This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2009 proceedings.

Achievable Rates of MIMO Bidirectional Broadcast Channels With Self-Interference Aided Channel Estimation

Jian Zhao, Marc Kuhn and Armin Wittneben
Communication Technology Laboratory, ETH Zurich
CH-8092 Zurich, Switzerland
Email: {zhao, kuhn, wittneben}@nari.ee.ethz.ch

Gerhard Bauch
DoCoMo Euro-Labs
D-80687 Munich, Germany
Email: bauch@docomolab-euro.com

***Abstract*—In this paper, we consider the broadcast (BRC) phase of two-way decode-and-forward (DF) relaying systems. The channel in that phase is called the *bidirectional broadcast channel*. Its achievable rates are calculated when the self-interference aided channel estimation scheme is applied. We consider a block-fading channel model in the BRC phase and exploit the self-interference that is inherent in two-way relaying techniques to get an initial estimate of the channel at the receiving terminals. This initial channel estimate is utilized to decode the data in the first several time slots of each coherence interval. Data-aided approaches are then employed to improve the channel estimates in the following time slots of each coherence interval. The spectral efficiency improvement for systems employing this self-interference aided channel estimation scheme is quantified by comparing its achievable rates to that of the traditional pilot-aided channel estimation scheme.**

I. Introduction

Two-way relaying [1] is a novel technique proposed to solve the problem of spectral efficiency loss in half-duplex relaying systems. This technique considers the scenario that two half-duplex wireless terminals, A and B, exchange data via another half-duplex wireless relay R, as shown in Fig. 1. Here we only consider decode-and-forward (DF) relaying systems. This two-way DF relaying technique consists of the following two phases: the multiple access (MAC) phase and the broadcast (BRC) phase, which can be separated in time (TDD) or in frequency (FDD). In the MAC phase, the two terminals transmit their data simultaneously to the relay and the relay decodes the received signals. After that, the relay combines the decoded data and retransmits them in the BRC phase. Upon receiving the signals from the relay, each terminal can cancel the known data that were transmitted by the the receiving terminal itself. Compared to traditional relaying techniques, the two-way relaying technique achieves bidirectional communication between the two terminals in two channel uses instead of four.

The transmission data rate pair of terminal A and B is an achievable rate pair of the two-way DF relaying system if it is achievable by both the MAC and BRC phases, i.e., we can determine the achievable rate pairs of the two-way DF relaying system by determining those for the MAC and BRC phases

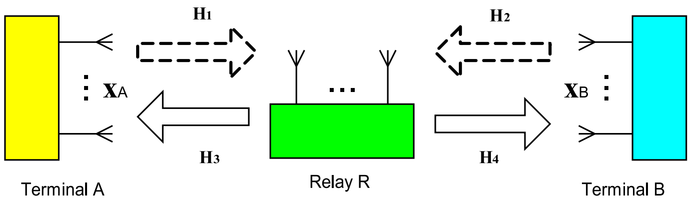

Fig. 1. MIMO two-way DF relaying system, where the dashed arrows represent the transmission in the MAC phase, and the solid arrows represent the transmission in the BRC phase.

separately. However, the MAC phase is a conventional multiple access channel, where the optimal coding schemes are known, e.g., ref. [2]. Compared to the MAC phase, the BRC phase is a new area of research. Leaving aside the constraints of the MAC phase, we assume the messages from the terminals A and B have been perfectly decoded at the relay. Thus in the BRC phase, the messages at the relay to be transmitted to terminal A are perfectly known to terminal B and vice versa for the messages intended for terminal B. Such a channel is called the *bidirectional broadcast channel* [3]. Another reason for us to consider the MAC and BRC phases independently is that the MAC and BRC phases may take place under different channel conditions, e.g., in an FDD scenario. In the following, we focus on transmission schemes in the BRC phase.

As we know, channel estimation is an integral part of wireless transmission schemes and it is particularly important for systems with multiple antennas. Traditional channel estimation schemes transmit orthogonal pilot sequences on different transmit antennas before sending data. The algorithms used for traditional pilot-aided channel estimation schemes are summarized in [4], where the linear least square (LS) and minimum mean square error (MMSE) approaches are the two major algorithms. The authors of [5] computed a lower bound on the capacity of a point-to-point multiple-input multiple-output (MIMO) channel that is estimated by the pilot-aided scheme. In [6], the authors proposed to transmit low-power orthogonal pilot signals concurrently with the data. A capacity lower bound for systems employing this pilot-embedding scheme has been shown in [7], [8]. All those channel estimation schemes rely on an initial estimate of the MIMO channel, based on the transmission of pilot sequences.

We consider the *superposition coding scheme* for relay retransmission in the BRC phase as discussed in [1]. In this scheme, the relay re-encodes the data from the two terminals separately after decoding, and retransmits a linear combination of them in the BRC phase. At each receiving terminal, the back-propagated signal part containing the known data from the receiving terminal itself is called *self-interference (SI)*, and it can be canceled before decoding. We proposed a practical SI-aided channel estimation scheme for two-way relaying systems in [9], where SI is used instead of pilot sequences to estimate the channels at the receivers in the BRC phase. There, the SI-aided channel estimation scheme is shown to be able to achieve similar bit-error rate (BER) performance as traditional pilot-aided channel estimation schemes in the considered scenario. Thus, the resources occupied by pilot sequences are freed, and the spectral efficiency can be further improved. Another way to re-encode the data at the relay is to combine the data on the bit level using the XOR operations (the *XOR scheme*) [10]. However, only conventional pilot-aided channel estimation schemes can be applied for this XOR scheme. So we only consider the superposition coding scheme in this paper.

