

VI. CSI ACQUISITION IN MP-MP SYSTEMS

In an MP-MP system with interference alignment (IA) [27]–[29], all CSIs are needed at each node. Therefore, a protocol is needed for CSI acquisition at various nodes. For MP-MP systems with disjoint resource allocation, CSI may also be needed at the node which performs resource allocation. In the following we first consider three scenarios for MP-MP IA systems which depend on the communication capabilities of the nodes $\{S_i\}$ and $\{D_\nu\}$. The three scenarios range from more limiting to less, so, for example, options for Scenario 1 may be used in all other scenarios, but not vice versa. The fourth scenario addresses for MP-MP systems with disjoint

TABLE I
PROTOCOL SEQUENCES FOR SECTION V-D: RELAY CSI SHARING

1A-TDD		
Step	Result	Notes
1. $\{S_i\}$ transmit $N_{S_{\text{TOT}}}L_1$ pilots.	Each R_k estimates CSI_{S-R_k} using (20) or (21).	Node multiplexing can be applied.
2. $\{D_m\}$ transmit $N_{D_{\text{TOT}}}L_2$ pilots.	Each R_k estimates CSI_{D-R_k} using (20) or (21).	Steps 1 and 2 may be performed simultaneously using node multiplexing.
3. Each R_k feedbacks quantized and encoded CSI_{S-R_k} and CSI_{D-R_k} disjointly. $\{R_k\}$ insert L pilot tones per transmit antenna.	$\{S_i\}$ and $\{D_m\}$ obtain CSI_{S-R} and CSI_{D-R}	Node multiplexing can be applied. If $\{S_i\}$ and $\{D_m\}$ have CSI_{R-S} and CSI_{R-D}, respectively, pilot insertion is not needed.
4. A predefined source or destination transmits quantized CSI_{S-R} and CSI_{D-R} to $\{R_k\}$.	Each R_k obtains CSI_{S-R} and CSI_{D-R}.	No pilot insertion is necessary.
1B-TDD		
3. Each R_k forwards received pilots from both $\{S_i\}$ and $\{D_m\}$ together with its inserted pilots (L tones per transmit antenna).	$\{S_i\}$ and $\{D_m\}$ estimate CSI_{S-R_k} and CSI_{D-R_k} using (20) or (21) and then (24) with (27) and (28).	Steps 1, 2 and 4 are the same as 1A-TDD. Node multiplexing can be applied.
2A-TDD		
1. $\{R_k\}$ transmit $N_{R_{\text{TOT}}}L$ pilots.	Each S_i and D_m estimate CSI_{R-S_i} and CSI_{R-D_m}, respectively, using (20) or (21).	Node multiplexing can be applied.
2. $\{S_i\}$ and $\{D_m\}$ quantize and encode the estimated channels and feedback them disjointly. Each S_i and D_m also insert their own pilots (L_1 and L_2 tones per transmit antenna respectively).	Each $\{R_k\}$ obtains CSI_{R-S} and CSI_{R-D}.	Node multiplexing can be applied. Pilot insertion at S_i and D_m is not needed if each R_k knows CSI_{S-R_k} and CSI_{D-R_k}, respectively.
2B-TDD		
2. $\{S_i\}$ and $\{D_m\}$ each forward received pilots uniquely, insert their own pilots (L_1 and L_2 tones per transmit antenna respectively), and transmit to $\{R_k\}$.	Each R_k estimates CSI_{S-R} and CSI_{D-R} using (20) or (21) and then (24) with (27) and (28).	Step 1 is the same as 2A-TDD. Node multiplexing can be applied.

resource allocation. In all four scenarios, when $\{S_i\}$ ($\{D_\nu\}$) transmit pilots or CSI to $\{D_\nu\}$ ($\{S_i\}$), the node multiplexing from Section III-D can be utilized to minimize overhead. Same as before, pilots are designed according to Section III.

Scenario 1: In this scenario, the sources can only transmit to destinations, and the destinations can only transmit to sources. There are two options for a TDD system, namely 1A-TDD and 1B-TDD, and for clarity of the protocol procedures, they are presented in Table II. For FDD, at step 1, both $\{S_i\}$ and $\{D_\nu\}$ are transmitting pilots on their separate subbands simultaneously. At step 2, both sets $\{S_i\}$ and $\{D_\nu\}$ simultaneously broadcast the CSIs of their receiving channels by either the explicit CSI feedforward described in Section V-C1 (Option 1A-FDD) or the pilot forwarding presented in Section III-C but without pilot insertion (Option 1B-FDD). After extracting or estimating CSIs using (24) with (27) and (28), each S_k now knows CSI_{S-D} and CSI_{D-S_k} while each D_m has CSI_{D-S} and CSI_{S-D_m}. Next, a predefined source and a predefined destination simultaneously transmit encoded CSI_{S-D} and encoded CSI_{D-S}, respectively. No pilot insertion is necessary. This transmission can use a predefined single antenna or multiple antennas with STBC/SFBC or SDM. The SDM approach can apply predefined multiple nodes as well and the choice consideration is the same as in TDD case. By now, each node knows all CSI_{S-D} and CSI_{D-S}.

Scenario 2: In this scenario, the sources can communicate with any node but the destinations can only transmit to and receive from the sources. If FDD is used, then the sources should have capability to transmit and receive on each frequency band. The options shown in Table II labeled 2A-TDD and 2B-TDD outline the procedure for systems using TDD. If FDD is used, the same initial steps are taken as in the TDD case. Then each D_ν would know all desired CSIs except $\text{CSI}_{S-D_l} \forall l \neq \nu$. Each source would know CSI_{D-S}. Therefore,

each D_ν can use either the explicit CSI feedforward approach described in Section V-C1 (Option 2A-FDD) or the forwarded pilots approach in Section III-C (Option 2B-FDD), together with the node multiplexing (described in Section III-D) to give each source CSI_{S-D_ν}. No pilot insertion is necessary. Then single or multiple predefined source nodes send back quantized and encoded $\text{CSI}_{S_i-D_\nu}$, $\forall i, \nu$ to the destinations (with no inserted pilots) from predefined single or multiple antennas with SFBC/STBC or SDM format. Furthermore, other source nodes need not perform the signal reception in this step. After this, all sources and destinations obtain all CSIs.

Scenario 3: In this scenario, any node can receive transmission from any other node. For FDD, each node should have capability to transmit and receive on each frequency band. For TDD, Options 3A-TDD and 3B-TDD are the same as Options 2A-TDD and 2B-TDD except they have the choice of whether $\{S_i\}$ or $\{D_\nu\}$ transmit in the first step. If $\{S_i\}$ transmits first, then the roles of the sources and destinations are reversed for the remaining procedure from 2A-TDD or 2B-TDD presented in Table II. In the following we will consider $\{D_\nu\}$ transmitting first as in Table II. If FDD is used, the system uses the chosen TDD procedure first and then sequentially each of $\{D_\nu\}$ broadcasts CSI_{S-D_ν}, along with $N_{D_\nu}L$ inserted pilots to S_i, $\forall i$ and D_l, $\forall l \neq \nu$. This broadcast can use either the explicit CSI feedforward or the pilot forwarding approach (denoted as Options 3A-FDD and 3B-FDD, respectively). Each receiving node I_l estimates its respective receiving channel $\text{CSI}_{D_\nu-I_l}$ and then obtains CSI_{S-D_ν}. After the broadcast of all $\{D_\nu\}$, every source and destination node obtain all CSIs.

Scenario 4: The final scenario is when the sources utilize separate resources and one node must acquire all CSIs present in the system to perform the resource allocation. We will assume that the node is S_{i^*}, and if it is a destination the

TABLE II
PROTOCOL SEQUENCES FOR SECTION VI: MP-MP SYSTEMS

1A-TDD		
Step	Result	Notes
1. $\{D_\nu\}$ transmit $N_{D_{TOT}}L$ pilots.	Each S_i estimates $\text{CSI}_{D\text{-}S_i}$ using (20) or (21).	Node multiplexing can be applied.
2. $\{S_i\}$ quantize and encode the estimated channels, insert their own pilots, and feedback the quantized CSIs disjointly.	$\{D_\nu\}$ estimate $\text{CSI}_{S\text{-}D_\nu}$ as well as decode the transmitted quantized $\text{CSI}_{D\text{-}S}$ ($= \text{CSI}_{S\text{-}D}$).	Node multiplexing can be used.
3. A predefined destination quantizes all $\text{CSI}_{S\text{-}D}$ and transmits these CSIs to $\{S_i\}$.	$\{S_i\}$ decode all desired CSIs.	No pilot insertion is necessary.
1B-TDD		
2. Each S_i forwards received pilots from $\{D_\nu\}$ uniquely, together with its own inserted pilots.	$\{D_\nu\}$ estimate $\text{CSI}_{S\text{-}D}$ using (24) with (27) and (28).	Steps 1 and 3 are the same as 1A-TDD. Node multiplexing can be used.
2A-TDD		
1. $\{D_\nu\}$ transmit $N_{D_{TOT}}L$ pilots.	Each S_i estimates $\text{CSI}_{D\text{-}S_i}$ using (20) or (21).	Node multiplexing can be applied.
2. Sequentially each S_i quantizes and encodes the estimated channels and transmits them disjointly together with its own inserted pilots.	All $\{D_\nu\}$ and S_l, $\forall\, l \neq i$ decode the CSIs from S_i.	Inserted pilots need not be designed jointly across all $\{S_i\}$ but just for individual S_i.
2B-TDD		
2. Sequentially each S_i forwards the received pilots together with its own inserted pilots.	All $\{D_\nu\}$ and S_l, $\forall\, l \neq i$ estimate the CSIs from S_i using (20) or (21) and then (24) with (27) and (28).	Step 1 is the same as 2A-TDD. Inserted pilots are designed just for individual S_i.

process is the same except their roles are reversed. For TDD, all sources, using node multiplexing, transmit $N_{S_{TOT}}L$ pilots to the destinations, where each D_ν estimates $\text{CSI}_{S_i\text{-}D_\nu}$, $\forall\, i$. $\{D_\nu\}$ then use either the explicit CSI feedforward approach (Option 4A-TDD) or the pilot forwarding (Option 4B-TDD) with their own inserted pilots to transmit to the sources. Now, S_{i*} retrieves the required CSIs.

