Oblivious Transfer Over Wireless Channels

Jithin Ravi, Bikash Kumar Dey, *Member, IEEE*, and Emanuele Viterbo, *Fellow, IEEE*

***Abstract*—We consider the problem of oblivious transfer (OT) over OFDM and MIMO wireless communication systems where only the receiver knows the channel state information. The sender and receiver also have unlimited access to a noise-free real channel. Using a physical layer approach, based on the properties of the noisy fading channel, we propose a scheme for honest-but-curious parties that enables the transmitter to send obliviously one-of-two files, i.e., without knowing which one has been actually requested by the receiver, while also ensuring that the receiver does not get any information about the other file.**

***Index Terms*—MIMO, oblivious transfer, OFDM, physical layer security, secure function computation.**

I. Introduction

CONSIDER a movie server, or a server of medical database. A subscriber wants a specific item (a movie, or information about a specific disease) without the server being able to know which item is desired by the subscriber. The subscriber is also not allowed to gain any significant information about any other item. This is an example of oblivious transfer.

In one-out-of-two string oblivious transfer (OT), one party, Alice, has two files and the other party, Bob, wants one of these files. Bob needs to obtain the required file without Alice finding out the identity of the file chosen by him. Bob should also not be able to recover any significant information about the other file. Alice and Bob are assumed to be “honest-but-curious”, or passive, participants - they follow the agreed protocol but are also curious to gain additional knowledge of the other’s data from their own observations during the protocol [1], [2].

OT has been studied in various forms for some time in cryptography [3], [4]. It is a special case of secure function computation problems, where multiple parties want to compute a function without revealing additional information about their data to other parties. It was shown by Kilian [5] that an OT protocol can be used as a subroutine to devise a protocol for two-party secure function computation for any function that is representable by a boolean circuit.

It is well known that OT can not be performed only by interactive communication over a noise-free channel. The OT is thus studied with a noisy channel as a critical resource in addition to unlimited access to a noise-free channel. The OT capacity is the largest length of file that can be transferred, per use of the noisy channel, between Alice and Bob. The OT capacity was defined in [6], [2] as an extension of the concept from [7]. In [1], [2], one-out-of-two string OT has been studied when the noisy channel between Alice and Bob is a Discrete Memoryless Channel (DMC). An upper bound for the OT capacity of a DMC was given in [1] and it was shown that the given upper bound is achievable by a simple scheme for binary erasure channels (BEC). Multi-user variants of OT have been studied over broadcast erasure channels in [8], [9]. OT has been considered with active or malicious participants in [10].

One-out-of-two string OT has been considered in the context of AWGN channels in [11], where a protocol was proposed. The case of fast fading wireless channels has also been discussed in [11], where the fading state varies in each transmission and is not known to the transmitter or the receiver. Under such assumption, the channel can be modeled by the conditional probability distribution $p_{Y|X}$ with the channel state marginalized. The fading state does not directly provide any additional advantage in OT here, other than through its influence on $p_{Y|X}$. The OT capacity is not known for many important channels including AWGN and binary symmetric channels.

In this paper, we consider OT under honest-but-curious setting over two classes of wireless slow-fading channels: orthogonal frequency division multiplexing (OFDM) channel and multiple input multiple output (MIMO) channel, where the fading state information is available only at the receiver (CSIR), [12]. Channels with CSIR (Fig. 1) have not been considered for OT before to the best of our knowledge. CSIR is a common assumption in wireless communication which can be made when the coherence block length n is sufficiently large. We allow an interactive protocol to run over n uses of the channel during which the channel state remains fixed, and in that period the noise-free channel can be used any finite number of times. In other words, we assume that one run of the OT protocol is completed in one coherence block. However, following common principle of rate-adaptation used in many wireless communication models, the OT rate may vary from block to block depending on the channel state. As we will see in our schemes, the knowledge of the state only at the receiver is the key to some interesting techniques for OT. Our techniques have the flavor of the protocol for BECs [1].

Communication under secrecy constraints has been studied by many authors (see [13]). In particular, private

Manuscript received July 20, 2015; revised November 25, 2015; accepted December 26, 2015. Date of publication January 8, 2016; date of current version March 15, 2016. This paper was presented in part at the Information Theory Workshop, ITW 2015, Jerusalem. The work of R. Jithin and B. K. Dey was supported in part by the Department of Science and Techonolgy, Govt. of India, under Grant SB/S3/EECE/057/2013, in part by the Information Technology Research Academy, India, under Grant ITRA/15(64)/Mobile/USEAADWN/01, and in part by the Bharti Centre for Communication at IIT Bombay. The work of E. Viterbo was supported by the Australian Research Council through the Discovery Project under Grant DP130100336. The associate editor coordinating the review of this paper and approving it for publication was V. Y. F. Tan.

R. Jithin and B. K. Dey are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India (e-mail: rjithin@ee.iitb.ac.in; bikash@ee.iitb.ac.in).

E. Viterbo is with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Vic. 3800, Australia (e-mail: emanuele.viterbo@monash.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2016.2515593

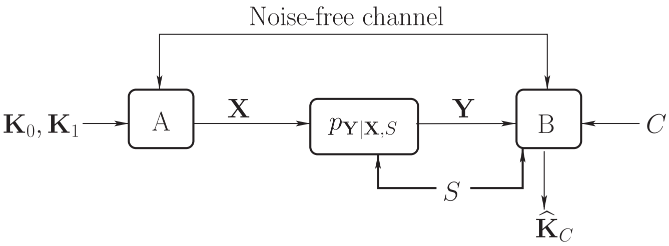

Fig. 1. Communication setup for oblivious transfer over channels with state.

communication over a wiretap channel in the presence of eavesdropper has been studied extensively [14], [15], [16], [17], [18], [19]. In this work, we make use of coding techniques for Gaussian wiretap channels as a building block for our achievability schemes.

In both OFDM and MIMO, we rely on the modeling of the channel as parallel fading channels. For the MIMO setup, this is done using the SVD precoder matrix that is communicated by Bob to Alice. The parallel channels are grouped in pairs. OT is performed independently at different rates over different pairs. For a MIMO channel, an apparently more general protocol would be to reduce the channel into a pair of MIMO channels, and then to perform OT over these parallel MIMO channels using codes for MIMO wiretap channels. However, we show that the rates achieved by such a protocol is upper bounded by the rates achieved using our technique.

We show (Theorem 1) that the best pairing of the parallel channels is that of the strongest channel with the weakest, and so on with the rest of the channels. The idea of pairing good and bad subchannels in OFDM and SVD-precoded MIMO was also used in [20], [21] with the aim of designing signal sets that minimize error probability or maximize mutual information. Here, we exploit subchannel pairing to guarantee that Alice is oblivious to which file is requested and that Bob only receives one of the two files. We also derive the optimal power allocation among the pairs of channels.

The paper is organized as follows. Section II presents the problem definition and the system model for both OFDM and MIMO channels. In Section III, we present protocols for OT over 2-channels OFDM, 2 × 2 MIMO and 2 × 1 MIMO channels. We present the general protocol for $2N$-channels OFDM and $2N \times n_B$ MIMO models in Section IV, following a common principle. Optimization of our protocol is discussed in Section VI. High SNR asymptotics of OT rate for our protocol is analyzed in Section VII. We provide simulation results of our OT scheme for simple OFDM and MIMO channels in Section VIII. Finally, we conclude the paper in Section X.

II. System Model

Alice (A) and Bob (B) are two parties in the system as shown in Fig. 1. Alice has two binary strings $\mathbf{K}_0, \mathbf{K}_1$ of equal length, and Bob wants one of these strings $\mathbf{K}_C$ where $C \in \{0, 1\}$ is Bob's *choice bit*. We assume that all the bits in $(\mathbf{K}_0, \mathbf{K}_1, C)$ are i.i.d. $\sim Ber(1/2)$. Alice can communicate with Bob over a channel $p_{Y|X,S}$ with state S, where the state remains fixed over a large block length n, and varies from block to block in an i.i.d. manner. The state is known to Bob at the beginning of a block. This models wireless communication setups, where in a large coherence block of length n, the fading state remains fixed, and the fading state is known (estimated) by the receiver. This is commonly known as the *quasi-static channel model* [12], [13]. In addition to this channel, there is also a noise-free channel over which Alice and Bob can communicate real numbers between each other without any error/distortion. During each block, the noise-free channel can be used any finite number of times. The length $L(S)$ of $\mathbf{K}_0, \mathbf{K}_1$ depends on S. Since Bob knows the state S at the beginning of a block, he is assumed to compute and communicate $L(S)$ to Alice over the noise-free channel. The goal of a protocol is to transfer $\mathbf{K}_C$ to Bob obliviously, within the current block, such that Bob has negligible knowledge about $\mathbf{K}_{\overline{C}}$, and Alice has no knowledge about C (*perfect secrecy* against Alice).

Our setup can also be used to transfer large files. We then need multiple coherence blocks to complete the OT session for one pair of files. The two files can be broken into multiple chunks to form one pair $(\mathbf{K}_{0i}, \mathbf{K}_{1i})$ for each block i. Then one run of the protocol is performed in each block, where the choice bit C of Bob remains the same over the whole session involving many runs of the protocol.

An $(n, L(\cdot))$ OT protocol is parameterized by the number n of channel uses and by a function $L(\cdot)$ of the state S. There are a total of k rounds of communication between Alice and Bob, including communication over both the noisy and noise-free channels. These are indexed by $1, 2, \cdots, k$, where k can be random and can be dependent on S. But for every S, it is required to be finite with probability 1. The noisy channel is used at rounds $i_1, i_2, \cdots, i_n \in \{1, \cdots, k\}$. At every round before round i_1, between consecutive i_j and i_{j+1}, and after round i_n, Alice and Bob exchange a sequence of real numbers over the noise-free channel. In the following, X_i and Y_i denote respectively the input and the output of the noisy channel at time index i. In the following description of the protocol, we denote $\mathbf{Y}^i := (Y_1, Y_2, \cdots, Y_i)$ for any positive integer i. $\mathbf{E}^i, \mathbf{F}^i$ are also similarly defined. In the rest of the paper, we also denote the transmitted length-n vector by $\mathbf{X}$. The length-n vector transmitted by the l-th antenna (in case of MIMO) or over the l-th subchannel (in case of OFDM) will be denoted by $\mathbf{X}_l = (X_{l1}, X_{l2}, \cdots, X_{ln})$.

A. The structure of an $(n, L(\cdot))$ Protocol

1) Alice has two bit-strings $\mathbf{K}_0, \mathbf{K}_1$ of length $L(S)$ each, and Bob has a choice bit C. $\mathbf{K}_0, \mathbf{K}_1$ can be substrings of two larger strings available with Alice, and their length $L(S)$ is computed by Alice based on some information about S sent by Bob during the protocol.
2) Alice and Bob generate private random variables W_A, W_B, respectively.
3) For $i_j < i < i_{j+1}$ for every $j = 0, 1, \cdots, n$ (assuming $i_0 = 0$ and $i_{n+1} = k + 1$), Alice sends $E_i = E_i(\mathbf{K}_0, \mathbf{K}_1, W_A, \mathbf{F}^{i-1})$ and Bob sends $F_i = F_i(C, S, W_B, \mathbf{E}^{i-1}, \mathbf{Y}^j)$ over the noise-free channel. Here $F^0 = E^0 = Y^0 = \emptyset$.
4) For $i = i_j$, Alice transmits $X_j = X_j(\mathbf{K}_0, \mathbf{K}_1, W_A, \mathbf{F}^{i_j - 1})$ over the noisy channel and Bob receives Y_j. There is

no communication over the noise-free channel in these rounds, and thus $E_i = F_i = \emptyset$.