The purpose of this paper is to quantify the spectral efficiency improvement by computing achievable rates of the bidirectional broadcast channel when the SI-aided channel estimation scheme is applied. The achievable rates serve as theoretical performance limits of the system and enable quantitative comparisons to traditional pilot-aided schemes. We consider a block-fading channel model for the BRC phase, i.e., the BRC phase contains a number of coherence intervals. Each coherence interval contains T time slots. The idea is to consider each time slot of the coherence interval as one use of T parallel channels with different signal-to-noise ratios (SNRs). The quality of the channel estimates determines the SNR for each of the parallel channels. Codes with different rates are allocated to those parallel channels according to the SNRs. We first exploit the SI to get an initial estimate of the channel and use it to decode the unknown data in the first T_s time slots of each coherence interval, where $T_\text{s} \geq N_\mathsf{R}$. After the unknown data in those time slots are fully decoded, the data-aided approach is applied, where the decoded data are utilized to re-estimate the channel and help to provide better channel estimates for the following time slots. As the channel estimation quality improves, codes with higher rates can be allocated in subsequent time slots of the coherence interval. We calculate the achievable rates for the bidirectional broadcast channel when the relay employs Gaussian codebooks or quadrature amplitude modulation (QAM) for retransmission.

The remainder of the paper is organized as follows: in Section II, we introduce the system model and briefly summarize the two-way relaying technique. The SI-aided channel estimation scheme is described in Section III. The achievable rates for systems employing SI-aided channel estimation schemes are compared with systems employing traditional pilot-aided schemes, and the simulation results are presented in Section IV. Finally, conclusions are drawn in Section V.

Notation: we use bold uppercase letters to denote matrices and bold lowercase letters to denote vectors. $\mathbf{I}_N$ is an $N \times N$ identity matrix. $\mathcal{CN}(0, \mathbf{K})$ denotes a circularly symmetric complex normal zero mean random vector with covariance matrix $\mathbf{K}$. Furthermore, $\mathbb{E}\{\cdot\}$, $\text{tr}(\cdot)$ and $(\cdot)^H$ stand for the expectation, the trace and the conjugate transpose, respectively. $\mathbb{E}_x\{y\}$ denotes the expectation of y with respect to the random variable x.

II. System Model

We consider a relaying system where two wireless terminals A and B exchange data via a half-duplex DF relay. We assume that there is no direct connection between terminal A and B (for example, due to shadowing or too large distance between them). The number of antennas at terminal A, the relay R and terminal B are denoted as N_A, N_R and N_B, respectively.

When the two-way relaying technique [1] is applied, the data of terminal A and B are exchanged in the MAC and BRC phases as shown in Fig. 1. In the MAC phase, terminal A and B transmit simultaneously to the relay. The data symbol vectors transmitted at terminal A and B in one time slot are denoted as $\mathbf{x}_\mathsf{A} \in \mathbb{C}^{N_\mathsf{A}\times 1}$ and $\mathbf{x}_\mathsf{B} \in \mathbb{C}^{N_\mathsf{B}\times 1}$, where $\mathbb{E}\{\mathbf{x}_\mathsf{A}\mathbf{x}_\mathsf{A}^H\} = \mathbf{I}_{N_\mathsf{A}}$ and $\mathbb{E}\{\mathbf{x}_\mathsf{B}\mathbf{x}_\mathsf{B}^H\} = \mathbf{I}_{N_\mathsf{B}}$. The received signal $\mathbf{y}_\mathsf{R}$ at the relay can be expressed as

$$\mathbf{y}_\mathsf{R} = \sqrt{\frac{Q_\mathsf{A}}{N_\mathsf{A}}}\mathbf{H}_1\mathbf{x}_\mathsf{A} + \sqrt{\frac{Q_\mathsf{B}}{N_\mathsf{B}}}\mathbf{H}_2\mathbf{x}_\mathsf{B} + \mathbf{v}_\mathsf{R} \tag{1}$$

where $\mathbf{H}_1 \in \mathbb{C}^{N_\mathsf{R}\times N_\mathsf{A}}$ and $\mathbf{H}_2 \in \mathbb{C}^{N_\mathsf{R}\times N_\mathsf{B}}$ are respectively the channel matrices from terminal A and B to the relay. Q_A and Q_B are the transmit power constraints at terminal A and B in the MAC phase. The additive noise vector at the relay is $\mathbf{v}_\mathsf{R} \sim \mathcal{CN}(0, \sigma^2\mathbf{I}_{N_\mathsf{R}})$. This is a conventional multiple access scenario. The receiver structure for decoding the data contained in $\mathbf{x}_\mathsf{A}$ and $\mathbf{x}_\mathsf{B}$ can be found in e.g., [11].

Since we focus on transmission schemes in the BRC phase, we assume the relay perfectly decodes what it receives in the MAC phase. In the BRC phase, we consider the superposition coding scheme [1]. The relay remodulates the decoded data from terminal A and B separately into symbol vectors $\mathbf{s}_\mathsf{A} \in \mathbb{C}^{N_\mathsf{R}\times 1}$ and $\mathbf{s}_\mathsf{B} \in \mathbb{C}^{N_\mathsf{R}\times 1}$, where $\mathbf{s}_\mathsf{A}$ and $\mathbf{s}_\mathsf{B}$ contain the same data as $\mathbf{x}_\mathsf{A}$ and $\mathbf{x}_\mathsf{B}$, respectively. Furthermore, we have $\mathbb{E}\{\mathbf{s}_\mathsf{A}\mathbf{s}_\mathsf{A}^H\} = \mathbf{I}_{N_\mathsf{R}}$ and $\mathbb{E}\{\mathbf{s}_\mathsf{B}\mathbf{s}_\mathsf{B}^H\} = \mathbf{I}_{N_\mathsf{R}}$. Here we assume the relay does not know the channel to the two terminals in the BRC phase. This is the case in FDD systems without channel feedback from the terminals. The relay then adds the two symbol vectors together and retransmits the sum vector:

$$\mathbf{s} = \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{s}_\mathsf{A} + \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{s}_\mathsf{B}. \tag{2}$$

In order to satisfy the power constraint, we require $P_\mathsf{A} + P_\mathsf{B} = P_\mathsf{R}$, where P_R is the transmit power constraint at the relay in the BRC phase. The modulation schemes and the power allocation of $\mathbf{s}_\mathsf{A}$ and $\mathbf{s}_\mathsf{B}$ used at the relay are known to both terminal A and B.