For FDD, both node banks first transmit their pilots using node multiplexing. $\{S_i\}$ estimate $\text{CSI}_{D\text{-}S_i}$ and $\{D_\nu\}$ estimate $\text{CSI}_{S\text{-}D_\nu}$. Then both use either the explicit CSI feedforward approach (Option 4A-FDD) or the pilot forwarding (Option 4B-FDD) without inserted pilots to transmit the estimated CSIs back to the other node bank. Then, S_{i*} and a predefined destination $D_{\nu*}$ retrieve the CSIs, namely, $\text{CSI}_{S\text{-}D}$ and $\text{CSI}_{D\text{-}S}$, respectively. Finally, $D_{\nu*}$ transmits $\text{CSI}_{D\text{-}S}$ to S_{i*}, where no pilot insertion is required, and S_{i*} obtains all the desired CSIs.

VII. OVERHEAD, PROTOCOL CHOICE AND PERFORMANCE

As mentioned in Section II, to reduce control signaling overhead for CSI acquisition, the proposed approach inserts into the initial link establishment process the number assignment to the nodes. This step is simply inserting the number into a control signaling message of the process. Thus, if a total of K nodes are involved, then the overhead would be $K\log_2(K)$ bits. In contrast, conventional approaches would need to send control signaling messages to indicate the resource assignment (for pilots and/or CSI feedback) of all the nodes during the CSI acquisition phase. The corresponding overhead would be $K(N_{ID} + N_{RA})$ where N_{ID} is the number of bits for a user's ID and N_{RA} is the number of bits for resource assignment for transmission of pilots or CSI feedback of each user. Note that $N_{ID} > \log_2(K)$ as the total number of users in the system will be much larger than K and $N_{RA} > \log_2(K)$ as the total number of resources in the resource assignment would also be larger than K. This shows additional overhead saving of the proposed approach in the control signaling on top of pilot overhead saving. The latter will be described in the following protocol choice section.

A. Protocol Choice

Our proposed CSI acquisition protocols include some options and the choice depends on the specifics of the system setup, requirements, and algorithms involved. A major factor is the overhead cost while the energy efficiency and CSI accuracy should also be taken into consideration. For example, regarding the choice between the quantized and encoded CSI transmission and the pilot transmission (Section III, V, and VI), one can keep the CSI accuracy and the energy efficiency of the two options to be at comparable level and select the one which has smaller overhead cost. The energy efficiency can be kept the same by adjusting the QAM size and the bit energy of the encoded CSI. The CSI accuracy of the pilot transmission approach is reflected by the MSE performance but that of the explicit CSI transmission depends on how it is quantized and encoded (codebook design), and such a design process is out of the scope of this paper.

Regarding the option as to whether sources or destinations should transmit pilots (Section V-B), the overhead cost, which depends on the number of nodes and antennas at each side, can be used to select which option. The overhead associated with various CSI acquisition options for the relay and MP-MP systems are given in Table III and IV, respectively, which depends on the system specifics and can be used in selecting the protocol options for a given system. In the tables of overhead, if pilots and data symbols are transmitted in separate OFDM symbol(s) (denoted as preamble format), then $\Theta(X)$ is $\lceil X/\bar{N} \rceil \bar{N}$ where $\bar{N}$ is the number of subcarriers available for transmission ($\bar{N} = N$ if all of the subcarriers are used). If pilots are transmitted together with data tones within an OFDM symbol in a pilot-data-multiplexed manner, then $\Theta(X) = X$.

Note that the overhead can be appropriately translated into delay of the CSI acquisition protocol. Considering a preamble format, each pilot or CSI feedback transmission can be converted to delay using $\lceil N_{over}/\bar{N} \rceil$ where N_{over} is the number of needed tones for each step in obtaining all desired CSI. Comparing the options from Scenarios 1 through 3 from Section VI, we look at the delays associated with each. For

TABLE III
OVERHEAD COMPARISON FOR RELAY SYSTEMS

Option	Overhead
Quantized CSI feedforward	$[\Theta(N_{S_{TOT}}L_1)]$ pilots $+[N_{S_{TOT}}N_{R_{TOT}}L_1 N_{bits}]$ bits
Pilot forwarding	$[\Theta(N_{S_{TOT}}L_1) + \Theta(N_{R_{TOT}}(N_{S_{TOT}}L_1 + L_2))]$ pilots
Section V-D: Relay CSI Sharing	
1A-TDD	$[\Theta(N_{S_{TOT}}L_1) + \Theta(N_{R_{TOT}}L) + \Theta(N_{D_{TOT}}L_2)]$ pilots $+[2N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + 2N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits
1B-TDD	$[\Theta(N_{S_{TOT}}L_1) + \Theta(N_{D_{TOT}}L_2) + \Theta(N_{R_{TOT}}(N_{S_{TOT}}L_1 + N_{D_{TOT}}L_2 + L))]$ pilots $+[N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits
1A-FDD	TDD Overhead $+[2N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + 2N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits
1B-FDD	TDD Overhead $+[2N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + 2N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits
2A-TDD	$[\Theta(N_{S_{TOT}}L_1) + \Theta(N_{R_{TOT}}L) + \Theta(N_{D_{TOT}}L_2)]$ pilots $+[N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + N_{D_{TOT}}N_{R_{TOT}}L_2 \bar{N}_{bits}]$ bits
2B-TDD	$[\Theta(N_{R_{TOT}}L) + \Theta(N_{S_{TOT}}(N_{R_{TOT}}L + L_1)) + \Theta(N_{D_{TOT}}(N_{R_{TOT}}L + L_2))]$ pilots
2A-FDD	TDD Overhead $+[2N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + 2N_{D_{TOT}}N_{R_{TOT}}L_2 \bar{N}_{bits}]$ bits
2B-FDD	TDD Overhead $+[2N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + 2N_{D_{TOT}}N_{R_{TOT}}L_2 \bar{N}_{bits}]$ bits
2A$'$-FDD	1A-FDD Overhead $-[N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits
2B$'$-FDD	1B-FDD Overhead $-[N_{S_{TOT}}N_{R_{TOT}}L_1 \bar{N}_{bits} + N_{R_{TOT}}N_{D_{TOT}}L_2 \bar{N}_{bits}]$ bits

Note: $\Theta(X) \triangleq \begin{cases} \left\lceil \frac{X}{\bar{N}} \right\rceil \bar{N}, & \text{preamble format} \\ X, & \text{pilot-data multiplexing format} \end{cases}$

TABLE IV
OVERHEAD COMPARISON FOR MP-MP SYSTEMS

Option	Overhead
1A - TDD	$[\Theta(N_{D_{TOT}}L) + \Theta(N_{S_{TOT}}L)]$ pilots $+[2N_{S_{TOT}}N_{D_{TOT}}L N_{bits}]$ bits
1B - TDD	$[\Theta(N_{D_{TOT}}L) + \Theta(N_{S_{TOT}}(N_{D_{TOT}}+1)L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
1A - FDD	TDD Overhead $+[2N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
1B - FDD	TDD Overhead $+[\Theta(N_{D_{TOT}}(N_{S_{TOT}}+1)L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
2A - TDD	$[\Theta(N_{D_{TOT}}L) + \sum_{i=0}^{\tilde{N}_S-1}\Theta(N_{S_i}L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
2B - TDD	$[\Theta(N_{D_{TOT}}L) + \sum_{i=0}^{\tilde{N}_S-1}\Theta(N_{S_i}L(N_{D_{TOT}}+1))]$ pilots
2A - FDD	TDD Overhead $+[2N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
2B - FDD	TDD Overhead $+[\Theta(N_{D_{TOT}}(N_{S_{TOT}}+1)L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
3A - TDD	$[\Theta(N_{D_{TOT}}L) + \sum_{i=0}^{\tilde{N}_S-1}\Theta(N_{S_i}L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
3B - TDD	$[\Theta(N_{D_{TOT}}L) + \sum_{i=0}^{\tilde{N}_S-1}\Theta(N_{S_i}L(N_{D_{TOT}}+1))]$ pilots
3A - FDD	TDD Overhead $+[\sum_{\nu=0}^{\tilde{N}_D-1}\Theta(N_{D_\nu}L)]$ pilots $+[N_{D_{TOT}}N_{S_{TOT}}L \bar{N}_{bits}]$ bits
3B - FDD	TDD Overhead $+[\sum_{\nu=0}^{\tilde{N}_D-1}\Theta(N_{D_i}L(N_{S_{TOT}}+1))]$ pilots
4A - TDD	$[\Theta(N_{S_{TOT}}L) + \Theta(N_{D_{TOT}}L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
4B - TDD	$[\Theta(N_{S_{TOT}}L) + \Theta(N_{S_{TOT}}(N_{D_{TOT}}+1)L)]$ pilots
4A - FDD	TDD Overhead $+[2N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits
4B - FDD	TDD Overhead $+[\Theta(N_{D_{TOT}}N_{S_{TOT}}L)]$ pilots $+[N_{S_{TOT}}N_{D_{TOT}}L \bar{N}_{bits}]$ bits

See Table III for the definition of $\Theta(X)$.
Note: The expressions assume the destinations transmit first. If the sources transmit first, then the expressions are the same except $N_{D_{TOT}}$ and $N_{S_{TOT}}$ switch roles.

example, let us consider a system with $N = \bar{N} = 256$, $L = 4$, $\tilde{N}_S = \tilde{N}_D = 4$, $N_{S_{TOT}} = N_{D_{TOT}} = 8$, $\beta = 1/2$, $N_{bits} = 8$, and a fixed M-QAM modulation used for CSI feedback with $M = 16$. For TDD, 2B-TDD and 3B-TDD have the lowest delay of 5 OFDM symbols. For FDD, 3B-FDD has the lowest delay of 9 OFDM symbols. If N is increased to 1024, the smallest delay is still 2B-TDD and 3B-TDD for TDD at 5 OFDM symbols, but for FDD, 1B-FDD now has the shortest delay at 6 OFDM symbols. The delays will be unique to each system setup as there are many variables which affect the total.