5) At the end of the protocol, Bob computes $\widehat{\mathbf{K}}_C = \widehat{\mathbf{K}}(C, S, W_B, \mathbf{E}^k, \mathbf{Y}^n)$.

The rate $L(s)/n$ of a protocol as described above is a function of the state s, and is denoted by $R(s)$.

Definition 1: A non-negative rate function $R(s)$ is said to be achievable if there is a sequence of $(n, L^{(n)}(\cdot))$-protocols such that for every state s, $\frac{L^{(n)}(s)}{n} \to R(s)$ as $n \to \infty$, and the protocols satisfy the conditions

- Reliability condition:

$$P\left(\widehat{\mathbf{K}}_C \neq \mathbf{K}_C\right) \to 0 \tag{1}$$

- Perfect secrecy condition on Bob's choice bit:

$$I\left(\mathbf{K}_0\mathbf{K}_1 W_A \mathbf{F}^k; C\right) = 0 \tag{2}$$

- Strong secrecy condition on the nonchosen string:

$$I\left(CSW_B\mathbf{Y}^n\mathbf{E}^k; \mathbf{K}_{\overline{C}}\right) \to 0. \tag{3}$$

The definition assumes the honest-but-curious (or passive) model. Here Alice's and Bob's views are respectively $\mathbf{K}_0\mathbf{K}_1 W_A \mathbf{F}^k$ and $CSW_B\mathbf{Y}^n\mathbf{E}^k$. The perfect secrecy condition (2) for the choice bit is necessary because the protocol may be used many times in a session to transfer one of two large files in smaller chunks, while keeping the choice bit the same. Even then the prefect secrecy of C guarantees that there is no leakage of information to Alice about C. On the other hand, requiring only strong secrecy would have resulted in accumulation of the leakage over a large number of runs of the protocol.

The average rate R is the expectation of $R(S)$. The OT capacity is the supremum of all achievable average OT rates.

B. Gaussian Wiretap Channel

Wiretap channel has been studied as a standard model for communication in the presence of an eavesdropper [14], [15]. We model our MIMO and OFDM channels as complex channels. If Alice and Bob are respectively the transmitter and receiver of a complex AWGN channel, and if Eve is a wiretapper, whose received symbol is more noisy than that of Bob (degraded channel assumption), then the *secrecy capacity* of the wiretapper channel is given by

$$\mathcal{C}\left(\frac{P}{\sigma_B^2}, \frac{P}{\sigma_E^2}\right) = \log_2\left(1 + \frac{P}{\sigma_B^2}\right) - \log_2\left(1 + \frac{P}{\sigma_E^2}\right) \tag{4}$$

where σ_B^2 and σ_E^2 are the variance of the noise at Bob and Eve, respectively, and P is the transmit power [15], [22]. Encoding for such channels involves mixing the message with some random bits (with rate equaling the capacity of the wiretapper) before encoding for the complex AWGN channels. Bob can decode both the message and the random bits as the total rate of these is below his capacity, whereas the random bits completely hide the message from Eve. Eve gets almost no information about the message [16]. We will denote this channel with power constraint P as $\mathcal{WT}(P, \frac{P}{\sigma_B^2}, \frac{P}{\sigma_E^2})$. Practical coding schemes approaching the secrecy capacity have been proposed for discrete memoryless channels using polar codes [23] and for the Gaussian channel based on lattice codes [24], under semantic security.

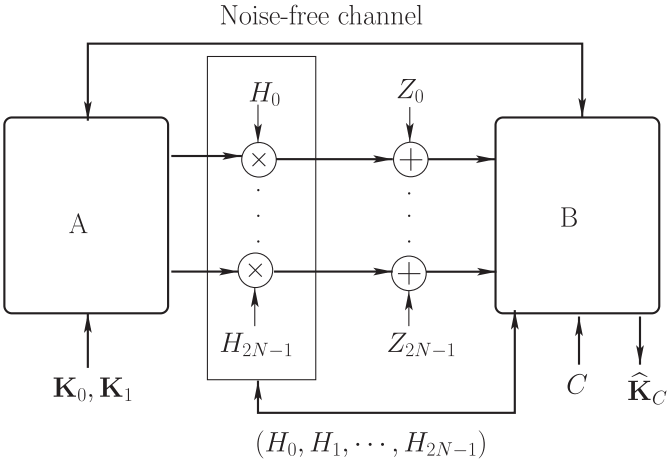

Fig. 2. The OT setup with independent parallel channels.

In this paper we consider two channels with states, OFDM and MIMO, as discussed below. The essential technique used for OT over both these setups is the same.

C. The OFDM Setup

The OFDM setup is modeled in Fig. 2 as $2N$ parallel fading AWGN channels between Alice and Bob. The channel states are given by independent fading coefficients $H_0, H_1, \cdots, H_{2N-1}$. If the vector $\mathbf{X}_l = (X_{l1}, X_{l2}, \cdots, X_{ln})$ is transmitted in n channel uses over the l-th channel for $l = 0, 1, \cdots, 2N-1$, then the received vector over the l-th channel is given by

$$\mathbf{Y}_l = H_l\mathbf{X}_l + \mathbf{Z}_l,$$

where $\mathbf{Z}_l$ is the noise with i.i.d. real and imaginary parts $\sim \mathcal{N}(0, 1/2)$. We assume that H_l are i.i.d. with Rayleigh distribution. The channel gains remain fixed for a block of length n, and change from block to block in an i.i.d. manner. We assume that they are known to Bob in the beginning of the block. The average transmitted power in any block is restricted to P, i.e., $\sum_{l=0}^{2N-1}\sum_{j=1}^{n}|X_{lj}|^2 \leq nP$.

D. The MIMO Setup

Let us consider the MIMO system with transmitter Alice and receiver Bob, as shown in Fig. 3. The transmitter has n_A antennas and the receiver has n_B antennas. We assume that n_A is even. Let $\mathbf{X} = (X_{lj})_{\substack{0 \leq l \leq n_A - 1 \\ 1 \leq j \leq n}}$ denote the complex matrix transmitted by Alice over n uses of the MIMO channel. The received matrix $\mathbf{Y}$ is given by

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{Z} \tag{5}$$

where $\mathbf{Z} \in \mathbb{C}^{n_B \times n}$ is the complex Gaussian noise matrix with all entries having i.i.d. real and imaginary parts $\sim \mathcal{N}(0, 1/2)$

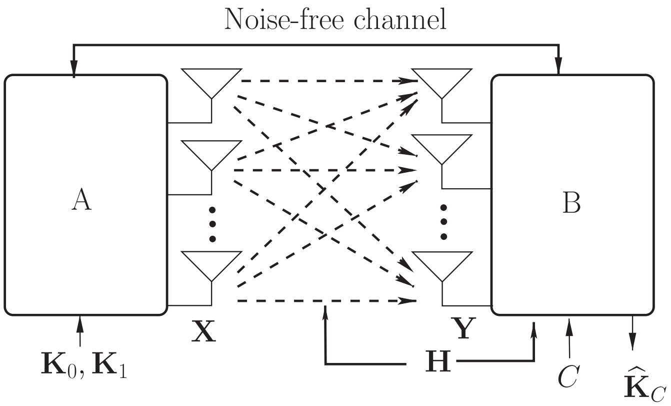

Fig. 3. MIMO system for oblivious transfer.

and $\mathbf{H} \in \mathbb{C}^{n_B \times n_A}$ represents the complex channel fading matrix. The entries of $\mathbf{H}$ are assumed to be i.i.d. complex random variables with independent real and imaginary parts $\sim \mathcal{N}(0, 1/2)$. $\mathbf{H}$ remains fixed over the block of length n, and changes in an i.i.d. manner from block to block. The average transmit power in any block is constrained to be P, i.e., $\sum_{l=0}^{n_A-1} \sum_{j=1}^{n} |X_{lj}|^2 \leq nP$. We assume that $\mathbf{H}$ is known only to Bob in the beginning of each block.

III. The Protocol: Some Examples

We now show our OT protocols for some simple examples to illustrate the basic principle. In all the three examples, Bob reveals some partial information about the channel state to Alice so that there are, in effect, two parallel channels with different SNRs, and Alice does not know which of them is the better channel. Bob reveals the channel over which each file is to be communicated – the desired file over the stronger channel, and the other file over the weaker channel. Alice uses encoding for a suitable wiretap channel so that Bob can decode the file transmitted over the stronger channel, but not the file transmitted over the weaker channel.

In Sec. III-A, we discuss our protocol for the simplest case of 2-channels OFDM. In Sec. III-B, we discuss our protocol for a 2×2 MIMO channel, where the MIMO channel is essentially reduced to a pair of parallel channels with order uncertainty of the gains for Alice. Finally, in Sec. III-C, we discuss our protocol for a 2×1 MISO channel, where the channel is reduced to a pair of parallel channels, with one of them having zero gain. In this case, we do not require encoding for wiretap channels. These will be extended in Sec. IV for a general MIMO channel.

A. 2-Channels OFDM

Let us consider an OFDM setup with 2 subchannels, each of which undergo independent and identical Rayleigh fading. For a state (H_0, H_1), let us define

$$B = \arg\max_j |H_j|.$$

In the following, we will show that the rate $\mathcal{C}(P|H_B|^2/2, P|H_{\overline{B}}|^2/2)$ is achievable for the state (H_0, H_1). Let us define

$$W = C \oplus B$$
$$R = \mathcal{C}\left(P|H_B|^2/2, P|H_{\overline{B}}|^2/2\right) - \epsilon$$

where $\oplus$ denotes the modulo-2 addition, $\mathcal{C}(\cdot, \cdot)$ is given in (4), and $\epsilon > 0$ is a pre-chosen constant.

The protocol:

1) Bob reveals $(W, |H_B|, |H_{\overline{B}}|)$ to Alice over the noise-free channel.
2) Alice takes strings $\mathbf{K}_0$ and $\mathbf{K}_1$ of length $L(|H_0|, |H_1|) := nR$ each. She encodes $\mathbf{K}_W$ and $\mathbf{K}_{\overline{W}}$ into two length-n codewords $\mathbf{X}_0$ and $\mathbf{X}_1$ respectively, such that each has an average power $P/2$. A code suitable for $\mathcal{WT}(\frac{P}{2}, \frac{P|H_B|^2}{2}, \frac{P|H_{\overline{B}}|^2}{2})$ is used to encode both the strings. $\mathbf{X}_0$ and $\mathbf{X}_1$ are transmitted over the respective channels. Note that $\mathbf{K}_C$ has been encoded into $\mathbf{X}_B$, and $\mathbf{K}_{\overline{C}}$ has been encoded into $\mathbf{X}_{\overline{B}}$.
3) Bob receives $\mathbf{Y}_0$ and $\mathbf{Y}_1$ with SNR $P|H_0|^2/2$ and $P|H_1|^2/2$ respectively. He decodes $\mathbf{K}_C$ from $\mathbf{Y}_B$ using the decoder for the wiretap channel referred above.