Note that $\mathbf{s}_\mathsf{A}$ is already known to terminal A, and $\mathbf{s}_\mathsf{B}$ is known to terminal B. The received signal part that contains the known data transmitted by the receiving terminal itself is the SI for the receiver. For example, the signal received at terminal A can be written as

$$\mathbf{y}_\mathsf{A} = \mathbf{H}_3\mathbf{s} + \mathbf{v}_\mathsf{A} \tag{3}$$

$$= \underbrace{\sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{H}_3\mathbf{s}_\mathsf{A}}_{\text{SI for terminal A}} + \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{H}_3\mathbf{s}_\mathsf{B} + \mathbf{v}_\mathsf{A}. \tag{4}$$

Similarly, the signal received at terminal B is

$$\mathbf{y}_\mathsf{B} = \mathbf{H}_4\mathbf{s} + \mathbf{v}_\mathsf{B} \tag{5}$$

$$= \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{H}_4\mathbf{s}_\mathsf{A} + \underbrace{\sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{H}_4\mathbf{s}_\mathsf{B}}_{\text{SI for terminal B}} + \mathbf{v}_\mathsf{B}. \tag{6}$$

Here $\mathbf{H}_3 \in \mathbb{C}^{N_\mathsf{A}\times N_\mathsf{R}}$ and $\mathbf{H}_4 \in \mathbb{C}^{N_\mathsf{B}\times N_\mathsf{R}}$ respectively denote the channel matrices from the relay to terminal A and B. $\mathbf{v}_\mathsf{A} \sim \mathcal{CN}(0, \sigma_\mathsf{A}^2\mathbf{I}_{N_\mathsf{A}})$ and $\mathbf{v}_\mathsf{B} \sim \mathcal{CN}(0, \sigma_\mathsf{B}^2\mathbf{I}_{N_\mathsf{B}})$ are the additive noise vectors at the receivers of terminal A and B, respectively.

The SI is "harmless" because it can be canceled at the receiver if the channel knowledge is available. The remaining part after canceling the SI only contains the unknown data transmitted from the other side. The decoding performance at the receivers is highly dependent on the accuracy of the channel knowledge. However, in reality the channel knowledge is not available at the receivers for free, and it has to be estimated. In traditional pilot-aided channel estimation schemes, the relay transmits orthogonal *pilot sequences* $\mathbf{S}_\mathrm{t}$ before transmitting data as shown in Fig. 2a. The pilot sequences occupy $T_\mathrm{t} \geq N_\mathsf{R}$ time slots, where $\mathbf{S}_\mathrm{t} \in \mathbb{C}^{N_\mathsf{R}\times T_\mathrm{t}}$ and $\mathbf{S}_\mathrm{t}\mathbf{S}_\mathrm{t}^H = T_\mathrm{t}\mathbf{I}_{N_\mathsf{R}}$. The receiver correlates the received signals with the pilot sequences and obtains the channel estimates.

Since the pilot sequences do not carry data information, pilot-aided channel estimation schemes waste part of system resources. On the other hand, we observe that the SI also contains the information about the channel. In the following, we derive an achievable rate of the system in the BRC phase when the SI-aided channel estimation scheme is applied. We only calculate the achievable rate at the receiver of terminal A, while the same discussions also apply to the receiver of terminal B.

III. Achievable Rates of SI-Aided Channel Estimation Scheme

We consider a block-fading channel model, i.e., the channels in the BRC phase remain constant for a coherence interval of T time slots and change independently between different coherence intervals. Each entry of the channel matrices $\mathbf{H}_3$ and $\mathbf{H}_4$ is an i.i.d. $\mathcal{CN}(0, 1)$ random variable. We assume that the transmitted data symbols from the relay are uncorrelated in space and time. This is easily satisfied by most communication systems since the data bits are usually interleaved before transmission to break the correlations between neighboring data. The idea of the SI-aided channel estimation scheme is

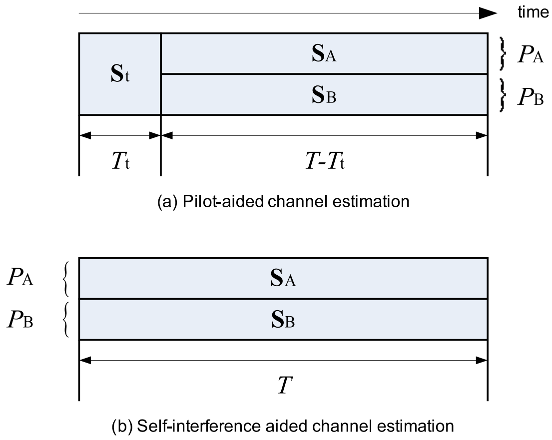

Fig. 2. Channel estimation schemes

to exploit the SI to get an initial estimate of the channel, and use the initial channel estimate to decode the unknown data in the first T_s time slots, where $T_\mathrm{s} \geq N_\mathsf{R}$. After the data in the first T_s time slots are decoded, the decoded data are utilized to re-estimate the channel and help to provide better channel estimates for decoding the data in the following time slots. As the channel estimates improve, higher rate codes can be allocated in subsequent time slots of the coherence interval.

A. Initial Channel Estimate Using SI

In the SI-aided channel estimation scheme as shown in Fig. 2b, no pilot sequence is used. To emphasize the time reference, we rewrite the received signal at terminal A as

$$\mathbf{y}_{\mathsf{A},k} = \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{H}_3\mathbf{s}_{\mathsf{A},k} + \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{H}_3\mathbf{s}_{\mathsf{B},k} + \mathbf{v}_{\mathsf{A},k} \tag{7}$$

where the index k denotes the kth time slot, $k \leq T_\mathrm{s}$. In order to estimate the channel $\mathbf{H}_3$ for decoding the data $\mathbf{s}_{\mathsf{B},k}$ at time slot k, we utilize the received signals in the remaining time slots $i \neq k, i \in \{1, \cdots, T\}$. As we will see later, estimating the channel only based on the received signals of the remaining time slots avoids the correlation between the estimated channel and the data $\mathbf{s}_{\mathsf{B},k}$, which facilitates the derivation of achievable rates of the system.