In the CSI feedback/feedforward, the choice of transmitting from a single antenna or multiple antennas with SFBC/STBC or SDM (Section V-B) depends on the system specifics such as MIMO size and power and rate constraint. If there is only one receiving node with a single antenna, transmitting from the best antenna can be utilized. If both transmitting and receiving nodes have multiple antennas and the nodes are power-constrained, then SFBC/STBC can be adopted. If the MIMO nodes are rate-optimizing, then SDM can be used.

With regard to the option of pilot forwarding from a single antenna or multiple antennas (Section III-C), if the receiving nodes need not know all of their receiving channels, then using a single transmit antenna can save some overhead, since only one set per transmit node of added pilots needs to be transmitted. If the receiving nodes know all of their receiving channels, then the overhead of pilot forwarding from single or multiple antennas is the same, but the latter will have advantage in terms of peak to average power ratio (PAPR) since the number of transmitted tones per radio frequency (RF) chain is smaller in this case.
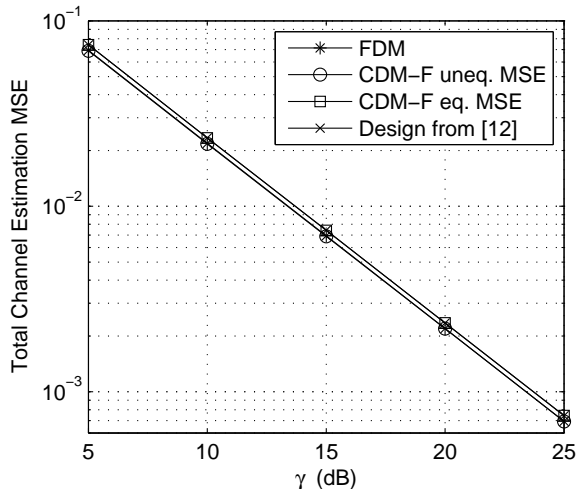
Fig. 5. Channel estimation MSE for all three channels in a node multiplexing

Fig. 6. Channel estimation MSE for individual channels of length $L_1 = 3$ and $L_2 = 9$

B. Channel Estimation for Channels with Different Lengths

In our simulations (Section VII-B-VII-E), we adopt OFDM with $N = 64$ subcarriers and 16 cyclic prefix samples. The average pilot SNR is defined as $\gamma = (E_{\text{total}}/N_{\text{total}})/(N\sigma_I^2)$ where N_{total} is the number of channels present in the node multiplexing scheme, E_{total} is the total pilot energy transmitted for all N_{total} channels, and σ_I^2 is the noise variance at the considered node I. For pilot forwarding scenarios, we will use γ_i to denote γ for the source pilots at the ith hop, and γ_{2r} to denote γ for the relay-inserted pilots. If not explicitly mentioned, we use $\gamma_2 = \gamma_{2r}$.

Here, we evaluate the proposed pilot designs for channels with different lengths. We consider a node multiplexing scenario where a source with two antennas and a single antenna destination simultaneously send pilots to a single antenna relay and the relay estimates the three channels. The channels from the source to relay have length $L_1 = 3$ and the channel between the relay and destination have $L_2 = 9$, and they have an exponential decay power delay profile (PDP), decreasing 3 dB per tap. We will collectively present the results in two perspectives: i) comparative performance of different pilot designs and ii) impacts of pilot designs with and without a constraint of equal individual channel estimation MSE. For the first perspective, we test three different pilot designs: the design from [12] and the two proposed designs, namely an approximately cyclically equi-spaced FDM design and the optimal CDM-F design presented in Section III-E without equal individual channel estimation MSE constraint. For the second perspective, we add the optimal CDM-F design with equal individual channel estimation MSE constraint. The same channel estimator in (20) is applied for all the pilot designs.

The total channel estimation MSE performance is shown in Fig. 5 while individual channel estimation MSEs are presented in Fig. 6. The theoretical MSE expressions match the simulation results and only the latter are presented for clarity. First, the design from [12], which is optimal under equal channel lengths and uses either 48 pilot tones (16 tones per channel) or 24 pilot tones (4 tones for each source channel and 16 tones for
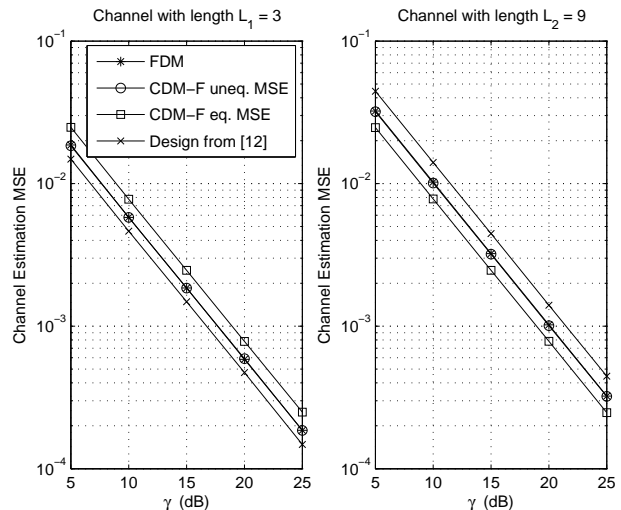
destination channel), is no longer optimal and suffers a slight total MSE degradation if compared with the proposed optimal CDM-F design without equal MSE constraint which uses 16 pilot tones. The proposed FDM uses minimum pilot amount of 15 tones and gives almost the same total MSE performance as the proposed optimal CDM-F design. Thus, the proposed designs provide better total MSE performance with smaller pilot overhead if compared with the existing design from [12]. In terms of MSEs of individual channels, the existing design from [12] yields a smaller MSE than the proposed designs for the channels with length $L_1 = 3$ and a larger MSE for the channel with length $L_2 = 9$. The reason is as follows. The channel with longer length experiences a larger MSE under equal pilot energy allocation due to larger dimension of the parameter space. Under the criterion of minimizing total MSE, the optimal design allocates more pilot energy to the channel with longer length to bring down the total MSE at the cost of some MSE increase for the shorter channels.

Next, we evaluate the proposed design with equal individual MSE constraint. Note that both our other proposed designs and the design from [12] do not provide equal individual channel estimation MSE. The proposed design yields equal MSE for individual channels, but its total MSE slightly increases (to the level similar to that of the design from [12]). The proposed minimum overhead FDM with equal MSE constraint gives almost the same MSE performance and hence is not shown in the figure for presentation clarity. The design with equal individual MSE constraint allocates more pilot energy to the channel with longer length than the optimal design without such constraint does, which decreases MSE of the longer channel and increases MSE of the shorter channel.

C. Tone Selection and Pilot Forwarding Schemes

First, we present the effect of forwarded pilot tone locations on the first-hop channel estimation performance. We set α_k as in [17] and $\gamma_1 = \gamma_2 = \gamma_{2r}$. For all channels, we assume
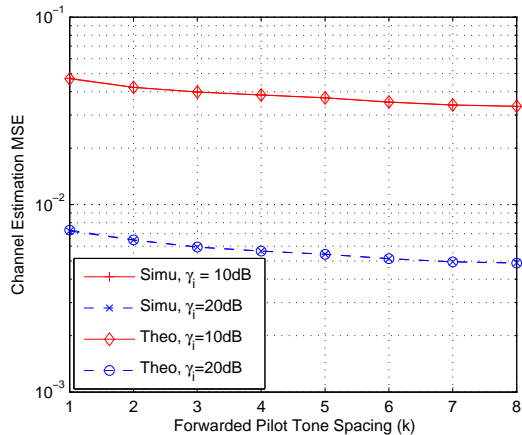
Fig. 7. Effect of tone selection for pilot forwarding ($\mathcal{J}_S = 8 \times \{0:7\}$, $\mathcal{J}_R = k \times \{0:7\}$)
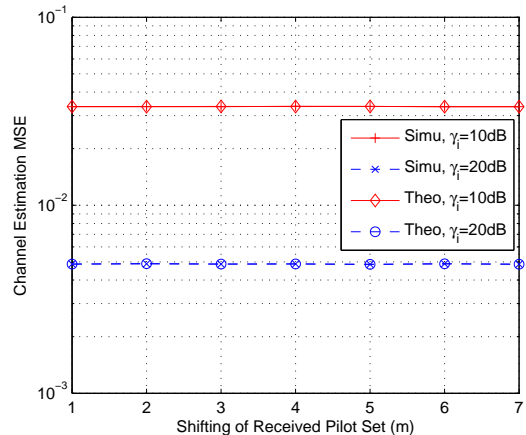