Correctness of the protocol: Note that $\mathbf{K}_C$ is transmitted over the stronger channel (B), and $\mathbf{K}_{\overline{C}}$ is transmitted over the weaker channel $(\overline{B})$. Bob's received SNR in the stronger channel is $P|H_B|^2/2$, whereas his received SNR in the weaker channel is $P|H_{\overline{B}}|^2/2$. Thus he can decode $\mathbf{K}_C$ with vanishing probability of error, whereas he can get negligible information about $\mathbf{K}_{\overline{C}}$ as his SNR is that of the wiretapper in this channel. Since $|H_0|$ and $|H_1|$ are independent and identically distributed, it is easy to check that $I(W; C) = 0$, thus Alice does not learn anything about Bob's choice C.

B. 2×2 MIMO

Consider a 2×2 fading MIMO channel between Alice and Bob. Alice and Bob each has 2 antennas. Let $\mathbf{H}$ denote the 2×2 complex fading matrix. The input-output relation for the channel is given by (5), where $\mathbf{Y}, \mathbf{X}, \mathbf{Z}$ are $2 \times n$ matrices.

Let the SVD decomposition of $\mathbf{H}$ be given by

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H,$$

where $\mathbf{\Lambda}$ is a diagonal matrix with diagonal elements λ_0, λ_1 such that $\lambda_0 \geq \lambda_1$. These are the (real) singular values of $\mathbf{H}$. We will show that the OT rate $\mathcal{C}(P\lambda_0^2/2, P\lambda_1^2/2)$ is achievable for the fading matrix $\mathbf{H}$. Let $\mathbf{V}_0, \mathbf{V}_1$ denote the columns of $\mathbf{V}$. We define

$$(\mathbf{W}_0, \mathbf{W}_1) = \left(\mathbf{V}_C, \mathbf{V}_{\overline{C}}\right)$$
$$\text{and } R = \mathcal{C}\left(P\lambda_0^2/2, P\lambda_1^2/2\right) - \epsilon \tag{6}$$

for some pre-decided ϵ, where the $\mathcal{C}(\cdot, \cdot)$ above is defined in (4). Note that $\mathbf{W}_0, \mathbf{W}_1$ are the same as $\mathbf{V}_0, \mathbf{V}_1$, but permuted depending on C. Bob shares $(\mathbf{W}_0, \mathbf{W}_1)$ with Alice in our protocol, and Alice uses it as the precoding matrix. Bob first pre-multiplies the received matrix by $\mathbf{U}^H$. The SVD precoding as shown in Fig. 4 transforms the MIMO channel into a parallel fading Gaussian channel, where Alice is unsure of which of

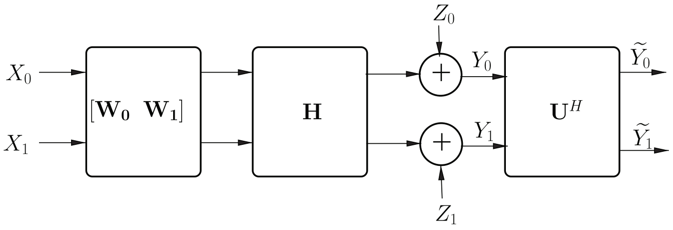

Fig. 4. MIMO precoding for OT.

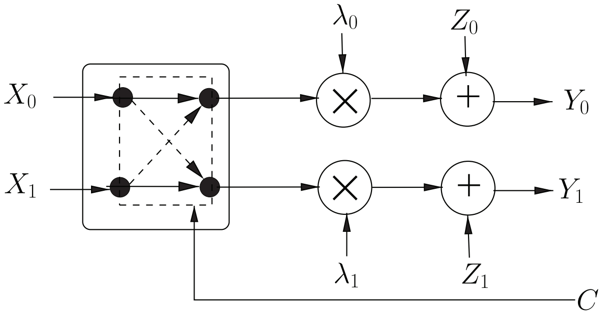

Fig. 5. The equivalent channel with a switch for 2 × 2 MIMO setup.

the two channels has the gain λ_0, and which has gain λ_1. The resulting end-to-end system is shown in Fig. 5 where a switch, controlled by Bob's choice bit C, determines which input of Alice passes through which channel to Bob. The firm lines and dotted lines show the two positions of the coupled switch.

The protocol:

1) Bob reveals $(\mathbf{W}_0, \mathbf{W}_1, \lambda_0, \lambda_1)$ to Alice over the noise-free channel.
2) The basic transmitter and receiver block diagram is shown in Fig. 4. Alice computes R using (6), and takes strings $\mathbf{K}_0$ and $\mathbf{K}_1$ of length $L(\lambda_0, \lambda_1) := nR$ each. She encodes $\mathbf{K}_0$ and $\mathbf{K}_1$ into two length-n codewords $\mathbf{X}_0$ and $\mathbf{X}_1$ respectively, such that each has an average power $P/2$. A code suitable for $\mathcal{WT}(\frac{P}{2}, \frac{P\lambda_0^2}{2}, \frac{P\lambda_1^2}{2})$ is used to encode both the strings. She then transmits the matrix

$$\begin{aligned}[\mathbf{W}_0\ \ \mathbf{W}_1]\begin{bmatrix}\mathbf{X}_0\\ \mathbf{X}_1\end{bmatrix} &= \mathbf{W}_0\mathbf{X}_0 + \mathbf{W}_1\mathbf{X}_1\\ &= \mathbf{V}_0\mathbf{X}_C + \mathbf{V}_1\mathbf{X}_{\overline{C}}\\ &= \mathbf{V}\begin{bmatrix}\mathbf{X}_C\\ \mathbf{X}_{\overline{C}}\end{bmatrix}.\end{aligned}$$

3) Bob first multiplies the received $2 \times n$ matrix by $\mathbf{U}^H$. The resulting end-to-end channel is given by

$$\begin{aligned}\widetilde{\mathbf{Y}} = \begin{bmatrix}\widetilde{\mathbf{Y}}_0\\ \widetilde{\mathbf{Y}}_1\end{bmatrix} &= \mathbf{U}^H\mathbf{H}\mathbf{V}\begin{bmatrix}\mathbf{X}_C\\ \mathbf{X}_{\overline{C}}\end{bmatrix} + \mathbf{U}^H\begin{bmatrix}\mathbf{Z}_0\\ \mathbf{Z}_1\end{bmatrix}\\ &= \begin{bmatrix}\lambda_0\mathbf{X}_C\\ \lambda_1\mathbf{X}_{\overline{C}}\end{bmatrix} + \mathbf{U}^H\begin{bmatrix}\mathbf{Z}_0\\ \mathbf{Z}_1\end{bmatrix}. \qquad (7)\end{aligned}$$

Bob gets $\widetilde{\mathbf{Y}}_0$ and $\widetilde{\mathbf{Y}}_1$ with SNR $P\lambda_0^2/2$ and $P\lambda_1^2/2$ respectively. He decodes $\mathbf{K}_C$ from $\mathbf{Y}_0$ using the decoder for the wiretap channel referred above.

Correctness of the protocol: First note that since $\widetilde{\mathbf{Y}}$ is obtained by a unitary (hence invertible) transformation on $\mathbf{Y}$, it contains exactly the same information as $\mathbf{Y}$. So we will henceforth treat $\widetilde{\mathbf{Y}}$ as Bob's received matrix. Since $\mathbf{U}$ is a unitary matrix, $\mathbf{U}^H\mathbf{Z}$ has the same distribution as that of $\mathbf{Z}$. Also note that $\mathbf{K}_C$ is encoded into $\mathbf{X}_C$, which is received as $\widetilde{\mathbf{Y}}_0$ with SNR $P\lambda_0^2/2$. Since this encoding is done by Alice for a complex Gaussian wiretap channel with the same receiver SNR, Bob can decode $\mathbf{K}_C$ with vanishing probability of error. On the other hand, $\mathbf{K}_{\overline{C}}$ is encoded into $\mathbf{X}_{\overline{C}}$, which is received as $\widetilde{\mathbf{Y}}_1$ with SNR $P\lambda_1^2/2$. Bob can get negligible information about $\mathbf{K}_{\overline{C}}$ as his SNR in $\widetilde{\mathbf{Y}}_1$ is that of the wiretapper. This ensures secrecy of Alice against Bob.

About the secrecy of Bob against Alice, first note that $\mathbf{H}$ is circularly symmetric, and thus $(\mathbf{V}_0, \mathbf{V}_1)$ and $(\mathbf{V}_1, \mathbf{V}_0)$ have the same distribution, that is, their joint distribution is symmetric in $\mathbf{V}_0$ and $\mathbf{V}_1$. Also, note that λ_0, λ_1 are independent of $C, \mathbf{V}_0, \mathbf{V}_1$. Thus

$$I(\mathbf{W}_0, \mathbf{W}_1, \lambda_0, \lambda_1; C) = I\left(\mathbf{V}_C, \mathbf{V}_{\overline{C}}; C\right) = 0.$$

This ensures the secrecy of Bob against Alice.

C. 2 × 1 MISO

Consider a 2 × 1 fading MISO channel between Alice and Bob. Let $\mathbf{H} = (H_0, H_1)$ denote the 1×2 fading matrix. The input-output relation for the channel is given by (5), where $\mathbf{X}$, $\mathbf{Z}$ are $2 \times n$ matrices, and $\mathbf{Y}$ is a $1 \times n$ vector.

Let the SVD of $\mathbf{H}$ be

$$\mathbf{H} = \mathbf{\Lambda}\mathbf{V}^H$$

where $\mathbf{\Lambda} = (\lambda, 0)$, $\lambda = \sqrt{|H_0|^2 + |H_1|^2}$, the first column of $\mathbf{V}$ is $\mathbf{V}_0 = (1/\lambda)\mathbf{H}^H$, and the second column of $\mathbf{V}$ is a unit vector $\mathbf{V}_1$ orthogonal to $\mathbf{H}$. We will show that the OT rate $\log_2\left(1 + \frac{P\lambda^2}{2}\right)$ is achievable for the fading matrix $\mathbf{H}$.

The best way to communicate messages (without any secrecy condition) is using SVD precoding wherein Alice multiplies her message symbol with the first column of $\mathbf{V}_0$ and transmits. Bob simply divides the received symbol by λ and chooses the message symbol nearest to the result. Note that if in addition, Alice added any scalar multiple of $\mathbf{V}_1$ to her transmission, it would not contribute to the received symbol as $\mathbf{V}_1$ is orthogonal to $\mathbf{H}$. Thus this dimension which is orthonormal to $\mathbf{H}$ (the null-space of $\mathbf{H}$) is not useful for communication, as it has zero gain. This reduces the MISO channel to a single fading AWGN channel with fading coefficient λ.