The signal received in the remaining time slots, i.e., the signals of the whole coherence interval except the kth time slot, can be expressed as

$$\bar{\mathbf{Y}}_\mathsf{A} = \mathbf{H}_3\left(\sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\bar{\mathbf{S}}_\mathsf{A} + \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\bar{\mathbf{S}}_\mathsf{B}\right) + \bar{\mathbf{V}}_\mathsf{A}, \tag{8}$$

$$= \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{H}_3\bar{\mathbf{S}}_\mathsf{A} + \underbrace{\sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{H}_3\bar{\mathbf{S}}_\mathsf{B} + \bar{\mathbf{V}}_\mathsf{A}}_{\mathbf{W}}. \tag{9}$$

Here $\bar{\mathbf{Y}}_\mathsf{A} = [\mathbf{y}_{\mathsf{A},1}, \cdots, \mathbf{y}_{\mathsf{A},k-1}, \mathbf{y}_{\mathsf{A},k+1}, \cdots, \mathbf{y}_{\mathsf{A},T}]$ is the $N_\mathsf{A} \times (T-1)$ received signal matrix. Similarly, $\bar{\mathbf{S}}_\mathsf{A}$, $\bar{\mathbf{S}}_\mathsf{B}$ and $\bar{\mathbf{V}}_\mathsf{A}$ are $N_\mathsf{R} \times (T-1)$ matrices denoting the signals transmitted from the relay and the noise matrices at terminal A in the whole coherence interval except the kth time slot, respectively.

In (9), $\bar{\mathbf{S}}_\mathsf{A}$ and $\bar{\mathbf{Y}}_\mathsf{A}$ are both known to the receiver of terminal A. Also based on the statistics of $\mathbf{W}$, we can get a first estimate of $\mathbf{H}_3$ and use it for decoding $\mathbf{s}_{\mathsf{B},k}$ in (7). By treating the unknown noisy symbol matrix $\mathbf{W}$ as noise and using the Bayesian Gauss-Markov Theorem in [12], we can obtain a linear minimum mean square error (LMMSE) estimate of the channel as

$$\begin{aligned}\bar{\mathbf{H}}_3 &= \bar{\mathbf{Y}}_\mathsf{A}\left[\frac{P_\mathsf{A}}{N_\mathsf{R}}\bar{\mathbf{S}}_\mathsf{A}^H\mathbf{R}_\mathbf{h}\bar{\mathbf{S}}_\mathsf{A}+\mathbf{R}_\mathbf{w}\right]^{-1}\left(\sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\bar{\mathbf{S}}_\mathsf{A}^H\right)\mathbf{R}_\mathbf{h} \\ &= \sqrt{\frac{N_\mathsf{R}}{P_\mathsf{A}}}\bar{\mathbf{Y}}_\mathsf{A}\left[\bar{\mathbf{S}}_\mathsf{A}^H\bar{\mathbf{S}}_\mathsf{A}+\frac{N_\mathsf{R}(P_\mathsf{B}+\sigma_\mathsf{A}^2)}{P_\mathsf{A}}\mathbf{I}_{T-1}\right]^{-1}\bar{\mathbf{S}}_\mathsf{A}^H. \quad (10)\end{aligned}$$

Here we have $\mathbf{R}_\mathbf{h} = \mathbb{E}\{\mathbf{h}\mathbf{h}^H\} = \mathbf{I}_{N_\mathsf{R}}$ and $\mathbf{R}_\mathbf{w} = \mathbb{E}\{\mathbf{w}\mathbf{w}^H\} = (P_\mathsf{B}+\sigma_\mathsf{A}^2)\mathbf{I}_{T-1}$, where $\mathbf{h}^H$ and $\mathbf{w}^H$ represent the rows of $\mathbf{H}_3$ and $\mathbf{W}$, respectively. In (10), we used the property that $\mathbf{H}_3$ and $\bar{\mathbf{S}}_\mathsf{B}$ are independent. In addition, we denote the estimation error as $\tilde{\mathbf{H}}_3 = \mathbf{H}_3 - \bar{\mathbf{H}}_3$. Let $\tilde{\mathbf{h}}^H$ denote the rows of $\tilde{\mathbf{H}}_3$. The covariance matrix of $\tilde{\mathbf{h}}$ is [12]

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{h}}} &= \mathbb{E}_{\mathbf{H}_3,\bar{\mathbf{S}}_\mathsf{A}}\left\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H\right\} && (11)\\ &= \mathbb{E}_{\bar{\mathbf{S}}_\mathsf{A}}\left\{\left[\mathbf{R}_\mathbf{h}^{-1}+\frac{P_\mathsf{A}}{N_\mathsf{R}}\bar{\mathbf{S}}_\mathsf{A}\mathbf{R}_\mathbf{w}^{-1}\bar{\mathbf{S}}_\mathsf{A}^H\right]^{-1}\right\} && (12)\\ &= \mathbb{E}_{\bar{\mathbf{S}}_\mathsf{A}}\left\{\left[\mathbf{I}_{N_\mathsf{R}}+\frac{P_\mathsf{A}}{N_\mathsf{R}(P_\mathsf{B}+\sigma_\mathsf{A}^2)}\bar{\mathbf{S}}_\mathsf{A}\bar{\mathbf{S}}_\mathsf{A}^H\right]^{-1}\right\}. && (13)\end{aligned}$$