Fig. 8. Shift-invariant characteristic of a forwarded pilot set: (For TDM/CDM-T source pilots, $\mathcal{J}_R = \mathcal{T}_0 + m$. For FDM/CDM-F source pilots, $\mathcal{J}_R = \{\mathcal{T}_0, \mathcal{T}_1 + m\}$.)

the ITU Pedestrian Channel Model B with a raised cosine filter with roll-off factor of 0.5 and $L = 8$ sample-spaced channel taps. To focus on the effect of forwarded pilot tone locations, we consider one source, one single-antenna relay and one destination. First, we investigate for a single-antenna source which transmits one set of 8 cyclically equi-spaced pilot tones ($\mathcal{J}_S = 8 \times \{0:7\}$) and the relay forwards the source pilots on $\mathcal{J}_R = k \times \{0:7\}$ for $k \in \{1:8\}$. The relay also transmits its own 8 pilots with cyclically equal spacing which are disjoint (in time or frequency) from the forwarded source pilots so that the destination can first estimate $\boldsymbol{H}_{R\text{-}D}$ and then $\boldsymbol{H}_{S\text{-}R}$. Fig. 7 shows the first-hop channel estimation MSE (using the proposed estimator) for different locations of forwarded pilots characterized by the spacing k of adjacent pilot tones. We observe that $k = 8$ (corresponding to cyclically equi-spaced pilots) gives the best results which corroborates the design analysis in Appendix B.

Next, we consider a source with two antennas which transmits two sets of 8 cyclically equi-spaced pilot tones where $\mathcal{J}_S = \mathcal{T}_0$ for TDM or CDM-T source pilots, and $\mathcal{J}_S = \{\mathcal{T}_0, \mathcal{T}_1\}$ for FDM or CDM-F source pilots where $\mathcal{T}_0 \triangleq 8 \times \{0:7\}$ and $\mathcal{T}_1 \triangleq \mathcal{T}_0 + 4$. We first check with $\mathcal{J}_R = k \times \{0:7\}$ for $k \in \{1:8\}$ for TDM and CDM-T source pilots (similarly for FDM while keeping disjointness of its two pilot sets) and obtain the same results as in Fig. 7. Next, we investigate the effect of shifting the locations of the pilot set(s), i.e., $\mathcal{J}_R = \mathcal{T}_0 + m$ for TDM or CDM-T source pilots and $\mathcal{J}_R = \{\mathcal{T}_0, \mathcal{T}_1 + m\}$ for FDM or CDM-F source pilots. Fig. 8 presents the first-hop channel estimation MSE for different shifting m. The MSE (the same for different proposed source pilot designs) is invariant to the pilot set shifting in pilot forwarding, confirming the analysis in Appendix B.

Now, we consider two antennas at the relay and evaluate FDM, TDM, CDM-T and CDM-F pilot forwarding schemes. Their channel estimation MSEs of $\boldsymbol{H}_{S_i\text{-}R}$ obtained at the destination are shown in Fig. 9. FDM, TDM, and CDM-T pilot forwarding schemes give the same first-hop channel estimation MSE as expected since forwarded pilot sets maintain orthogonality (the pilot sets can be decoupled without
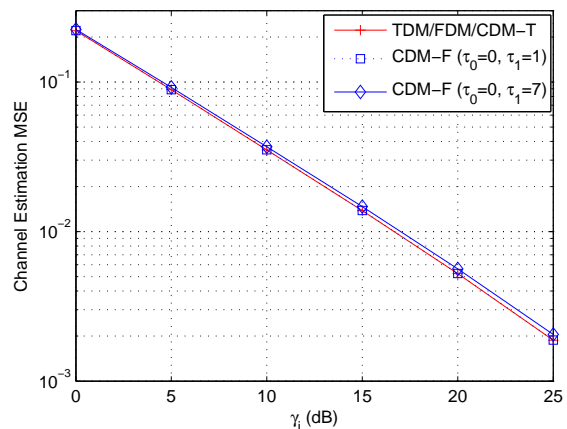
Fig. 9. Comparison of several pilot forwarding schemes

knowing $\{H_{R\text{-}D}[k]\}$). The proposed CDM-F pilot forwarding scheme ($\tau_0 = 0, \tau_1 = 1$) also gives practically the same MSE. However, with an unrecommended setting ($\tau_0 = 0, \tau_1 = 7$), the CDM-F scheme shows a slight MSE degradation. This is expected since the orthogonality of the CDM-F forwarded pilot sets is affected by the channel estimation errors $\{\Delta H_{R\text{-}D}[k]\}$ at $\mathcal{J}_{S_i}$ and $\mathcal{J}_{S_i} + \tau_1$, and such effect increases with more random (less correlated) channel estimation errors (i.e. larger τ_1). Nevertheless, the MSE degradation is insignificant and hence CDM-F with suboptimal setting of $\{\tau_i\}$ could still be considered if other constraints prohibit the use of optimal setting of $\{\tau_i\}$.

D. Estimator Performance and Effects of Pilot SNRs

Here, we will present performance of the proposed MMSE estimator in estimating $\boldsymbol{h}_{S\text{-}R_m}$ at the destination and the impact of different pilot SNR settings. We consider one source, two relays, and one destination, all with a single antenna. The source transmits pilots, and the relays forward received source pilots together with inserted own pilots to the destination. The pilots and their forwarding mechanisms are according to Section III. The system and channel models are the same
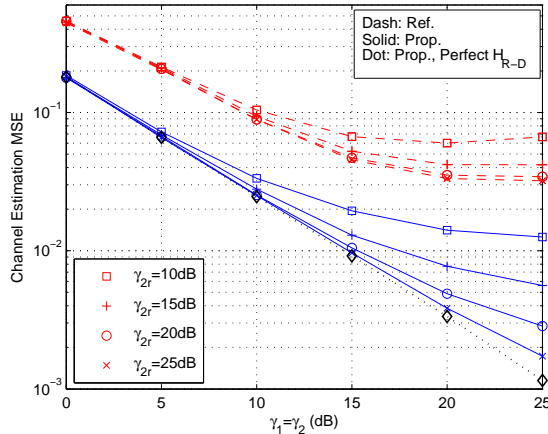
Fig. 10. Effect of the second-hop channel estimation accuracy on the first-hop channel estimation performance

Fig. 11. The first-hop channel estimation performance under fixed second-hop SNR

as in Section VII-C. The destination estimates $\{\boldsymbol{h}_{R_m\text{-}D}\}$ first and then $\{\boldsymbol{h}_{S\text{-}R_m}\}$. We include the MMSE estimator from [16] as reference. We obtain practically the same result for all proposed pilot forwarding mechanisms.

The channel estimation MSE for $\{\boldsymbol{h}_{S\text{-}R_m}\}$ is shown in Fig. 10 for various accuracy levels of $\hat{\boldsymbol{h}}_{R_i\text{-}D}$ (indicated by γ_{2r}). The MSE performances for $\boldsymbol{h}_{S\text{-}R_0}$ and $\boldsymbol{h}_{S\text{-}R_1}$ are identical, so only one is shown. We observe that the estimators could experience MSE floors depending on γ_{2r}. It is due to the nature of using the second hop channel estimate in estimating the first hop channel. To explain more, we can observe from (23) that any error in $\{\boldsymbol{H}_{R_m\text{-}D}\}$ will be multiplied with the pilot signal causing an interference in estimation of $\boldsymbol{h}_{S_i\text{-}R_m}$ which increases with pilot power. If the estimation error of $\{\boldsymbol{H}_{R_m\text{-}D}\}$ is not small enough compared to the noise level affecting the source pilots, the above mentioned interference will become dominant and cause the MSE floor. The proposed method outperforms the reference method due to two reasons: i) the proposed method incorporates the effect of second-hop channel estimation errors in estimating the first-hop channel while the reference method ignores that effect[2] and ii) the exploitation of channel frequency correlation property in the reference method is less fruitful than the proposed method.

For the proposed estimator with $\gamma_1 = \gamma_2$, we observe γ_{2r} should be at least the same as or larger than γ_i to avoid MSE floor for the first-hop channel estimation, and boosting γ_{2r} by 5 dB above γ_i would yield the MSE performance very close to that with perfect $\{\boldsymbol{H}_{R_m\text{-}D}\}$.