We now give an OT protocol for this channel when only Bob has the knowledge of $\mathbf{H}$ at the beginning of a block. We define

$$(\mathbf{W}_0, \mathbf{W}_1) = (\mathbf{V}_C, \mathbf{V}_{\overline{C}}) \qquad (8)$$

$$\text{and } R = \log_2\left(1 + \frac{P\lambda^2}{2}\right) - \epsilon \qquad (9)$$

for some pre-decided ϵ. Bob shares $(\mathbf{W}_0, \mathbf{W}_1)$ with Alice in our protocol, and Alice uses it as the precoding matrix. The resulting channel is equivalent to what is shown in Fig. 6 where a switch, controlled by Bob's choice bit C, determines which input of Alice passes through the channel to Bob.

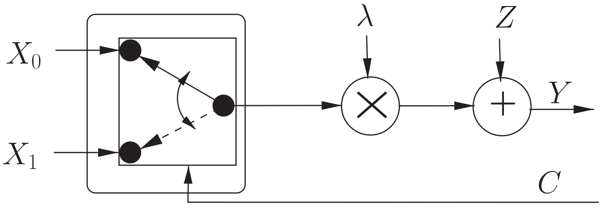

Fig. 6. The equivalent channel with a switch for 2 × 1 MIMO setup.

The protocol

1) Bob reveals $(\mathbf{W}_0, \mathbf{W}_1, \lambda)$ to Alice over the noise-free channel. He sets $(\mathbf{W}_0, \mathbf{W}_1)$ as in (8).
2) Both Alice and Bob compute $L(\lambda) := Rn$ with R given in (9). Alice encodes each of $\mathbf{K}_0$ and $\mathbf{K}_1$ (of length $L(\lambda)$ each) into a n-length vector. She uses a code suitable for a complex AWGN channel with SNR $\frac{P}{2}\lambda^2$. Let these encoded vectors be $\mathbf{X}_0$ and $\mathbf{X}_1$ respectively. Over n uses of the channel, Alice transmits the $2 \times n$ matrix $\mathbf{W}_0\mathbf{X}_0 + \mathbf{W}_1\mathbf{X}_1$.
3) Bob receives

$$\begin{aligned}\mathbf{Y} &= \mathbf{H}(\mathbf{W}_0\mathbf{X}_0 + \mathbf{W}_1\mathbf{X}_1) + \mathbf{Z}\\ &= \lambda\mathbf{X}_C + \mathbf{Z}.\end{aligned}$$

Bob now decodes $\mathbf{K}_C$ from $\mathbf{Y}$ with probability of error going to zero as $n \to \infty$.

Correctness of the protocol: Since $\mathbf{X}_{\overline{C}}$ is transmitted in the null-space of $\mathbf{H}$, it does not contribute to Bob's received vector. Thus Bob has no information about $\mathbf{K}_{\overline{C}}$. Since $\mathbf{H}$ has i.i.d. Gaussian entries, $(\mathbf{V}_0, \mathbf{V}_1)$ has a distribution which is symmetric in $\mathbf{V}_0$ and $\mathbf{V}_1$, and λ is independent of $(\mathbf{V}_0, \mathbf{V}_1)$. Thus, $I(\mathbf{W}_0, \mathbf{W}_1, \lambda; C) = 0$. Thus the secrecy of Bob against Alice is met.

IV. The General Protocol

In this section, we present a protocol for the general $2N$-channels OFDM and $2N \times n_B$-MIMO models. Here we assume that Alice has more $(2N)$ antennas than Bob has (n_B). The case $n_B > 2N$ is similar, and is discussed briefly later.

For the MIMO setup, we first discuss how Bob can reveal some partial information about the channel matrix to reduce the channel to a parallel channel. We will then treat both OFDM and MIMO models as parallel channels and present a common OT protocol. The OT protocol will group the parallel channels into pairs and perform OT over each pair using similar technique as in the previous section.

A. Reducing MIMO Setup to Parallel Channels

Let the SVD decomposition of $\mathbf{H}$ be given by

$$\mathbf{H} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H,$$

where $\mathbf{\Lambda}$ is a $n_B \times 2N$ diagonal[1] matrix with diagonal elements $\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n_B-1}$. Let $\mathbf{P}$ be a random $2N \times 2N$ permutation matrix chosen by Bob. Note that a permutation matrix is unitary, and thus $\mathbf{P}^T = \mathbf{P}^{-1}$. Let us add $(2N - n_B)$ zero rows with $\mathbf{U}^H$ to define the $2N \times n_B$ matrix

$$\widetilde{\mathbf{U}} = \begin{bmatrix}\mathbf{U}^H\\ \mathbf{0}\end{bmatrix}.$$

Bob sends $\mathbf{W} = \mathbf{VP}$ over the noise-free channel, and Alice uses it as the precoding matrix to transmit $\mathbf{VPX}$. Bob first multiplies the received vector $\mathbf{Y}$ by $\mathbf{P}^T\widetilde{\mathbf{U}}$ to get

$$\begin{aligned}\widetilde{\mathbf{Y}} &= \mathbf{P}^T\widetilde{\mathbf{U}}\mathbf{Y}\\ &= \mathbf{P}^T\begin{bmatrix}\mathbf{\Lambda}\mathbf{PX} + \mathbf{U}^H\mathbf{Z}\\ \mathbf{0}\end{bmatrix}\\ &= \mathbf{P}^T\begin{bmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{bmatrix}\mathbf{PX} + \mathbf{P}^T\begin{bmatrix}\mathbf{U}^H\mathbf{Z}\\ \mathbf{0}\end{bmatrix}\end{aligned}$$

Let us denote $\underline{\boldsymbol{\lambda}} := (\lambda_0, \lambda_1, \cdots, \lambda_{2N-1})^T$ as the $2N$ length vector of diagonal elements of $\begin{bmatrix}\mathbf{\Lambda}\\ \mathbf{0}\end{bmatrix}$ where $\lambda_l = 0$ for $l \geq n_B$. Let us also denote $\widetilde{\mathbf{Z}} := \begin{bmatrix}\mathbf{U}^H\mathbf{Z}\\ \mathbf{0}\end{bmatrix}$. Let π denote the permutation induced on a vector by pre-multiplication by $\mathbf{P}^T$, that is, $\mathbf{P}^T\underline{\boldsymbol{\lambda}} = (\lambda_{\pi(0)}, \lambda_{\pi(1)}, \cdots, \lambda_{\pi(2N-1)})$ in particular. Then

$$\widetilde{Y}_l = \lambda_{\pi(l)}X_l + \widetilde{Z}_{\pi(l)}.$$

We note that for $\pi(l) \geq n_B$, $\lambda_{\pi(l)} = \widetilde{Z}_{\pi(l)} = 0$. This gives a set of parallel channels such that $2N - n_B$ of them have zero gain and zero noise. These channels are completely useless for communication. Since $\mathbf{U}^H$ is unitary, $\mathbf{U}^H\mathbf{Z}$ is also i.i.d. with independent real and imaginary components $\sim \mathcal{N}(0, 1/2)$. Since Bob knows $\mathbf{P}$ (and so π), he will neglect the channels l for which $\pi(l) \geq n_B$. To reduce this model to a standard parallel AWGN channels model with constant noise variance in all channels but different channel gains, we assume that Bob adds some independent noise with real and imaginary parts $\sim \mathcal{N}(0, 1/2)$ to each of the channels for which $\pi(l) \geq n_B$.

We now prove a lemma which states that in the resulting parallel channels, Alice can not know the order of the channel gains.

Lemma 1: Let $\mathbf{H}$ be the channel matrix and $\mathbf{P}$ is a permutation matrix chosen uniformly at random. Let $\mathbf{W} = \mathbf{VP}$ denote the precoding matrix sent to Alice by Bob, and $\underline{\boldsymbol{\lambda}}$ be the zero-padded vector of ordered singular values. Then for any $\mathbf{W}$ and $\underline{\boldsymbol{\lambda}}$, and for any two permutations $\mathbf{P}$ and $\mathbf{P}'$, we have $Pr(\mathbf{P}|\mathbf{W}, \underline{\boldsymbol{\lambda}}) = Pr(\mathbf{P}'|\mathbf{W}, \underline{\boldsymbol{\lambda}}) = \frac{1}{(2N)!}$.

Proof: $\mathbf{V}$ is uniformly distributed over the set of $2N \times 2N$ unitary matrices (see [29], Lemma 5). Since $\mathbf{P}$ is a unitary matrix $\mathbf{W} = \mathbf{VP}$ is also unitary and both $\mathbf{VP}$ and $\mathbf{VP}'$ are Haar matrices with the same uniform distribution over the set of $2N \times 2N$ unitary matrices. Hence $f_{\mathbf{W},\underline{\boldsymbol{\lambda}}|\mathbf{P}}(\mathbf{W}, \underline{\boldsymbol{\lambda}}|\mathbf{P}) = f_{\mathbf{V},\underline{\boldsymbol{\lambda}}}(\mathbf{W}\mathbf{P}^T, \underline{\boldsymbol{\lambda}}) = f_{\mathbf{V},\underline{\boldsymbol{\lambda}}}(\mathbf{W}, \underline{\boldsymbol{\lambda}})$, and also $f_{\mathbf{W},\underline{\boldsymbol{\lambda}}}(\mathbf{W}, \underline{\boldsymbol{\lambda}}) = f_{\mathbf{V},\underline{\boldsymbol{\lambda}}}(\mathbf{W}, \underline{\boldsymbol{\lambda}})$. So we have $Pr(\mathbf{P}|\mathbf{W}, \underline{\boldsymbol{\lambda}}) = \frac{1}{(2N)!}$. ■

We have now reduced the MIMO channel to a standard parallel AWGN channels with different gains (singular values) in different subchannels. The above lemma says that from the partial channel state information given to Alice, she still would

[1] That is, its (i, j)-th element $\lambda_{ij} = 0$ for $i \neq j$, and λ_{ii} is denoted by λ_i.

be 'completely uncertain' about the association of the singular values to the resulting subchannels.

The case of $n_B > 2N$: When $n_B > 2N$, $\mathbf{U}$ is an $n_B \times n_B$ matrix and $\mathbf{\Lambda}$ is a $n_B \times 2N$ diagonal matrix with $(n_B - 2N)$ zero rows. Let the last $n_B - 2N$ rows of $\mathbf{U}^H$, $\mathbf{\Lambda}$ and $\mathbf{U}^H\mathbf{Z}$ be removed to obtain respectively $\widetilde{\mathbf{U}}$, $\widetilde{\mathbf{\Lambda}}$ and $\widetilde{\mathbf{Z}}$. As before, Alice transmits $\mathbf{VPX}$. Bob first multiplies $\mathbf{P}^T\widetilde{\mathbf{U}}$ to the received vector to obtain

$$\begin{aligned}\widetilde{\mathbf{Y}} &= \mathbf{P}^T\widetilde{\mathbf{U}}\mathbf{Y}\\ &= \mathbf{P}^T\widetilde{\mathbf{\Lambda}}\mathbf{P}\mathbf{X} + \mathbf{P}^T\widetilde{\mathbf{Z}}.\end{aligned}$$

The protocol now continues with the $2N$ components of $\widetilde{\mathbf{Y}}$ which constitute the output of the $2N$ parallel channels as before.