The mean square error (MSE) of the estimated channel is $\sigma^2_{\tilde{\mathbf{H}}_3} = \operatorname{tr}\mathbf{R}_{\tilde{\mathbf{h}}}/N_\mathsf{R}$. When the entries of $\bar{\mathbf{S}}_\mathsf{A}$ and $\bar{\mathbf{S}}_\mathsf{B}$ are i.i.d. Gaussian, i.e., when we use a Gaussian codebook for transmitting data symbols, the MSE of the channel estimate can be expressed as [13]

$$\begin{aligned}\sigma^2_{\tilde{\mathbf{H}}_3} &= \frac{1}{N_\mathsf{R}}\operatorname{tr}\mathbf{R}_{\tilde{\mathbf{h}}} \\ &= \frac{1}{N_\mathsf{R}}\int_0^\infty\left[1+\frac{P_\mathsf{A}}{N_\mathsf{R}(P_\mathsf{B}+\sigma_\mathsf{A}^2)}x\right]^{-1} \\ &\quad\times\sum_{k=0}^{N_\mathsf{R}-1}\frac{k!\left[L_k^{T-1-N_\mathsf{R}}(x)\right]^2}{(k+T-1-N_\mathsf{R})!}x^{T-1-N_\mathsf{R}}e^{-x}dx \quad (14)\end{aligned}$$

where $L_j^i(x)$ are the associated Laguerre polynomials.

After we get the channel estimate $\bar{\mathbf{H}}_3$, we remove the SI from the received signal matrix as if the channel estimate $\bar{\mathbf{H}}_3$ is the real channel matrix. The remaining signal at time slot k can be expressed as

$$\begin{aligned}\bar{\mathbf{y}}_{\mathsf{A},k} &= \mathbf{y}_{\mathsf{A},k} - \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\bar{\mathbf{H}}_3\mathbf{s}_{\mathsf{A},k} && (15)\\ &= \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\bar{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k} \\ &\quad+\underbrace{\sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\tilde{\mathbf{H}}_3\mathbf{s}_{\mathsf{A},k}+\sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\tilde{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k}+\mathbf{v}_{\mathsf{A},k}}_{\mathbf{n}}. && (16)\end{aligned}$$

Since $\bar{\mathbf{H}}_3$ is the LMMSE estimation of the channel $\mathbf{H}$, the channel estimation error $\tilde{\mathbf{H}}_3$ has zero-mean entries and is uncorrelated with $\bar{\mathbf{H}}_3$ and $\mathbf{s}_{\mathsf{B},k}$. Thus, the noise term $\mathbf{n}$ in (16) is uncorrelated with the data $\mathbf{s}_{\mathsf{B},k}$ and has zero-mean entries. The noise variance at time slot k is

$$\begin{aligned}\sigma^2_\mathbf{n} &= \frac{1}{N_\mathsf{A}}\operatorname{tr}\mathbb{E}\left\{\mathbf{n}\mathbf{n}^H\right\} \\ &= \frac{1}{N_\mathsf{A}}\mathbb{E}\operatorname{tr}\left[\frac{P_\mathsf{A}}{N_\mathsf{R}}\tilde{\mathbf{H}}_3^H\tilde{\mathbf{H}}_3\mathbf{s}_{\mathsf{A},k}\mathbf{s}_{\mathsf{A},k}^H\right] \\ &\quad+\frac{1}{N_\mathsf{A}}\mathbb{E}\operatorname{tr}\left[\frac{P_\mathsf{B}}{N_\mathsf{R}}\tilde{\mathbf{H}}_3^H\tilde{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k}\mathbf{s}_{\mathsf{B},k}^H\right]+\frac{1}{N_\mathsf{A}}\mathbb{E}\operatorname{tr}\left[\mathbf{v}_{\mathsf{A},k}\mathbf{v}_{\mathsf{A},k}^H\right] \\ &= \frac{P_\mathsf{A}}{N_\mathsf{A}N_\mathsf{R}}\operatorname{tr}\left[\mathbb{E}\left(\tilde{\mathbf{H}}_3^H\tilde{\mathbf{H}}_3\right)\mathbb{E}\left(\mathbf{s}_{\mathsf{A},k}\mathbf{s}_{\mathsf{A},k}^H\right)\right] \\ &\quad+\frac{P_\mathsf{B}}{N_\mathsf{A}N_\mathsf{R}}\operatorname{tr}\left[\mathbb{E}\left(\tilde{\mathbf{H}}_3^H\tilde{\mathbf{H}}_3\right)\mathbb{E}\left(\mathbf{s}_{\mathsf{B},k}\mathbf{s}_{\mathsf{B},k}^H\right)\right]+\sigma_\mathsf{A}^2 \\ &= (P_\mathsf{A}+P_\mathsf{B})\sigma^2_{\tilde{\mathbf{H}}_3}+\sigma_\mathsf{A}^2. \quad (17)\end{aligned}$$

Here we used the fact that the entries in $\mathbf{s}_{\mathsf{A},k}$ and $\mathbf{s}_{\mathsf{B},k}$ are uncorrelated and also $\tilde{\mathbf{H}}_3$ is uncorrelated with $\mathbf{s}_{\mathsf{A},k}$ and $\mathbf{s}_{\mathsf{B},k}$. This is because $\bar{\mathbf{H}}_3$ is obtained using the received signals in time slots $i \neq k, i \in \{1,\cdots,T\}$. Eq. (16) describes a system with a known channel $\bar{\mathbf{H}}_3$ and noise $\mathbf{n}$ with variance $\sigma^2_\mathbf{n}$. According to [5, Theorem 1], the worst case noise $\mathbf{n}$ has a zero-mean Gaussian distribution. An achievable rate of such a system can be calculated by substituting $\mathbf{n}$ with a Gaussian noise with the same variance $\sigma^2_\mathbf{n}$.

B. Improving Channel Estimation by Data-Aided Approach

The initial channel estimation using SI is subject to the residual error due to the unknown data part $\bar{\mathbf{S}}_\mathsf{B}$. Thus the MSE of the channel estimate $\bar{\mathbf{H}}_3$ can still be high. On the other hand, after decoding the symbols in the first several time slots, we can re-estimate the channel by exploiting the decoded data. This is the commonly used *data-aided* approach. Note that the data-aided approach can only be started after the unknown data in the first $T_\mathrm{s} \geq N_\mathsf{R}$ time slots have been decoded. An initial channel estimate is indispensable to the data-aided approach.