Next, we present the effect of fixed second-hop SNR (γ_2 and γ_{2r}) on the first-hop channel estimation performance of the proposed estimator in Fig. 11. The results obtained with perfect $\boldsymbol{H}_{R\text{-}D}$ (i.e., $\gamma_{2r} = \infty$) show the effect of fixed γ_2 while

---

[2] The MSE curve of the reference method for $\gamma_{2r} = 10$dB shows an intriguing result of a slight increase at high SNR. By neglecting the effect of estimation error for $\{\boldsymbol{H}_{R_k\text{-}D}\}$ and hence its interference effect, the reference method induces a model mismatch in its MMSE formulation which grows with increased first-hop pilot power and its effect becomes dominant if the level of the neglected errors (dependent on γ_{2r}) is significant compared to the noise level affecting the source pilots (dependent on γ_i). This explains the nature of the MSE curves of the reference method.
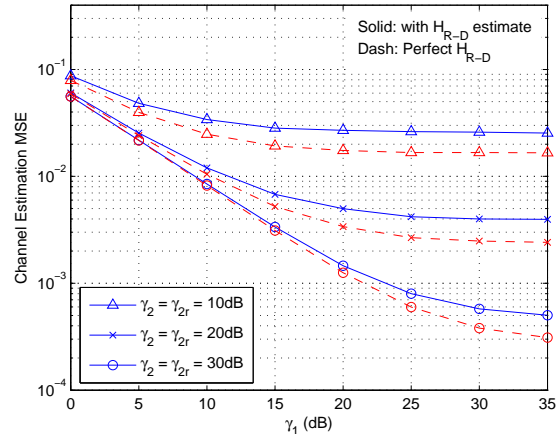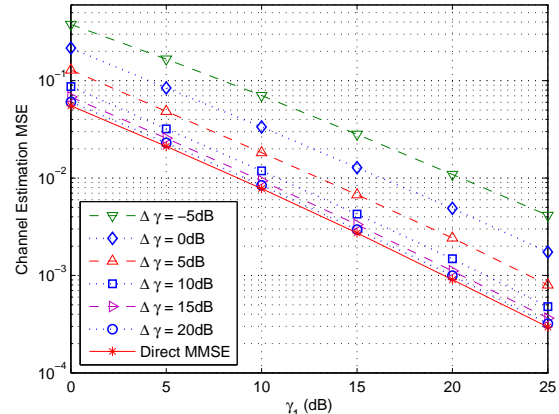

Fig. 12. Effect of boosting the second-hop SNR on the first-hop channel estimation ($\gamma_2 = \gamma_{2r} = \gamma_1 + \Delta\gamma$)

those with $\boldsymbol{H}_{R\text{-}D}$ estimate represent the effects of both fixed γ_2 and γ_{2r}. There are MSE floors at high γ_1 and the starting point and level of the floor depend on values of γ_2 and γ_{2r}. We observe that at a fixed γ_2, there is little improvement on first-hop estimation performance by increasing γ_1 beyond 5dB above γ_2, and $\gamma_1 = \gamma_2$ would be an energy-efficient setting.

In Fig. 12, we show the effect of boosting the second-hop SNR on the first-hop channel estimation performance where $\gamma_2 = \gamma_{2r} = \gamma_1 + \Delta\gamma$. As an ideal reference, we include the performance of directly estimating $\boldsymbol{h}_{S\text{-}R_m}$ by the MMSE estimator at R_m (denoted "Direct MMSE"). In general, boosting the second-hop SNR can improve the fisrt-hop channel estimation performance subject to the power budget of the second hop, and 20dB boost yields almost the same performance as Direct MMSE. However, boosting beyond 10dB will not be energy-efficient as the improvement becomes quite incremental.

To evaluate the proposed MMSE estimator's performance in scenarios with mismatched channel covariance knowledge, we assume a uniform PDP for both hops in the development of the proposed MMSE estimator while the actual channels correspond to the ITU Pedestrian Model B. The other settings
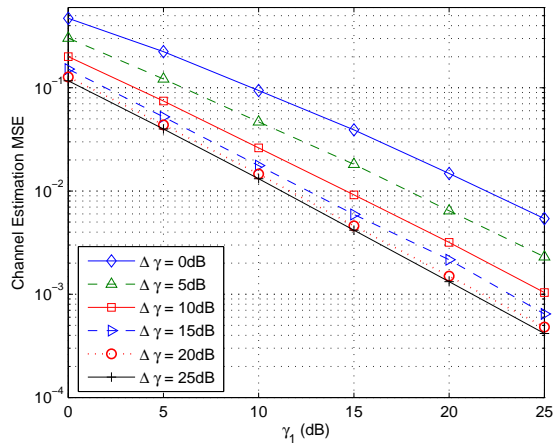
Fig. 13. Channel estimation MSE of $\boldsymbol{h}_{S\text{-}R_m}$ using the proposed MMSE estimator with mismatched statistics

Fig. 14. Impact of CSI accuracy on BER of a relay system with selection diversity

are the same as in Fig. 12. The simulation results are presented in Fig. 13 which shows similar MSE behaviors but with degraded MSE performance if compared to those in Fig. 12 obtained with perfect knowledge of the channel covariance matrices. We also observe that the performance degradation due to the mismatched knowledge can be offset by increasing the pilot SNRs of both hops.

E. BER Simulation

To illustrate the impact of CSI acquisition accuracy on relay systems, we evaluate the bit error rate (BER) performance of a relay system (with the same setting as in Section VII-D) with selection diversity[3]. The OFDM subcarriers are divided into 8 subchannels containing 8 consecutive subcarriers each, and only one subchannel is assigned to the considered S-R-D group. Data is transmitted using QPSK and γ_i is set the same value as the data SNR on a transmitted tone of the ith hop (denoted by SNR_i). After the CSI acquisition phase, the system selects a specific relay and a specific subchannel (same for S-R and R-D links) (named relay and subchannel selection) for data transmission to the destination or a specific relay on a fixed subchannel (named relay selection). Channel estimates are used in selection as well as data detection. The system computes the cascaded channel power gain summed across the subcarriers of each subchannel and selects the one with maximum value. The proposed estimator and the reference method are the same as in Section VII-D, and the performance with perfect channel knowledge of both hops is also included as an ideal reference. The exact closed-form analytical BER expression for the above system is intractable due to the involvement of ordered statistics of correlated random variables. Thus, we include a quasi-analytic approximate BER evaluation for the proposed method (see Appendix D).

Fig. 14 shows the BER performance. The relay and subchannel selection gives better BER than the relay selection

---

[3]Both selection diversity and maximum ratio combining are well-received approaches; here we arbitrarily adopt selection diversity to illustrate the BER advantage of the proposed method over the reference method.
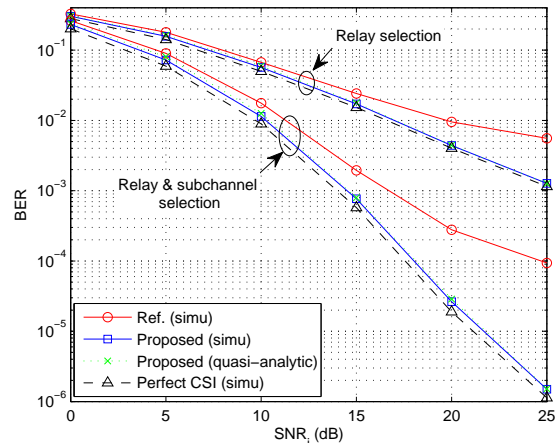
alone due to its higher diversity order. The quasi-analytic BER results match well with the simulation results. Our proposed estimator's performance is close to that with perfect channel knowledge, whereas the reference method again sees a degradation due to its inaccurate channel estimation.

VIII. CONCLUSIONS

Relay and MP-MP systems rely heavily on channel knowledge at various nodes. We have presented CSI acquisition protocols, pilot designs, pilot forwarding mechanisms, and channel estimation methods for various scenarios of relay and MP-MP systems using MIMO OFDM with TDD or FDD. Our proposed new pilot designs offer overhead saving and the same optimal estimation performance for channels with equal length and better estimation performance and smaller overhead for channels with unequal lengths than the existing designs. We also presented pilot designs (in particular pilot energy allocation) for minimum total MSE without the constraint of equal MSE of individual channels as well as with the constraint. We have developed an MMSE estimator to estimate the first hop channels at the receiving node of the second hop and shown its significant advantages over an existing method.

We have analyzed tone selection for pilot forwarding and proposed several pilot forwarding schemes for relay nodes with multiple antennas. We showed that the tones for forwarding a set of cyclically equi-spaced source pilots should maintain the same relative positions within the set while the relative positions of different sets can be changed. We investigated the effects of the first-hop and second-hop SNR of the forwarded pilots (γ_1 and γ_2) and the pilot SNR for the second-hop channel estimation (γ_{2r}) on the first-hop channel estimation at the end node of the second hop. We observed that $\gamma_{2r} = \gamma_2 + 5$dB gives performance close to that with perfect channel knowledge of the second hop while $\gamma_2 = \gamma_{2r} = \gamma_1 + 20$dB yields performance close to that obtained by directly estimating the first-hop channel at the end node of the first hop. At a fixed value of $\gamma_2 = \gamma_{2r}$, an energy-efficient setting for the first-hop is $\gamma_1 = \gamma_2$. But with a fixed

γ_1, boosting $\gamma_2 = \gamma_{2r}$ up to 10dB above γ_1 offers energy-efficient performance improvement. We have introduced a node multiplexing concept which reduces overhead and delay of the CSI acquisition. We also discussed how to choose among various CSI acquisition approaches in both relay and MP-MP systems with different CSI needs. The best protocol would depend on the specifics of the system and channel parameters, constraints, and CSI needs.

APPENDIX A
PROOF OF OPTIMALITY OF PILOT DESIGN IN (3) AND (4):

Using $\boldsymbol{\Omega}_{c_{i,\nu}}[0] = \sqrt{N}\boldsymbol{F}^H \boldsymbol{\Lambda}_{C_{i,\nu}^{(0)}} \boldsymbol{F}_{L_1}$, the optimal conditions satisfying (2) can be expressed as

$$\sum_{k=0}^{N-1} |C_{i,\nu}^{(0)}[k]|^2 e^{\frac{j2\pi dk}{N}} = \begin{cases} 0, & d = \pm 1, \cdots, \pm(L_1-1) \\ E_{\mathrm{av}}, & d = 0, \end{cases} \tag{31}$$

$$\sum_{k=0}^{N-1} C_{i,\nu_1}^{(0)*}[k] \, C_{i,\nu_2}^{(0)}[k] e^{\frac{j2\pi dk}{N}} = 0, \ d = 0, \pm 1, \cdots, \pm(L_1-1), \tag{32}$$

where $\nu_1 \neq \nu_2$. By substituting (4), the left hand side (LHS) of (31) and (31) become

$$\sum_{l=0}^{L_1'-1} \frac{E_{\mathrm{av}}}{L_1'} e^{\frac{j2\pi dl}{L_1'}} = E_{\mathrm{av}}\delta[d], \ d = -L_1'+1 : L_1'-1, \tag{33}$$

$$\sum_{l=0}^{L_1'-1} \frac{E_{\mathrm{av}}}{L_1'} e^{\frac{j2\pi l\{d+d_{12}\}}{L_1'}} = E_{\mathrm{av}}\delta[d+d_{12}], \tag{34}$$

where $d_{12} \triangleq (n_{i,\nu_2} - n_{i,\nu_1})L_1'$. As $L_1' > L_1$, (33) satisfies (31). To see (34) satisfies (32), recall that $|n_{i,\nu_2} - n_{i,\nu_1}| \in \{1 : N_{S_i} - 1\}$, thus for $d = 0, \pm 1, \cdots, \pm(L_1-1)$, we have $d + d_{12} \neq 0$ and $\delta[d+d_{12}] = 0$ which completes the proof.