In the following, we consider a set of parallel channels indexed by $0, 1, \cdots, 2N-1$, as depicted in Fig. 2. Such a model could have resulted from an OFDM channel or a MIMO channel under the scheme discussed above. To treat MIMO and OFDM in a unified manner in the following, we also assume $\lambda_l = |H_l|$ to be the channel gains in case of OFDM as they provide the same performance. For OFDM, we assume that $\lambda_0, \lambda_1, \cdots, \lambda_{2N-1}$ are i.i.d. and Rayleigh distributed. We now define an OT-pairing of the channels and a power allocation under a given total power constraint.

Definition 2: An OT-pairing of the $2N$ channels is defined using two maps $\ell, k : \{0, 1, \cdots, N-1\} \to \{0, 1, \cdots, 2N-1\}$ such that

1) ℓ, k are $1-1$
2) $Im(\ell) \cap Im(k) = \emptyset$
3) $\lambda_{\ell(l)} > \lambda_{k(l)}\ \forall\, l$.

The ordered pairs of the channels are then $(\ell(l), k(l)); l = 0, 1, \cdots, N-1$.

B. Power Allocation

Alice divides the total average transmit power P between the subchannels. In our OT protocol, Alice transmits the same power over the subchannels in a pair. Let P_l the average power transmitted on each of the subchannels in pair l, that is, in the subchannels $\ell(l)$ and $k(l)$, be P_l. Then $P_l \geq 0$ and

$$\sum_{l=0}^{N-1} P_l \leq \frac{P}{2}. \tag{10}$$

The rates for the pairs are taken as

$$R_l = \mathcal{C}\left(P_l\lambda^2_{\ell(l)}, P_l\lambda^2_{k(l)}\right) - \epsilon \tag{11}$$

for an arbitrarily small fixed constant $\epsilon > 0$. We denote $\underline{\mathbf{R}} = (R_0, R_1, \cdots, R_{N-1})$. Note that R_l is close to the capacity of the wiretap channel $\mathcal{WT}(P_l, P_l\lambda^2_{\ell(l)}, P_l\lambda^2_{k(l)})$. Our OT protocol for the 2-channels OFDM can be used with average power constraint $2P_l$ to achieve a rate R_l for each pair of subchannels. The total rate achieved is thus

$$R = \sum_{l=0}^{N-1} \mathcal{C}\left(P_l\lambda^2_{\ell(l)}, P_l\lambda^2_{k(l)}\right) - \epsilon N. \tag{12}$$

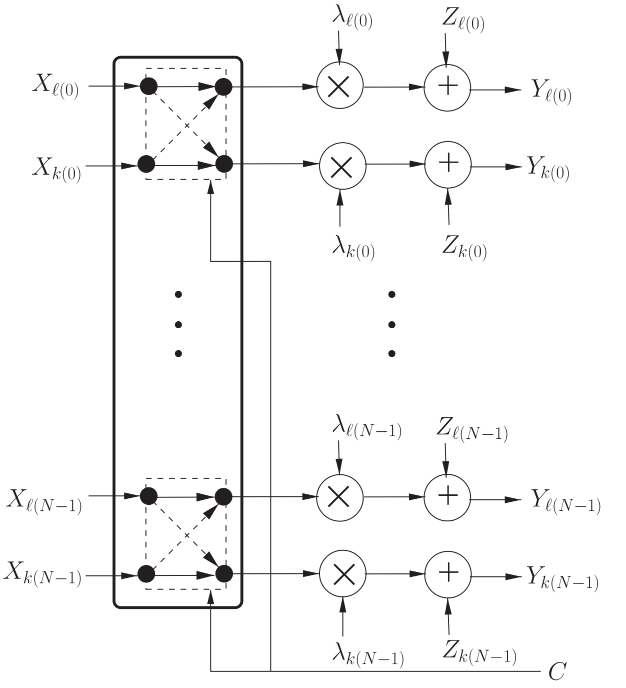

Fig. 7. The equivalent channel with a switch.

For simplicity, we assume that nR_l is an integer for each l. We define for $l = 0, 1, \cdots, N-1$,

$$\tilde{\gamma}_l = (\gamma_{l0}, \gamma_{l1}) = (\ell(l), k(l)) \tag{13}$$

$$\tilde{\lambda}_l = \left(\lambda_{\ell(l)}, \lambda_{k(l)}\right), \tag{14}$$

and denote $\underline{\tilde{\gamma}} := (\tilde{\gamma}_0, \tilde{\gamma}_1, \cdots, \tilde{\gamma}_{N-1})$ and $\underline{\tilde{\lambda}} = (\tilde{\lambda}_0, \tilde{\lambda}_1, \cdots, \tilde{\lambda}_{N-1})$.

Let $\mathbf{T}$ denote the $2N \times 2N$ permutation matrix representing the transposition of consecutive pairs. $\mathbf{T}$ consists of N diagonal 2×2 blocks $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. We define

$$\underline{\gamma} = \begin{cases} \underline{\tilde{\gamma}} & \text{if } C = 0 \\ \underline{\tilde{\gamma}}T & \text{if } C = 1 \end{cases} \tag{15}$$

Bob shares $(\underline{\gamma}, \underline{\tilde{\lambda}})$ with Alice. From Alice's point of view, the parallel channels appear to be associated with the gains shown in Fig. 7. The association of the gains to the channels has one bit of uncertainty as depicted by the two possible positions of the coupled switches. The position of the switches is controlled by C, and is not known to Alice. We give the protocol below.

C. The Protocol

1) In case of a MIMO setup, Bob first reveals $\mathbf{W}$ to Alice, and Alice uses it as the precoding matrix. Bob also does appropriate pre-processing as discussed in Sec. IV-A to reduce the channel to a set of parallel channels.
2) Bob selects an OT pairing ℓ, k and reveals $(\underline{\gamma}, \underline{\tilde{\lambda}})$ to Alice over the noise-free channel. He computes these using (15) and (14) respectively.
3) Both Alice and Bob compute R_l using (11) and $L_l = R_l n$ for $l = 0, 1, \cdots, N-1$. Let us denote $L = \sum_{l=0}^{N-1} L_l$.

For each $j = 0, 1$, Alice breaks $\mathbf{K}_j$ (of length L) into N substrings $\mathbf{K}_{jl}; l = 0, 1, \cdots, N-1$ of lengths L_l respectively. For each $j = 0, 1$, and $l = 0, 1, \cdots, N-1$, she encodes $\mathbf{K}_{jl}$ into a n-length vector $\mathbf{X}_{jl}$ of average power P_l using a code for the wiretap channel $\mathcal{WT}(P_l, P_l\lambda_{\ell(l)}^2, P_l\lambda_{k(l)}^2)$. Alice transmits this vector over n uses of the channel γ_{lj}.

4) Note that from (15), $\gamma_{lC} = \ell(l)$ and $\gamma_{l\overline{C}} = k(l)$ for each $l = 0, 1, \cdots, N-1$. Thus Bob receives

$$\mathbf{Y}_{\ell(l)} = \lambda_{\ell(l)}\mathbf{X}_{Cl} + \mathbf{Z}_{\ell(l)}.$$

Bob now decodes $\mathbf{K}_{Cl}$ from $\mathbf{Y}_{\ell(l)}$ with probability of error going to zero as $n \to \infty$.

Correctness of the protocol: Bob can decode $\mathbf{K}_{Cl}$ from $\mathbf{Y}_{\ell(l)}$ for each l with arbitrarily small probability of error. This follows from standard results in Gaussian wiretap channels [15]. It also follows that he gets only an arbitrarily small amount of information about $\mathbf{K}_{\overline{C}}$ from $\mathbf{Y}_{k(l)}$ in the sense of (3), [16].

Alice knows that $\underline{\tilde{\boldsymbol{\gamma}}} \in \{\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T\}$. Since $\underline{\boldsymbol{\gamma}}$ and $\underline{\tilde{\boldsymbol{\lambda}}}$ are revealed to Alice during the protocol, the uncertainty in C is equivalent to the uncertainty in which of $\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T$ is the value of $\underline{\tilde{\boldsymbol{\gamma}}}$.

Now, let us first consider an OFDM channel. From the point of view of Alice,

$$\begin{aligned} Pr\left(C = 0 \middle| \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}\right) &= Pr\left(\underline{\tilde{\boldsymbol{\gamma}}} = \underline{\boldsymbol{\gamma}} \middle| \underline{\tilde{\boldsymbol{\gamma}}} \in \left\{\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T\right\}, \underline{\tilde{\boldsymbol{\lambda}}}\right) \\ &= Pr\left(\underline{\tilde{\boldsymbol{\gamma}}} = \underline{\boldsymbol{\gamma}}T \middle| \underline{\tilde{\boldsymbol{\gamma}}} \in \left\{\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T\right\}, \underline{\tilde{\boldsymbol{\lambda}}}\right) \qquad (16) \\ &= Pr\left(C = 1 | \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}\right). \end{aligned}$$

Here (16) follows as we have assumed that the channel gains of the parallel channels are i.i.d. This implies that $I(C; \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}) = 0$.

Similarly, if the parallel channels have resulted from a MIMO channel, then Alice has also learned the precoding matrix $\mathbf{W}$. Now,

$$\begin{aligned} Pr\left(C = 0 \middle| \mathbf{W}, \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}\right) &= Pr\left(\underline{\tilde{\boldsymbol{\gamma}}} = \underline{\boldsymbol{\gamma}} | \mathbf{W}, \underline{\tilde{\boldsymbol{\gamma}}} \in \left\{\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T\right\}, \underline{\tilde{\boldsymbol{\lambda}}}\right) \\ &= Pr\left(\underline{\tilde{\boldsymbol{\gamma}}} = \underline{\boldsymbol{\gamma}}T \middle| \mathbf{W}, \underline{\tilde{\boldsymbol{\gamma}}} \in \left\{\underline{\boldsymbol{\gamma}}, \underline{\boldsymbol{\gamma}}T\right\}, \underline{\tilde{\boldsymbol{\lambda}}}\right) \\ &= Pr\left(C = 1 \middle| \mathbf{W}, \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}\right). \qquad (17) \end{aligned}$$

Here (17) follows from Lemma 1. Thus we have $I(C; \mathbf{W}, \underline{\boldsymbol{\gamma}}, \underline{\tilde{\boldsymbol{\lambda}}}) = 0$. This proves that Alice does not gain any information about C from what she learns during the protocol.

We now discuss the optimal OT-pairing and the optimal power allocation.