Suppose the data in the first $k-1$ time slots have been decoded. Let

$$\mathbf{S} = \sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\mathbf{S}_{\mathsf{A}[k-1]}+\sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\mathbf{S}_{\mathsf{B}[k-1]} \quad (18)$$

where $\mathbf{S}_{\mathsf{A}[k-1]} = [\mathbf{s}_{\mathsf{A},1},\cdots,\mathbf{s}_{\mathsf{A},k-1}]$ and $\mathbf{S}_{\mathsf{B}[k-1]} = [\mathbf{s}_{\mathsf{B},1},\cdots,\mathbf{s}_{\mathsf{B},k-1}]$ denote the matrices composed of the transmitted signals in the first $k-1$ time slots. The received signal matrix in the first $k-1$ time slots is denoted as $\mathbf{Y} = [\mathbf{y}_{\mathsf{A},1},\cdots,\mathbf{y}_{\mathsf{A},k-1}]$. The MMSE channel estimate based on $\mathbf{S}$ and $\mathbf{Y}$ can be written as

$$\hat{\mathbf{H}}_3 = \mathbf{Y}\left(\sigma_\mathsf{A}^2\mathbf{I}_{k-1}+\mathbf{S}^H\mathbf{S}\right)^{-1}\mathbf{S}^H. \quad (19)$$

We also define the channel estimation error as $\check{\mathbf{H}}_3 = \mathbf{H}_3 - \hat{\mathbf{H}}_3$. Let $\check{\mathbf{h}}^H$ denote the rows of $\check{\mathbf{H}}_3$. The covariance matrix $\mathbf{R}_{\check{\mathbf{h}}}$ of $\check{\mathbf{h}}$ is

$$\mathbf{R}_{\check{\mathbf{h}}} = \mathbb{E}\left\{\check{\mathbf{h}}\check{\mathbf{h}}^H\right\} = \mathbb{E}\left\{\left(\mathbf{I}_{N_\mathsf{R}}+\frac{1}{\sigma_\mathsf{A}^2}\mathbf{S}\mathbf{S}^H\right)^{-1}\right\} \quad (20)$$

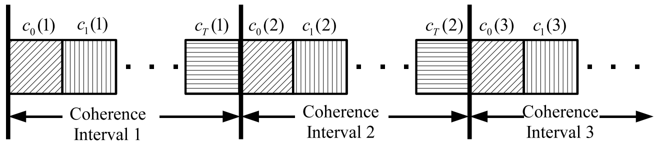

Fig. 3. Rate allocation in the time slots. Different hatching represents different code rate. Note coding is spread across many different coherence intervals.

where we used the property that each entry in $\mathbf{S}_{\mathsf{A}[k-1]}$ and $\mathbf{S}_{\mathsf{B}[k-1]}$ is uncorrelated in space and time. So the MSE of the channel estimation error is

$$\sigma^2_{\check{\mathbf{H}}_3} = \frac{1}{N_\mathsf{R}} \operatorname{tr} \mathbb{E}\left\{\check{\mathbf{h}}\check{\mathbf{h}}^H\right\}. \tag{21}$$

After canceling the SI, the received signal vector at time slot k is

$$\hat{\mathbf{y}}_{\mathsf{A},k} = \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\hat{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k} + \underbrace{\sqrt{\frac{P_\mathsf{A}}{N_\mathsf{R}}}\check{\mathbf{H}}_3\mathbf{s}_{\mathsf{A},k} + \sqrt{\frac{P_\mathsf{B}}{N_\mathsf{R}}}\check{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k} + \mathbf{v}_{\mathsf{A},k}}_{\check{\mathbf{n}}}. \tag{22}$$

This describes a system with known channel $\hat{\mathbf{H}}_3$ and noise term $\check{\mathbf{n}}$. We have the covariance matrix of the noise vector $\check{\mathbf{n}}$ as

$$\mathbb{E}\left(\check{\mathbf{n}}\check{\mathbf{n}}^H\right) = \left[(P_\mathsf{A}+P_\mathsf{B})\sigma^2_{\check{\mathbf{H}}_3} + \sigma^2_\mathsf{A}\right]\mathbf{I}_{N_\mathsf{A}}. \tag{23}$$

So the noise variance is

$$\sigma^2_{\check{\mathbf{n}}} = (P_\mathsf{A}+P_\mathsf{B})\sigma^2_{\check{\mathbf{H}}_3} + \sigma^2_\mathsf{A}. \tag{24}$$

Again, for a system described in (22), the worst case for noise term $\check{\mathbf{n}}$ is when $\check{\mathbf{n}}$ is Gaussian. An achievable rate can be calculated by assuming $\check{\mathbf{n}} \sim \mathcal{CN}\left(0, \sigma^2_{\check{\mathbf{n}}}\mathbf{I}_{N_\mathsf{A}}\right)$.

C. Achievable Rates of Bidirectional Broadcast Channel

In this section, we derive achievable rates of the bidirectional broadcast channel when the SI-aided channel estimation scheme is applied. The idea is that we can allocate codes with different rates on different time slots of the coherence intervals, as shown in Fig. 3. Each time slot of a coherence interval can be considered as one use of T parallel channels with different SNRs. At the beginning of each coherence interval, only SI is available for channel estimation and the system model is shown in (16). Here the effective SNR is low. Low rate codes are allocated at those time slots so that they can be fully decoded after many channel uses, i.e., coherence intervals. After the data at the beginning time slots of each coherence interval are decoded, the data-aided approach can be used to improve the channel estimation and the effective SNR, where the system model is shown in (22). Thus higher rate codes can be allocated at those time slots. Since we assume the channel changes independently between different coherence intervals, each coherence interval can be considered as a realization of the T parallel channels. According to channel coding theorem in a fast fading channel, those codes can be decoded *without error* after many independent realizations of those parallel channels. This method is also used in [8] to derive an achievable rate of the pilot-embedding schemes when data-aided approach is applied. Note this decoding scheme is different from [9], where we were interested in a practical scheme for improving the BER performance and the coded bits were spread across the whole coherence interval. There we utilized the error-correcting capability of convolutional codes in each iteration to correct the errors contained in the unknown data, and the decoded data might contain errors in each iteration.