APPENDIX B
PILOT FORWARDING DESIGN:

Suppose the source S_i with N_{S_i} antennas transmits pilots on $\mathcal{J}_{S_i} = \{\mathcal{T}_\nu; \nu = 0:N_{S_i}-1\}$ where $\mathcal{T}_\nu$ consists of L_1 cyclically equi-spaced pilot tone indexes. In general, a relay $R_{m,n}$ can forward the received source pilots on any $|\mathcal{J}_{S_i}|$ tones. The destination $D_{l,k}$ uses the forwarded pilots and $\hat{\boldsymbol{h}}_{R_{m,n}-D_{l,k}}$ to estimate $\boldsymbol{h}_{S_i-R_{m,n}}$. We formulate the design of tone selection for pilot forwarding by minimizing the MSE in estimating $\boldsymbol{h}_{S_i-R_{m,n}}$ as

$$\arg\min_{\mathcal{J}_n} \mathrm{MSE}_{\boldsymbol{h}_{S_i-R_{m,n}}} \tag{35}$$

where the elements of $\mathcal{J}_n \triangleq \{\mathcal{J}_{n,\nu}; \nu = 0:N_{S_i} - 1\}$ can be any $|\mathcal{J}_{S_i}|$ distinct numbers from $\{0 : N-1\}$ if forwarded in one OFDM symbol (if $\mathcal{J}_{n,\nu}$ and $\mathcal{J}_{n,l}$ are in different OFDM symbols, they can be the same) and $\mathcal{J}_{n,\nu}$ represents the tone index set on which the source pilots received on $\mathcal{T}_\nu$ are forwarded. $\mathrm{MSE}_{\boldsymbol{h}_{S_i-R_{m,n}}}$ is given by (30) obtained for the FDM pilot forwarding mechanism with only one pilot index set $\mathcal{J}_n$ (i.e., with $N_{R_m} = 1$, $\boldsymbol{P} = \alpha\boldsymbol{I}_{\tilde{L}_{S_i}}$, $\bar{\boldsymbol{\Lambda}}_{H_{R_{m,n}-D_{l,k}}} = \boldsymbol{\Lambda}_{H_{R_{m,n}-D_{l,k}}} = \mathrm{diag}\{\boldsymbol{\Phi}_{\mathcal{J}_n} \boldsymbol{H}_{R_{m,n}-D_{l,k}}\}$, and $\boldsymbol{Q}_{W_{D_l}} = \sigma_D^2 \boldsymbol{I}$).

A straight-forward manipulation yields $\mathrm{MSE}_{\boldsymbol{h}_{S_i-R_{m,n}}}$ as in (36) (shown at the top of the next page) where (a) is according to (eq. 177 of [37]), $\tilde{\boldsymbol{\Psi}}_{R_{m,n}-D_{l,k}} \triangleq \boldsymbol{Q}_{\Delta\bar{H}_{R_{m,n}-D_{l,k}}} \odot \boldsymbol{Q}_Y + \frac{\sigma_D^2}{\alpha^2}\boldsymbol{I}$, $\boldsymbol{\Psi}_{\mathcal{J}_n} \triangleq \boldsymbol{Q}_{\hat{H}_{R_{m,n}-D_{l,k}}^{-1}} \odot \tilde{\boldsymbol{\Psi}}_{R_{m,n}-D_{l,k}}$, and $\hat{H}_{R_{m,n}-D_{l,k}}^{-1}$ is the vector $\{\hat{H}_{R_{m,n}-D_{l,k}}^{-1}[p]; p \in \mathcal{J}_n\}$. $\boldsymbol{\Psi}_{\mathcal{J}_n}$ depends on $\mathcal{J}_n$ through $\boldsymbol{Q}_{\hat{H}_{R_{m,n}-D_{l,k}}^{-1}}$ and $\boldsymbol{Q}_{\Delta\bar{H}_{R_{m,n}-D_{l,k}}}$. Dropping the terms irrelevant to $\mathcal{J}_n$, the design becomes

$$\min_{\mathcal{J}_n} \mathrm{Tr}\left\{\boldsymbol{Q}_{\boldsymbol{h}_{S_i-R_{m,n}}} \bar{\boldsymbol{\Omega}}_{c_i}^\dagger \boldsymbol{Q}_Y^{-1} \boldsymbol{\Psi}_{\mathcal{J}_n} \boldsymbol{Q}_Y^{-1} \bar{\boldsymbol{\Omega}}_{c_i} \boldsymbol{Q}_{\boldsymbol{h}_{S_i-R_{m,n}}}\right\}. \tag{37}$$

With optimal source pilot design for source to relay channels with L_1 taps each, we have $\bar{\boldsymbol{\Omega}}_{c_i} \bar{\boldsymbol{\Omega}}_{c_i}^\dagger = E_i\boldsymbol{I}$ and $\bar{\boldsymbol{\Omega}}_{c_i}^\dagger \boldsymbol{Q}_Y^{-1} = (\boldsymbol{Q}_{\boldsymbol{h}_{S_i-R_{m,n}}} + \frac{\sigma_{W_R}^2}{E_i}\boldsymbol{I})^{-1}\bar{\boldsymbol{\Omega}}_{c_i}^\dagger$. Let $\boldsymbol{G} \triangleq \boldsymbol{Q}_{\boldsymbol{h}_{S_i-R_{m,n}}}(\boldsymbol{Q}_{\boldsymbol{h}_{S_i-R_{m,n}}} + \frac{\sigma_{W_R}^2}{E_i}\boldsymbol{I})^{-1}$. Then, the design becomes

$$\min_{\mathcal{J}_n} \mathrm{Tr}\left\{\boldsymbol{G}\bar{\boldsymbol{\Omega}}_{c_i}^\dagger \boldsymbol{\Psi}_{\mathcal{J}_n} \bar{\boldsymbol{\Omega}}_{c_i} \boldsymbol{G}^\dagger\right\}. \tag{38}$$

For $\boldsymbol{H}_{R_{m,n}-D_{l,k}}$ with wide-sense stationarity in frequency domain (e.g., with diagonal $\boldsymbol{Q}_{\boldsymbol{h}_{R_{m,n}-D_{l,k}}}$), $\boldsymbol{Q}_{\hat{H}_{R_{m,n}-D_{l,k}}}$ and $\boldsymbol{Q}_{\Delta H_{R_{m,n}-D_{l,k}}}$ are $N \times N$ Hermitian Toeplitz matrices, and the statistics of the $|\mathcal{J}_n| \times 1$ complex Gaussian vectors $\hat{\boldsymbol{H}}_{R_{m,n}-D_{l,k}}$ and $\Delta\boldsymbol{h}_{R_{m,n}-D_{l,k}}$ selected according to $\mathcal{J}_n$ remain the same when $\mathcal{J}_n$ is shifted by any fixed constant. Consequently, $\boldsymbol{Q}_{\hat{H}_{R_{m,n}-D_{l,k}}^{-1}}$ and $\boldsymbol{Q}_{\Delta\bar{H}_{R_{m,n}-D_{l,k}}}$ are shift-invariant to $\mathcal{J}_n$, and so is the channel estimation MSE in (36).

Define $\bar{\boldsymbol{\Omega}}_{k,l}$ and $\boldsymbol{\Psi}_{\mathcal{J}_n}^{k,l}$ to be the $L_1 \times L_1$ (k,l)-th submatrices of $\bar{\boldsymbol{\Omega}}_{c_i}$ and $\boldsymbol{\Psi}_{\mathcal{J}_n}$, respectively. For independent source channels, we have $\boldsymbol{G} = \mathrm{diag}\{\boldsymbol{G}_0, \cdots, \boldsymbol{G}_{N_{S_i}-1}\}$ with $\boldsymbol{G}_\nu \triangleq \boldsymbol{Q}_{\boldsymbol{h}_{S_{i,\nu}-R_{m,n}}}(\boldsymbol{Q}_{\boldsymbol{h}_{S_{i,\nu}-R_{m,n}}} + \frac{\sigma_{W_R}^2}{E_i}\boldsymbol{I})^{-1}$. With the above submatrices, (38) can be given as

$$\min_{\mathcal{J}_n} \mathrm{Tr}\left\{\sum_{\nu=0}^{N_{S_i}-1} \boldsymbol{G}_\nu \left(\sum_{k=0}^{N_{S_i}-1} \bar{\boldsymbol{\Omega}}_{k,\nu}^\dagger \boldsymbol{\Psi}_{\mathcal{J}_n}^{k,k} \bar{\boldsymbol{\Omega}}_{k,\nu}\right) \boldsymbol{G}_\nu^\dagger\right\} \tag{39}$$

which does not contain the terms corresponding to $\{\boldsymbol{\Psi}_{\mathcal{J}_n}^{k,l} : k \neq l\}$ as they sum up to zero due to orthogonal source pilots. Consequently, the design is only affected by individual sets $\{\mathcal{J}_{n,\nu}\}$ (i.e., relative index positions within each set) but not the relative shifting between the sets.