V. A Protocol Using Parallel MIMO Channels

In this section, we consider a possible generalization of the approach discussed in the last section for OT over a MIMO channel. In general, we can use suitable unitary precoding and equalization matrices to reduce the MIMO channel to a pair of parallel MIMO channels. Here Alice knows the channel matrices $\mathbf{H}_0$ and $\mathbf{H}_1$. The received vectors are given by

$$\begin{bmatrix} \mathbf{Y}_0 \\ \mathbf{Y}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0\mathbf{X}_C \\ \mathbf{H}_1\mathbf{X}_{\overline{C}} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_0 \\ \mathbf{Z}_1 \end{bmatrix}.$$

A protocol similar to that for 2-channels OFDM can be used over this channel. The achieved OT rate R is given by the capacity of the MIMO wiretap channel [25]:

$$\max_{\mathbf{K}_X}\left(\log\det\left(I + \mathbf{H}_0\mathbf{K}_X\mathbf{H}_0^H\right) - \log\det\left(I + \mathbf{H}_1\mathbf{K}_X\mathbf{H}_1^H\right)\right) =$$
$$\max_{\mathbf{K}_X}\left(\log\det\left(I + \mathbf{K}_X\mathbf{H}_0^H\mathbf{H}_0\right) - \log\det\left(I + \mathbf{K}_X\mathbf{H}_1^H\mathbf{H}_1\right)\right)$$

where the maximization is over all input covariance matrices $\mathbf{K}_X$ with $\mathrm{Tr}(\mathbf{K}_X) \leq P/2$. We note that, the singular values of the original MIMO channel is the union of the singular values of $\mathbf{H}_0$ and $\mathbf{H}_1$. Let the singular values of $\mathbf{H}$ be $\lambda_0, \lambda_1, \cdots, \lambda_{2N-1}$ in decreasing order of magnitude. Let the eigenvalues of $\mathbf{K}_X$ be $P_0, P_1, \cdots, P_{N-1}$ in decreasing order.

Lemma 2: If A and B are two $N \times N$ positive semidefinite Hermitian matrices with eigenvalues $\alpha_0, \alpha_1, \cdots, \alpha_{N-1}$ and $\beta_0, \beta_1, \cdots, \beta_{N-1}$, ordered in decreasing order, then

$$\prod_i (1 + \alpha_i\beta_{N-1-i}) \leq \det(I + AB) \leq \prod_i (1 + \alpha_i\beta_i)$$

Proof: First let us assume that $\max\{\alpha_{N-1}, \beta_{N-1}\} > 0$. Then, without loss of generality, let us assume that $\beta_{N-1} > 0$, that is, $\det(B) \neq 0$. Then

$$\det(I + AB) = \det(B)\det\left(B^{-1} + A\right)$$

Now, for any two $N \times N$ positive semidefinite Hermitian matrices C and D with eigenvalues $\gamma_0, \gamma_1, \cdots, \gamma_{N-1}$ and $\nu_0, \nu_1, \cdots, \nu_{N-1}$, ordered in decreasing order, it is known [26] that

$$\prod_i (\gamma_i + \nu_i) \leq \det(C + D) \leq \prod_i (\gamma_i + \nu_{N-1-i}).$$

Using this inequality for B^{-1} and A, we get the desired inequality.

Now, if $\alpha_{N-1} = \beta_{N-1} = 0$, then we can perturb and replace all the zero eigenvalues of A and B by $\epsilon > 0$ and leave the unitary matrices in their eigenvalue decomposition unchanged. For such perturbed matrices, the inequality holds. Now, since both the determinant and its bounds are continuous functions of the eigenvalues, the bounds also hold for A and B (that is, at $\epsilon = 0$). ■

Using the lemma, we have

$$\det\left(I + \mathbf{H}_0\mathbf{K}_X\mathbf{H}_0^H\right) \leq \prod_{i=0}^{N-1}\left(1 + \lambda_i^2 P_i\right)$$

and

$$\det\left(I + \mathbf{H}_1\mathbf{K}_X\mathbf{H}_1^H\right) \geq \prod_{i=0}^{N-1}\left(1 + \lambda_{2N-i-1}^2 P_i\right),$$

and thus

$$R \leq \max_{\mathbf{P}} \sum_{i=0}^{N-1}\left(\log\left(1 + \lambda_i^2 P_i\right) - \log\left(1 + \lambda_{2N-i-1}^2 P_i\right)\right), \qquad (18)$$

where the maximization is over all power allocations $\mathbf{P} = (P_0, P_1, \cdots, P_{N-1})$ with $\sum_i P_i \leq P/2$. Our earlier protocol (Sec. IV) which uses suitably paired parallel SISO channels, achieves equality in (18). Thus using a MIMO wiretap channel code for the two parallel MIMO channels can not give better rates than full diagonalization of the channel.

We also note that if we are given an OFDM channel, the channel matrix is already diagonal. Even then, (18) suggests an optimal pairing, namely the best with the worst and so on.

VI. Optimization of the Protocol

Let us first consider the simple setup where equal power is allocated in all pairs of subchannels, i.e.,

$$P_l = \frac{P}{2N} \qquad \forall l.$$

The capacity for this power allocation is

$$R = \sum_{l=0}^{N-1} \log\left(1 + \frac{P\lambda^2_{\ell(l)}}{2N}\right) - \sum_{l=0}^{N-1} \log\left(1 + \frac{P\lambda^2_{k(l)}}{2N}\right)$$

Clearly, this is maximized if $\lambda^2_{\ell(l)} > \lambda^2_{k(j)}$ for all l, j. That is, provided the best half of the channels form the stronger channels of the pairs, the achieved rate is independent of the actual pairing. However, this is not true if we have the freedom to pair the channels as well as to allocate variable power P_l to different pairs. In general, we would like to choose an *optimal pairing* $(\ell(l), k(l)); 0 \leq l \leq N-1$ and power allocation $P_l; 0 \leq l \leq N-1$ so as to maximize

$$R = \sum_{l=0}^{N-1} \log\left(1 + \frac{P_l\lambda^2_{\ell(l)}}{2N}\right) - \sum_{l=0}^{N-1} \log\left(1 + \frac{P_l\lambda^2_{k(l)}}{2N}\right). \quad (19)$$

For any given power allocation $P_l; 0 \leq l \leq N-1$ arranged in non-increasing order, i.e., $P_0 \geq P_1 \geq \cdots \geq P_{N-1}$, (18) directly implies the following result, which states that an optimal OT pairing couples the best channel with the worst, and so on with the remaining channels.

Theorem 1: A pairing which combines the best channel with the worst channel and continues similarly with the remaining channels is optimal. That is, the pairing given by $\ell(l) = l$ and $k(l) = 2N - l - 1$ for $l = 0, \cdots, N-1$ is optimal when the channel gains are ordered in non-increasing order, i.e., $\lambda_l^2 \geq \lambda_{l+1}^2$ for $0 \leq l < 2N - 1$.

This result reduces the problem of joint optimization of (19) for the best pairing and power allocation to separate optimization of the pairing and the power allocation among the pairs of channels. With high probability, all the gains $(\lambda_0, \cdots, \lambda_{2N-1})$ are distinct. Under this high probability event, Theorem 1 gives a unique optimal pairing. We now find the optimal power allocation.

Optimal Power Allocation: In light of Theorem 1, we assume that the channels are ordered such that

$$\lambda_l \geq \lambda_{l+1} \quad \text{for } 0 \leq l < 2N - 1$$

and the channel with gain λ_l is paired with the channel with gain λ'_l, where $\lambda'_l = \lambda_{2N-l-1}$. Then for a given power allocation $P_l; 0 \leq l \leq N-1$, the achieved rate is

$$R(P_0, \cdots, P_{N-1}) = \sum_{l=0}^{N-1} \log\left(1 + P_l\lambda_l^2\right) - \sum_{l=0}^{N-1} \log\left(1 + P_l\lambda_l'^2\right).$$

We need to maximize this with respect to the P_ls under the condition

$$\sum_{l=0}^{N-1} P_l \leq \frac{P}{2}.$$

Similar optimization was needed for power allocation over different fading states for block fading wiretap channel [27]. This can be solved by defining the Lagrangian objective function

$$J = R(P_0, \cdots, P_{N-1}) - \eta\left(\sum_{l=0}^{N-1} P_l - \frac{P}{2}\right).$$

The optimal power allocation is given by

$$P_l = \begin{cases} \left(\left(f\left(\lambda_l, \lambda'_l, \eta\right)\right)^{1/2} - \frac{1}{2}\left(\frac{1}{\lambda_l^2} + \frac{1}{\lambda_l'^2}\right)\right)^+ & \text{if } \lambda'_l \neq 0 \\ \left(\frac{1}{\eta} - \frac{1}{\lambda_l^2}\right)^+ & \text{if } \lambda'_l = 0 \end{cases}$$

where

$$f(\lambda_l, \lambda'_l, \eta) = \frac{1}{4}\left(\frac{1}{\lambda_l'^2} - \frac{1}{\lambda_l^2}\right)\left[\left(\frac{1}{\lambda_l'^2} - \frac{1}{\lambda_l^2}\right) + \frac{4}{\eta}\right],$$

and η is determined by the condition

$$\sum_{l=1}^{N} P_l = \frac{P}{2}.$$

Power allocation across coherence blocks: If variable amount of average power is allowed to be transmitted in different blocks under a long term average power constraint, then potentially higher rates are achievable. Let $\underline{\boldsymbol{\lambda}}$ denote the random vector that represents the ordered (non-increasing) channel vector in a block, and let $P_l(\underline{\boldsymbol{\lambda}})$ denote the power allocated to the l-th pair of channels. The optimum pairing in each block is still as given by Theorem 1. The optimal power allocation is the maximizer of the expected rate $\bar{R}$:

$$E\left[\sum_{l=0}^{N-1}\left(\log\left(1 + P_l\left(\underline{\boldsymbol{\lambda}}\right)\lambda_l^2\right) - \log\left(1 + P_l\left(\underline{\boldsymbol{\lambda}}\right)\lambda_{2N-l-1}^2\right)\right)\right]$$

under the average power constraint

$$E\left[\sum_{l=0}^{N-1} P_l(\underline{\boldsymbol{\lambda}})\right] \leq \frac{P}{2}.$$

By similar steps as before, the solution is given by

$$P_l\left(\underline{\lambda}\right) = \begin{cases} \left(\left(f\left(\lambda_l, \lambda_l', \eta\right)\right)^{1/2} - \frac{1}{2}\left(\frac{1}{\lambda_l^2} + \frac{1}{\lambda_l'^2}\right)\right)^+ & \text{if } \lambda_l' \neq 0 \\ \left(\frac{1}{\eta} - \frac{1}{\lambda_l^2}\right)^+ & \text{if } \lambda_l' = 0. \end{cases}$$

where η is a global constant determined by the condition

$$E\left[\sum_{l=0}^{N-1} P_l(\underline{\lambda})\right] = \frac{P}{2}. \tag{20}$$

Here η depends only on the channel statistics and P.