Assuming the initial SI estimated channel $\bar{\mathbf{H}}_3$ is used to decode the data in the first T_s time slots, the following expression gives an average achievable rate expression for the data decoded at terminal A in the BRC phase

$$C_\mathsf{A} = \frac{0.5}{T}\left\{\sum_{k=1}^{T_\mathrm{s}} I(\mathbf{y}_{\mathsf{A},k};\mathbf{s}_{\mathsf{B},k}|\bar{\mathbf{H}}_3) + \sum_{k=T_\mathrm{s}+1}^{T} I(\mathbf{y}_{\mathsf{A},k};\mathbf{s}_{\mathsf{B},k}|\hat{\mathbf{H}}_3)\right\}$$

where the factor 0.5 is due to the fact that the rate can only be achieved after both the MAC and BRC phases, and the length of the two phases are equal. In the expression of C_A, the first term is the mutual information conveyed by the data symbols decoded by using the channel estimated purely by SI, and the second term represents the mutual information conveyed by the data symbols decoded by using the data-aided approach. When a Gaussian codebook is used, we can write

$$C_\mathsf{A}^{\mathrm{Gau}} = \frac{0.5T_\mathrm{s}}{T}\mathbb{E}\log\det\left(\mathbf{I}_{N_\mathsf{A}} + \rho_0\frac{\bar{\mathbf{H}}_3\bar{\mathbf{H}}_3^H}{N_\mathsf{R}}\right) + \frac{0.5}{T}\sum_{k=T_\mathrm{s}+1}^{T}\mathbb{E}\log\det\left(\mathbf{I}_{N_\mathsf{A}} + \rho_k\frac{\hat{\mathbf{H}}_3\hat{\mathbf{H}}_3^H}{N_\mathsf{R}}\right) \tag{25}$$

where $\rho_0 = P_\mathsf{B}/\sigma^2_\mathbf{n}$ and $\rho_k = P_\mathsf{B}/\sigma^2_{\check{\mathbf{n}}}$ according to (17) and (24). However, the optimum choice of codebook of $\mathbf{s}_{\mathsf{A},k}$ and $\mathbf{s}_{\mathsf{B},k}$ is still an open question. This is because Gaussian codebooks maximize the mutual information only if the channel is perfectly known at the receiver. When the channel has to be estimated, Gaussian codebooks do not necessarily lead to the lowest $\sigma^2_{\tilde{\mathbf{H}}_3}$ in the channel estimation in (16).

For QAM modulations, no explicit expression for mutual information is available. However, it can be obtained by Monte Carlo simulations by using the following expression

$$C_\mathsf{A}^{\mathrm{QAM}} = \frac{0.5T_\mathrm{s}}{T}\mathbb{E}\left\{\log\frac{p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\bar{\mathbf{H}}_3)}{\sum_{\mathbf{s}_{\mathsf{B},k}} p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\bar{\mathbf{H}}_3)\cdot p(\mathbf{s}_{\mathsf{B},k})}\right\} + \frac{0.5}{T}\sum_{k=T_\mathrm{s}+1}^{T}\mathbb{E}\left\{\log\frac{p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\hat{\mathbf{H}}_3)}{\sum_{\mathbf{s}_{\mathsf{B},k}} p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\hat{\mathbf{H}}_3)\cdot p(\mathbf{s}_{\mathsf{B},k})}\right\}.$$

Here it is assumed that $\mathbf{s}_{\mathsf{B},k}$ is chosen from a QAM constellation with equal probabilities. According to (16) and (22), we have

$$p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\bar{\mathbf{H}}_3) = \frac{1}{(\pi\sigma^2_\mathbf{n})^{N_\mathsf{A}}}\exp\left(-\frac{1}{\sigma^2_\mathbf{n}}\|\mathbf{y}_{\mathsf{A},k} - \bar{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k}\|^2\right),$$

$$p(\mathbf{y}_{\mathsf{A},k}|\mathbf{s}_{\mathsf{B},k},\hat{\mathbf{H}}_3) = \frac{1}{(\pi\sigma^2_{\check{\mathbf{n}}})^{N_\mathsf{A}}}\exp\left(-\frac{1}{\sigma^2_{\check{\mathbf{n}}}}\|\mathbf{y}_{\mathsf{A},k} - \hat{\mathbf{H}}_3\mathbf{s}_{\mathsf{B},k}\|^2\right).$$

Here we choose the noise vectors $\mathbf{n} \sim \mathcal{CN}\left(0, \sigma^2_\mathbf{n}\mathbf{I}_{N_\mathsf{A}}\right)$ and $\check{\mathbf{n}} \sim \mathcal{CN}\left(0, \sigma^2_{\check{\mathbf{n}}}\mathbf{I}_{N_\mathsf{A}}\right)$ in order to calculate the achievable rates following the discussions in Section III.

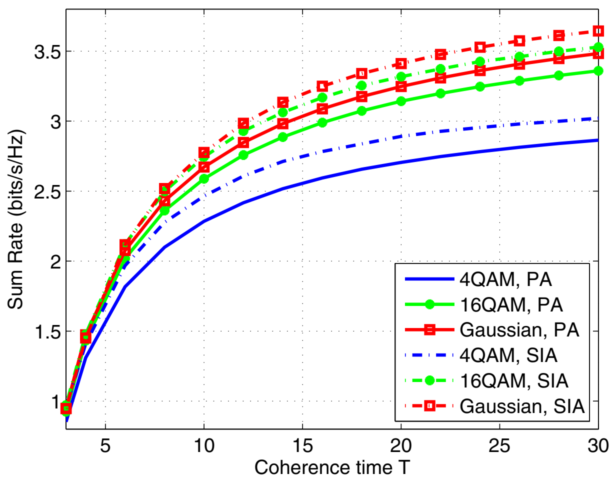

Fig. 4. Average achievable sum rates in the BRC phase of two-way relaying systems ($N_{\mathsf{A}} = N_{\mathsf{R}} = N_{\mathsf{B}} = 2$, SNR $= P_{\mathsf{R}}/\sigma_{\mathsf{A}}^2 = P_{\mathsf{R}}/\sigma_{\mathsf{B}}^2 = 10$dB). "PA" and "SIA" denote pilot-aided and SI-aided channel estimation, respectively.