Furthermore, $\boldsymbol{\Psi}_{\mathcal{J}_n}$ maintains properties of a covariance matrix, and for channels with uncorrelated taps its main diagonal elements are the same regardless of $\mathcal{J}_n$ while the off-diagonal elements depend on $\mathcal{J}_n$. Thus, from (39) a judicious design choice is to select $\{\mathcal{J}_{n,\nu}\}$ such that the off-diagonal elements of $\{\boldsymbol{\Psi}_{\mathcal{J}_n}^{\nu\nu}\}$ (equivalently those of the main diagonal submatrices of $\boldsymbol{Q}_{\hat{H}_{R_{m,n}-D_{l,k}}^{-1}}$ and $\boldsymbol{Q}_{\Delta\bar{H}_{R_{m,n}-D_{l,k}}}$) have as small magnitudes as possible. This implies that the spacing of any adjacent tone indexes in $\mathcal{J}_{n,\nu}$ should be as large as possible (the indexes in $\mathcal{J}_{n,\nu}$ should be cyclically (mod-N) equi-spaced) while $\mathcal{J}_{n,\nu}$ and $\mathcal{J}_{n,l}$ within the same OFDM symbol are disjoint for $\nu \neq l$. With optimal source pilot indexes $\mathcal{J}_{S_i} = \{\mathcal{T}_\nu; \nu = 0:N_{S_i} - 1\}$, the above design can be simply given as $\mathcal{J}_{n,\nu} = \mathcal{T}_\nu + \tau_\nu$ (mod N) for any $\{\tau_\nu\}$ such that $\{\mathcal{J}_{n,\nu}\}$ within the same OFDM symbol are disjoint.

$$\text{MSE}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} = \text{Tr}\left\{ E\left[\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} - \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}}^{\dagger} \left(\boldsymbol{Q_Y} + \boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}_{R_{m,n}\text{-}D_{l,k}}}^{-1} \tilde{\boldsymbol{\Psi}}_{R_{m,n}\text{-}D_{l,k}} \boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}_{R_{m,n}\text{-}D_{l,k}}}^{-\dagger} \right)^{-1} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}} \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \right]\right\}$$

$$\overset{(a)}{\approx} \text{Tr}\left\{ E\left[\boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} - \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}}^{\dagger} \left(\boldsymbol{Q_Y}^{-1} - \boldsymbol{Q_Y}^{-1}\boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}_{R_{m,n}\text{-}D_{l,k}}}^{-1} \tilde{\boldsymbol{\Psi}}_{R_{m,n}\text{-}D_{l,k}} \boldsymbol{\Lambda}_{\hat{\boldsymbol{H}}_{R_{m,n}\text{-}D_{l,k}}}^{-\dagger} \boldsymbol{Q_Y}^{-1} \right) \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}} \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \right]\right\}$$

$$\approx \text{Tr}\left\{ \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} - \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}}^{\dagger} \left(\boldsymbol{Q_Y}^{-1} - \boldsymbol{Q_Y}^{-1}\boldsymbol{\Psi}_{\mathcal{J}_n}\boldsymbol{Q_Y}^{-1} \right) \bar{\boldsymbol{\Omega}}_{\boldsymbol{c_i}} \boldsymbol{Q}_{\boldsymbol{h}_{S_i\text{-}R_{m,n}}} \right\} \tag{36}$$

APPENDIX C
PROOF OF OPTIMALITY OF (14,15,17,18) FOR CHANNELS WITH DIFFERENT LENGTHS:

Let $\boldsymbol{\lambda}$ and $\boldsymbol{d}$ denote the vectors consisting of the eigenvalues and the diagonal elements of $\boldsymbol{\Omega}_{\boldsymbol{c}}^{\dagger}\boldsymbol{\Omega}_{\boldsymbol{c}}$, respectively. Then, $\boldsymbol{d}$ majorizes $\boldsymbol{\lambda}$ [36] and consequently, $\text{MSE}/\sigma_{\mathcal{J}_\nu}^2 = \sum_{k=0}^{L_{\text{total}}-1} 1/\lambda_k \geq \sum_{k=0}^{L_{\text{total}}-1} 1/d_k$. For fixed $\{E_i\}$, $\boldsymbol{d}$ is always fixed at $[E_0\mathbf{1}_{L_0}^T, E_1\mathbf{1}_{L_1}^T, \cdots, E_{K-1}\mathbf{1}_{L_{K-1}}^T]^T$ while $\boldsymbol{\lambda}$ can vary depending on the pilots. Thus, for fixed $\{E_i\}$, the MSE minimizing eigenvalues are given by $\boldsymbol{\lambda} = [E_0\mathbf{1}_{L_0}^T, E_1\mathbf{1}_{L_1}^T, \cdots, E_{K-1}\mathbf{1}_{L_{K-1}}^T]^T$ which is straight-forwardly obtained by (14) and the MSE becomes $\sigma_{\mathcal{J}_\nu}^2 \sum_{i=0}^{K-1} L_i/E_i$. The remaining optimization is to find the set of $\{E_i\}$ that minimizes MSE under the constraint of constant $\sum_{i=0}^{K-1} E_i$. Using the Lagrange multiplier method yields the optimal pilot energy allocation as given in (15).

Next, using $\boldsymbol{\Omega}_{\boldsymbol{c_i}} = \sqrt{N}\boldsymbol{F}^H\boldsymbol{\Lambda}_{C_i}\boldsymbol{F}_{L_i}$, the optimal condition can be expressed as

$$\sum_{k=0}^{N-1} |C_i[k]|^2 e^{\frac{j2\pi dk}{N}} = \begin{cases} 0, & d = \pm 1, \cdots, \pm(L_i-1) \\ E_i, & d = 0 \end{cases} \tag{40}$$

$$\sum_{k=0}^{N-1} C_i^*[k]\, C_\nu[k] e^{\frac{j2\pi dk}{N}} = 0, \ d = -L_\nu+1 : L_i-1. \tag{41}$$

By substituting the pilots from (16) and (17), the LHS of (40) and (41) become

$$\sum_{m=0}^{\tilde{L}_{\text{total}}-1} \frac{E_i}{\tilde{L}_{\text{total}}} e^{\frac{j2\pi dm}{\tilde{L}_{\text{total}}}} = E_i\delta[d], \ d = -\tilde{L}_{\text{total}}+1 : \tilde{L}_{\text{total}}-1 \tag{42}$$

$$\sum_{m=0}^{\tilde{L}_{\text{total}}-1} \sqrt{E_iE_\nu} e^{\frac{j2\pi m(d+\tau_\nu-\tau_i)}{\tilde{L}_{\text{total}}}} = \sqrt{E_iE_\nu}\delta[d+\tau_\nu-\tau_i)]. \tag{43}$$

Since $\tilde{L}_{\text{total}} > L_i$, (42) satisfies (40). For $d = -L_\nu+1:L_i-1$, when $\nu > i$, we have $\tau_\nu - \tau_i = \sum_{m=i+1}^{\nu}(L_m + \Delta_m)$ and $0 < d+\tau_\nu-\tau_i < \tilde{L}_{\text{total}}$, while when $i > \nu$, we have $\tau_\nu - \tau_i = -\sum_{m=\nu+1}^{i}(L_m + \Delta_m)$ and $-\tilde{L}_{\text{total}} < d+\tau_\nu-\tau_i < 0$. Thus, $\delta[d + \tau_\nu - \tau_i)] = 0$ in (43) which completes the proof. The proof for the design in (18) is obtained in the same way.

APPENDIX D
QUASI-ANALYTIC BER EXPRESSION:

As the receiver at the destination uses channel estimates for data detection, the received data on subcarrier m can be written as

$$Z[m] = \alpha\hat{H}_{R\text{-}D}[m]\hat{H}_{S\text{-}R}[m]B[m] + \tilde{I}[m] \tag{44}$$

where $B[m]$ is the zero-mean transmitted data with $E[|B[m]|^2] = \sigma_B^2$ and $\tilde{I}[m]$ is the effective noise due to channel estimation errors and AWGN noises and is given as

$$\tilde{I}[m] = \alpha(H_{R\text{-}D}[m]H_{S\text{-}R}[m] - \hat{H}_{R\text{-}D}[m]\hat{H}_{S\text{-}R}[m])B[m]$$
$$+ \alpha H_{R\text{-}D}[m]W_R[m] + W_D[m] \tag{45}$$
$$= \hat{H}_{S\text{-}R}[m]\Delta H_{R\text{-}D}[m] + \hat{H}_{R\text{-}D}[m]\Delta H_{S\text{-}R}[m] + W_D[m]+$$
$$\Delta H_{R\text{-}D}[m]\Delta H_{S\text{-}R}[m] + \alpha(\hat{H}_{R\text{-}D}[m] + \Delta H_{R\text{-}D}[m])W_R[m]. \tag{46}$$

Given MMSE estimates $\{\hat{H}_{R\text{-}D}[m], \hat{H}_{S\text{-}R}[m]\}$, the dominant terms of $\tilde{I}[m]$ are complex Gaussian and hence $\tilde{I}[m]$ can be well approximated as complex Gaussian. Then, the BER for QPSK can be given as

$$\text{BER} \approx E\left[Q\left(\sqrt{\frac{|\alpha\hat{H}_{R\text{-}D}[m]\hat{H}_{S\text{-}R}[m]|^2\sigma_B^2}{E[|\tilde{I}[m]|^2]}} \right) \right] \tag{47}$$

where $E[|\tilde{I}[m]|^2]$ is computed from (45) as $\alpha^2\sigma_B^2|H_{R\text{-}D}[m]H_{S\text{-}R}[m] - \hat{H}_{R\text{-}D}[m]\hat{H}_{S\text{-}R}[m]|^2 + \sigma_{W_D}^2 + \alpha^2|H_{R\text{-}D}[m]|^2\sigma_{W_R}^2$, while $Q(\cdot)$ is the Q-function and its expectation is carried out over $\{H_{R\text{-}D}[m], H_{S\text{-}R}[m], \hat{H}_{R\text{-}D}[m], \hat{H}_{S\text{-}R}[m]\}$ via simulation.