VII. High SNR Asymptotics

Let us consider a set of parallel channels. We want to study the asymptotic expected rate. Let us consider a fixed ordered channel vector $(\lambda_0, \lambda_1, \cdots, \lambda_{2N-1})$ to start with. Note that in the case of a $(2N \times n_B)$ MIMO system with precoding, there are $2N$ channels. If $n_B \leq N$, then there are n_B useful pairs of channels with channel gains (λ_0, λ_0'), (λ_1, λ_1'), $\cdots$, $(\lambda_{n_B-1}, \lambda'_{n_B-1})$, where $\lambda_l' = \lambda_{2N-l-1} = 0$, for $l = 0, 1, \cdots, n_B - 1$. If $N < n_B < 2N$, then there are N pairs. $(2N - n_B)$ of them have the second channel gain zero, more specifically, $\lambda_0' = \cdots = \lambda'_{(2N-n_B-1)} = 0$.

Clearly, $\eta \to 0$ as $P \to \infty$. So, $P_l \to \infty$ as $P \to \infty$. Now, for a pair of channels with $\lambda_l' = 0$, the rate contributed by the pair is[2]

$$\begin{aligned} R_l &= \log\left(1 + P_l\lambda_l^2\right) \\ &\to \log(P_l\lambda_l^2). \end{aligned} \tag{21}$$

For such a channel pair,

$$\begin{aligned} P_l &= \frac{1}{\eta}\left(1 - \frac{\eta}{\lambda_l^2}\right) \\ &\Rightarrow \eta P_l \to 1 \qquad \text{as } \eta \to 0 \end{aligned} \tag{22}$$

When $\lambda_l' \neq 0$ and $\lambda_l \neq \lambda_l'$, as $\eta \to 0$,

$$\sqrt{\eta} P_l \to \left(\frac{1}{\lambda_l'^2} - \frac{1}{\lambda_l^2}\right)^{\frac{1}{2}}. \tag{23}$$

So, for such channel pairs,

$$\begin{aligned} R_l &= \log\left(1 + P_l\lambda_l^2\right) - \log\left(1 + P_l\lambda_l'^2\right) \\ &\to \log\left(\frac{\lambda_l^2}{\lambda_l'^2}\right) \qquad \text{as } P \to \infty. \end{aligned} \tag{24}$$

Now, using (22) and (23), the power constraint gives

$$\eta P \to 2(2N - n_B) \qquad \text{as } P \to \infty. \tag{25}$$

Inspired by similar concepts for communication over MIMO channels, it is reasonable to define the *OT-multiplexing gain* as

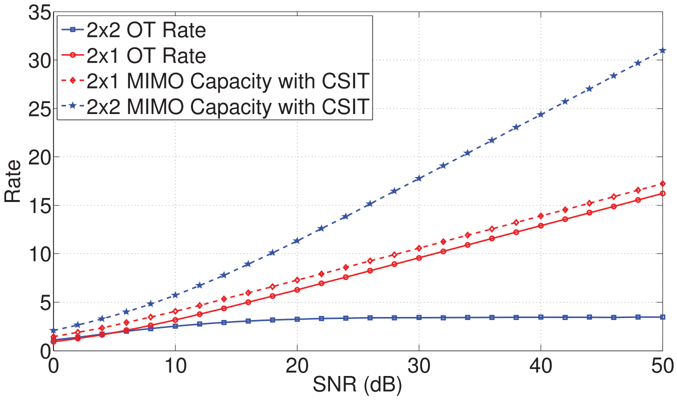

Fig. 8. OT Rate and MIMO capacity versus SNR for 2 × 1, 2 × 2 MIMO.

$$\mu_{OT} = \lim_{P \to \infty} \frac{E\left[\sum_i R_i\right]}{\log P}.$$

So,

$$\begin{aligned} \mu_{OT} &= \lim_{P \to \infty} \frac{E\left[\sum_{l:\lambda_l'=0} R_l\right]}{\log P} \qquad \text{(using24))} \\ &= \lim_{P \to \infty} \frac{E\left[\sum_{l:\lambda_l'=0} \log(P_l)\right]}{\log P} \qquad \text{(using21))} \\ &= \lim_{P \to \infty} \frac{E\left[\sum_{l:\lambda_l'=0} (\log(P_l) - \log(\eta P_l))\right]}{\log P - E(\log(\eta P))} \\ &= \lim_{P \to \infty} \frac{E\left[\sum_{l:\lambda_l'=0} (-\log(\eta))\right]}{-E(\log(\eta))} \\ &= E\left[\left|\{l : \lambda_l' = 0\}\right|\right] \end{aligned} \tag{26}$$

Here (26) follows from (22) and (25). Thus our protocol achieves the OT-multiplexing gain of

$$\mu_{OT} = \begin{cases} n_B & \text{if } n_B \leq N \\ 2N - n_B & \text{if } N < n_B \leq 2N \\ 0 & \text{if } n_B \geq 2N. \end{cases}$$

In contrast, for communication over a $2N \times n_B$ MIMO channel, the multiplexing gain is $\min\{n_B, 2N\}$. For $n_B \geq 2N$, the average OT rate converges to a constant as $P \to \infty$. This can be seen as a consequnce of the fact that the secrecy capacity of the Gaussian wiretap channel goes to a constant as $P \to \infty$.

VIII. Numerical Results

In this section, we provide numerical results of our OT protocols for some simple MIMO and OFDM channels which include the examples discussed in Section III.

In Fig. 8, we plot the average OT rate of our protocol for 2×1 and 2×2 MIMO channels. The average OT rate is numerically evaluated using Monte Carlo simulation methods for SNR varying from 0 dB to 50 dB. The channel capacities for these channels with CSIT are also numerically evaluated and shown. It can be seen that the average OT rate of 2×1 MIMO

[2] Here we mean $R_l - \log(P_l\lambda_l^2) \to 0$ as $P \to \infty$

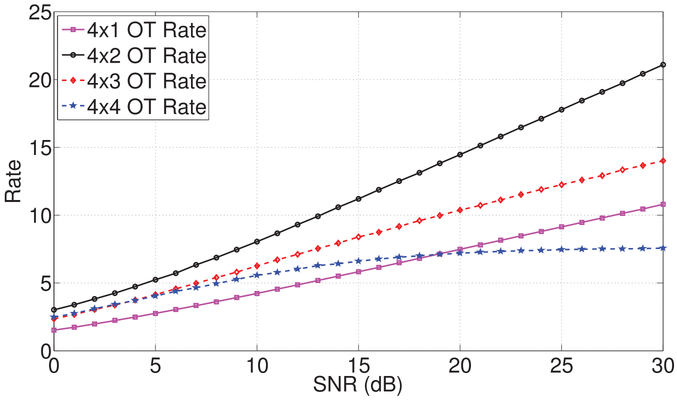

Fig. 9. OT Rates for MIMO with $n_A = 4$ transmit antennas, and $n_B = 1, 2, 3, 4$ receive antennas.

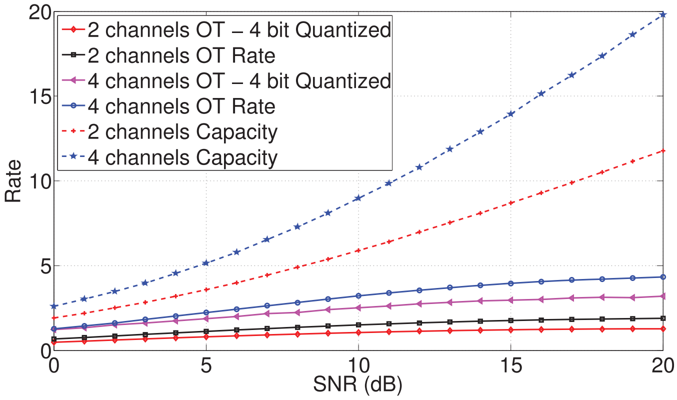

Fig. 10. OT Rate and OFDM capacity versus SNR for 2, 4 Channels OFDM.

channel at SNR P dB is approximately equal to the capacity of 2×1 MIMO channel with CSIT at 3 dB lower transmit power. This is due to the fact that in our OT protocol, half of the power is given to the null-space of $\mathbf{H}$ which is useless for communication. The average OT rate of 2×1 MIMO channel increases at the rate of 1 bit/3dB, as $\mu_{OT} = 1$.

Using (24) we see that at very high SNR, the average OT rate for 2×2 MIMO system is given by $\bar{R} \approx E\left[\log\left(\frac{\lambda_0^2}{\lambda_1^2}\right)\right]$. Recall that λ_0^2, λ_1^2 are the eigenvalues of the Wishart matrix $\mathbf{H}\mathbf{H}^{\dagger}$. The joint p.d.f. of the ordered eigenvalues, $\gamma_0 = \lambda_0^2, \gamma_1 = \lambda_1^2$, is given by $e^{-(\gamma_0+\gamma_1)}(\gamma_0 - \gamma_1)^2$ [30, Theorem 2.17]. The asymptotic value of the average OT rate is thus

$$E\left[\log\left(\frac{\gamma_0}{\gamma_1}\right)\right] = \int_0^{\infty}\int_0^{\gamma_0} \log\left(\frac{\gamma_0}{\gamma_1}\right) e^{-(\gamma_0+\gamma_1)}(\gamma_0 - \gamma_1)^2 d\gamma_1 d\gamma_0$$
$$= 1 + 2\ln(2) \text{ nats} \approx 3.45 \text{ bits}.$$

In Fig. 9, the average OT rates for MIMO with $n_A = 4$ and $1 \le n_B \le 4$ are shown as a function of SNR. As expected from Section VII, the best average OT rate is achieved when $n_B = n_A/2 = 2$, with asymptotic slope of 2 bits/3dB ($\mu_{OT} = 2$). The asymptotic slope for $n_B = 1$ and $n_B = 3$ is 1 bit/3dB ($\mu_{OT} = 1$). For $n_B \ge 4$, $\mu_{OT} = 0$, and the rate is bounded.

In Fig. 10, we show the average OT rate for 2-channels OFDM and 4-channels OFDM, along with the capacities of the corresponding channels. The average OT rate of 2-channel OFDM converges to a constant as SNR increases, since $\mu_{OT} = 0$. To find this constant, we note that $|H_0|$ and $|H_1|$ are i.i.d. with Rayleigh distribution. So $|H_0|^2$ and $|H_1|^2$ have exponential distribution. Let $S = \max(|H_0|^2, |H_1|^2)$ and $T = \min(|H_0|^2, |H_1|^2)$. Then the probability density functions of S and T are $2(1 - e^{-s})e^{-s}$ and $2e^{-2t}$ respectively. As SNR increases, the average OT rate for our protocol converges to

$$E[\log(S/T)] = \int_0^{\infty}\int_0^{\infty} \log(s/t)2(1 - e^{-s})e^{-s}2e^{-2t} ds dt$$
$$= 2\ln(2) \text{ nats} = 2 \text{ bits}.$$

The average OT rate of 4-channels OFDM also converges to a constant and $\mu_{OT} = 0$. The figure also shows the average OT rates achieved for 2 and 4 channels OFDM, when the fedback channel gains are uniformly quantized with 4 bits in the range 0-3.72 (where the cdf value is 0.999). The OT rate plots for 8-bits quantization were observed to be indistinguishable from those without quantization.