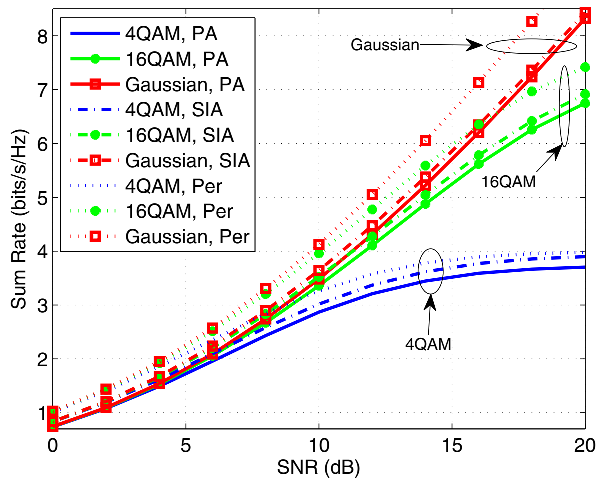

Fig. 5. Average achievable sum rates in the BRC phase of two-way relaying systems ($N_{\mathsf{A}} = N_{\mathsf{R}} = N_{\mathsf{B}} = 2$, the coherence time $T = 30$). "Per" denotes the achievable rates when channel knowledge is perfectly available at the receivers.

IV. Simulation Results

In this section, we compare the average achievable rates of the SI-aided channel estimation scheme with that of the traditional pilot-aided channel estimation scheme by using Monte Carlo simulations. The data-aided approach is utilized to improve the channel estimation in both cases. We consider a two-way relaying system where $N_{\mathsf{A}} = N_{\mathsf{R}} = N_{\mathsf{B}} = 2$. We only consider the BRC phase and equal power is allocated for transmission to terminal A and B, i.e., $P_{\mathsf{A}} = P_{\mathsf{B}} = P_{\mathsf{R}}/2$. The first N_{R} rows of the Hadamard matrix are taken as the pilot sequences in the pilot-aided scheme. In Fig. 4, SNR $= P_{\mathsf{R}}/\sigma_{\mathsf{A}}^2 = P_{\mathsf{R}}/\sigma_{\mathsf{B}}^2 = 10$dB, and we plot the average sum rate of terminal A and B. We can observe that the achievable rate increases with the coherence time. On the one hand, this is because the initial channel estimate gets better in the SI-aided channel estimation when the coherence interval gets longer; on the other hand, better channel estimate can be obtained at the end of each coherence interval in the data-aided approach as the coherence time increases. In Fig. 5, we show how the average achievable rate changes with the SNR for $T = 30$, where SNR $= P_{\mathsf{R}}/\sigma_{\mathsf{A}}^2 = P_{\mathsf{R}}/\sigma_{\mathsf{B}}^2$. We can observe that the gains of the SI-aided channel estimation scheme remain nearly as constants in the considered SNR range. This is due to the fact that the received power of both SI and the noise term $\mathbf{W}$ in (9) increase as SNR increases. In 4QAM modulations, we can also observe that the SI-aided channel estimation can achieve nearly the same rates at high SNR as when perfect channel knowledge is available at the receivers.

V. Conclusions

We derived achievable rate expressions for bidirectional broadcast channels of two-way relaying systems when the SI-aided channel estimation scheme is applied. This scheme utilizes the SI to obtain an initial estimate of the channel and use data-aided approaches to improve the channel estimates. Simulation results showed that our scheme can achieve higher rates compared to traditional pilot-aided channel estimation schemes. The gain is relatively high in the low SNR regime. In high SNR regime, the performance of the SI-aided channel estimation scheme for 4QAM modulations is nearly as good as when perfect channel knowledge is available at the receivers.

References

[1] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels," *IEEE J. Select. Areas Commun.*, vol. 25, pp. 379–389, Feb. 2007.

[2] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–702, June 2003.

[3] T. J. Oechtering, C. Schnurr, I. Bjelakovic, and H. Boche, "Broadcast capacity region of two-phase bidirectional relaying," *IEEE Trans. Inform. Theory*, vol. 54, pp. 454–458, Jan. 2008.

[4] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Processing*, vol. 54, pp. 884–893, Mar. 2006.

[5] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links," *IEEE Trans. Inform. Theory*, vol. 49, pp. 951–963, Apr. 2003.

[6] H. Zhu, R.-R. Chen, and B. Farhang-Boroujeny, "Capacity of pilot-aided MIMO communication systems," in *Proc. IEEE Int. Symposium on Inf. Theory*, (Chicago, IL), p. 544, June 27 – July 2, 2004.

[7] M. Coldrey and P. Bohlin, "Training-based MIMO systems – part I: Performance comparison," *IEEE Trans. Signal Processing*, vol. 55, pp. 5464 – 5476, Nov. 2007.

[8] M. Coldrey and P. Bohlin, "Training-based MIMO systems: Part II – improvements using detected symbol information," *IEEE Trans. Signal Processing*, vol. 56, pp. 296 – 303, Jan. 2008.

[9] J. Zhao, M. Kuhn, A. Wittneben, and G. Bauch, "Self-interference aided channel estimation in two-way relaying systems," in *Proc. IEEE Global Comm. Conf. (GLOBECOM)*, (New Orleans, LA), Nov. 30 – Dec. 4, 2008.

[10] Y. Wu, P. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *Proc. of 39th Annual Conf. on Information Sciences and Systems (CISS)*, (Baltimore, MD), Mar. 2005.

[11] S. Verdú, *Multiuser Detection*. Cambridge University Press, Aug. 1998.

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

[13] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.