REFERENCES

[1] H. Park and J. Chun, "A two-stage antenna subset selection scheme for amplify-and-forward MIMO relay systems," *IEEE Signal Processing Letters*, vol. 17, no. 11, Nov. 2010.
[2] D. Munoz, B. Xie, and H. Minn, "An adaptive MIMO-OFDMA relay system," *IEEE Wireless Comm. Letters*, vol. 1, no. 5, pp. 496–499, Oct. 2012.
[3] T. Abe, H. Shi, T. Asai, and H. Yoshino, "A relaying scheme for MIMO wireless networks with multiple source and destination pairs," *Asia-Pacific Conf. Commun.*, Oct. 2005.
[4] B. Khoshnevis, W. Yu, and R. Adve, "Grassmannian beamforming for MIMO amplify-and-forward relaying," *IEEE Journ. Select. Areas Commun.*, vol. 26, no. 8, Oct. 2008.
[5] A. Talebi and W. A. Krzymien, "Cooperative MIMO multiple-relay system with optimized beamforming and power allocation," *IET Commun*, vol. 4, no. 14, pp. 1677-1686, Mar. 2010.
[6] T. Taniguchi, Y. Karasawa, and N. Nakajima, "Weight design of AF MIMO relay system with perfect channel state information," *IEEE ITS Telecommun., 9th Intl. Conf.*, pp. 164-169, Oct. 2009.
[7] E. Park, K. J. Lee, and I. Lee, "Joint MMSE transceiver design for MIMO amplify-and-forward relay systems with multiple relays," *IEEE VTC Fall*, Sept. 2009.
[8] N. Wu, W. Huang, and H. Li, "A novel relay selection algorithm for relaying networks," *IEEE VTC Fall*, Sept. 2009.
[9] R. H. Louie, Y. Li, H.A. Suraweera, and B. Vucetic, "Beamforming with antenna correlation in two hop amplify and forward relay networks," *IEEE WCNC*, April 2009.
[10] R. Yin, Y. Zhang, J. Zhang, G. Yu, and Z. Zhang, "Joint resource allocation in multiple channels, multiple relays systems," *IEEE WCNC*, April 2010.
[11] A. El-Keyi and B. Champagne, "Adaptive training-based collaborative MIMO beamforming for multiuser relay networks," *IEEE VTC Spring*, April 2009.

[12] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1158–1168, May 2006.

[13] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Proc.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[14] H. Minn, N. Al-Dhahir, and Y. Li, "Optimal training signals for MIMO OFDM channel estimation in the presence of frequency offset and phase noise," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1754–1759, Oct. 2006.

[15] X. Ma, L. Yang, and G. B. Giannakis, "Optimal training for MIMO frequency-selective fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 453–466, Mar. 2005.

[16] M. Liu, J. Zhang, Y. Zhang, and Y. Liu, "A channel estimation scheme for amplify-and-forward OFDM relay networks," *IEEE VTC Fall*, Sept. 2009.

[17] F. Gao, T. Cui, and A. Nallanathan, "Optimal training design for channel estimation in decode-and-forward relay networks with individual and total power constraints," *IEEE Trans. Signal Proc.*, vol. 56, no. 12, pp. 5937–5949, 2008.

[18] F. Gao, R. Zhang, and Y. C. Liang, "Optimal channel estimation and training design for two-way relay networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3024–3033, 2009.

[19] F. Gao, B. Jiang, X. Gao, and X. Zhang, "Superimposed training based channel estimation for OFDM modulated amplify-and-forward relay networks," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2029–2039, 2011.

[20] T. Kong and Y. Hua, "Optimal design of source and relay pilots for MIMO relay channel estimation," *IEEE Trans. Signal Proc.*, vol. 59, no. 9, pp. 4438–4446, 2011.

[21] S. Sun, Y. Jing, "Channel training design in amplify-and-forward MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3380–3391, 2011.

[22] J. Ma, P. Orlik, J. Zhang, and G. Y. Li, "Pilot matrix design for estimating cascaded channels in two-hop MIMO amplify-and-forward relay systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 6, pp. 1956–1965, 2011.

[23] P. Lioliou, M. Viberg, and M. Matthaiou, "Bayesian approach to channel estimation for AF MIMO relaying systems," *IEEE Journ. Select. Areas Commun.*, vol. 30, no. 8, pp. 1440-1451, 2012.

[24] J. Ma, P. Orlik, J. Zhang, and G. Y. Li, "Pilot matrix design for interim channel estimation in two-hop MIMO AF relay systems," *IEEE ICC*, 2009.

[25] D. Neves, C. Ribeiro, A. Silva, and A. Gameiro, "Channel estimation schemes for OFDM relay-assisted systems," *IEEE VTC Spring*, 2009.

[26] Z. Zhang, W. Zhang, and C. Tellambura, "Cooperative OFDM channel estimation in the presence of frequency offsets," *IEEE Trans. Vehicular Tech.*, vol. 58, no. 7, pp. 3447–3459, 2009.

[27] H. Huang, V. K. N. Lau, Y. Du, and S. Liu, "Robust approximate lattice alignment design for K-pairs quasi-static MIMO interference channels with imperfect CSI," *IEEE ISIT*, pp. 2298–2302, June 2010.

[28] S. W. Peters and R. W. Heath Jr, "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Vehicular Tech.*, vol. 60, no. 1, pp. 206-218, Jan. 2011.

[29] M. Pischella and E. Vivier, "Comparison of distributed space and frequency interference alignment," *IEEE PIMRC*, pp. 532–537, Sept. 2010.

[30] B. Xie, Y. Li, H. Minn and A. Nosratinia, "Interference alignment under training and feedback constraints," *IEEE Globecom*, 2011, Houston, Texas.

[31] B. Xie, Y. Li, H. Minn, and A. Nosratinia, "Adaptive interference alignment with CSI uncertainty," *IEEE Trans. Commun.*, vol. 61, no. 2, pp. 792–801, Feb. 2013.

[32] B. Kecicioglu, H. Minn, and C.C. Chong, "Intra-frame transmission adaptation for fast fading MIMO-OFDMA systems," *IEEE Trans. Commun.*, vol. 59, no. 9, pp. 2585–2594, Sept. 2011.

[33] IEEE LAN/MAN Standards Committee, "IEEE Standard 802.16j: Air Interface for Fixed and Mobile Broadband Wireless Access: Mobile Multihop Relay", *IEEE*, 2009.

[34] 3GPP, "Overview of 3GPP Release 10 V0.1.10", www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases /Rel-10_description_20130924.zip, Sept. 2013.

[35] S. M. Kay, "Fundamentals of Statistical Signal Processing: Estimation Theory," *Prentice Hall PTR*, 1993.

[36] R. A. Horn and C. R. Johnson, "Matrix Analysis," *Cambridge University Press*, 1996.

[37] K. B. Petersen and M. S. Pedersen, "The Matrix Cookbook", version: 14 Nov. 2008. http://matrixcookbook.com.
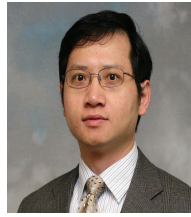
Hlaing Minn (S'99, M'01, SM'07) received the B.E. degree in electronics from the Yangon institute of Technology, Yangon, Myanmar, in 1995, the M.Eng. degree in telecommunications from the Asian Institute of Technology (AIT), Pathumthani, Thailand, in 1997, and the Ph.D. degree in electrical engineering from the University of Victoria, Victoria, BC, Canada, in 2001. He was a post-doctoral fellow at the University of Victoria during Jan.-Aug. 2002. He has been with the Erik Jonsson School of Engineering and Computer Science, the University of Texas at Dallas, Richardson since Sept. 2002, and currently is an Associate Professor. His research interests include wireless communications, statistical signal processing, error control, detection, estimation, synchronization, signal design, cross-layer design, cognitive radios, and wireless health-care applications. Prof. Minn is an Editor for the *IEEE Transactions on Communications* and *International Journal of Communications and Networks*. He has served as a Technical Program Co-Chair of the Wireless Communications Symposium of the IEEE GLOBECOM 2014 and the Wireless Access Track of the IEEE VTC 2009 (Fall), and a TPC member in several IEEE conferences.

Daniel Munoz (S'05) received his B.S.E.E. from Washington State University in the Spring of 2007. After two years at University of Texas at Dallas (UTD), he obtained an M.S.E.E. with an emphasis on Telecommunications. In the Fall of 2012, he received his Ph.D. degree in electrical engineering from UTD. He has had internships at Agilent Technologies as well as Idaho National Laboratories. Currently he is a Patent Examiner at the U.S. Patent and Trademark Office.