IX. Discussion

In AWGN channels, the noise realization is used to perform OT in [11], [28]. Following similar principle, the noise realization can potentially be further utilized in our setup to achieve better rate. In particular, for a single point-to-point fading channel or for parallel fading channels with the same fading coefficient, an obvious scheme is for Bob to first reveal the channel state to Alice over the noise-free channel. Then they can follow a protocol suitable for the resulting AWGN channel. However, as pointed out in [28], the OT rate saturates to a constant as $P \to \infty$ in AWGN channels. Thus further utilization of the noise realization in our protocol will not only result in a much more complex protocol, but it will also not provide any additional asymptotic OT-multiplexing gain.

With an odd number of OFDM channels, or an odd number of transmit antennas in a MIMO system, we have an odd number of parallel channels. In such a case, our protocol will leave one channel of middle rank in strength unused. That channel-state can be revealed to Alice by Bob, and the OT protocol of [28] can be used in the resulting AWGN channel. This also does not give any asymptotic ($P \to \infty$) improvement in terms of multiplexing gain.

In the presence of a more practical noise-free *finite rate* channel instead of a noise-free real channel, our OT protocols can be used with some modifications. For parallel fading channels, directly applicable to OFDM setup, this can be done as below. We illustrate this for a pair of parallel fading channels. The pair of fading coefficients can be quantized before sending them to Alice. Let $\lceil h \rceil$ and $\lfloor h \rfloor$ denote the q-bits quantization reconstruction levels above and below a fading magnitude h respectively. For an ordered pair of channel magnitudes (λ_0, λ_1) with $\lambda_0 \ge \lambda_1$, Bob sends $(\lfloor \lambda_0 \rfloor, \lceil \lambda_1 \rceil)$. This requires the transmission of $2q$ bits over the noise-free channel. Alice decides the OT rate assuming these to be the true ordered states. If $\lfloor \lambda_0 \rfloor \le \lambda_1 \le \lambda_0 \le \lceil \lambda_1 \rceil$, then Alice sees that the first quantized state is not larger than the second, and thus the OT rate is taken as zero by her in that block. The OT rate achieved for a pair of channels using this scheme is given by

$$R = \left(\log\left(1 + \frac{\lfloor\lambda_0\rfloor^2 P}{2}\right) - \log\left(1 + \frac{\lceil\lambda_1\rceil^2 P}{2}\right)\right)^+$$

where $(x)^+ := \max\{x, 0\}$. In simulations, we observed that an 8-bit quantized feedback in the range 0-3.72 provided almost no difference in average OT rate from the unquantized feedback. The rates for 4-bits quantization are plotted in Fig. 10.

For a MIMO channel, the precoding matrix needs to be also sent in a quantized form if the noise-free channel has finite rate. This will not result in an exact set of parallel SISO channels, and will leave some amount of mixing between the channels. For such channels, there will be rate loss due to several issues. First, like in OFDM, there will be a rate loss due to inaccurate estimate of the channels available for Alice. Second, there will be a rate loss due to inter-channel interference. Lastly, the leakage of information encoded in one channel into other channels need to be handled using appropriate additional encoding, resulting in additional rate loss. Designing OT protocols over such channels is outside the scope of this paper, and will be considered in future work.

X. Conclusion

We presented a technique for OT for honest-but-curious parties over parallel fading AWGN channels with receiver CSI with application to OFDM and MIMO. For privacy of Bob against Alice, our techniques use primarily Bob's exclusive knowledge of the fading states, whereas the additive noise is utilized for privacy of Alice against Bob. Altogether, the technique proposed in this paper can be an important tool for performing OT efficiently over wireless channels.

Acknowledgments

The authors are grateful to the Associate Editor and the anonymous referees for their constructive comments. In addition to many other technical comments which helped improve the manuscript, Sec. V was added due to a comment of the referees.

References

[1] R. Ahlswede and I. Csiszar, "On oblivious transfer capacity," in *Information Theory, Combinatorics and Search Theory*. Berlin, Heidelberg: Springer, 2013, pp. 145–166.

[2] A. C. A. Nascimento and A. Winter, "On the oblivious-transfer capacity of noisy resources," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2572–2581, Jun. 2008.

[3] M. Rabin, "How to exchange secrets by oblivious transfer," Aiken Comput. Lab., Harvard Univ., Cambridge, MA, USA, Tech. Memo TR-81, 1981.

[4] C. Crépeau, "Efficient cryptographic protocols based on noisy channels," in *Proc. EUROCRYPT*, 1997, vol. 1233, pp. 306–317.

[5] J. Kilian, "Founding cryptography on oblivious transfer," in *Proc. 20th Symp. Theory Comput.*, 1988, pp. 20–31.

[6] A. C. A. Nascimento and A. Winter, "On the oblivious transfer capacity of noisy correlations," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, USA, Jul. 2006, pp. 1871–1875.

[7] A. Winter, A. C. A. Nascimento, and H. Imai, "Commitment capacity of discrete memoryless channels," in *Proc. IMA Int. Conf. Cryptogr. Coding*, Cirencester, U.K., Dec. 2003, pp. 35–51.

[8] M. Mishra, B. K. Dey, V. M. Prabhakaran, and S. Diggavi, "The oblivious transfer capacity of the wiretapped binary erasure channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 1539–1543.

[9] M. Mishra, B. K. Dey, V. M. Prabhakaran, and S. Diggavi, "On the oblivious transfer capacity region of the binary erasure broadcast channel," in *Proc. IEEE Inf. Theory Workshop*, Hobart, TAS, Australia, Nov. 2014, pp. 237–241.

[10] A. C. Pinto, R. Dowsley, K. Morozov, and A. C. A. Nascimento, "Achieving oblivious transfer capacity of generalized erasure channels in the malicious model," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 5566–5571, Aug. 2011.

[11] M. Isaka, "On unconditionally secure oblivious transfer from continuous channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 2617–2621.

[12] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[13] M. Bloch and J. Barros, *Physical Layer Security From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, Oct. 2011.

[14] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, pp. 1355–1387, Oct. 1975.

[15] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[16] V. Y. F. Tan and M. R. Bloch, "Information spectrum approach to strong converse theorems for degraded wiretap channels," in *Proc. Allerton Conf. Commun. Control Comput.*, Oct. 2014, pp. 747–754.

[17] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

[18] Z. Rezki, A. Khisti, and M. S. Alouini, "On the secrecy capacity of the wiretap channel with imperfect main channel estimation," *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3652–3664, Oct. 2014.

[19] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008.

[20] S. K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "Precoding by pairing subchannels to increase MIMO capacity with discrete input alphabets," *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4156–4169, Jul. 2011.

[21] S. K. Mohammed, E. Viterbo, Y. Hong, and A. Chockalingam, "MIMO precoding with X- and Y-codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3542–3566, Jun. 2011.

[22] J. Barros and M. Bloch, "Strong secrecy for wireless channels," in *Proc. Int. Conf. Inf. Theor. Sec. (ICITS)*, 2008, pp. 40–53.

[23] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes,"*IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.

[24] C. Ling, L. Luzzi, J. C. Belfiore, and D. Stehlé, "Semantically secure lattice codes for the Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. 60, no. 10, pp. 6399–6416, Oct. 2014.

[25] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[26] M. Fiedler, "Bounds for the determinant of the sum of hermitian matrices," *Proc. Amer. Math. Soc.*, vol. 30, no. 1, pp. 27–31, Sep. 1971.

[27] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[28] M. Isaka, "Unconditionally secure oblivious transfer from algebraic signaling over the Gaussian channel," *IEICE Trans. Fundam.*, vol. E93-A, no. 11, pp. 2017–2025, Nov. 2010.

[29] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, 1999.

[30] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*. Delft, The Netherlands: now Publishers, 2004.

Jithin Ravi received the B.Tech. degree in electronics and communication engineering from S.C.T. College of Engineering, Thiruvananthapuram, India, in 2006. From 2006 to 2008, he worked as a software enginner at Infosys Technologies Ltd. He has been pursuing the M.Tech.+Ph.D. dual degree in electrical engineering at Indian Institute of Technology Bombay, since 2009. His research interests include information theory, coding theory, network coding, and wireless communications.

Bikash Kumar Dey (S'00–M'04) received the B.E. degree in electronics and telecommunication engineering from Bengal Engineering College, Howrah, India, in 1996, the M.E. degree in signal processing from the Indian Institute of Science, Karnataka, India, in 1999, and the Ph.D. degree in electrical communication engineering from the Indian Institute of Science, in 2003. From August 1996 to June 1997, he worked with the Wipro Infotech Global R&D. In February 2003, he joined Hellosoft India Pvt. Ltd., Hyderabad, India, as a Technical Member. In June 2003, he joined the International Institute of Information Technology, Hyderabad, India, as Assistant Professor. In May 2005, he joined the Department of Electrical Engineering of Indian Institute of Technology Bombay, Mumbai, India, where he is currently a Professor. His research interests include information theory, coding theory, and wireless communications. Dr. Dey was awarded the Prof. I.S.N. Murthy Medal from IISc as the best M.E. student in the Department of Electrical Communication Engineering and Electrical Engineering for 1998-1999 and Alumni Medal for the best Ph.D. thesis in the Division of Electrical Sciences for 2003-2004.

Emanuele Viterbo (M'95–SM'04–F'11) was born in Torino, Italy, in 1966. He received the degree (Laurea) and Ph.D. degree in electrical engineering from the Politecnico di Torino, Torino, Italy, in 1989 and in 1995. From 1990 to 1992, he was with the European Patent Office, The Hague, The Netherlands, as a Patent Examiner in the field of dynamic recording and error-control coding. Between 1995 and 1997, he held a post-doctoral position in the Dipartimento di Elettronica of the Politecnico di Torino, Torino, Italy, in Communications Techniques over Fading Channels. He became Associate Professor with the Politecnico di Torino, Dipartimento di Elettronica in 2005 and a Full Professor in DEIS at Università della Calabria, Italy, in 2006. Since 2010, he is a Full Professor with the Department of Electrical and Computer Systems Engineering, Monash University, Clayton, Australia, and the Associate Dean Graduate Research of the Faculty of Engineering at Monash University. In 1993, he was Visiting Researcher in the Communications Department of DLR, Oberpfaffenhofen, Germany. In 1994 and 1995, he was Visiting the E.N.S.T., Paris. In 1998, he was Visiting Researcher in the Information Sciences Research Center of AT\&T Research, Florham Park, NJ, USA. In 2003, he was Visiting Researcher with the Maths Department of EPFL, Lausanne, Switzerland. In 2004, he was Visiting Researcher with the Telecommunications Department of UNICAMP, Campinas, Brazil. In 2005, he was Visiting Researcher with the ITR of UniSA, Adelaide, Australia. He was Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY, European Transactions on Telecommunications and Journal of Communications and Networks; and is now an Editor of Foundations and Trends in Communications and Information Theory. His main research interests are in lattice codes for the Gaussian and fading channels, algebraic coding theory, algebraic space-time coding, digital terrestrial television broadcasting, and digital magnetic recording. Dr. Emanuele Viterbo was awarded a NATO Advanced Fellowship in 1997 from the Italian National Research Council.